Shahab Sahraee
Peter Wriggers

Tensor Calculus and Differential Geometry for Engineers

With Solved Exercises

Springer

Tensor Calculus and Differential Geometry for Engineers

Shahab Sahraee · Peter Wriggers

Tensor Calculus and Differential Geometry for Engineers

With Solved Exercises

Springer

Shahab Sahraee
Hannover, Germany

Peter Wriggers
Institut für Kontinuumsmechanik
Leibniz University Hannover
Garbsen, Germany

ISBN 978-3-031-33952-3 ISBN 978-3-031-33953-0 (eBook)
https://doi.org/10.1007/978-3-031-33953-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Tensors enable one to formulate mathematical ideas in the most elegant way. They are used to express physical phenomena mathematically. The concept of tensor dates back to the pioneering works of Bernhard Riemann (1826–1866) and Elwin Bruno Christoffel (1829–1900) on the development of modern differential geometry. And tensor calculus, also known as Ricci calculus, was established by Gregorio Ricci-Curbastro (1853–1925) and his student Tullio Levi-Civita (1873–1941). Tensor calculus has nowadays proven to be a universal language among many scientists and engineers. It manifests itself as the mathematical underpinning in many branches of physics and engineering such as electromagnetism and continuum mechanics. The most important twentieth-century development in science, i.e. Einstein's general theory of relativity, could not be developed without the language of tensor calculus. And no one can work on theoretical as well as computational mechanics without the concept of tensor.

This book offers a comprehensive treatment of tensor algebra and calculus (in both Cartesian and curvilinear coordinates) and also provides a comprehensive introduction to differential geometry of surfaces and curves (only in its local aspect). This book contains three parts. The first part of this text (Chaps. 1–5) deals with tensor algebra. Chapter 1 briefly discusses some fundamentals of vector algebra, for the sake of self-containedness. Although the rules and identities introduced here are extensively used in later chapters, the readers that are already familiar with algebra of vectors may preferably skip this chapter. Chapter 2 introduces algebra of second-order tensors. It will be shown that how these mathematical entities, that aim at describing linear relation between vectors, are constructed from vectors in a certain way. Chapter 3 presents an introduction to algebra of higher-order tensors. Similar discussions will be followed in this chapter. Here, higher-order tensors are defined by appealing to the notion of linear mapping and their important relationships are characterized. The main focus will be on the fourth-order tensors which are extensively used, for instance, in general relativity and nonlinear solid mechanics. Chapter 4 offers an introduction to eigenvalues and eigenvectors of second-order tensors due to their great importance in scientific and engineering problems. Spectral decompositions and eigenprojections of tensors are also studied. This requires some important theorems which have also been introduced. Chapter 5 deals with representation of tensorial variables in general class of curvilinear coordinates. In the previous chapters, tensorial variables have been expressed with respect to only a Cartesian coordinate frame. But, the goal here is to work with the old art of curvilinear coordinates which enables one to represent the most general form of tensorial variables and address their fundamental properties in a general framework. The reader should now be ready to start by studying tensor calculus.

The second part of this text (Chaps. 6–8) deals with the treatment of tensor calculus. Chapter 6 studies differentiation of tensor functions and representation theorems. It contains two sections. In this first section, the fundamental rules of differentiation for tensor functions are introduced. Their gradients are then represented by means of a first-order Taylor series expansion. Some analytical derivatives are finally approximated by means of finite difference method. In the second section, some recent developments in representation theorems are studied. Chapter 7 deals with the gradient of tensor fields and its related operators. The main goal here is to study the actions of gradient, divergence and curl operators on vectors and tensors. Needless to say that these differential operators are the workhorses of vector and tensor calculus. Chapter 8 introduces integral theorems and differential forms. This chapter contains two sections. The first section deals with the well-known integral theorems of Gauss and Stokes. This finishes our discussion of vector and tensor calculus started from Chap. 6. The second section aims at introducing what are known as differential forms which are used, for instance, in electromagnetic field theory. The main goal here is to introduce the so-called generalized Stokes' theorem which unifies the four fundamental theorems of calculus.

The last part of this work (Chap. 9) deals with differential geometry. It only contains one chapter which offers an introduction to differential geometry of embedded surfaces and curves in the three-dimensional Euclidean space. Chapter 9 studies the local theory of embedded two-dimensional surfaces as well as one-dimensional curves in the three-dimensional Euclidean flat space and introduces their fundamental properties. Different concepts of curvature (i.e. intrinsic curvature and extrinsic curvature) are thoroughly discussed. Three crucially important differential operators in this context (i.e. surface covariant differentiation, Lie derivative and invariant time differentiation) are also introduced. At the end of this chapter, the well-known application of the surface theory in structural mechanics (i.e. shell structures) is provided.

The authors deeply believe that the theoretical material cannot be well understood unless the reader solves a lot of exercises. We fully understand the needs and concerns of students. Numerous solved problems are thus provided at the end of each chapter for in-depth learning. Moreover, all identities in this book are verified line by line for the reader's convenience. We have taken pain to show the derivations in this elegant form. And this significantly saves the book space. Finally, for the sake of fast learning, each figure in this text is accompanied usually by a text which summarizes the corresponding theoretical arguments.

This book is intended for advanced mathematically minded students in engineering fields and research engineers. It is also addressed to advanced undergraduates and beginning graduate students of applied mathematics and physics. To read this text, one needs to have a background in linear algebra and calculus of several variables. However, a few exercises rely on some basic concepts of solid mechanics. With regard to this, a first course in continuum mechanics should suffice for the interested reader.

Shahab Sahraee is grateful to the Department of Mechanical Engineering at University of Guilan for having provided a pleasant environment during the writing of this book. His greatest debt goes to Farzam Dadgar-Rad for reviewing all chapters of the manuscript and many fruitful discussions on representation theorems. His valuable comments on each chapter of the manuscript are really appreciated. Shahab Sahraee would also like to thank all students who helped him plot the figures with PGF/TikZ in LaTeX. Finally, his special thanks go to his sisters and brother for their continuous support, patience and understanding during the development of this book.

Last but not least, we express our thanks to Springer Verlag for accepting to publish this book.

Rasht, Iran Shahab Sahraee
Hannover, Germany Peter Wriggers
March 2023

Contents

1 Algebra of Vectors 1
1.1 Three-Dimensional Vector Space 3
1.2 Basis Vectors 5
1.3 Inner Product 5
1.3.1 Vector Projection and Rejection 8
1.4 Complexification 9
1.5 Three-Dimensional Euclidean Point Space 10
1.6 Vector Representation in Cartesian Coordinates 11
1.7 Indicial Notation and Summation Convention 12
1.8 Matrix Notation 15
1.9 Cross Product 17
1.10 Exercises 23
References 33

2 Algebra of Tensors 35
2.1 Tensor as a Linear Transformation 35
2.2 Tensor Product and Representation 38
2.3 Tensor Operations 42
2.3.1 Composition 42
2.3.2 Transposition 46
2.3.3 Decomposition into Symmetric and Skew-Symmetric Parts 47
2.3.4 Contraction and Trace 50
2.4 Tensor in Matrix Notation 53
2.5 Determinant, Inverse and Cofactor of a Tensor 55
2.5.1 Determinant of a Tensor 55
2.5.2 Inverse of a Tensor 56
2.5.3 Cofactor of a Tensor 58
2.6 Positive Definite and Negative Definite Tensors 60
2.7 Orthogonal Tensor 62
2.8 Projection, Spherical and Deviatoric Tensors 65

2.9 Transformation Laws 67
2.9.1 Change of Cartesian Basis 69
2.9.2 Change of Cartesian Components 71
2.9.3 Isotropic Vectors and Tensors 72
2.10 Exercises 74
References 90

3 Algebra of Higher-Order Tensors 93
3.1 Tensors of Order Three 93
3.2 Tensors of Order Four 98
3.2.1 Construction and Representation of Fourth-Order Tensors 101
3.2.2 Operations with Fourth-Order Tensors 108
3.2.3 Major and Minor Symmetries of Fourth-Order Tensors 115
3.2.4 Fourth-Order Tensor in Matrix Notation 117
3.2.5 Determinant and Inverse of Fourth-Order Tensors 120
3.2.6 Positive Definite and Negative Definite Fourth-Order Tensors 122
3.3 Exercises 122
References 130

4 Eigenvalues, Eigenvectors and Spectral Decompositions of Tensors 131
4.1 Eigenvalue Problem 131
4.2 Characteristic Equation and Principal Scalar Invariants 133
4.3 Cayley-Hamilton Equation 139
4.4 Spectral Decomposition 140
4.5 Eigenprojections of a Tensor 145
4.6 Spectral Decomposition of Symmetric Tensors 147
4.7 Exercises 151
References 167

5 Representation of Tensorial Variables in Curvilinear Coordinates 169
5.1 Basis Vectors 169
5.1.1 Covariant Basis Vectors 170
5.1.2 Contravariant Basis Vectors 174
5.2 Metric Coefficients 177
5.2.1 Covariant Metric Coefficients 178
5.2.2 Contravariant Metric Coefficients 179
5.3 Tensor Property of Basis Vectors and Metric Coefficients 181
5.4 Contravariant and Covariant Components of First-Order Tensors 183
5.4.1 Dot Product and Cross Product Between Two Vectors 184

5.4.2 Matrix Notation 185
5.5 Contravariant, Mixed and Covariant Components of Second-Order Tensors 186
5.5.1 Basic Tensor Relationships 189
5.5.2 Matrix Notation 191
5.6 Contravariant, Mixed and Covariant Components of Higher-Order Tensors 193
5.7 Tensor Property of Components 195
5.8 Line, Surface and Volume Elements 196
5.9 Exercises 199
References 207

6 Differentiation of Tensor Functions and Representation Theorems 209
6.1 Tensor Functions and Their Derivatives 209
6.1.1 Gradient of a Scalar-Valued Function 212
6.1.2 Gradient of a Vector-Valued Function 218
6.1.3 Gradient of a Tensor-Valued Function 221
6.1.4 Proof of Tensor Calculus Identities by Index Notation 228
6.1.5 Numerical Differentiation 230
6.2 Representation Theorems 233
6.2.1 Mathematical Preliminaries 234
6.2.2 Representation Theorem for an Isotropic Scalar-Valued Function of a Vector 242
6.2.3 Representation Theorem for an Isotropic Scalar-Valued Function of a Symmetric Tensor 243
6.2.4 Representation Theorem for an Isotropic Vector-Valued Function of a Vector 248
6.2.5 Representation Theorem for an Isotropic Vector-Valued Function of a Symmetric Tensor 250
6.2.6 Representation Theorem for an Isotropic Symmetric Tensor-Valued Function of a Vector 252
6.2.7 Representation Theorem for an Isotropic Symmetric Tensor-Valued Function of a Symmetric Tensor 254
6.3 Exercises 259
References 311

7 Gradient and Related Operators 315
7.1 Differentiation of Fields 315
7.1.1 Christoffel Symbols of First and Second Kind 317
7.1.2 First Covariant Differentiation 321
7.1.3 Invariance of Covariant Derivative 322
7.1.4 Ricci's Theorem 324
7.1.5 Second Covariant Differentiation 326

7.1.6 Riemann-Christoffel Curvature Tensor 328
7.2 Gradient of Scalar, Vector and Tensor Fields 332
7.3 Divergence of Vector and Tensor Fields 336
7.4 Curl of Vector and Tensor Fields 339
7.5 Laplacian and Hessian of Scalar, Vector and Tensor Fields 342
7.6 Exercises 346
References 364

8 Integral Theorems and Differential Forms 365
8.1 Integral Theorems 365
8.1.1 Divergence Theorem 365
8.1.2 Stokes' Theorem 370
8.2 Differential Forms 374
8.2.1 Differential 1-Form 375
8.2.2 Differential 2-Form 377
8.2.3 Differential 3-Form 380
8.2.4 Hodge Star Operator 383
8.2.5 Exterior Derivatives 387
8.2.6 Hodge Star Operator in Exterior Calculus 392
8.2.7 Differential Forms Integration 400
8.2.8 Generalized Stokes' Theorem 405
8.3 Exercises 409
References 431

9 Differential Geometry of Surfaces and Curves 433
9.1 Representation of Tensorial Variables on Surfaces 436
9.1.1 Surfaces in Three-Dimensional Euclidean Space 436
9.1.2 Surface Basis Vectors and Metric Coefficients 440
9.1.3 Tensor Property of Surface Metric Coefficients 445
9.1.4 Shift Tensors and Related Identities 446
9.1.5 Tensor Property of Shift Tensors 452
9.1.6 Surface Vectors and Tensors 453
9.2 Gauss and Weingarten Formulas 455
9.2.1 Surface Christoffel Symbols of Second Kind 455
9.2.2 Surface Second-Order Curvature Tensor 456
9.2.3 Mean and Gaussian Curvatures 458
9.2.4 Surface Christoffel Symbols of First Kind and Levi-Civita Connection 459
9.3 Surface Covariant Derivative 462
9.3.1 Surface First Covariant Differentiation 462
9.3.2 Surface Second Covariant Differentiation 473
9.3.3 Invariance of Surface Covariant Differentiation 474
9.3.4 Surface Covariant Differentiation of Invariant Objects with Surface and Ambient Indices 475
9.4 Surface Riemann-Christoffel Curvature Tensor 478

9.5 Fundamental Forms of Surfaces ... 486
9.5.1 First Fundamental Form ... 486
9.5.2 Second Fundamental Form ... 491
9.5.3 Third Fundamental Form ... 495
9.5.4 Relation Between First, Second and Third Fundamental Forms ... 495
9.6 Embedded Curves in Three-Dimensional Euclidean Space and Ruled Surfaces ... 497
9.6.1 Mathematical Description of Space Curves ... 497
9.6.2 Curvature and Torsion of Space Curves ... 500
9.6.3 Frenet Formulas for Space Curves ... 506
9.6.4 Ruled Surfaces ... 508
9.7 Curvature ... 515
9.7.1 Curvature of Surface Curve ... 515
9.7.2 Geodesics ... 524
9.7.3 Principal Curvatures ... 537
9.7.4 Lines of Curvature and Geodesic Torsion ... 545
9.7.5 Gaussian Curvature ... 549
9.7.6 Ricci Curvature Tensor and Scalar ... 557
9.8 Lie Derivatives ... 564
9.8.1 Lie Derivative of Vector Fields ... 565
9.8.2 Lie Derivative of Covector Fields ... 572
9.8.3 Space Symmetry and Killing Vector ... 573
9.8.4 Lie Derivative of Higher-Order Tensor Fields ... 575
9.8.5 Lie Derivative of Differential Forms ... 580
9.8.6 Commutator ... 583
9.9 Calculus of Moving Surfaces ... 585
9.9.1 Mathematical Description of Moving Surfaces ... 586
9.9.2 Geometric Approach to Invariant Time Derivative ... 589
9.9.3 Invariant Time Derivative of Objects with Surface Indices ... 593
9.9.4 Invariant Time Differentiation of Objects with Ambient Indices ... 599
9.9.5 Invariant Time Differentiation of Surface Mixed Curvature Tensor and Its Principal Invariants ... 602
9.10 Application to Shell Structures ... 604
9.10.1 Shell Geometry ... 604
9.10.2 External and Internal Forces and Moments ... 607
9.10.3 Equilibrium Equations ... 609
9.10.4 Basic Shell Mathematical Model ... 611
9.11 Exercises ... 613
References ... 657

Index ... 661

List of Figures

Fig. 1.1 Scalar multiplication and vector addition 4
Fig. 1.2 Projection (rejection) of a vector **u** along (from) a unit vector $\hat{\mathbf{e}}$ 8
Fig. 1.3 A vector **u** with its rectangular components u_1, u_2 and u_3 in a right-handed Cartesian coordinate frame 12
Fig. 1.4 Cross product of **u** and **v** 18
Fig. 1.5 The scalar triple product of **u**, **v** and **w** with the resulting parallelepiped 22
Fig. 2.1 Orthogonal transformation 63
Fig. 2.2 Projection tenors 66
Fig. 2.3 Rotations of a right-handed Cartesian coordinate frame about its origin 69
Fig. 2.4 A counterclockwise rotation by an angle θ about the x_3-axis and a reflection across the x_2-axis in a right-handed Cartesian coordinate frame 82
Fig. 2.5 Finite rotation of a vector **u** in $\mathscr{E}_{\text{r}}^{\text{o}3}$ based on the Rodrigues rotation formula 82
Fig. 5.1 Curvilinear coordinate system 170
Fig. 5.2 Covariant basis $\{\mathbf{g}_i\}$ with its companion contravariant basis $\{\mathbf{g}^i\}$ 176
Fig. 5.3 Projection of a position increment vector in an arbitrary curvilinear coordinate system 197
Fig. 6.1 Reference, two intermediate and current configurations 292
Fig. 7.1 Gradient of a scalar field along with its unit vector at an arbitrary point of a level surface 334
Fig. 8.1 Domain V (with the volume element $\triangle V$) and its bounded closed surface A (with the differential surface element dA and associated unit normal field $\hat{\mathbf{n}}$) 366
Fig. 8.2 Open surface with positive orientation 371
Fig. 8.3 Area and signed area 379

Fig. 8.4 Gradient of a scalar function in the language of exterior calculus 395
Fig. 8.5 Divergence of a vector field in the language of exterior calculus 396
Fig. 8.6 Curl of a vector field in the language of exterior calculus 397
Fig. 8.7 Divergence of curl of a vector field in the language of exterior calculus 398
Fig. 8.8 Curl of gradient of a scalar field in the language of exterior calculus 399
Fig. 8.9 Laplacian of a scalar field in the language of exterior calculus 400
Fig. 8.10 Infinitesimal tetrahedron with edge lengths dx_1, dx_2, dx_3 and outwardly oriented unit vectors on its faces in a right-handed Cartesian coordinate frame 409
Fig. 9.1 Ellipsoid 437
Fig. 9.2 Sphere 437
Fig. 9.3 Torus 438
Fig. 9.4 Elliptic paraboloid 439
Fig. 9.5 Hyperboloid of one sheet 439
Fig. 9.6 Hyperboloid of two sheets 440
Fig. 9.7 Hyperbolic paraboloid 440
Fig. 9.8 Ambient basis at an arbitrary point of a surface encompassed by three-dimensional Euclidean space 441
Fig. 9.9 Surface area element 449
Fig. 9.10 Projection of the partial derivative $\partial \mathbf{a}_\alpha / \partial t^\beta$ onto the ambient basis $\{\mathbf{a}_1, \mathbf{a}_2, \hat{\mathbf{n}}\}$ 458
Fig. 9.11 Surface covariant derivative 462
Fig. 9.12 Parallel transport 463
Fig. 9.13 Parallel transport of vector field 467
Fig. 9.14 Surface Riemann-Christoffel curvature 479
Fig. 9.15 Surface curves $\boldsymbol{\gamma}$, $\bar{\boldsymbol{\gamma}}$ with corresponding tangent vectors $\mathbf{a}_t$, $\bar{\mathbf{a}}_t$ and surface patch Ω 489
Fig. 9.16 Local structure of surface 494
Fig. 9.17 Helix 498
Fig. 9.18 Tangent vector, principal normal vector, osculating plane and osculating circle of a smooth curve with natural parametrization 502
Fig. 9.19 Space curves with their local properties 506
Fig. 9.20 Ruled surface 508
Fig. 9.21 Two families of rulings on one-sheeted hyperboloid 509
Fig. 9.22 Generalized cylinder and generalized cone 512
Fig. 9.23 Tangent developable 512
Fig. 9.24 Normal section 519
Fig. 9.25 Normal and geodesic curvature vectors 520
Fig. 9.26 Surface of revolution 534

Fig. 9.27 Geodesics on torus and one-sheeted hyperboloid 537
Fig. 9.28 Surface local shape 541
Fig. 9.29 Parallel transport of a vector along two paths defined by that vector and a generic one 562
Fig. 9.30 Area change along geodesics on sphere 563
Fig. 9.31 Flow lines 566
Fig. 9.32 Lie derivative 566
Fig. 9.33 Space symmetry 573
Fig. 9.34 Commutator 585
Fig. 9.35 Invariant time derivative 589
Fig. 9.36 Infinitesimal shell element 605
Fig. 9.37 Catenoid 619
Fig. 9.38 Twisted cubic with its curvature and torsion 628
Fig. 9.39 Parallel transport along two (geodesic and non-geodesic) curves on unit sphere 634
Fig. 9.40 Geodesics on funnel surface 638
Fig. 9.41 Monkey saddle 641
Fig. 9.42 Catenoid and its parallel surfaces 643
Fig. 9.43 Deviation of the area of a small ball in curved space from that of standard ball in flat space 654

List of Tables

Table 6.1 List of isotropic **scalar** invariants 246
Table 6.2 List of isotropic **vector** invariants 250
Table 6.3 List of isotropic **symmetric tensor** invariants 258
Table 6.4 List of isotropic **skew-symmetric tensor** invariants 260

Chapter 1
Algebra of Vectors

Tensors are characterized by their orders. *Scalars* are zeroth-order tensors with only magnitude and no other characteristics. They are extensively used in physical relations and usually designated by small and capital lightface Latin or Greek letters; examples of which include mass, time and temperature. In this text, their magnitudes are elements of the field of real numbers R.

In contrast, *vectors* are physical quantities that are completely characterized by both magnitude and direction. These mathematical objects, that are first-order tensors, obey the parallelogram law of vector addition; examples of which include force, velocity and momentum. In some books or handwritten documents, they are usually represented by bar above or below a character or arrow over it in order to denote boldface setting. In this text, small (and rarely capital) boldface Latin letters are set to denote these quantities.

Vectors play an essential role in establishing mathematical frameworks in almost all branches of physics and engineering. For the sake of self-containedness, this chapter briefly represents some fundamentals of vector algebra. Although the rules, properties and identities introduced here are extensively used in later chapters, the readers that are already familiar with algebra of vectors may preferably skip this chapter and leave it as a reference when it is needed.

To describe mathematical entities, specific font styles need to be adopted. Here, it has been tried to keep the most standard notation in the literature. Throughout this textbook, unless otherwise stated, the following notation will be utilized:

✓ Lowercase (Uppercase) italic lightface Latin letters $a, b, \ldots$ ($A, B, \ldots$) denote scalars or scalar-valued functions.
✓ Lowercase and uppercase lightface Greek letters α, β, ..., $\varPsi$, $\varOmega$ symbolize scalars or scalar-valued functions.
✓ Lowercase boldface Latin letters $\mathbf{a}$, $\mathbf{b}$, ... stand for vectors or vector-valued functions. The Cartesian components of these objects, identified by indices, are shown in italic lightface form as a_i, b_j, ...

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_1

Hint: Some exceptions should be highlighted here. First, the letters $\mathbf{o}$, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ are set aside for points in three-dimensional Euclidean space. Then, some spatial second-order tensors are represented by lowercase boldface Latin letters. Some examples will be the Finger tensor $\mathbf{b}$, the left stretch tensor $\mathbf{v}$ and the surface covariant curvature tensor $\underline{b}_{\alpha\beta}$.

- ✓ Uppercase boldface Latin letters $\mathbf{A}$, $\mathbf{B}$, ... represent second-order tensors or (second-order) tensor-valued functions. And their Cartesian components, with two indices, are written in italic lightface form as A_{ij}, B_{kl}, ...
  Hint: The letters $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are reserved for points in three-dimensional Euclidean space.
- ✓ Lowercase boldface Greek letters $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ with the Cartesian components σ_{ij} and ε_{kl} present second-order tensors or (second-order) tensor-valued functions.
- ✓ Uppercase boldface Latin text symbols $\textsf{\textbf{A}}$, $\textsf{\textbf{B}}$, ... indicate third-order tensors. And their components with respect to the Cartesian basis vectors are denoted by A_{ijk}, B_{mno}, ...
- ✓ Lowercase (uppercase) blackboard Latin letters $\mathbb{a}$, $\mathbb{c}$, ... ($\mathbb{A}$, $\mathbb{C}$, ...) with the Cartesian components $\mathbb{a}_{ijkl}$, $\mathbb{c}_{mnop}$, ... ($\mathbb{A}_{ijkl}$, $\mathbb{C}_{mnop}$, ...) designate fourth-order tensors.
- ✓ Uppercase lightface calligraphic Latin letters $\mathscr{A}$, $\mathscr{B}$, ... exhibit spaces. Moreover, R (C) renders the field of real (complex) numbers and $\mathcal{S}$ ($\mathcal{C}$) demonstrates the set of points constructing a two-dimensional surface (one-dimensional curve) embedded in the three-dimensional Euclidean space.

Note that the components of a tensor of an arbitrary order in the Cartesian coordinate system were denoted by Latin and Greek lightface letters. In the literature, the covariant form of components in curvilinear coordinates are also usually represented in this way. To avoid any confliction and keep the notation as simple as possible, throughout this text, uppercase and lowercase lightface Latin and Greek letters with **underline** beneath and

- ✔ subindices denote the covariant form of components, e.g., $\underline{C}_{AB}$ and $\underline{b}_{\alpha\gamma}$.
- ✔ superindices stand for the contravariant components of a tensor, e.g., $\underline{u}^{i}$ and $\underline{\tau}^{kl}$.
- ✔ mixed subindices and superindices represent the co-contravariant and contra-covariant components, e.g., $\underline{b}_{\alpha}^{\cdot\beta}$ and $\underline{F}^{i}_{\cdot A}$.

In this text, almost all subindices and superindices which identify the components of a general tensor are designated by lowercase lightface Latin and Greek letters. But, only at the end of Chap. 6 one needs to distinguish between different configurations in the context of nonlinear continuum mechanics wherein small as well as capital lightface Latin letters have been used.

Consider the following equations

$$\mathbf{q} = \alpha\mathbf{r} = \beta\mathbf{s} = \gamma\,\mathbf{t}\ . \tag{1.1}$$

The first expression, i.e. $\mathbf{q} = \alpha \mathbf{r}$, will be referred to by $(1.1)_1$ and accordingly $(1.1)_2$ refers to $\mathbf{q} = \beta \mathbf{s}$. Now, suppose one is given

$$\text{curl}\hat{\mathbf{h}} = \underbrace{-J^{-1}\varepsilon^{ijk} g_{jl} \; \underline{\hat{h}}^l \Big|_k \mathbf{g}_i}_{= -g_{jl} \; \underline{\hat{h}}^l \big|_k \mathbf{g}^j \times \mathbf{g}^k} = \underbrace{-J^{-1}\varepsilon^{ijk} \; \underline{\hat{h}}_j \Big|_k \mathbf{g}_i}_{= - \; \underline{\hat{h}}_j \big|_k \mathbf{g}^j \times \mathbf{g}^k} \; . \tag{1.2}$$

The expression $(1.2)_2$ then renders $\text{curl}\hat{\mathbf{h}} = -g_{jl} \; \underline{\hat{h}}^l \Big|_k \mathbf{g}^j \times \mathbf{g}^k$ and $(1.2)_3$ refers to $\text{curl}\hat{\mathbf{h}} = -J^{-1}\varepsilon^{ijk} \; \underline{\hat{h}}_j \Big|_k \mathbf{g}_i$. At the end, consider

$$\underbrace{\text{curl}\mathbf{u} = \left(* d \overset{1}{\omega}_{\mathbf{u}} \right)^{\sharp}}_{\text{or } \overset{1}{\omega}_{\text{curl}\mathbf{u}} = (\text{curl}\mathbf{u})^{\flat} = * d \overset{1}{\omega}_{\mathbf{u}}} \; . \tag{1.3}$$

Here, $(1.3)_3$ means $\overset{1}{\omega}_{\text{curl}\mathbf{u}} = * d \overset{1}{\omega}_{\mathbf{u}}$.

There is a large amount of literature on tensor algebra and calculus. The reader here is referred to, for example, the books by Synge and Schild [1], McConnell [2], Sokolnikoff [3], Borisenko and Tarapov [4], Akivis and Goldberg [5], Flügge [6], Danielson [7], Simmonds [8], Heinbockel [9], Dimitrienko [10], Talpaert [11], Akivis and Goldberg [12], Lebedev and Cloud [13], Spain [14], Ruíz-Tolosa and Castillo [15], Bowen and Wang [16], Schade and Neemann [17], Schouten [18], Nayak [19], Itskov [20] and Irgens [21].

1.1 Three-Dimensional Vector Space

Let $\mathscr{V}^3$ be a set of directed line elements in the three-dimensional space. The set $\mathscr{V}^3$ is called a *vector space over the field* R if it remains closed with respect to the following mathematical operations[1]

- $\underbrace{\mathbf{u} + \mathbf{v}}_{\text{this is called } \textit{vector addition} \text{ or sum of } \mathbf{u} \text{ and } \mathbf{v}} \in \mathscr{V}^3, \quad \text{for any } \mathbf{u}\,,\ \mathbf{v} \in \mathscr{V}^3\,,$
- $\underbrace{\alpha \mathbf{u}}_{\text{this is called the } \textit{scalar multiplication} \text{ of } \mathbf{u} \text{ by } \alpha \text{ or the product of } \alpha \text{ and } \mathbf{u}} \in \mathscr{V}^3, \quad \text{for any } \alpha \in \text{R and } \mathbf{u} \in \mathscr{V}^3\,,$

such that the following rules hold for any $\mathbf{u}\,,\ \mathbf{v}\,,\ \mathbf{w} \in \mathscr{V}^3$ and $\alpha\,,\ \beta \in$ R:

---

[1] Although there is no limitation to consider finite-dimensional vector spaces, almost all vectors (and also higher-order tensors) in this text belong to the three-dimensional spaces. And it is assumed that all these objects are real. Note that in the literature, the three-dimensional vector space is sometimes denoted by R^3.

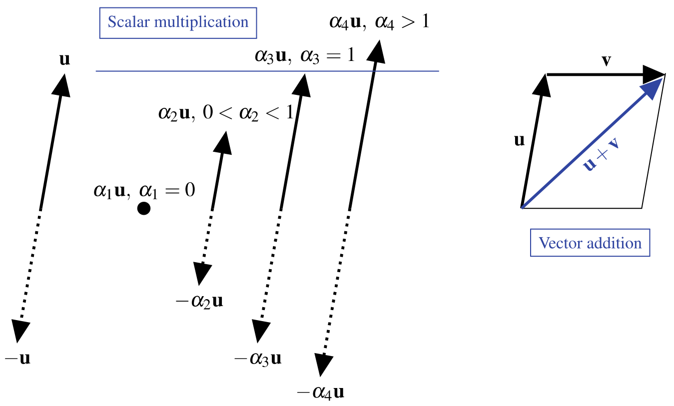


Fig. 1.1 Scalar multiplication and vector addition

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}\ , \quad \longleftarrow \textit{commutative}\ \text{property} \tag{1.4a}$$

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})\ , \quad \longleftarrow \textit{associative}\ \text{property} \tag{1.4b}$$

for any $\mathbf{u} \in \mathscr{V}^3$ there exists $-\mathbf{u} \in \mathscr{V}^3$ such that

$$\mathbf{u} + (-\mathbf{u}) = \mathbf{0}\ , \quad \longleftarrow \begin{array}{c}\text{here } -\mathbf{u} \text{ is said to be}\\ \text{the } \textit{additive inverse} \text{ of } \mathbf{u}\end{array} \tag{1.4c}$$

a unique **zero vector** with zero magnitude and unspecified direction exists in the set $\mathscr{V}^3$ such that

$$\mathbf{u} + \mathbf{0} = \mathbf{u}\ , \quad \longleftarrow \textit{additive identity}\ , \tag{1.4d}$$

$$1\mathbf{u} = \mathbf{u}\ , \quad \longleftarrow \begin{array}{c}\text{here 1 presents the}\\ \textit{multiplicative identity} \text{ in the field}\end{array} \tag{1.4e}$$

$$(\alpha\beta)\,\mathbf{u} = \alpha\,(\beta\mathbf{u})\ , \quad \longleftarrow \begin{array}{c}\text{associative property for scalar}\\ \text{multiplication of a vector by scalars}\end{array} \tag{1.4f}$$

$$(\alpha + \beta)\,\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}\ , \quad \longleftarrow \begin{array}{c}\text{scalar multiplication of a vector}\\ \text{by a field addition is } \textit{distributive}\end{array} \tag{1.4g}$$

$$\alpha\,(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}\ . \quad \longleftarrow \begin{array}{c}\textit{distributive} \text{ property also holds for scalar}\\ \text{multiplication of a vector addition by a scalar}\end{array} \tag{1.4h}$$

Note that **vector subtraction** is implied by vector addition and scalar multiplication. For subsequent developments, the difference between two vectors **u** and **v** is indicated by

$$\boxed{\mathbf{u} + (-\mathbf{v}) = \mathbf{u} - \mathbf{v}\ .} \quad \longleftarrow \text{see (f) in (1.76)} \tag{1.5}$$

See Fig. 1.1 for a geometrical interpretation of the scalar multiplication and vector addition.

1.2 Basis Vectors

A *basis* is basically an *independent spanning* set. A set is called *independent* if there is no redundant element in that set. For an independent set of three vectors in $\mathscr{V}^3$, this states that no vector can be built as a **linear combination** of the remaining ones, that is,

$$\text{if} \quad \underbrace{\alpha\mathbf{u} + \beta\mathbf{v} + \gamma\,\mathbf{w} = \mathbf{0}}_{\text{for any } \alpha\,,\ \beta\,,\ \gamma\,\in\,\mathrm{R} \text{ and any } \mathscr{G}\,=\,\{\mathbf{u}\,,\ \mathbf{v}\,,\ \mathbf{w}\}\,\subset\,\mathscr{V}^3} \quad \text{then} \quad \alpha = \beta = \gamma = 0\,. \tag{1.6}$$

The subset $\mathscr{G}$ of the vector space $\mathscr{V}^3$, written in the above expression, is said to be *linearly independent* when all scalars constructing that linear combination are zero. Otherwise, its is referred to as *linearly dependent* if such scalars are not all zero. In this case the subset may contain the zero vector.

A set *spans* if any element of the space (say for any $\mathbf{u} \in \mathscr{V}^3$ in the present context) can be represented as a linear combination of the elements in that set. It should be noted that a spanning set may have redundant terms which can be eliminated from the set but still spanning the space.

The notion of basis vectors can now be understood; either a collection of linearly independent vectors that spans a vector space or any spanning set for that vector space without any redundant vector.

Let $\mathscr{G}_b = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a basis for the three-dimensional vector space $\mathscr{V}^3$. Then, any vector $\mathbf{r} \in \mathscr{V}^3$ can be expressed as

$$\boxed{\mathbf{r} = \alpha\,\mathbf{u} + \beta\,\mathbf{v} + \gamma\,\mathbf{w}\,,} \tag{1.7}$$

where $\alpha\,,\ \beta\,,\ \gamma\,\in \mathrm{R}$ are called the *components* of $\mathbf{r}$ with respect to $\mathscr{G}_b$. These scalars can **uniquely** be determined and, therefore, the expression (1.7) renders a **unique representation** for the given vector.

The number of elements in a basis of space is termed the *dimension* of space. For instance, the dimension of $\mathscr{V}^3$, denoted by $\dim \mathscr{V}^3$, is 3. That is why the subset $\mathscr{G}$ in (1.6) contains three different nonzero vectors. The vector space $\mathscr{V}^3$ can have infinitely many bases but the dimension of each should be the same.

1.3 Inner Product

The vector space $\mathscr{V}^3$ basically has a linear structure, i.e. it admits the **vector addition** and **scalar multiplication**. This important feature was used to make the definition of $\mathscr{V}^3$. There are other important features such as the notions of *length* and *angle* that are set in a concept called *inner product* (also known as the *scalar product* or *dot product*). Its application in establishing vector (or tensor) algebraic as well as calculus identities is of crucial importance. The way that is defined and the fact that whether a vector space is equipped with the inner product or not greatly affect the properties of the vector space.

The inner product of two arbitrary vectors **u** and **v**; indicated by,

$$\boxed{g^{\mathrm{ip}}(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \ ,} \tag{1.8}$$

is a scalar-valued function from $\mathscr{V}^3 \times \mathscr{V}^3$ to R satisfying the following axioms

$$\underbrace{g^{\mathrm{ip}}(\mathbf{u}, \mathbf{v}) = g^{\mathrm{ip}}(\mathbf{v}, \mathbf{u})}_{\text{or } \mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} \text{ representing } \textit{commutative} \text{ or } \textit{symmetry} \text{ property}} \ , \tag{1.9a}$$

$$\underbrace{g^{\mathrm{ip}}(\mathbf{u} + \mathbf{v}, \mathbf{w}) = g^{\mathrm{ip}}(\mathbf{u}, \mathbf{w}) + g^{\mathrm{ip}}(\mathbf{v}, \mathbf{w})}_{\text{or } (\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w} \text{ rendering } \textit{distributive} \text{ property in the first argument}} \ , \tag{1.9b}$$

$$\underbrace{g^{\mathrm{ip}}(\alpha\mathbf{u}, \mathbf{v}) = \alpha g^{\mathrm{ip}}(\mathbf{u}, \mathbf{v})}_{\text{or } (\alpha\mathbf{u}) \cdot \mathbf{v} = \alpha (\mathbf{u} \cdot \mathbf{v}) \text{ presenting } \textit{associative} \text{ property in the first argument}} \ , \tag{1.9c}$$

$$\underbrace{\left.\begin{aligned} g^{\mathrm{ip}}(\mathbf{u}, \mathbf{u}) > 0 \quad \text{if } \mathbf{u} \neq \mathbf{0} \\ g^{\mathrm{ip}}(\mathbf{u}, \mathbf{u}) = 0 \quad \text{if } \mathbf{u} = \mathbf{0} \end{aligned}\right\}}_{\text{or } \mathbf{u} \cdot \mathbf{u} \geq 0, \text{ for which } \mathbf{u} \cdot \mathbf{u} = 0 \Longleftrightarrow \mathbf{u} = \mathbf{0}, \text{ expressing } \textit{positive-definite} \text{ property}} \ , \tag{1.9d}$$

for any $\alpha \in$ R and $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathscr{V}^3$. Note that the second and third rules are often unified according to

$$g^{\mathrm{ip}}(\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}) = \alpha g^{\mathrm{ip}}(\mathbf{u}, \mathbf{w}) + \beta g^{\mathrm{ip}}(\mathbf{v}, \mathbf{w}) \ , \tag{1.10}$$

for any $\alpha, \beta \in$ R and $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathscr{V}^3$. This is referred to as the *linearity property* in the first argument. But, the first rule then implies linearity in the second argument. Hence, the inner product is a *symmetric bilinear form* on the three-dimensional **real** vector space.[2]

Consider a vector space $\mathscr{V}^3$ equipped with the inner product satisfying the rules listed above. This is called *three-dimensional Euclidean vector space*, designated here by $\mathscr{E}_{\mathrm{r}}^3$.

The inner product is an algebraic operation which determines the **length** (or **magnitude**) of a vector according to

[2] The rules (1.9a)–(1.9d) are generally listed as conjugate symmetry, linearity in the first argument and positive-definiteness, although there is some disagreement on the second rule. Note that, in general, the inner product is not a symmetric bilinear form, see Sect. 1.4.

$$|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}} \ . \tag{1.11}$$

This is also known as the *Euclidean norm* (or simply *norm*) of $\mathbf{u}$.[3] Such an algebraic operation also determines the angle $\theta(\mathbf{u}, \mathbf{v})$ between two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ with an identical origin via

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} \ , \quad \text{for } \theta(\mathbf{u}, \mathbf{v}) \in [0, \pi] \ . \tag{1.12}$$

From the above expression, one trivially obtains

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}|\,|\mathbf{v}| \cos\theta \ , \tag{1.13}$$

which presents the **geometrical definition** of the inner product introduced in many texts on linear algebra in advance.

For subsequent developments, some useful definitions are given below:

Unit vector. A vector of unit length is said to be a *unit vector*. Unit vectors in this text are designated by boldface Latin letters with widehat. Hence, if $\widehat{\mathbf{e}}$ is a unit vector then $|\widehat{\mathbf{e}}| = 1$. In this regard, any arbitrary nonzero vector $\mathbf{u}$ in space can be normalized according to $\widehat{\mathbf{u}} = \mathbf{u}/\,|\mathbf{u}|$.

Equal vectors. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be *equal* when they have the same direction and magnitude. Technically, for all vectors $\mathbf{w}$,[4]

$$\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} \quad \text{if and only if} \quad \mathbf{u} = \mathbf{v} \ . \tag{1.14}$$

Orthogonal vectors. Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* if their inner product vanishes:

$$\mathbf{u} \cdot \mathbf{v} = 0 \quad \text{if and only if} \quad \cos\theta(\mathbf{u}, \mathbf{v}) = \frac{\pi}{2} \ . \tag{1.15}$$

And $\mathbf{u}$ is said to be **perpendicular** to $\mathbf{v}$ and vice versa.

Orthonormal basis. A basis, say $\mathscr{G}_{\mathrm{b}} = \{\widehat{\mathbf{u}}, \widehat{\mathbf{v}}, \widehat{\mathbf{w}}\} \subset \mathscr{E}_{\mathrm{r}}^3$, is called *orthonormal* if

$$\widehat{\mathbf{u}} \cdot \widehat{\mathbf{v}} = \widehat{\mathbf{u}} \cdot \widehat{\mathbf{w}} = \widehat{\mathbf{v}} \cdot \widehat{\mathbf{w}} = 0 \quad \text{and} \quad |\widehat{\mathbf{u}}| = |\widehat{\mathbf{v}}| = |\widehat{\mathbf{w}}| = 1 \ . \tag{1.16}$$

This basically represents a set of three mutually orthogonal unit vectors.

---

[3] In some texts, the Euclidean length of a vector $\mathbf{u}$ is denoted by $\|\mathbf{u}\|$.

[4] The result (1.14) can easily be verified. Suppose that $\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w}$ holds for all vectors $\mathbf{w}$. Then, by invoking $(-1)\,\mathbf{v} = -\mathbf{v}$ from (1.76) and making use of (1.5) and (1.10), one can arrive at the relation $(\mathbf{u} - \mathbf{v}) \cdot \mathbf{w} = 0$. Now, choosing $\mathbf{w} = \mathbf{u} - \mathbf{v}$ presents $g^{\mathrm{ip}}(\mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v}) = 0$. Finally, by means of (1.9d), one can obtain the desired result $\mathbf{u} = \mathbf{v}$. The converse is immediate.

1.3.1 Vector Projection and Rejection

The vector projection of $\mathbf{u}$ onto $\mathbf{v}$, denoted by $\mathbf{proj_v u}$, is a vector having identical direction with $\mathbf{v}$. Its magnitude, called *scalar projection* of $\mathbf{u}$ on $\mathbf{v}$, is defined by

$$\mathrm{proj_v}\mathbf{u} = \mathbf{u} \cdot \widehat{\mathbf{v}} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|} , \tag{1.17}$$

where $\widehat{\mathbf{v}}$ denotes the unit vector in the direction of $\mathbf{v}$. Thus,

$$\boxed{\mathbf{proj_v u} = (\mathrm{proj_v}\mathbf{u})\,(\widehat{\mathbf{v}}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|}\right)\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right) = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v} .} \tag{1.18}$$

The projection of $\mathbf{u}$ onto the plane perpendicular to $\mathbf{v}$ is called the *vector rejection* of $\mathbf{u}$ from $\mathbf{v}$ and designated here by $\mathbf{reje_v u}$. The sum of $\mathbf{proj_v u}$ and $\mathbf{reje_v u}$ is again $\mathbf{u}$. And this enables one to obtain

$$\boxed{\mathbf{reje_v u} = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v} .} \tag{1.19}$$

For a geometrical interpretation, see Fig. 1.2.

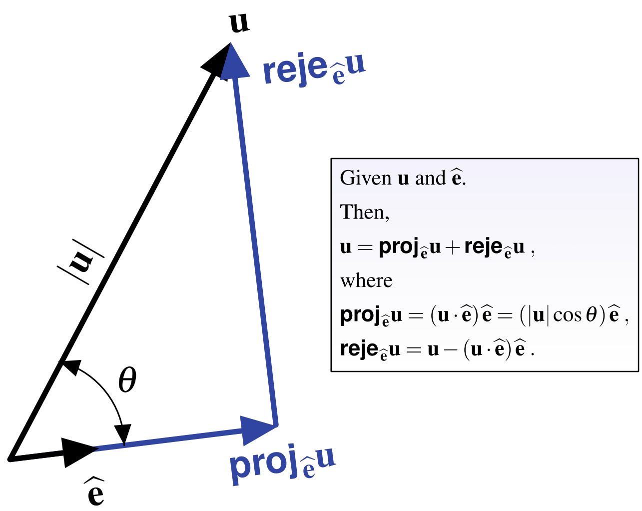


Fig. 1.2 Projection (rejection) of a vector $\mathbf{u}$ along (from) a unit vector $\widehat{\mathbf{e}}$

1.4 Complexification

In general, a vector space is not only defined over the field of real numbers R but also over the field of **complex** numbers C. This new vector space cannot be equipped with the inner product (1.8) properly and, therefore, this scalar-valued function along with its properties should consistently be modified. Similarly to $\mathscr{E}_{\mathrm{r}}^3$, one can also define an *three-dimensional Euclidean complex vector space* $\mathscr{E}_{\mathrm{c}}^3$ in such a way that its *complex vector space* $\mathscr{C}^3$, with axioms analogous to those of $\mathscr{V}^3$, is furnished by a consistent inner product $g_{\mathrm{c}}^{\mathrm{ip}}$. The goal is thus to define this new algebraic operation and identify its properties.

First, one can introduce a **complex number** $c^* \in$ C as $c^* = \alpha + \mathrm{i}\beta$ where $\alpha, \beta \in$ R and **complex vector** $\mathbf{z}^* \in \mathscr{E}_{\mathrm{c}}^3$ by $\mathbf{z}^* = \mathbf{u} + \mathrm{i}\mathbf{v}$ where $\mathbf{u}, \mathbf{v} \in \mathscr{E}_{\mathrm{r}}^3$. Here, i denotes the **imaginary unit** satisfying $\mathrm{i}^2 + 1 = 0$. As can be seen, $\mathscr{E}_{\mathrm{r}}^3$ renders a subspace of $\mathscr{E}_{\mathrm{c}}^3$. Let $\mathbf{z}_1^* = \mathbf{u}_1 + \mathrm{i}\mathbf{v}_1$ and $\mathbf{z}_2^* = \mathbf{u}_2 + \mathrm{i}\mathbf{v}_2$ be two arbitrary complex vectors. The vector addition $\mathbf{z}_1^* + \mathbf{z}_2^*$ and scalar multiplication $c^*\mathbf{z}^*$ for $\mathscr{C}^3$ are accordingly defined by the rules

$$\begin{aligned} \mathbf{z}_1^* + \mathbf{z}_2^* &= (\mathbf{u}_1 + \mathrm{i}\mathbf{v}_1) + (\mathbf{u}_2 + \mathrm{i}\mathbf{v}_2) \\ &= (\mathbf{u}_1 + \mathbf{u}_2) + \mathrm{i}\,(\mathbf{v}_1 + \mathbf{v}_2) \ , \qquad (1.20\mathrm{a}) \\ c^*\mathbf{z}^* &= (\alpha + \mathrm{i}\beta)\,(\mathbf{u} + \mathrm{i}\mathbf{v}) \\ &= (\alpha\mathbf{u} - \beta\mathbf{v}) + \mathrm{i}\,(\alpha\mathbf{v} + \beta\mathbf{u}) \ . \qquad (1.20\mathrm{b}) \end{aligned}$$

Denoting by $\overline{\mathbf{z}}^*$ ($\overline{c}^*$) the complex conjugate of $\mathbf{z}^*$ (c^*) according to $\overline{\mathbf{z}}^* = \mathbf{u} - \mathrm{i}\mathbf{v}$ ($\overline{c}^* = \alpha - \mathrm{i}\beta$), the inner product $g_{\mathrm{c}}^{\mathrm{ip}}$ is now defined via the relation[5]

$$\begin{aligned} g_{\mathrm{c}}^{\mathrm{ip}}\left(\mathbf{z}_1^*, \mathbf{z}_2^*\right) &= \overline{\mathbf{z}}_1^* \cdot \mathbf{z}_2^* \\ &= (\mathbf{u}_1 - \mathrm{i}\mathbf{v}_1) \cdot (\mathbf{u}_2 + \mathrm{i}\mathbf{v}_2) \\ &= \mathbf{u}_1 \cdot \mathbf{u}_2 + \mathrm{i}\,(\mathbf{u}_1 \cdot \mathbf{v}_2 - \mathbf{v}_1 \cdot \mathbf{u}_2) + \mathbf{v}_1 \cdot \mathbf{v}_2 \ . \qquad (1.21) \end{aligned}$$

Notice that $(1.21)_3$ will reduce to (1.8) when the imaginary parts of $\mathbf{z}_1^*$ and $\mathbf{z}_2^*$ become zero, i.e. $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{0}$. This scalar-valued function then satisfies the following axioms for any c^*, $d^* \in$ C and $\mathbf{z}_1^*$, $\mathbf{z}_2^*$, $\mathbf{z}^* \in \mathscr{E}_{\mathrm{c}}^3$:

$$\underbrace{g_{\mathrm{c}}^{\mathrm{ip}}\left(\mathbf{z}_1^*, \mathbf{z}_2^*\right) = \overline{g_{\mathrm{c}}^{\mathrm{ip}}\left(\mathbf{z}_2^*, \mathbf{z}_1^*\right)}}_{\text{or } \overline{\mathbf{z}}_1^* \cdot \mathbf{z}_2^* = \mathbf{z}_2^* \cdot \overline{\mathbf{z}}_1^* \text{ representing conjugate symmetry}} \ , \qquad (1.22\mathrm{a})$$

[5] Note that the property (1.9d) does not hold in general for complex vectors. For instance, consider a vector $\mathbf{z}^* = \widehat{\mathbf{u}} + \sqrt{3}\widehat{\mathbf{v}} - \mathrm{i}\,(2\widehat{\mathbf{w}})$ where $\{\widehat{\mathbf{u}}, \widehat{\mathbf{v}}, \widehat{\mathbf{w}}\}$ forms an orthonormal basis according to (1.16). It follows that $g^{\mathrm{ip}}\left(\mathbf{z}^*, \mathbf{z}^*\right) = 0$. Such a nonzero complex vector, which is orthogonal to itself based on (1.8), is called *isotropic vector*. That is why the inner product (1.8) along with its properties in (1.9a) to (1.9d) need to be reformulated. The definition $(1.21)_1$ is basically an inner product with positive definite property (without symmetry and bilinearity properties). This can easily be verified for the given example.

$$\underbrace{g_c^{ip}\left(c^*\mathbf{z}_1^* + d^*\mathbf{z}_2^*, \mathbf{z}^*\right) = \overline{c}^* g_c^{ip}\left(\mathbf{z}_1^*, \mathbf{z}^*\right) + \overline{d}^* g_c^{ip}\left(\mathbf{z}_2^*, \mathbf{z}^*\right)}_{\text{or } \left(\overline{c^*\mathbf{z}_1^* + d^*\mathbf{z}_2^*}\right)\cdot\mathbf{z}^* = \overline{c}^*\left(\overline{\mathbf{z}}_1^*\cdot\mathbf{z}^*\right) + \overline{d}^*\left(\overline{\mathbf{z}}_2^*\cdot\mathbf{z}^*\right) \text{ rendering linearity in the first argument}} \quad , \tag{1.22b}$$

$$\underbrace{\left.\begin{aligned} g_c^{ip}\left(\mathbf{z}^*, \mathbf{z}^*\right) > 0 \quad \text{if } \mathbf{z}^* \neq \mathbf{0} \\ g_c^{ip}\left(\mathbf{z}^*, \mathbf{z}^*\right) = 0 \quad \text{if } \mathbf{z}^* = \mathbf{0} \end{aligned}\right\}}_{\text{or } \overline{\mathbf{z}}^*\cdot\mathbf{z}^* \geq 0, \text{ for which } \overline{\mathbf{z}}^*\cdot\mathbf{z}^* = 0 \Longleftrightarrow \mathbf{z}^* = \mathbf{0}, \text{ expressing positive-definiteness}} \quad . \tag{1.22c}$$

The inner product $g_c^{ip}\left(\mathbf{z}_1^*, \mathbf{z}_2^*\right)$ with the above properties is basically a *sesquilinear form*. And it is usually denoted by $\left\langle\mathbf{z}_1^*, \mathbf{z}_2^*\right\rangle$ in physics community.

1.5 Three-Dimensional Euclidean Point Space

It is worth mentioning that there is no unified definition for Euclidean space and the previous definition relied on the notion of vector space wherein concept of **point** was excluded. Points contain the concept of what is meant by the notion of precise location in Euclidean space without any dimensional characteristic.

Intuitively, every ordered pair of distinct points in space can be associated with a vector that seems to act as a connecting tool. So, if a space of vectors exists then a space of points must exist consistently. Moreover, representing Euclidean space with only vectors without such a fundamental object upon which Euclidean geometry is built may produce a great concern. These motivate to define a space with closely related properties to $\mathscr{E}_r^3$.

An *three-dimensional Euclidean point space*, denoted by $\mathscr{E}_p^3$, is a set of points satisfying the following rules

for every ordered pair of points $(\mathbf{y}, \mathbf{x}) \in \mathscr{E}_p^3$,

there exists a vector $\mathbf{v} \in \mathscr{E}_r^3$ such that

$$\mathbf{v} = \mathbf{x} - \mathbf{y} \ , \quad \leftarrow \text{a vector is defined here by the difference of two arbitrary points} \tag{1.23a}$$

for any three points $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathscr{E}_p^3$,

$$(\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z}) = \mathbf{x} - \mathbf{z} \ , \quad \leftarrow \text{this expresses parallelogram law of vector addition and clearly shows that points cannot be summed up (only their differences are allowed to be summed up)} \tag{1.23b}$$

for any point $\mathbf{y} \in \mathscr{E}_p^3$ and vector $\mathbf{u} \in \mathscr{E}_r^3$,

there exists a unique point $\mathbf{x} \in \mathscr{E}_p^3$ such that

$$\mathbf{x} = \mathbf{y} + \mathbf{u} \ . \quad \leftarrow \text{this states that the sum of a point and a vector will be a point} \tag{1.23c}$$

By defining $\mathscr{E}_p^3$, the three-dimensional Euclidean vector space $\mathscr{E}_r^3$ is now referred to as *translation space* of $\mathscr{E}_p^3$.

1.6 Vector Representation in Cartesian Coordinates

A coordinate system is a way for identifying the location of a point in space. Indeed, uniquely defining any point or any other geometric element in space are enabled by coordinate systems. Denoting by n the space dimension, this is done by assigning a n-tuple of real numbers, or *coordinates*, to every point. For the space under consideration, each point is identified by three coordinates, e.g. a point $\mathbf{x}$ is shown by the ordered triplet $(x_1\,,\, x_2\,,\, x_3)$. Each coordinate indicates the distance between the point and an arbitrary reference point, called *origin*. The coordinate systems can be thought of as grid of points with fixed coordinates in a particular system. For three-dimensional spaces, a *coordinate line* is defined by keeping any two of the coordinates constant and varying the remaining one. And *coordinate axes* are directed lines (or curves) intersecting at the origin. A three-dimensional coordinate system for which the coordinate axes obey the right-hand rule is called a *right-handed coordinate system* and one that satisfies the left-hand rule is referred to as a *left-handed coordinate system*.

Cartesian coordinate system. A broadly used *Cartesian coordinate system* - in physics and engineering and also many other scientific fields - is one for which all coordinate lines passing through each point are mutually orthogonal. This simple structure makes it greatly preferable among many different coordinate systems.

Cartesian coordinate frame. Any arbitrary origin $\mathbf{o} \in \mathscr{E}_{\mathrm{p}}^3$ together with a *positively oriented* orthonormal basis $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{e}}_3\} \subset \mathscr{E}_{\mathrm{r}}^3$ defines a *Cartesian coordinate frame*.[6] Such a frame will extensively be used in this text.

By setting a Cartesian coordinate frame, any vector $\mathbf{u} \in \mathscr{E}_{\mathrm{r}}^3$ in space pointing from $\mathbf{o}$ to an specific point can be represented, in view of (1.7), by

$$\boxed{\mathbf{u} = u_1\,\widehat{\mathbf{e}}_1 + u_2\,\widehat{\mathbf{e}}_2 + u_3\,\widehat{\mathbf{e}}_3\ ,} \tag{1.24}$$

where $u_i\,,\ i = 1, 2, 3$ denote the **Cartesian** (or **rectangular**) **components** of $\mathbf{u}$. They are uniquely determined by

$$u_1 = \underbrace{\mathbf{u}\cdot\widehat{\mathbf{e}}_1}_{=\,\mathrm{proj}_{\widehat{\mathbf{e}}_1}\mathbf{u}}\ ,\quad u_2 = \underbrace{\mathbf{u}\cdot\widehat{\mathbf{e}}_2}_{=\,\mathrm{proj}_{\widehat{\mathbf{e}}_2}\mathbf{u}}\ ,\quad u_3 = \underbrace{\mathbf{u}\cdot\widehat{\mathbf{e}}_3}_{=\,\mathrm{proj}_{\widehat{\mathbf{e}}_3}\mathbf{u}}\ . \quad \longleftarrow \text{see (1.17)} \tag{1.25}$$

The expression (1.24) is known as the *coordinate representation* of a vector. With the aid of (1.25), it can be rephrased as

$$\mathbf{u} = \underbrace{(\mathbf{u}\cdot\widehat{\mathbf{e}}_1)\,\widehat{\mathbf{e}}_1}_{=\,\mathbf{proj}_{\widehat{\mathbf{e}}_1}\mathbf{u}} + \underbrace{(\mathbf{u}\cdot\widehat{\mathbf{e}}_2)\,\widehat{\mathbf{e}}_2}_{=\,\mathbf{proj}_{\widehat{\mathbf{e}}_2}\mathbf{u}} + \underbrace{(\mathbf{u}\cdot\widehat{\mathbf{e}}_3)\,\widehat{\mathbf{e}}_3}_{=\,\mathbf{proj}_{\widehat{\mathbf{e}}_3}\mathbf{u}}\ . \quad \longleftarrow \text{see (1.18)} \tag{1.26}$$

[6] For a **positively oriented basis**, the basis vectors can be either right-handed or left-handed which is a matter of convention, see (1.63). But, the right-handed bases are commonly acknowledged to be positively oriented.

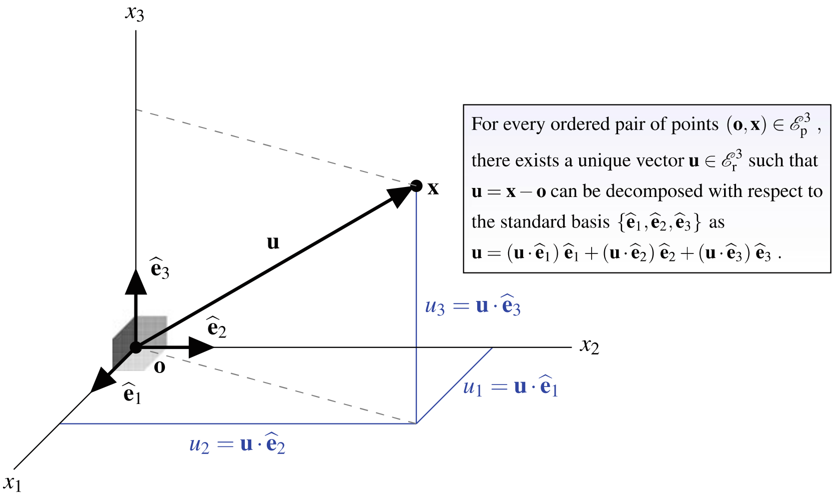


Fig. 1.3 A vector $\mathbf{u}$ with its rectangular components u_1, u_2 and u_3 in a right-handed Cartesian coordinate frame

Figure 1.3 schematically shows a Cartesian coordinate frame with right-handed coordinate axes as well as orthonormal bases wherein $\mathbf{u}$ in (1.26) renders the difference between an arbitrary point $\mathbf{x}$ and origin $\mathbf{o}$. For brevity, the Cartesian basis vectors from now on are denoted by

$$\{\widehat{\mathbf{e}}_i\} := \{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{e}}_3\} \ , \tag{1.27}$$

collectively. In a similar manner, for instance, $\{\mathbf{g}_i\}$ refers to $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$.

With reference to Fig. 1.3, each basis vector $\widehat{\mathbf{e}}_i$ is eventually the tangent vector that is tangent to x_i-axis. These tangent vectors are given by[7]

$$\boxed{\widehat{\mathbf{e}}_i = \frac{\partial \mathbf{x}}{\partial x_i} \ , \quad i = 1, 2, 3 \ .} \tag{1.28}$$

1.7 Indicial Notation and Summation Convention

Appearance of vectors with entirely boldface setting in an expression basically represents *direct* (or *absolute* or *symbolic* or *tensorial*) *notation*. As discussed before, once a specific coordinate system is chosen for practical reasons, a large number

[7] The expression (1.28) is well defined because the difference of two points delivers a vector. From the consistency point of view, it should be written as $\widehat{\mathbf{e}}_i = \partial \mathbf{x} / \partial x^i$. The reason is that they are basically the general basis vectors of an arbitrary curvilinear coordinate system, see (5.3). Here, they are reduced to the standard basis of a Cartesian coordinate frame.

of components will appear which refer a vector to the corresponding basis vectors. Now, one may ask if there is a notation to completely express the equations in terms of components. Indeed, such a notation is invented and called *indicial* (or *index* or *subscript* or *suffix*) *notation* but now the question is why one should apply it. At the first glance, the direct notation is superior since it treats vectors (or tensors) as single mathematical objects rather than collections of components. And this helps better understand the geometrical and/or physical meaning of equations. But, this notation requires much more effort to derive/prove huge number of vector (or tensor) algebraic and calculus identities. This becomes even worse when higher-order tensors come to the problem. The indicial notation as a promising tool, on the other hand, is capable of simplifying the writing of complicated expressions and performing the mathematical operations more conveniently. It is often the preferred notation for the theoretical and computational aspects of physical problems. In this regard, derivation of formulations in this text is mostly carried out by using this powerful and handy notation.

Index notation is quite simple and can be understood by some examples. For instance, the following system of equations

$$\left.\begin{aligned} A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = y_1 \\ A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = y_2 \\ A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = y_3 \end{aligned}\right\} , \tag{1.29}$$

in indicial notation can be written as

$$\sum_{j=1}^{3} A_{ij}x_j = y_i \ . \tag{1.30}$$

Latin indices take on various integer numbers from 1 to 3 in this book to refer a specific component in a vector (or tensor) identity. If the range of an index is not written, e.g. i in the above expression, one should realize that it is implicitly there and can be any of $1, 2, 3$ which totally presents 3 equations.

Another example will be the change of notation for the vector addition from the direct notation $\mathbf{u} = \mathbf{v} + \mathbf{w}$ to the index notation $u_i = v_i + w_i$. Since the basis vectors are not written, this representation is often referred to as the *component form* of $\mathbf{u} = \mathbf{v} + \mathbf{w}$. It is sometimes written as $(\mathbf{u})_i = (\mathbf{v})_i + (\mathbf{w})_i$.

In summary, for any vector (or tensor) identity such as the vector addition, one needs to distinguish between the following representations

$$\left.\begin{aligned} \text{Tensorial form (or direct form)}: \quad & \mathbf{u} = \mathbf{v} + \mathbf{w} \\ \text{Indicial form (or component form)}: \quad & u_i = v_i + w_i \quad \text{or} \quad (\mathbf{u})_i = (\mathbf{v})_i + (\mathbf{w})_i \end{aligned}\right\} . \tag{1.31}$$

The Eq. (1.24) can now be written as

$$\mathbf{u} = \sum_{i=1}^{3} u_i \,\widehat{\mathbf{e}}_i \ , \tag{1.32}$$

where

$$u_i = \mathbf{u} \cdot \widehat{\mathbf{e}}_i = \widehat{\mathbf{e}}_i \cdot \mathbf{u} = (\mathbf{u})_i \ . \quad \longleftarrow \text{see (1.25)} \tag{1.33}$$

If one writes the expression (1.32) without the summation symbol as

$$\boxed{\mathbf{u} = u_i \,\widehat{\mathbf{e}}_i \overset{\text{or}}{=\!=} (\mathbf{u} \cdot \widehat{\mathbf{e}}_i)\,\widehat{\mathbf{e}}_i \overset{\text{or}}{=\!=} (\widehat{\mathbf{e}}_i \cdot \mathbf{u})\,\widehat{\mathbf{e}}_i \overset{\text{or}}{=\!=} (\mathbf{u})_i \,\widehat{\mathbf{e}}_i \ ,} \tag{1.34}$$

the so-called *Einstein summation convention* (or simply *summation convention*) has been adopted. According to Einstein notation, if an index in a single term of an expression appears **twice**, the summation over all of the values of that index is implied. As an example, the relation (1.30) should now be rephrased as $A_{ij} x_j = y_i$. To apply this convention appropriately, the following rules must be obeyed:

- ✍ Each index is allowed to repeat once or twice, e.g. $u_i = v_i + a_j b_j c_j w_i$ is not true (clearly, it is acceptable to write $u_i = v_i + w_i \sum_{j=1}^{3} a_j b_j c_j$).
- ✍ Non-repeated indices should appear in each term identically, e.g. an expression of the form $u_i = v_j + w_i$ has no meaning.

An index that is summed over is called **dummy** (or **summation**) index since the summation is independent of the letter selected for this index, e.g. $u_i v_i = u_j v_j = u_k v_k$ all results in an identical real number. An index that only appears once in a term (if it is replaced by a new index in a single term, this should be carried out throughout the entire expression) is referred to as **free** (or **live**) index, e.g. $u_i = v_i + w_i$ and $u_j = v_j + w_j$ both have the same meaning. Note that in each term of an expression, an index should be either free or dummy.

It is worth noting that real advantage of the Einstein notation manifests itself when higher-order tensors and their relationships present in various terms of an expression, see Chaps. 3 and 6.

The expression (1.16) can be unified according to[8]

$$\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \ , \tag{1.35}$$

where δ_{ij} is called *Kronecker delta*. The Kronecker delta $\delta_{ij} = \delta_{ji}$ plays a major role in vector and tensor algebra (or calculus) mainly because of its **replacement property**. Basically, the Kronecker delta acts as a **substitution operator** since, for instance,

[8] For the three independent coordinates (x_1, x_2, x_3), the Kronecker delta δ_{ij} can also be indicated by $\delta_{ij} = \partial x_i / \partial x_j$.

$$\delta_{ij}\widehat{\mathbf{e}}_j = \widehat{\mathbf{e}}_i \ , \ \delta_{ij}u_j = u_i \ , \ \delta_{ij}A_{jk} = A_{ik} \ , \ \delta_{ij}\varepsilon_{jkl} = \varepsilon_{ikl} \ , \ \delta_{ij}\mathbb{A}_{jklm} = \mathbb{A}_{iklm} \ . \tag{1.36}$$

These identities can simply be verified by the direct expansion of the left hand sides. The above expressions show that if one index of δ_{ij}, say j, is repeated in a vector (or tensor) variable of an expression, the other index i is replaced with j in that variable and subsequently the Kronecker delta disappears.

The following identities hold true:

$$\left.\begin{aligned} \delta_{ii} &= 3 \\ \delta_{ii}\delta_{jj} &= 9 \end{aligned}\right\} \quad , \quad \delta_{ik}\delta_{kj} = \delta_{ij} \quad , \quad \left.\begin{aligned} \delta_{jk}\delta_{jk} &= 3 \\ \delta_{km}\delta_{mn}\delta_{nl} &= \delta_{kl} \end{aligned}\right\} \ . \tag{1.37}$$

It is worth mentioning that the summation convention does not have any conflict with the inner product since this operation is distributive with respect to the vector addition. Hence, the **coordinate representation** $(1.34)_1$ can help compute the inner product of any two vectors **u** (with u_i , $i = 1, 2, 3$) and **v** (with v_j , $j = 1, 2, 3$). This will be carried out by means of (1.35) and (1.36) along with the linearity property of the inner product according to

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (u_i \widehat{\mathbf{e}}_i) \cdot \left(v_j \widehat{\mathbf{e}}_j\right) = u_i \left(\widehat{\mathbf{e}}_i\right) \cdot \left(v_j \widehat{\mathbf{e}}_j\right) = u_i \left(v_j \widehat{\mathbf{e}}_j\right) \cdot \left(\widehat{\mathbf{e}}_i\right) \\ &= u_i v_j \left(\widehat{\mathbf{e}}_j\right) \cdot \left(\widehat{\mathbf{e}}_i\right) = u_i v_j \left(\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_j\right) = u_i v_j \left(\delta_{ij}\right) \\ &= \boxed{u_i v_i \ .} \end{aligned} \tag{1.38}$$

This result - which shows that the inner product of any two vectors is equal to the sum of the products of their components - is often used to present the **algebraic definition** of the inner product. The relations (1.11) and $(1.38)_7$ now help obtain the Euclidean length of **u**, in component form, according to

$$\boxed{|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_i u_i} = \sqrt{u_1^2 + u_2^2 + u_3^2} \ .} \tag{1.39}$$

Note that **0** is a vector whose components are zero in any coordinate system. Its scalar product with any arbitrary vector **u** thus trivially renders

$$\mathbf{u} \cdot \mathbf{0} = \mathbf{0} \cdot \mathbf{u} = 0 \ . \tag{1.40}$$

1.8 Matrix Notation

In introductory texts on linear algebra, vectors are represented by single-column matrices containing their ordered set of Cartesian components. This notation is very useful in computer programming required, for instance, in computational mechanics since a vector (or tensor) is treated as an array of numbers. In the following, the goal is thus to represent the matrix of a vector.

The Cartesian basis vectors in matrix notation are written as

$$[\,\widehat{\mathbf{e}}_1\,] = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad [\,\widehat{\mathbf{e}}_2\,] = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad , \quad [\,\widehat{\mathbf{e}}_3\,] = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \; , \tag{1.41}$$

or, collectively as,

$$[\,\widehat{\mathbf{e}}_i\,] = \begin{bmatrix} \delta_{i1} \\ \delta_{i2} \\ \delta_{i3} \end{bmatrix} \; . \tag{1.42}$$

Here, the square bracket $[\bullet]$ has been introduced to denote a matrix. This notation admits summation convention and, therefore, any vector $\mathbf{u} = u_i\,\widehat{\mathbf{e}}_i$ will take the following form

$$[\,\mathbf{u}\,] = u_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \text{or} \quad [\,\mathbf{u}\,] = \begin{bmatrix} u_i\delta_{i1} \\ u_i\delta_{i2} \\ u_i\delta_{i3} \end{bmatrix} \; . \tag{1.43}$$

It is often indicated by

$$[\,\mathbf{u}\,] = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^{\mathrm{T}} \; , \tag{1.44}$$

where $^{\mathrm{T}}$ denotes the **transpose operator**. Note that $[\,\mathbf{u}\,]$ is not the only form of $\mathbf{u}$ in matrix notation, see (5.70) and (5.71).

The inner product of $\mathbf{u}$ and $\mathbf{v}$ is in accordance with the matrix multiplication, that is,

$$\mathbf{u} \cdot \mathbf{v} = [\mathbf{u}]^{\mathrm{T}} [\mathbf{v}] = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = u_1 v_1 + u_2 v_2 + u_3 v_3 \; . \tag{1.45}$$

Guided by (1.35), the **identity matrix** renders

$$[\,\mathbf{I}\,] = \begin{bmatrix} \delta_{ij} \end{bmatrix} = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{12} & \delta_{22} & \delta_{23} \\ \delta_{13} & \delta_{23} & \delta_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \; . \tag{1.46}$$

And an arbitrary matrix $[\,\mathbf{A}\,]$ with its transpose $[\,\mathbf{A}\,]^{\mathrm{T}}$ are represented by

$$[\,\mathbf{A}\,] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \quad , \quad [\,\mathbf{A}\,]^{\mathrm{T}} = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \; . \tag{1.47}$$

One often needs to save an array of numbers (not necessarily the components of a vector or tensor with respect to a basis) for convenience. This handy notation, for instance, helps collect the diagonal elements of the identity matrix as

$$\left[\, \mathbf{1}_{\mathrm{p}} \,\right] := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} . \tag{1.48}$$

This single-column matrix is useful in fast computations within the context of matrix algebra. Some more examples will be introduced in later chapters.

1.9 Cross Product

The *cross* (or *vector*) product of any two vectors $\mathbf{u}$ and $\mathbf{v}$, designated by $\mathbf{u} \times \mathbf{v}$, is again a vector. That cross product $\times : \mathscr{E}_{\mathrm{r}}^3 \times \mathscr{E}_{\mathrm{r}}^3 \to \mathscr{E}_{\mathrm{r}}^3$ satisfies the following properties for any $\mathbf{u}\,,\ \mathbf{v}\,,\ \mathbf{w} \in \mathscr{E}_{\mathrm{r}}^3$ and $\alpha\,,\ \beta \in \mathrm{R}$:

$$\underbrace{\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}}_{\text{representing anti-commutative property}} \quad , \tag{1.49a}$$

$$\underbrace{(\alpha\mathbf{u} + \beta\mathbf{v}) \times \mathbf{w} = \alpha\,(\mathbf{u} \times \mathbf{w}) + \beta\,(\mathbf{v} \times \mathbf{w})}_{\substack{\text{or } \mathbf{u} \times (\alpha\mathbf{v} + \beta\mathbf{w}) = \alpha\,(\mathbf{u} \times \mathbf{v}) + \beta\,(\mathbf{u} \times \mathbf{w}) \text{ rendering linearity} \\ \text{in each argument, i.e. it is } \textbf{bilinear} \text{ for any ordered pair of vectors}}} \quad , \tag{1.49b}$$

$$\underbrace{\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0}_{\substack{\text{presenting the orthogonality of } \mathbf{u} \text{ (or } \mathbf{v}\text{) and } \mathbf{u} \times \mathbf{v}\,, \text{ i.e.} \\ \text{the cross product of two vectors is perpendicular to each of them}}} \quad , \tag{1.49c}$$

$$\underbrace{(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{u} \times \mathbf{v}) = (\mathbf{u} \cdot \mathbf{u})\,(\mathbf{v} \cdot \mathbf{v}) - (\mathbf{u} \cdot \mathbf{v})^2}_{\text{indicating the Euclidean length } |\mathbf{u} \times \mathbf{v}|^2 = (\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{u} \times \mathbf{v})} \quad . \quad \longleftarrow \text{see (1.78a)} \tag{1.49d}$$

The property (1.49c) is visualized in Fig. 1.4 wherein $\mathbf{u} \times \mathbf{v}$ is perpendicular to a plane containing the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$. The last property with the aid of (1.13) delivers

$$|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}|\,|\mathbf{v}|\sin\theta\,(\mathbf{u}, \mathbf{v}) = \mathsf{A}\,, \quad \text{for } \theta\,(\mathbf{u}, \mathbf{v}) \in [0, \pi]\,, \tag{1.50}$$

where A denotes the area of parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$. For the special cases $\theta = 0, \pi$, i.e. when these vectors are **linearly dependent**, one finds that $|\mathbf{u} \times \mathbf{v}| = 0$. It is then a simple exercise to verify that

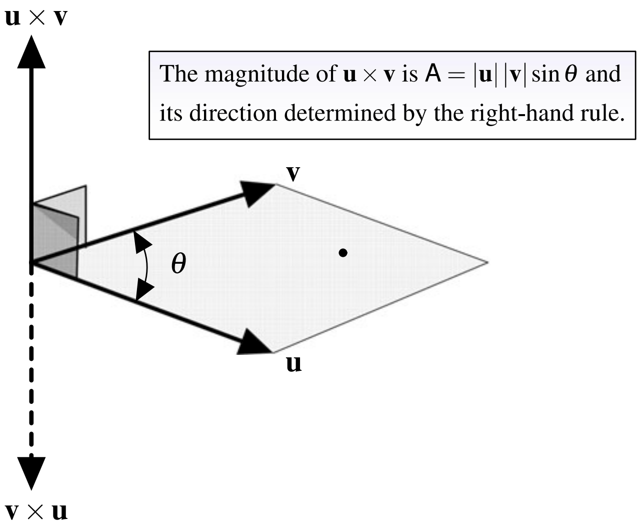


Fig. 1.4 Cross product of **u** and **v**

$$\boxed{\mathbf{u} \times \mathbf{v} = \mathbf{0} \quad \text{if and only if} \quad \mathbf{u} = \alpha \mathbf{v} \ .} \tag{1.51}$$

Since $\mathbf{u} \times \mathbf{v}$ carries the information regarding the area defined by **u** and **v**, it is often referred to as the *area vector*. The correctness of the above properties for vector product can readily be shown by algebraic relations in Cartesian coordinates. This requires the introduction of a symbol

$$\varepsilon_{ijk} = \begin{cases} +1 \ , & \text{for } \textbf{even} \text{ permutations of } ijk, \text{ i.e. } 123, 231, 312 \\ -1 \ , & \text{for } \textbf{odd} \text{ permutations of } ijk, \text{ i.e. } 132, 321, 213 \\ 0 \ , & \text{if there is a repeated index, i.e. } i = j \ , \ i = k \ , \ j = k \end{cases} \ , \tag{1.52}$$

called the *permutation* (or *alternating* or *antisymmetric* or *Levi–Civita*) *symbol*. Note that this definition is not unique and there are some alternatives such as

$$\boxed{\varepsilon_{ijk} = \frac{1}{2} (i - j) (j - k) (k - i) \ ,} \tag{1.53}$$

satisfying

$$\boxed{\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij} \quad , \quad \varepsilon_{ikj} = \varepsilon_{kji} = \varepsilon_{jik} = -\varepsilon_{ijk} \ .} \tag{1.54}$$

The permutation symbol is expressible in terms of the Kronecker delta as follows:

$$
\begin{aligned}
\varepsilon_{ijk} &= \delta_{im}\delta_{jn}\delta_{kl}\varepsilon_{mnl} \\
&= \underbrace{\delta_{i1}\delta_{jn}\delta_{kl}\varepsilon_{1nl}}_{=\delta_{i1}\delta_{j2}\delta_{kl}\varepsilon_{12l}+\delta_{i1}\delta_{j3}\delta_{kl}\varepsilon_{13l}=\delta_{i1}\delta_{j2}\delta_{k3}\varepsilon_{123}+\delta_{i1}\delta_{j3}\delta_{k2}\varepsilon_{132}=\delta_{i1}\delta_{j2}\delta_{k3}-\delta_{i1}\delta_{j3}\delta_{k2}} \\
&\quad + \underbrace{\delta_{i2}\delta_{jn}\delta_{kl}\varepsilon_{2nl}}_{=\delta_{i2}\delta_{j1}\delta_{kl}\varepsilon_{21l}+\delta_{i2}\delta_{j3}\delta_{kl}\varepsilon_{23l}=\delta_{i2}\delta_{j1}\delta_{k3}\varepsilon_{213}+\delta_{i2}\delta_{j3}\delta_{k1}\varepsilon_{231}=-\delta_{i2}\delta_{j1}\delta_{k3}+\delta_{i2}\delta_{j3}\delta_{k1}} \\
&\quad + \underbrace{\delta_{i3}\delta_{jn}\delta_{kl}\varepsilon_{3nl}}_{=\delta_{i3}\delta_{j1}\delta_{kl}\varepsilon_{31l}+\delta_{i3}\delta_{j2}\delta_{kl}\varepsilon_{32l}=\delta_{i3}\delta_{j1}\delta_{k2}\varepsilon_{312}+\delta_{i3}\delta_{j2}\delta_{k1}\varepsilon_{321}=\delta_{i3}\delta_{j1}\delta_{k2}-\delta_{i3}\delta_{j2}\delta_{k1}} \\
&= \delta_{i1}\left(\delta_{j2}\delta_{k3}-\delta_{j3}\delta_{k2}\right) - \delta_{j1}\left(\delta_{i2}\delta_{k3}-\delta_{i3}\delta_{k2}\right) + \delta_{k1}\left(\delta_{i2}\delta_{j3}-\delta_{i3}\delta_{j2}\right) \ . \qquad (1.55)
\end{aligned}
$$

In $(1.55)_1$, observe that the free indices of a variable in a single term can be replaced by new indices making use of the Kronecker delta but the old ones should still be kept for consistency with the other terms. The result $(1.55)_3$ can be written in a more convenient form as

$$
\begin{aligned}
\varepsilon_{ijk} &= \underbrace{\det\begin{bmatrix}\delta_{i1} & \delta_{j1} & \delta_{k1}\\ \delta_{i2} & \delta_{j2} & \delta_{k2}\\ \delta_{i3} & \delta_{j3} & \delta_{k3}\end{bmatrix} = \det\begin{bmatrix}\delta_{i1} & \delta_{i2} & \delta_{i3}\\ \delta_{j1} & \delta_{j2} & \delta_{j3}\\ \delta_{k1} & \delta_{k2} & \delta_{k3}\end{bmatrix}}_{\text{recall that the determinant of a matrix equals to the determinant of its transpose}} \\
&= \delta_{i1}\left(\delta_{j2}\delta_{k3}-\delta_{k2}\delta_{j3}\right) \\
&\quad - \delta_{i2}\left(\delta_{j1}\delta_{k3}-\delta_{k1}\delta_{j3}\right) + \delta_{i3}\left(\delta_{j1}\delta_{k2}-\delta_{k1}\delta_{j2}\right) \ , \qquad (1.56)
\end{aligned}
$$

where $\det[\bullet]$ stands for the determinant of a matrix. The above expression can be viewed as another definition for the permutation symbol.

The expressions $(1.56)_{1-2}$ help establish the product $\varepsilon_{ijk}\varepsilon_{lmn}$ according to

$$
\varepsilon_{ijk}\varepsilon_{lmn} = \underbrace{\det\left[\begin{bmatrix}\delta_{i1} & \delta_{i2} & \delta_{i3}\\ \delta_{j1} & \delta_{j2} & \delta_{j3}\\ \delta_{k1} & \delta_{k2} & \delta_{k3}\end{bmatrix}\begin{bmatrix}\delta_{1l} & \delta_{1m} & \delta_{1n}\\ \delta_{2l} & \delta_{2m} & \delta_{2n}\\ \delta_{3l} & \delta_{3m} & \delta_{3n}\end{bmatrix}\right] = \det\begin{bmatrix}\delta_{il} & \delta_{im} & \delta_{in}\\ \delta_{jl} & \delta_{jm} & \delta_{jn}\\ \delta_{kl} & \delta_{km} & \delta_{kn}\end{bmatrix}}_{\text{note that }\det[\bullet]\det[\circ]=\det[\bullet\circ]\text{ and, e.g. }\delta_{i1}\delta_{1l}+\delta_{i2}\delta_{2l}+\delta_{i3}\delta_{3l}=\delta_{io}\delta_{ol}=\delta_{il}} \ . \qquad (1.57)
$$

Three special cases can then be deduced:

$$
\varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \ , \qquad (1.58a)
$$

$$
\begin{aligned}
\varepsilon_{ijk}\varepsilon_{ljk} &= \delta_{il}\delta_{jj} - \delta_{ij}\delta_{jl} = 3\delta_{il} - \delta_{il} \\
&= 2\delta_{il} \ , \qquad (1.58b)
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_{ijk}\varepsilon_{ijk} &= 2\delta_{ii} \\
&= 6 \ . \qquad (1.58c)
\end{aligned}
$$

One can finally establish

$$
\boxed{\varepsilon_{ijk}\delta_{mn} = \varepsilon_{mjk}\delta_{in} + \varepsilon_{imk}\delta_{jn} + \varepsilon_{ijm}\delta_{kn} \ ,} \qquad (1.59)
$$

since

$$
\begin{aligned}
2\varepsilon_{ijk}\delta_{mn} &\underset{(1.58b)}{\overset{\text{from}}{=\!=}} \varepsilon_{mop}\left[\varepsilon_{ijk}\varepsilon_{nop}\right] \\
&\underset{(1.57)}{\overset{\text{from}}{=\!=}} \varepsilon_{mop}\left[\delta_{in}\left(\delta_{jo}\delta_{kp}-\delta_{ko}\delta_{jp}\right)\right. \\
&\quad \left.-\delta_{io}\left(\delta_{jn}\delta_{kp}-\delta_{kn}\delta_{jp}\right)+\delta_{ip}\left(\delta_{jn}\delta_{ko}-\delta_{kn}\delta_{jo}\right)\right] \\
&\underset{(1.36)}{\overset{\text{from}}{=\!=}} \varepsilon_{mjk}\delta_{in}-\varepsilon_{mkj}\delta_{in}-\varepsilon_{mik}\delta_{jn}+\varepsilon_{mij}\delta_{kn}+\varepsilon_{mki}\delta_{jn}-\varepsilon_{mji}\delta_{kn} \\
&\underset{(1.54)}{\overset{\text{from}}{=\!=}} 2\varepsilon_{mjk}\delta_{in}+2\varepsilon_{imk}\delta_{jn}+2\varepsilon_{ijm}\delta_{kn} \ .
\end{aligned}
$$

The way that the permutation symbol helps characterize the cross product in terms of the components of vectors is explained in the following.

Basically, the cross product of vectors leads to the cross product of their basis vectors which is enabled by using the property (1.49b). Let $\{\widehat{\mathbf{e}}_i^{\bullet}\}$ be an the orthonormal basis which can be either right-handed or left-handed. Consider, for example, the cross product $\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}$. By use of (1.50), the norm is $|\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}| = 1$ and the direction by means of (1.49c) will be either $\widehat{\mathbf{e}}_3^{\bullet}$ or $-\widehat{\mathbf{e}}_3^{\bullet}$, that is,

$$\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet} = \pm\widehat{\mathbf{e}}_3^{\bullet} \ . \tag{1.60}$$

This reveal the fact that the presented axioms for the cross product is unable to uniquely determine the direction of $\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}$. To proceed, the vector $\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}$ is expressed in terms of its components as

$$
\begin{aligned}
\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet} &= \alpha_m^{\bullet}\widehat{\mathbf{e}}_m^{\bullet} \\
&= \left[\left(\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}\right)\cdot\widehat{\mathbf{e}}_m^{\bullet}\right]\widehat{\mathbf{e}}_m^{\bullet} \\
&= \left[\left(\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}\right)\cdot\widehat{\mathbf{e}}_3^{\bullet}\right]\widehat{\mathbf{e}}_3^{\bullet} \ .
\end{aligned} \tag{1.61}
$$

Thus, by comparing (1.60) and $(1.61)_3$, one finds that

$$\left(\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}\right)\cdot\widehat{\mathbf{e}}_3^{\bullet} = \pm 1 \ . \tag{1.62}$$

This motivates to define *orientation* for bases. A basis $\{\widehat{\mathbf{e}}_1^{\bullet}, \widehat{\mathbf{e}}_2^{\bullet}, \widehat{\mathbf{e}}_3^{\bullet}\}$, not necessarily orthonormal in general, is said to be *positively oriented* if

$$\left(\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}\right)\cdot\widehat{\mathbf{e}}_3^{\bullet} > 0 \ , \tag{1.63}$$

and it is referred to as negatively oriented if $\left(\widehat{\mathbf{e}}_1^{\bullet} \times \widehat{\mathbf{e}}_2^{\bullet}\right)\cdot\widehat{\mathbf{e}}_3^{\bullet} < 0$. As commonly accepted, let's declare the right-handed bases as positively oriented ones and consider (1.63) as a trivial condition for basis vectors that obey the right-hand screw rule. A three-dimensional Euclidean vector space equipped with the vector product and positively (or negatively) oriented bases is called an *oriented three-dimensional Euclidean vector space*. Here, the positive one is denoted by $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$.

Hereafter, all Cartesian coordinate frames in this text contain the origin $\mathbf{o} \in \mathscr{E}_{\mathrm{p}}^{3}$ and the right-handed orthonormal basis $\{\widehat{\mathbf{e}}_i\} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$. Now, the cross product $\widehat{\mathbf{e}}_1 \times \widehat{\mathbf{e}}_2$ in (1.60) becomes $\widehat{\mathbf{e}}_1 \times \widehat{\mathbf{e}}_2 = \widehat{\mathbf{e}}_3$. There are nine possible choices for $\widehat{\mathbf{e}}_i \times \widehat{\mathbf{e}}_j$ which can all be presented in the following unified form

$$\boxed{\widehat{\mathbf{e}}_i \times \widehat{\mathbf{e}}_j = \varepsilon_{ijk} \widehat{\mathbf{e}}_k \ .} \tag{1.64}$$

This demonstrates the major role of permutation tensor in evaluating the cross product, see Exercise 1.5.

From (1.64), one immediately obtains

$$\boxed{\varepsilon_{ijk} = \det \left[\widehat{\mathbf{e}}_i \ \widehat{\mathbf{e}}_j \ \widehat{\mathbf{e}}_j \right] = \widehat{\mathbf{e}}_i \cdot \left(\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k \right) \ .} \tag{1.65}$$

As a result,

$$\widehat{\mathbf{e}}_l = \underbrace{\frac{1}{2} \varepsilon_{ljk} \widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k}_{\text{since } \widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_l = \frac{1}{2} 2\delta_{il} = \frac{1}{2}\varepsilon_{ijk}\varepsilon_{ljk} = \frac{1}{2}\varepsilon_{ljk}\widehat{\mathbf{e}}_i \cdot (\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k) = \widehat{\mathbf{e}}_i \cdot \left(\frac{1}{2}\varepsilon_{ljk}\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k\right)} \ . \tag{1.66}$$

For any two Cartesian vectors $\mathbf{u} = u_i \widehat{\mathbf{e}}_i$ and $\mathbf{v} = v_j \widehat{\mathbf{e}}_j$, the cross product $\mathbf{u} \times \mathbf{v}$ as a bilinear operator, with the aid of (1.64), admits the coordinate representation

$$\underbrace{\mathbf{w} = \mathbf{u} \times \mathbf{v} = (u_i \widehat{\mathbf{e}}_i) \times \left(v_j \widehat{\mathbf{e}}_j\right) = u_i v_j \left(\widehat{\mathbf{e}}_i \times \widehat{\mathbf{e}}_j\right) = u_i v_j \varepsilon_{ijk} \widehat{\mathbf{e}}_k}_{\text{with } w_k = u_i v_j \varepsilon_{ijk} \text{ or } w_k = \varepsilon_{kij} u_i v_j} \ . \tag{1.67}$$

By direct expansion, the above result takes the form

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \underbrace{(u_2 v_3 - u_3 v_2)}_{= w_1} \widehat{\mathbf{e}}_1 + \underbrace{(u_3 v_1 - u_1 v_3)}_{= w_2} \widehat{\mathbf{e}}_2 + \underbrace{(u_1 v_2 - u_2 v_1)}_{= w_3} \widehat{\mathbf{e}}_3 \ . \tag{1.68}$$

This is often represented in a more convenient form as

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \widehat{\mathbf{e}}_1 & \widehat{\mathbf{e}}_2 & \widehat{\mathbf{e}}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix} \ . \tag{1.69}$$

It is worthwhile to point out that the cross product is **not associative** but it satisfies the so-called *Jacobi identity*:

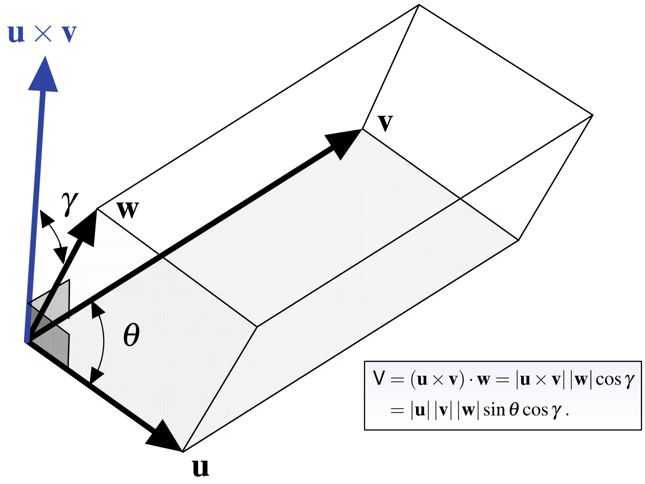


Fig. 1.5 The scalar triple product of **u**, **v** and **w** with the resulting parallelepiped

$$\boxed{\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0}\ ,} \quad \longleftarrow \text{the proof is given in Exercise 1.4} \tag{1.70}$$

where each term basically presents a *triple vector product* satisfying

$$\boxed{\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\, \mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\, \mathbf{w}\ .} \tag{1.71}$$

In general, the triple vector product is **not associative** due to

$$\boxed{(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w})\, \mathbf{v} - (\mathbf{v} \cdot \mathbf{w})\, \mathbf{u}\ .} \quad \longleftarrow \text{the proof is given in Exercise 1.3} \tag{1.72}$$

For any three vectors **u**, **v** and **w** in $\mathscr{E}_{\mathrm{r}}^{o3}$, *scalar triple* (or *mixed* or *box*) *product* is defined as

$$\boxed{\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \mathsf{V}\ ,} \tag{1.73}$$

where V presents the volume of the parallelepiped constructed by **u**, **v** and **w** as shown in Fig. 1.5.

Having in mind the bilinearity property of the inner product and cross product, by use of $(1.34)_1$, $(1.35)_1$, (1.54) and (1.64), the volume V takes the form

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= (u_i \widehat{\mathbf{e}}_i) \cdot \left(v_j \widehat{\mathbf{e}}_j \times w_k \widehat{\mathbf{e}}_k \right) = u_i v_j w_k \left(\widehat{\mathbf{e}}_i \right) \cdot \left(\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k \right) \\ &= u_i v_j w_k \left(\widehat{\mathbf{e}}_i \right) \cdot \left(\varepsilon_{jkl} \widehat{\mathbf{e}}_l \right) = u_i v_j w_k \varepsilon_{jkl} \delta_{il} \\ &= \boxed{u_i v_j w_k \varepsilon_{jki} = \varepsilon_{ijk} u_i v_j w_k\ .} \end{aligned} \tag{1.74}$$

This can also be written in the handy determinant form

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} . \tag{1.75}$$

From the above relation, one can deduce the following statement:

> If the volume V of the parallelepiped defined by three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is nonzero, then these vectors are linearly independent and, therefore, the triad $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ forms a basis for $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$.

1.10 Exercises

Exercise 1.1

Use the vector space properties (given on Sect. 1.1) to verify that

$$\left.\begin{array}{ll} \text{(a)} & \alpha\mathbf{0} = \mathbf{0} \\ \text{(b)} & 0\mathbf{u} = \mathbf{0} \end{array}\right\} , \quad \left.\begin{array}{ll} \text{(c)} & 0\mathbf{0} = \mathbf{0} \\ \text{(d)} & \mathbf{0} = -\mathbf{0} \end{array}\right\} , \quad \left.\begin{array}{ll} \text{(e)} & (-\alpha)\,\mathbf{u} = \alpha\,(-\mathbf{u}) \\ \text{(f)} & (-1)\,\mathbf{u} = -\mathbf{u} \end{array}\right\} . \tag{1.76}$$

Solution.

(a) :

$$\begin{aligned} \alpha\mathbf{0} &\xlongequal[\text{(1.4d)}]{\text{from}} \alpha\mathbf{0} + \mathbf{0} \xlongequal[\text{(1.4c)}]{\text{from}} \alpha\mathbf{0} + \alpha\mathbf{u} + (-\alpha\mathbf{u}) \\ &\xlongequal[\text{(1.4h)}]{\text{from}} \alpha\,(\mathbf{0} + \mathbf{u}) + (-\alpha\mathbf{u}) \xlongequal[\text{(1.4a)}]{\text{from}} \alpha\,(\mathbf{u} + \mathbf{0}) + (-\alpha\mathbf{u}) \\ &\xlongequal[\text{(1.4d)}]{\text{from}} \alpha\mathbf{u} + (-\alpha\mathbf{u}) \xlongequal[\text{(1.4c)}]{\text{from}} \mathbf{0} \,. \end{aligned}$$

(b) :

$$\begin{aligned} 0\mathbf{u} &\xlongequal[\text{(1.4d)}]{\text{from}} 0\mathbf{u} + \mathbf{0} \xlongequal[\text{(1.4c)}]{\text{from}} 0\mathbf{u} + 0\mathbf{u} + (-0\mathbf{u}) \\ &\xlongequal[\text{(1.4g)}]{\text{from}} (0 + 0)\,\mathbf{u} + (-0\mathbf{u}) \xlongequal[\text{(1.4c)}]{\text{from}} \mathbf{0} \,. \end{aligned}$$

(c) :

$$0\mathbf{0} \xlongequal[(1.4\text{d})]{\text{from}} 0\mathbf{0} + \mathbf{0} \xlongequal[(1.4\text{c})]{\text{from}} 0\mathbf{0} + 0\mathbf{0} + (-0\mathbf{0})$$
$$\xlongequal[(1.4\text{g})]{\text{from}} (0+0)\,\mathbf{0} + (-0\mathbf{0}) \xlongequal[(1.4\text{c})]{\text{from}} \mathbf{0} \; .$$

The identities (a)–(c) collectively represent

$$\alpha\mathbf{u} = \begin{cases} \mathbf{0} & \text{if } \alpha \neq 0 \;,\; \mathbf{u} = \mathbf{0} \\ \mathbf{0} & \text{if } \alpha = 0 \;,\; \mathbf{u} \neq \mathbf{0} \\ \mathbf{0} & \text{if } \alpha = 0 \;,\; \mathbf{u} = \mathbf{0} \end{cases} .$$

(d) :

$$\mathbf{0} \xlongequal[(1.4\text{c})]{\text{from}} \mathbf{0} + (-\mathbf{0}) \xlongequal[(1.4\text{a})]{\text{from}} (-\mathbf{0}) + \mathbf{0} \xlongequal[(1.4\text{d})]{\text{from}} (-\mathbf{0}) = -\mathbf{0} \; .$$

(e) :

$$(-\alpha)\,\mathbf{u} \xlongequal[(1.4\text{d})]{\text{from}} (-\alpha)\,\mathbf{u} + \mathbf{0} \xlongequal[\text{(a) in } (1.76)]{\text{from}} (-\alpha)\,\mathbf{u} + \alpha\mathbf{0}$$
$$\xlongequal[(1.4\text{c})]{\text{from}} (-\alpha)\,\mathbf{u} + \alpha\,(\mathbf{u} + (-\mathbf{u})) \xlongequal[(1.4\text{h})]{\text{from}} (-\alpha)\,\mathbf{u} + \alpha\mathbf{u} + \alpha\,(-\mathbf{u})$$
$$\xlongequal[(1.4\text{g})]{\text{from}} (-\alpha + \alpha)\,\mathbf{u} + \alpha\,(-\mathbf{u}) \xlongequal[\text{(b) in } (1.76)]{\text{from}} \mathbf{0} + \alpha\,(-\mathbf{u})$$
$$\xlongequal[(1.4\text{a})]{\text{from}} \alpha\,(-\mathbf{u}) + \mathbf{0} \xlongequal[(1.4\text{d})]{\text{from}} \alpha\,(-\mathbf{u}) \; .$$

(f) : By substituting $\alpha = 1$ into the above result and making use of the property (1.4e), one can arrive at the last identity:

$$(-1)\,\mathbf{u} = 1\,(-\mathbf{u}) = -\mathbf{u} \; .$$

Exercise 1.2

By means of the properties

$$(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w} \;\; (\text{or } \mathbf{w} \times (\mathbf{u} + \mathbf{v}) = \mathbf{w} \times \mathbf{u} + \mathbf{w} \times \mathbf{v}) \; , \tag{1.77a}$$
$$(\alpha\mathbf{u}) \times \mathbf{w} = \mathbf{u} \times (\alpha\mathbf{w}) = \alpha\,(\mathbf{u} \times \mathbf{w}) \; , \tag{1.77b}$$

derive the property (1.49b). Note that these two properties and (1.49b) are basically equivalent but (1.49b) is often written for brevity.

Solution. For any three vectors $\alpha\mathbf{u}$, $\beta\mathbf{v}$ and $\mathbf{w}$, one will have

$$(\alpha\mathbf{u}+\beta\mathbf{v})\times\mathbf{w} \xlongequal[\text{(1.77a)}]{\text{from}} (\alpha\mathbf{u})\times\mathbf{w}+(\beta\mathbf{v})\times\mathbf{w} \xlongequal[\text{(1.77}b\text{)}]{\text{from}} \alpha\,(\mathbf{u}\times\mathbf{w})+\beta\,(\mathbf{v}\times\mathbf{w}) \ .$$

Exercise 1.3

Verify the triple vector product $(\mathbf{u}\times\mathbf{v})\times\mathbf{w}=(\mathbf{u}\cdot\mathbf{w})\,\mathbf{v}-(\mathbf{v}\cdot\mathbf{w})\,\mathbf{u}$, according to (1.72), and show that the vector $(\mathbf{u}\times\mathbf{v})\times\mathbf{w}$ lies in the plane defined by the two vectors $\mathbf{u}$ and $\mathbf{v}$.

Solution. This exercise will be solved by using the permutation symbol and considering the coordinate representations $\mathbf{u}=u_i\widehat{\mathbf{e}}_i$, $\mathbf{v}=v_j\widehat{\mathbf{e}}_j$ and $\mathbf{w}=w_l\widehat{\mathbf{e}}_l$ (see Fig. 1.3). Having in mind the bilinearity of the cross product, one can show that

$$\begin{aligned}
[\mathbf{u}\times\mathbf{v}]\times(\mathbf{w}) &= \left[u_i v_j\left(\widehat{\mathbf{e}}_i\times\widehat{\mathbf{e}}_j\right)\right]\times\left(w_l\widehat{\mathbf{e}}_l\right)\\
&\xlongequal[\text{(1.64)}]{\text{from}} \left[u_i v_j\left(\varepsilon_{ijk}\widehat{\mathbf{e}}_k\right)\right]\times\left(w_l\widehat{\mathbf{e}}_l\right)=\left[u_i v_j w_l\varepsilon_{ijk}\right]\left(\widehat{\mathbf{e}}_k\times\widehat{\mathbf{e}}_l\right)\\
&\xlongequal[\text{(1.64)}]{\text{from}} \left(u_i v_j w_l\varepsilon_{ijk}\right)\left(\varepsilon_{klm}\widehat{\mathbf{e}}_m\right)\\
&\xlongequal[\text{(1.54)}]{\text{from}} u_i v_j w_l\left(\varepsilon_{ijk}\varepsilon_{lmk}\right)\widehat{\mathbf{e}}_m\\
&\xlongequal[\text{(1.58a)}]{\text{from}} u_i v_j w_l\left(\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}\right)\widehat{\mathbf{e}}_m\\
&\xlongequal[\text{(1.36)}]{\text{from}} u_i v_j w_i\widehat{\mathbf{e}}_j-u_i v_j w_j\widehat{\mathbf{e}}_i=\left(u_i w_i\right)\left(v_j\widehat{\mathbf{e}}_j\right)-\left(v_j w_j\right)\left(u_i\widehat{\mathbf{e}}_i\right)\\
&\xlongequal[\text{(1.34) and (1.38)}]{\text{from}} (\mathbf{u}\cdot\mathbf{w})\,\mathbf{v}-(\mathbf{v}\cdot\mathbf{w})\,\mathbf{u} \ .
\end{aligned}$$

The above result will be used to verify the next part. Here, one needs to show that the vectors $(\mathbf{u}\times\mathbf{v})\times\mathbf{w}$ and $\mathbf{u}\times\mathbf{v}$ are orthogonal knowing that the latter itself is perpendicular to both $\mathbf{u}$ and $\mathbf{v}$ (see Fig. 1.4):

$$\begin{aligned}
[(\mathbf{u}\times\mathbf{v})\times\mathbf{w}]\cdot(\mathbf{u}\times\mathbf{v}) &\xlongequal[\text{(1.72)}]{\text{from}} [(\mathbf{u}\cdot\mathbf{w})\,\mathbf{v}-(\mathbf{v}\cdot\mathbf{w})\,\mathbf{u}]\cdot(\mathbf{u}\times\mathbf{v})\\
&\xlongequal[\text{(1.9b) and (1.9c)}]{\text{from}} (\mathbf{u}\cdot\mathbf{w})\,[\mathbf{v}\cdot(\mathbf{u}\times\mathbf{v})]-(\mathbf{v}\cdot\mathbf{w})\,[\mathbf{u}\cdot(\mathbf{u}\times\mathbf{v})]\\
&\xlongequal[\text{(1.49c)}]{\text{from}} (\mathbf{u}\cdot\mathbf{w})\,0-(\mathbf{v}\cdot\mathbf{w})\,0=0 \ .
\end{aligned}$$

Exercise 1.4

First, prove the Jacobi identity (1.70), i.e. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$. Then, verify that

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \ , \tag{1.78a}$$

$$\begin{aligned}(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{b} - [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{a} \\ &= [\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{c} - [\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{d} \ ,\end{aligned} \tag{1.78b}$$

$$\begin{aligned}[(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})] \cdot (\mathbf{e} \times \mathbf{f}) &= [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, [\mathbf{b} \cdot (\mathbf{e} \times \mathbf{f})] \\ &\quad - [\mathbf{a} \cdot (\mathbf{e} \times \mathbf{f})]\, [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})] \ ,\end{aligned} \tag{1.78c}$$

$$\begin{aligned}\mathbf{0} &= [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{a} - [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{b} \\ &\quad + [\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{c} - [\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{d} \ ,\end{aligned} \tag{1.78d}$$

$$[(\mathbf{a} \times \mathbf{b}) \times (\mathbf{b} \times \mathbf{c})] \cdot (\mathbf{c} \times \mathbf{a}) = [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]^2 \ . \tag{1.78e}$$

Solution. The proof is similar to what followed in the previous exercise. The verification will be carried out step by step for each desired result below.

The Jacobi identity:

$$\begin{aligned}&\underbrace{\mathbf{a} \times (\mathbf{b} \times \mathbf{c})}_{= [(\mathbf{a} \cdot \mathbf{c})\, \mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\, \mathbf{c}] , \text{ according to } (1.71)} + \underbrace{\mathbf{b} \times (\mathbf{c} \times \mathbf{a})}_{= [(\mathbf{b} \cdot \mathbf{a})\, \mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\, \mathbf{a}]} + \underbrace{\mathbf{c} \times (\mathbf{a} \times \mathbf{b})}_{= [(\mathbf{c} \cdot \mathbf{b})\, \mathbf{a} - (\mathbf{c} \cdot \mathbf{a})\, \mathbf{b}]} \\ &= \left[\cancel{(\mathbf{a} \cdot \mathbf{c})\, \mathbf{b}} + \cancel{(-\mathbf{a} \cdot \mathbf{b})\, \mathbf{c}} \right] + \left[\cancel{(\mathbf{b} \cdot \mathbf{a})\, \mathbf{c}} + \cancel{(-\mathbf{b} \cdot \mathbf{c})\, \mathbf{a}} \right] \\ &\quad + \left[\cancel{(\mathbf{c} \cdot \mathbf{b})\, \mathbf{a}} + \cancel{(-\mathbf{c} \cdot \mathbf{a})\, \mathbf{b}} \right] \\ &= \mathbf{0} \ .\end{aligned}$$

The identity (1.78a):

$$\begin{aligned}(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &\overset{\text{from}}{\underset{(1.67)}{=\!=}} \underbrace{\left(a_i b_j \varepsilon_{ijk} \widehat{\mathbf{e}}_k\right) \cdot \left(c_l d_m \varepsilon_{lmn} \widehat{\mathbf{e}}_n\right)}_{= \left(a_i b_j c_l d_m\right) \left(\varepsilon_{ijk} \varepsilon_{lmn}\right) \left(\widehat{\mathbf{e}}_k \cdot \widehat{\mathbf{e}}_n\right)} \\ &\overset{\text{from}}{\underset{(1.35)}{=\!=}} \left(a_i b_j c_l d_m\right) \left(\varepsilon_{ijk} \varepsilon_{lmn}\right) \delta_{kn} \\ &\overset{\text{from}}{\underset{(1.36)}{=\!=}} \left(a_i b_j c_l d_m\right) \left(\varepsilon_{ijk} \varepsilon_{lmk}\right) \\ &\overset{\text{from}}{\underset{(1.58a)}{=\!=}} \left(a_i b_j c_l d_m\right) \left(\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}\right) \\ &\overset{\text{from}}{\underset{(1.36)}{=\!=}} a_i c_i b_j d_j - a_i d_i b_j c_j \\ &\overset{\text{from}}{\underset{(1.38)}{=\!=}} (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \ .\end{aligned}$$

Notice that the property (1.49d) - which represents the magnitude of the cross product - can be recovered from the above identity by substituting $\mathbf{c} = \mathbf{a}$ and $\mathbf{d} = \mathbf{b}$.

The identity (1.78b):

$$\text{Recall from (1.72) that } (\mathbf{a} \times \mathbf{b}) \times \mathbf{u} = [\mathbf{a} \cdot (\mathbf{u})]\, \mathbf{b} - [\mathbf{b} \cdot (\mathbf{u})]\, \mathbf{a}$$
$$\text{and let } \mathbf{u} = \mathbf{c} \times \mathbf{d} \; .$$
$$\text{Then, } (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{b} - [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{a} \; .$$

This identity has another representation:

$$\begin{aligned}
(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &\xlongequal[(1.67)]{\text{from}} \left(a_i b_j \varepsilon_{ijk} \widehat{\mathbf{e}}_k\right) \times \left(c_l d_m \varepsilon_{lmn} \widehat{\mathbf{e}}_n\right) \\
&\xlongequal[(1.54) \text{ and } (1.64)]{\text{from}} \left(a_i b_j c_l d_m\right) \left(-\varepsilon_{ijk} \varepsilon_{lmn} \varepsilon_{kon} \widehat{\mathbf{e}}_o\right) \\
&\xlongequal[(1.58\text{a})]{\text{from}} \left(a_i b_j c_l d_m\right) \varepsilon_{ijk} \left(-\delta_{lk}\delta_{mo} + \delta_{lo}\delta_{mk}\right) \widehat{\mathbf{e}}_o \\
&\xlongequal[(1.36)]{\text{from}} \underbrace{-\left(a_i b_j c_k d_m\right) \varepsilon_{ijk} \widehat{\mathbf{e}}_m + \left(a_i b_j c_l d_k\right) \varepsilon_{ijk} \widehat{\mathbf{e}}_l}_{= -\left(a_i b_j c_k \varepsilon_{ijk}\right)\left(d_m \widehat{\mathbf{e}}_m\right) + \left(a_i b_j d_k \varepsilon_{ijk}\right)\left(c_l \widehat{\mathbf{e}}_l\right)} \\
&\xlongequal[(1.74)]{\text{from}} [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{d})]\, \mathbf{c} - [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\, \mathbf{d} \\
&\xlongequal[(1.73)]{\text{from}} [\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{c} - [\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{d} \; .
\end{aligned}$$

The identity (1.78c):

$$\begin{aligned}
\{(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})\} \cdot (\mathbf{e} \times \mathbf{f}) &\xlongequal[(1.78\text{b})]{\text{from}} \{[\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{b} - [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{a}\} \cdot (\mathbf{e} \times \mathbf{f}) \\
&= [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, [\mathbf{b} \cdot (\mathbf{e} \times \mathbf{f})] - [\mathbf{a} \cdot (\mathbf{e} \times \mathbf{f})]\, [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})] \; .
\end{aligned}$$

The identity (1.78d):

$$\begin{aligned}
&\xRightarrow[\text{and } (1.5)]{\text{from } (1.4\text{c})} \{(\mathbf{c} \times \mathbf{d}) \times (\mathbf{a} \times \mathbf{b})\} - \{(\mathbf{c} \times \mathbf{d}) \times (\mathbf{a} \times \mathbf{b})\} = \mathbf{0} \\
&\xRightarrow[(1.49\text{a})]{\text{from}} -\{(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})\} - \{(\mathbf{c} \times \mathbf{d}) \times (\mathbf{a} \times \mathbf{b})\} = \mathbf{0} \\
&\xRightarrow[(1.78\text{b})]{\text{from}} -\{[\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{b} - [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{a}\} - \{[\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{d} - [\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{c}\} = \mathbf{0} \\
&\Rightarrow \{[\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{a} - [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\, \mathbf{b}\} + \{[\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{c} - [\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})]\, \mathbf{d}\} = \mathbf{0} \; .
\end{aligned}$$

The identity (1.78e):

$$
\begin{aligned}
[(\mathbf{a} \times \mathbf{b}) \times (\mathbf{b} \times \mathbf{c})] \cdot (\mathbf{c} \times \mathbf{a}) &\xlongequal[(1.78b)]{\text{from}} [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\, [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})] \\
&\quad - [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{a})]\, [\mathbf{b} \cdot (\mathbf{b} \times \mathbf{c})] \\
&\xlongequal[(1.49c)]{\text{from}} [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\, [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})] \\
&\xlongequal[(1.73)]{\text{from}} [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\, [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})] \\
&= [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]^2 \ .
\end{aligned}
$$

Exercise 1.5

The alternating symbol is not only used for computing the cross product but also helps obtain the determinant of a square matrix $[\,\mathbf{A}\,]$ given in $(1.47)_1$. First, show that

$$\det [\,\mathbf{A}\,] = \underbrace{\varepsilon_{ijk} A_{i1} A_{j2} A_{k3}}_{= \varepsilon_{ijk} A_{1i} A_{2j} A_{3k} = \det \left[\mathbf{A}^{\mathrm{T}}\right]} \ . \tag{1.79}$$

Then, use the above result to obtain the alternative form

$$\det [\,\mathbf{A}\,] = \underbrace{\frac{1}{6} \varepsilon_{ijk} \varepsilon_{lmn} A_{il} A_{jm} A_{kn}}_{= \frac{1}{6} \varepsilon_{ijk} \varepsilon_{lmn} A_{li} A_{mj} A_{nk} = \det \left[\mathbf{A}^{\mathrm{T}}\right]} \ . \tag{1.80}$$

Finally, verify that

$$\frac{\varepsilon_{ijk} \varepsilon_{qrs}}{6} \varepsilon_{lmn} A_{lq} A_{mr} A_{ns} = \varepsilon_{lmn} A_{li} A_{mj} A_{nk} \ . \tag{1.81}$$

Solution. The determinant of a 3 by 3 square matrix

$$
\begin{aligned}
\det [\,\mathbf{A}\,] = A_{11} \left(A_{22} A_{33} - A_{32} A_{23} \right) \\
- A_{21} \left(A_{12} A_{33} - A_{32} A_{13} \right) + A_{31} \left(A_{12} A_{23} - A_{22} A_{13} \right) \ ,
\end{aligned} \tag{1.82}
$$

is expressible in terms of the permutation symbol (1.52) as follows:

$$\det [\,\mathbf{A}\,] = A_{11} \left(\varepsilon_{1jk} A_{j2} A_{k3} \right) + A_{21} \left(\varepsilon_{2jk} A_{j2} A_{k3} \right) + A_{31} \left(\varepsilon_{3jk} A_{j2} A_{k3} \right) \ . \tag{1.83}$$

The above expression by means of the summation convention simply provides the desired result $(1.79)_1$.

To show (1.80)$_1$, first recall from (1.58c)$_2$ that $\varepsilon_{lmn}\varepsilon_{lmn} = 6$. This motivates to multiply both sides of (1.79)$_1$ by ε_{lmn} to arrive at

$$
\begin{aligned}
\det[\mathbf{A}]\,\varepsilon_{lmn} &= \left(A_{i1}A_{j2}A_{k3}\right)\varepsilon_{ijk}\varepsilon_{lmn} \\
&\xlongequal[\text{(1.36) and (1.57)}]{\text{from}} \underbrace{\left(A_{i1}A_{j2}A_{k3}\right)\delta_{il}\left(\delta_{jm}\delta_{kn} - \delta_{km}\delta_{jn}\right)}_{= A_{l1}A_{m2}A_{n3} - A_{l1}A_{n2}A_{m3}} \\
&\quad - \underbrace{\left(A_{i1}A_{j2}A_{k3}\right)\delta_{im}\left(\delta_{jl}\delta_{kn} - \delta_{kl}\delta_{jn}\right)}_{= A_{m1}A_{l2}A_{n3} - A_{m1}A_{n2}A_{l3}} \\
&\quad + \underbrace{\left(A_{i1}A_{j2}A_{k3}\right)\delta_{in}\left(\delta_{jl}\delta_{km} - \delta_{kl}\delta_{jm}\right)}_{= A_{n1}A_{l2}A_{m3} - A_{n1}A_{m2}A_{l3}} \\
&= A_{l1}\left(\varepsilon_{1jk}A_{mj}A_{nk}\right) + A_{l2}\left(\varepsilon_{2jk}A_{mj}A_{nk}\right) + A_{l3}\left(\varepsilon_{3jk}A_{mj}A_{nk}\right) \\
&= \varepsilon_{ijk}\left(A_{li}A_{mj}A_{nk}\right) \ .
\end{aligned}
\tag{1.84}
$$

This result, by means of $\det[\mathbf{A}] = \det[\mathbf{A}]^{\mathrm{T}}$ and $A^{\mathrm{T}}_{uv} = A_{vu}$, can also be represented by

$$
\det[\mathbf{A}]\,\varepsilon_{lmn} = \varepsilon_{ijk}\left(A_{il}A_{jm}A_{kn}\right) \ . \tag{1.85}
$$

By multiplying both sides of (1.85) by ε_{lmn}, taking into account the identity (1.58c)$_2$, one can arrive at the required result (1.80)$_1$.

It should not be difficult now to see that [4]

$$
\begin{aligned}
\frac{\varepsilon_{ijk}\varepsilon_{qrs}}{6}\varepsilon_{lmn}A_{lq}A_{mr}A_{ns} &= \frac{1}{6}\varepsilon_{lmn}A_{lq}A_{mr}A_{ns}\left[\delta_{iq}\left(\delta_{jr}\delta_{ks} - \delta_{kr}\delta_{js}\right)\right. \\
&\quad \left. - \delta_{ir}\left(\delta_{jq}\delta_{ks} - \delta_{kq}\delta_{js}\right) + \delta_{is}\left(\delta_{jq}\delta_{kr} - \delta_{kq}\delta_{jr}\right)\right] \\
&= \frac{1}{6}\varepsilon_{lmn}\left[A_{li}\left(A_{mj}A_{nk} - A_{mk}A_{nj}\right)\right. \quad \begin{array}{l} \longleftarrow \text{note that } -\varepsilon_{lmn}A_{mk}A_{nj} \\ = -\varepsilon_{lnm}A_{nk}A_{mj} = +\varepsilon_{lmn}A_{nk}A_{mj} \end{array} \\
&\quad \left. - A_{mi}\left(A_{lj}A_{nk} - A_{lk}A_{nj}\right) + A_{ni}\left(A_{lj}A_{mk} - A_{lk}A_{mj}\right)\right] \\
&= \varepsilon_{lmn}A_{li}A_{mj}A_{nk} \ .
\end{aligned}
$$

Exercise 1.6

The goal of this exercise is to provide a deeper insight into indicial notation as well as matrix notation which are extensively used, for instance, in nonlinear solid mechanics and numerical procedures such as the finite element method.

For each part given below, express the given variable in index notation as well as matrix notation.

1. $s = a_1 x_1 x_2 + a_2 x_2 x_2 + a_3 x_3 x_2$.

Solution.

$$\text{Index notation:} \quad s = a_i x_i x_2 \ , \tag{1.86a}$$

$$\text{Matrix notation:} \quad s = \begin{bmatrix} a_1 \ a_2 \ a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} x_2 \ . \tag{1.86b}$$

2. $\bar{s} = a_1 x_1 x_1 + a_2 x_2 x_2 + a_3 x_3 x_3$.

Solution.

$$\text{Index notation:} \quad \bar{s} = \sum_{i=1}^{3} a_i x_i^2 \ , \tag{1.87a}$$

$$\text{Matrix notation:} \quad \bar{s} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \ . \tag{1.87b}$$

Note that the matrix form (1.87b) may also be written by

$$\text{Matrix notation:} \quad \bar{s} = \begin{bmatrix} a_1 \ a_2 \ a_3 \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{bmatrix} \ . \tag{1.88}$$

3. $\tilde{s} = \frac{\partial^2 \Psi}{\partial x_1^2} + \frac{\partial^2 \Psi}{\partial x_2^2} + \frac{\partial^2 \Psi}{\partial x_3^2}$.

Solution.

$$\text{Index notation:} \quad \tilde{s} = \frac{\partial^2 \Psi}{\partial x_i \partial x_i} \ , \tag{1.89a}$$

$$\text{Matrix notation:} \quad \tilde{s} = \begin{bmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \end{bmatrix} \begin{bmatrix} \dfrac{\partial \Psi}{\partial x_1} \\ \dfrac{\partial \Psi}{\partial x_2} \\ \dfrac{\partial \Psi}{\partial x_3} \end{bmatrix} \ . \tag{1.89b}$$

4. The total differential df of a function $f = \hat{f}\,(x_1, x_2, x_3)$.

Solution.

$$\text{Index notation:} \quad df = \frac{\partial \hat{f}}{\partial x_i} dx_i \ , \tag{1.90a}$$

$$\text{Matrix notation:} \quad df = \begin{bmatrix} \dfrac{\partial \hat{f}}{\partial x_1} & \dfrac{\partial \hat{f}}{\partial x_2} & \dfrac{\partial \hat{f}}{\partial x_3} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} . \tag{1.90b}$$

5. The differential change in the functions $v_i = \hat{v}_i\,(x_1, x_2, x_3)\,,\ i = 1, 2, 3$, and the sum $dv_i dv_i$.

Solution.

$$\text{Index notation:} \quad dv_i = \frac{\partial \hat{v}_i}{\partial x_j} dx_j \ , \tag{1.91a}$$

$$\text{Matrix notation:} \quad \begin{bmatrix} dv_1 \\ dv_2 \\ dv_3 \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \hat{v}_1}{\partial x_1} & \dfrac{\partial \hat{v}_1}{\partial x_2} & \dfrac{\partial \hat{v}_1}{\partial x_3} \\ \dfrac{\partial \hat{v}_2}{\partial x_1} & \dfrac{\partial \hat{v}_2}{\partial x_2} & \dfrac{\partial \hat{v}_2}{\partial x_3} \\ \dfrac{\partial \hat{v}_3}{\partial x_1} & \dfrac{\partial \hat{v}_3}{\partial x_2} & \dfrac{\partial \hat{v}_3}{\partial x_3} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} . \tag{1.91b}$$

And,

$$\text{Index notation:} \quad dv_i dv_i = \frac{\partial \hat{v}_i}{\partial x_j} \frac{\partial \hat{v}_i}{\partial x_k} dx_j dx_k \ , \tag{1.92a}$$

$$\text{Matrix notation:} \quad dv_i dv_i = \begin{bmatrix} [\mathbf{v}]^{\mathrm{T}} & [\mathbf{v}]^{\mathrm{T}} & [\mathbf{v}]^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{A}_3 \end{bmatrix} \begin{bmatrix} [\mathbf{v}] \\ [\mathbf{v}] \\ [\mathbf{v}] \end{bmatrix} , \tag{1.92b}$$

where $[\mathbf{O}]$ is a 3 by 3 zero matrix and

$$[\mathbf{A}_i] = \begin{bmatrix} \dfrac{\partial \hat{v}_i}{\partial x_1}\dfrac{\partial \hat{v}_i}{\partial x_1} & \dfrac{\partial \hat{v}_i}{\partial x_1}\dfrac{\partial \hat{v}_i}{\partial x_2} & \dfrac{\partial \hat{v}_i}{\partial x_1}\dfrac{\partial \hat{v}_i}{\partial x_3} \\ \dfrac{\partial \hat{v}_i}{\partial x_2}\dfrac{\partial \hat{v}_i}{\partial x_1} & \dfrac{\partial \hat{v}_i}{\partial x_2}\dfrac{\partial \hat{v}_i}{\partial x_2} & \dfrac{\partial \hat{v}_i}{\partial x_2}\dfrac{\partial \hat{v}_i}{\partial x_3} \\ \dfrac{\partial \hat{v}_i}{\partial x_3}\dfrac{\partial \hat{v}_i}{\partial x_1} & \dfrac{\partial \hat{v}_i}{\partial x_3}\dfrac{\partial \hat{v}_i}{\partial x_2} & \dfrac{\partial \hat{v}_i}{\partial x_3}\dfrac{\partial \hat{v}_i}{\partial x_3} \end{bmatrix} , \ [\mathbf{v}] = \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} . \tag{1.93}$$

6. The first- and second-order total differentials, dA_{ij} and d^2A_{ij}, of six independent functions $A_{ij} = \hat{A}_{ij}\,(x_1, x_2, x_3)$ for which $A_{12} = A_{21}\ ,\ A_{23} = A_{32}\ ,\ A_{13} = A_{31}$. Assume that the functions and the respective higher-order derivatives are continuous.

Solution.

$$\text{Index notation:}\quad dA_{ij} = \frac{\partial \hat{A}_{ij}}{\partial x_k} dx_k \ , \tag{1.94a}$$

$$\text{Matrix notation:}\quad \begin{bmatrix} dA_{11} \\ dA_{22} \\ dA_{33} \\ dA_{23} \\ dA_{13} \\ dA_{12} \end{bmatrix} = \begin{bmatrix} \frac{\partial \hat{A}_{11}}{\partial x_1} & \frac{\partial \hat{A}_{11}}{\partial x_2} & \frac{\partial \hat{A}_{11}}{\partial x_3} \\ \frac{\partial \hat{A}_{22}}{\partial x_1} & \frac{\partial \hat{A}_{22}}{\partial x_2} & \frac{\partial \hat{A}_{22}}{\partial x_3} \\ \frac{\partial \hat{A}_{33}}{\partial x_1} & \frac{\partial \hat{A}_{33}}{\partial x_2} & \frac{\partial \hat{A}_{33}}{\partial x_3} \\ \frac{\partial \hat{A}_{23}}{\partial x_1} & \frac{\partial \hat{A}_{23}}{\partial x_2} & \frac{\partial \hat{A}_{23}}{\partial x_3} \\ \frac{\partial \hat{A}_{13}}{\partial x_1} & \frac{\partial \hat{A}_{13}}{\partial x_2} & \frac{\partial \hat{A}_{13}}{\partial x_3} \\ \frac{\partial \hat{A}_{12}}{\partial x_1} & \frac{\partial \hat{A}_{12}}{\partial x_2} & \frac{\partial \hat{A}_{12}}{\partial x_3} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix} . \tag{1.94b}$$

Accordingly,

$$\text{Index notation:}\quad d^2 A_{ij} = \frac{\partial^2 \hat{A}_{ij}}{\partial x_l \partial x_k} dx_l dx_k \ , \tag{1.95a}$$

$$\text{Matrix notation:}\quad \left[d^2 \mathbf{A} \right] = \left[\frac{\partial^2 \hat{\mathbf{A}}}{\partial \mathbf{x} \partial \mathbf{x}} \right] \left[d\mathbf{x} d\mathbf{x} \right] \ , \tag{1.95b}$$

where

$$\left[d^2 \mathbf{A} \right] = \left[d^2 A_{11} \ d^2 A_{22} \ d^2 A_{33} \ d^2 A_{23} \ d^2 A_{13} \ d^2 A_{12} \right]^{\mathrm{T}} \ , \tag{1.96a}$$

$$\left[\frac{\partial^2 \hat{\mathbf{A}}}{\partial \mathbf{x} \partial \mathbf{x}} \right] = \begin{bmatrix} \frac{\partial^2 \hat{A}_{11}}{\partial x_1 \partial x_1} & \frac{\partial^2 \hat{A}_{11}}{\partial x_2 \partial x_2} & \frac{\partial^2 \hat{A}_{11}}{\partial x_3 \partial x_3} & \frac{\partial^2 \hat{A}_{11}}{\partial x_2 \partial x_3} & \frac{\partial^2 \hat{A}_{11}}{\partial x_1 \partial x_3} & \frac{\partial^2 \hat{A}_{11}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \hat{A}_{22}}{\partial x_1 \partial x_1} & \frac{\partial^2 \hat{A}_{22}}{\partial x_2 \partial x_2} & \frac{\partial^2 \hat{A}_{22}}{\partial x_3 \partial x_3} & \frac{\partial^2 \hat{A}_{22}}{\partial x_2 \partial x_3} & \frac{\partial^2 \hat{A}_{22}}{\partial x_1 \partial x_3} & \frac{\partial^2 \hat{A}_{22}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \hat{A}_{33}}{\partial x_1 \partial x_1} & \frac{\partial^2 \hat{A}_{33}}{\partial x_2 \partial x_2} & \frac{\partial^2 \hat{A}_{33}}{\partial x_3 \partial x_3} & \frac{\partial^2 \hat{A}_{33}}{\partial x_2 \partial x_3} & \frac{\partial^2 \hat{A}_{33}}{\partial x_1 \partial x_3} & \frac{\partial^2 \hat{A}_{33}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \hat{A}_{23}}{\partial x_1 \partial x_1} & \frac{\partial^2 \hat{A}_{23}}{\partial x_2 \partial x_2} & \frac{\partial^2 \hat{A}_{23}}{\partial x_3 \partial x_3} & \frac{\partial^2 \hat{A}_{23}}{\partial x_2 \partial x_3} & \frac{\partial^2 \hat{A}_{23}}{\partial x_1 \partial x_3} & \frac{\partial^2 \hat{A}_{23}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \hat{A}_{13}}{\partial x_1 \partial x_1} & \frac{\partial^2 \hat{A}_{13}}{\partial x_2 \partial x_2} & \frac{\partial^2 \hat{A}_{13}}{\partial x_3 \partial x_3} & \frac{\partial^2 \hat{A}_{13}}{\partial x_2 \partial x_3} & \frac{\partial^2 \hat{A}_{13}}{\partial x_1 \partial x_3} & \frac{\partial^2 \hat{A}_{13}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \hat{A}_{12}}{\partial x_1 \partial x_1} & \frac{\partial^2 \hat{A}_{12}}{\partial x_2 \partial x_2} & \frac{\partial^2 \hat{A}_{12}}{\partial x_3 \partial x_3} & \frac{\partial^2 \hat{A}_{12}}{\partial x_2 \partial x_3} & \frac{\partial^2 \hat{A}_{12}}{\partial x_1 \partial x_3} & \frac{\partial^2 \hat{A}_{12}}{\partial x_1 \partial x_2} \end{bmatrix} , \tag{1.96b}$$

$$\left[d\mathbf{x} d\mathbf{x} \right] = \left[(dx_1)^2 \ (dx_2)^2 \ (dx_3)^2 \ 2dx_2 dx_3 \ 2dx_1 dx_3 \ 2dx_1 dx_2 \right]^{\mathrm{T}} . \tag{1.96c}$$

The representations (1.94b) and (1.95b) are very suitable for implementations in computer codes. Observe that how the number of free indices is decreased in these representations. As can be seen, $d^2 A_{ij}$ with two indices that could be set in a matrix is now collected in an object having only one index. Moreover, observe that a variable of the form $\left(\partial^2 \hat{\mathbf{A}}/\partial\mathbf{x}\partial\mathbf{x}\right)_{ijkl} = \partial^2 \hat{A}_{ij}/\partial x_k \partial x_l$, with four indices in (1.95a), which possesses the following *minor symmetries*

$$\underbrace{\frac{\partial^2 \hat{A}_{ij}}{\partial x_k \partial x_l} = \frac{\partial^2 \hat{A}_{ji}}{\partial x_k \partial x_l}}_{\substack{\text{symmetry in the first two indices} \\ \text{is called minor (left) symmetries}}} \quad , \quad \underbrace{\frac{\partial^2 \hat{A}_{ij}}{\partial x_k \partial x_l} = \frac{\partial^2 \hat{A}_{ij}}{\partial x_l \partial x_k}}_{\substack{\text{symmetry in the last two indices} \\ \text{is termed minor (right) symmetries}}} \quad , \tag{1.97}$$

is now expressed as a matrix in (1.96b) possessing two indices. See Sect. 3.2.4 for more elaborations.

References

1. Synge, J. L., & Schild, A. (1956). *Tensor calculus*. Toronto: University of Toronto Press.
2. McConnell, A. J. (1957). *Applications of tensor analysis*. New York: Dover Publications.
3. Sokolnikoff, I. S. (1964). *Tensor analysis: Theory and applications to geometry and mechanics of continua*. Hoboken: Wiley.
4. Borisenko, A. I., & Tarapov, I. E. (1968). *Vector and tensor analysis with applications*. New York: Dover Publications.
5. Akivis, M. A., & Goldberg, V. V. (1972). *An introduction to linear algebra and tensors*. New York: Dover Publications.
6. Flügge, W. (1972). *Tensor analysis and continuum mechanics*. Berlin, Heidelberg: Springer.
7. Danielson, D. A. (1992). *Vectors and tensors in engineering and physics*. Boston: Addison-Wesley.
8. Simmonds, J. G. (1994). *A brief on tensor analysis* (2nd ed.). New York: Springer.
9. Heinbockel, J. H. (2001). *Introduction to tensor calculus and continuum mechanics*. Bloomington: Trafford Publishing.
10. Dimitrienko, Y. I. (2002). *Tensor analysis and nonlinear tensor functions*. Dordrecht: Springer.
11. Talpaert, Y. R. (2002). *Tensor analysis and continuum mechanics*. Dordrecht: Springer.
12. Akivis, M. A., & Goldberg, V. V. (2003). *Tensor calculus with applications*. Singapore: World Scientific Publishing.
13. Lebedev, L. P., & Cloud, M. J. (2003). *Tensor analysis*. Singapore: World Scientific Publishing.
14. Spain, B. (2003). *Tensor calculus: A concise course*. Mineola: Dover Publications.
15. Ruíz-Tolosa, J. R., & Castillo, E. (2005). *From vectors to tensors*. Berlin, Heidelberg: Springer.
16. Bowen, R. M., & Wang, C.-C. (2009). *Introduction to vectors and tensors: Second edition-two volumes bound as one*. Mineola: Dover Publications.
17. Schade, H., & Neemann, K. (2009). *Tensoranalysis*. Berlin, New York: Walter de Gruyter.
18. Schouten, J. A. (2011). *Tensor analysis for physicists* (2nd ed.). Mineola: Dover Publications.
19. Nayak, P. K. (2012). *Textbook of tensor calculus and differential geometry*. New Delhi: PHI Learning.
20. Itskov, M. (2015). *Tensor algebra and tensor analysis for engineers: With applications to continuum mechanics* (4th ed.). Cham: Springer.
21. Irgens, F. (2019). *Tensor analysis*. Cham: Springer.

Chapter 2
Algebra of Tensors

Vector algebra was briefly discussed in the previous chapter not only to introduce the concept of vector and represent its important relationships but also as a primary mathematics for an introduction to *tensor algebra*. Consistent with vectors, *second-order tensors* or simply *tensors* are geometric objects that aim at describing linear relation between vectors. These mathematical entities are constructed from vectors in a certain way. They are designated here by capital and some specific small boldface Latin and Greek letters. Some familiar examples include strain and stress tensors in continuum mechanics, conductivity tensor in electromagnetic field theory and curvature tensor in differential geometry.

2.1 Tensor as a Linear Transformation

There is no unified definition of tensor in the literature. In this textbook, the term tensor should be thought of as a **linear transformation** (or **linear map** or **linear mapping** or **linear function**) from the oriented three-dimensional Euclidean vector space to the same space. Denoting by $\mathscr{T}_{\text{so}}$ the set of all these linear transformations,[1] a linear mapping $\mathbf{A} \in \mathscr{T}_{\text{so}}$ assigns to each vector $\mathbf{u} \in \mathscr{E}_{\text{r}}^{\text{o3}}$ generally another vector $\mathbf{v} \in \mathscr{E}_{\text{r}}^{\text{o3}}$ according to

$$\boxed{\mathbf{v} = \mathbf{A}\mathbf{u}\ ,} \tag{2.1}$$

and satisfies the linearity requirement

$$\left.\begin{aligned} \mathbf{A}(\mathbf{u} + \mathbf{v}) &= \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{v} \\ \mathbf{A}(\alpha\mathbf{u}) &= \alpha(\mathbf{A}\mathbf{u}) \end{aligned}\right\} \quad \text{or} \quad \boxed{\mathbf{A}(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha(\mathbf{A}\mathbf{u}) + \beta(\mathbf{A}\mathbf{v})\ ,} \tag{2.2}$$

[1] In the literature, the set $\mathscr{T}_{\text{so}}$ is often denoted by Lin.

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_2

for all scalars $\alpha, \beta \in \mathbb{R}$. The vectorial variable $\mathbf{Au}$ can be termed *right mapping* since the tensor $\mathbf{A}$ is postmultiplied by the vector $\mathbf{u}$. In fact, $\mathbf{A}$ operates on $\mathbf{u}$ to generate $\mathbf{v}$. It is important to note that similarly to vectors, and as can be seen from (2.1), tensors are **independent** of any coordinate system. Indeed, they are coordinate free entities.

For $\alpha = 1$, $\beta = -1$ and $\mathbf{u} = \mathbf{v}$, the introduced linear transformation gives $\mathbf{A}\,(\,1\mathbf{u} + (-1)\,\mathbf{u}\,) = 1\,(\mathbf{Au}) + (-1)\,(\mathbf{Au})$. This with the aid of the identity (f) in (1.76) and the rule (1.4e) yields $\mathbf{A}\,(\,\mathbf{u} + (-\mathbf{u})\,) = \mathbf{Au} + (-\mathbf{Au})$. This result along with the axiom (1.4c) delivers

$$\boxed{\mathbf{A0} = \mathbf{0}\ .} \tag{2.3}$$

The following definitions are required for the subsequent developments:

Zero tensor. The *zero tensor* $\mathbf{O}$ is defined by

$$\boxed{\mathbf{Ou} = \mathbf{0}\ , \quad \text{for all } \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o}}3.} \tag{2.4}$$

Unit tensor. The *unit* (or *identity*) tensor $\mathbf{I}$ is extensively used in tensor algebra and calculus. It is a special linear transformation with identical input and output, that is,

$$\boxed{\mathbf{Iu} = \mathbf{u}\ , \quad \text{for all } \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}\ .} \tag{2.5}$$

Equal tensors. Two tensors $\mathbf{A}$ and $\mathbf{B}$ are said to be *equal* if they identically transform all vectors $\mathbf{u}$, more precisely,

$$\boxed{\mathbf{A} = \mathbf{B} \quad \text{if and only if} \quad \mathbf{Au} = \mathbf{Bu}\ .} \tag{2.6}$$

or, equivalently,[2]

$$\boxed{\mathbf{A} = \mathbf{B} \quad \text{if and only if} \quad \mathbf{v} \cdot \mathbf{Au} = \mathbf{v} \cdot \mathbf{Bu}\ .} \tag{2.7}$$

This equality condition plays a major role in verifying many upcoming identities.

By appealing to the operations of addition and multiplication for a vector space, the **sum**, $\mathbf{A} + \mathbf{B}$, of two tensors $\mathbf{A}$ and $\mathbf{B}$ as well as the **scalar multiplication**, $\alpha\mathbf{A}$, of a tensor $\mathbf{A}$ by a scalar number α are defined via the following rules

[2] It is not difficult to show that $\mathbf{Au} = \mathbf{Bu}$ can be deduced from $\mathbf{v} \cdot \mathbf{Au} = \mathbf{v} \cdot \mathbf{Bu}$. By considering the inner product as a symmetric bilinear form and setting $\mathbf{v} = \mathbf{Au} - \mathbf{Bu}$, one will have $|\mathbf{Au} - \mathbf{Bu}| = 0$. This result along with $(1.9\text{d})_2$ implies that $\mathbf{Au} = \mathbf{Bu}$.

$$(\mathbf{A}+\mathbf{B})\,\mathbf{u}=\mathbf{A}\mathbf{u}+\mathbf{B}\mathbf{u}\;, \tag{2.8a}$$

$$(\alpha\mathbf{A})\,\mathbf{u}=\alpha\,(\mathbf{A}\mathbf{u})=\mathbf{A}\,(\alpha\mathbf{u})\;\;, \tag{2.8b}$$

for all $\mathbf{A}\,,\;\mathbf{B}\in\mathscr{T}_{\text{so}}$, $\mathbf{u}\in\mathscr{E}_{\text{r}}^{\text{o3}}$ and $\alpha\in\text{R}$.

For convenience, the difference between two tensors **A** and **B** is indicated by

$$\mathbf{A}+(-\mathbf{B})=\mathbf{A}-\mathbf{B}\;. \quad \longleftarrow \text{see (1.5)} \tag{2.9}$$

It is not then difficult to deduce that

$$(\mathbf{A}-\mathbf{B})\,\mathbf{u}=\mathbf{A}\mathbf{u}-\mathbf{B}\mathbf{u}\;. \tag{2.10}$$

The rules (2.8a) and (2.8b) states that the set $\mathscr{T}_{\text{so}}$ remains closed under addition and scalar multiplication. This immediately implies that the set $\mathscr{T}_{\text{so}}$ of all tensors constitutes a new vector space over the field of real numbers, if the rules (1.4a) to (1.4h) can be applied to its elements. In this regard, the goal is to consistently rewrite the properties of vector space for the problem at hand.

Let's first see whether the axiom (1.4a) is applicable here or not, i.e. a property of the form $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ is valid or not. This can be verified according to

$$\left.\begin{aligned}(\mathbf{A}+\mathbf{B})\,\mathbf{u}&\overset{\text{from}}{\underset{(2.8\text{a})}{=\!=\!=}}\mathbf{A}\mathbf{u}+\mathbf{B}\mathbf{u}\\&\overset{\text{from}}{\underset{(1.4\text{a})\text{ and }(2.1)}{=\!=\!=\!=\!=}}\mathbf{B}\mathbf{u}+\mathbf{A}\mathbf{u}\\&\overset{\text{from}}{\underset{(2.8\text{a})}{=\!=\!=}}(\mathbf{B}+\mathbf{A})\,\mathbf{u}\end{aligned}\right\}\;\overset{\text{from}}{\underset{(2.6)}{\Longrightarrow}}\;\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}\;.$$

The second property, i.e. $(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$ can be verified by following a similar procedure.

The third property, i.e. $\mathbf{A}+(-\mathbf{A})=\mathbf{O}$, can be shown as follows:

$$\left.\begin{aligned}\mathbf{0}&\overset{\text{from}}{\underset{(2.3)}{=\!=\!=}}\mathbf{A}\mathbf{0}\\&\overset{\text{from}}{\underset{(1.4\text{c})}{=\!=\!=}}\mathbf{A}\,(\,\mathbf{u}+(-\mathbf{u})\,)\\&\overset{\text{from}}{\underset{(\text{f})\text{ in }(1.76)\text{ and }(2.8\text{a})\text{-}(2.10)}{=\!=\!=\!=\!=\!=\!=\!=}}(\,\mathbf{A}+(-\mathbf{A})\,)\,\mathbf{u}\end{aligned}\right\}\;\overset{\text{from}}{\underset{(2.4)}{\Longrightarrow}}\;\mathbf{A}+(-\mathbf{A})=\mathbf{O}\;.$$

The last property considered here regards $\mathbf{A}+\mathbf{O}=\mathbf{A}$. This can be verified as follows:

$$\left.\begin{aligned}\mathbf{Au} &\overset{\text{from}}{\underset{(1.4\text{d})}{=\!=}} \mathbf{Au}+\mathbf{0} \\ &\overset{\text{from}}{\underset{(2.4)}{=\!=}} \mathbf{Au}+\mathbf{Ou} \\ &\overset{\text{from}}{\underset{(2.8\text{a})}{=\!=}} (\mathbf{A}+\mathbf{O})\,\mathbf{u}\end{aligned}\right\} \overset{\text{from}}{\underset{(2.6)}{\Longrightarrow}} \mathbf{A}+\mathbf{O}=\mathbf{A}\ .$$

The remaining properties can be verified in a straightforward manner. The set $\mathscr{T}_{\text{so}}$ should thus be recognized as a new vector space. In summary, the following properties hold:

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}\ , \tag{2.11a}$$

$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})\ , \tag{2.11b}$$

$$\mathbf{A}+(-\mathbf{A})=\mathbf{O}\ , \tag{2.11c}$$

$$\mathbf{A}+\mathbf{O}=\mathbf{A}\ , \tag{2.11d}$$

$$1\mathbf{A}=\mathbf{A}\ , \tag{2.11e}$$

$$(\alpha\beta)\,\mathbf{A}=\alpha\,(\beta\mathbf{A})\ , \tag{2.11f}$$

$$(\alpha+\beta)\,\mathbf{A}=\alpha\mathbf{A}+\beta\mathbf{A}\ , \tag{2.11g}$$

$$\alpha\,(\mathbf{A}+\mathbf{B})=\alpha\mathbf{A}+\alpha\mathbf{B}\ . \tag{2.11h}$$

It is then a simple exercise to represent the tensorial analogues of the identities (1.76) as

$$\left.\begin{aligned}&\text{(a)}\quad \alpha\mathbf{O}=\mathbf{O}\\&\text{(b)}\quad 0\mathbf{A}=\mathbf{O}\end{aligned}\right\},\quad \left.\begin{aligned}&\text{(c)}\quad 0\mathbf{O}=\mathbf{O}\\&\text{(d)}\quad \mathbf{O}=-\mathbf{O}\end{aligned}\right\},\quad \left.\begin{aligned}&\text{(e)}\quad (-\alpha)\,\mathbf{A}=\alpha\,(-\mathbf{A})\\&\text{(f)}\quad (-1)\,\mathbf{A}=-\mathbf{A}\end{aligned}\right\}. \tag{2.12}$$

2.2 Tensor Product and Representation

The *tensor* (or *direct* or *dyadic*) *product* of the vectors $\mathbf{u} \in \mathscr{E}_{\text{r}}^{\text{o3}}$ and $\mathbf{v} \in \mathscr{E}_{\text{r}}^{\text{o3}}$, is designated by $\mathbf{u} \otimes \mathbf{v}$ (or sometimes $\mathbf{uv}$). The **dyad** $\mathbf{u} \otimes \mathbf{v}$ is a linear transformation that maps any vector $\mathbf{w} \in \mathscr{E}_{\text{r}}^{\text{o3}}$ onto a scalar multiple of $\mathbf{u}$ by the rule

$$\underbrace{(\mathbf{u}\otimes\mathbf{v})\,\mathbf{w}=(\mathbf{v}\cdot\mathbf{w})\,\mathbf{u}}_{\text{using (1.9a), one can also write }(\mathbf{u}\otimes\mathbf{v})\,\mathbf{w}=(\mathbf{u}\otimes\mathbf{w})\,\mathbf{v}}\ , \tag{2.13}$$

where $\otimes$ designates the tensor product. Note that the above rule eventually represents a **right mapping** in alignment with (2.1).

The map $\boldsymbol{g}_{\text{so}}^{\text{tp}}(\mathbf{u}, \mathbf{v}) = \mathbf{u} \otimes \mathbf{v}$ from $\mathscr{E}_{\text{r}}^{\text{o3}} \times \mathscr{E}_{\text{r}}^{\text{o3}}$ to $\mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ truly represents a tensor because when it acts on $\mathbf{w}$, a vector (in the direction of $\mathbf{u}$) will be generated and also it fulfills the linearity condition

$$\begin{aligned}(\mathbf{u}\otimes\mathbf{v})\left(\alpha\mathbf{w}+\beta\bar{\mathbf{w}}\right) &\xlongequal[(2.13)]{\text{from}} \left[\mathbf{v}\cdot\left(\alpha\mathbf{w}+\beta\bar{\mathbf{w}}\right)\right]\mathbf{u} \\ &\xlongequal[(1.9a)\text{ to }(1.9c)]{\text{from}} \left[\alpha\left(\mathbf{v}\cdot\mathbf{w}\right)+\beta\left(\mathbf{v}\cdot\bar{\mathbf{w}}\right)\right]\mathbf{u} \\ &\xlongequal[(1.4f)\text{ and }(1.4g)]{\text{from}} \alpha\left[\left(\mathbf{v}\cdot\mathbf{w}\right)\mathbf{u}\right]+\beta\left[\left(\mathbf{v}\cdot\bar{\mathbf{w}}\right)\mathbf{u}\right] \\ &\xlongequal[(2.13)]{\text{from}} \alpha\left(\mathbf{u}\otimes\mathbf{v}\right)\mathbf{w}+\beta\left(\mathbf{u}\otimes\mathbf{v}\right)\bar{\mathbf{w}}\ . \end{aligned} \tag{2.14}$$

The bilinearity of $\boldsymbol{g}^{\text{tp}}_{\text{so}}$ is then implied via the relations

$$\begin{aligned}\left[\left(\alpha\mathbf{u}+\beta\mathbf{v}\right)\otimes\mathbf{w}\right]\bar{\mathbf{w}} &\xlongequal[(2.13)]{\text{from}} \left(\mathbf{w}\cdot\bar{\mathbf{w}}\right)\left(\alpha\mathbf{u}+\beta\mathbf{v}\right) \\ &\xlongequal[(1.4f)\text{ and }(1.4h)]{\text{from}} \alpha\left(\left(\mathbf{w}\cdot\bar{\mathbf{w}}\right)\mathbf{u}\right)+\beta\left(\left(\mathbf{w}\cdot\bar{\mathbf{w}}\right)\mathbf{v}\right) \\ &\xlongequal[(2.13)]{\text{from}} \alpha\left(\mathbf{u}\otimes\mathbf{w}\right)\bar{\mathbf{w}}+\beta\left(\mathbf{v}\otimes\mathbf{w}\right)\bar{\mathbf{w}} \\ &\xlongequal[(2.8a)\text{ and }(2.8b)]{\text{from}} \left[\alpha\left(\mathbf{u}\otimes\mathbf{w}\right)+\beta\left(\mathbf{v}\otimes\mathbf{w}\right)\right]\bar{\mathbf{w}}\ , \end{aligned} \tag{2.15a}$$

$$\begin{aligned}\left[\mathbf{u}\otimes\left(\alpha\mathbf{v}+\beta\mathbf{w}\right)\right]\bar{\mathbf{w}} &\xlongequal[(2.13)]{\text{from}} \left[\left(\alpha\mathbf{v}+\beta\mathbf{w}\right)\cdot\bar{\mathbf{w}}\right]\mathbf{u} \\ &\xlongequal[(1.10)]{\text{from}} \left[\alpha\left(\mathbf{v}\cdot\bar{\mathbf{w}}\right)+\beta\left(\mathbf{w}\cdot\bar{\mathbf{w}}\right)\right]\mathbf{u} \\ &\xlongequal[(1.4f)\text{ and }(1.4g)]{\text{from}} \alpha\left[\left(\mathbf{v}\cdot\bar{\mathbf{w}}\right)\mathbf{u}\right]+\beta\left[\left(\mathbf{w}\cdot\bar{\mathbf{w}}\right)\mathbf{u}\right] \\ &\xlongequal[(2.13)]{\text{from}} \alpha\left[\left(\mathbf{u}\otimes\mathbf{v}\right)\bar{\mathbf{w}}\right]+\beta\left[\left(\mathbf{u}\otimes\mathbf{w}\right)\bar{\mathbf{w}}\right] \\ &\xlongequal[(2.8a)\text{ and }(2.8b)]{\text{from}} \left[\alpha\left(\mathbf{u}\otimes\mathbf{v}\right)+\beta\left(\mathbf{u}\otimes\mathbf{w}\right)\right]\bar{\mathbf{w}}\ , \end{aligned} \tag{2.15b}$$

since, by means of (2.6), one can deduce that

$$\boldsymbol{g}^{\text{tp}}_{\text{so}}\left(\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{w}\right)=\left(\alpha\mathbf{u}+\beta\mathbf{v}\right)\otimes\mathbf{w}=\alpha\left(\mathbf{u}\otimes\mathbf{w}\right)+\beta\left(\mathbf{v}\otimes\mathbf{w}\right)\ , \tag{2.16a}$$

$$\boldsymbol{g}^{\text{tp}}_{\text{so}}\left(\mathbf{u},\alpha\mathbf{v}+\beta\mathbf{w}\right)=\mathbf{u}\otimes\left(\alpha\mathbf{v}+\beta\mathbf{w}\right)=\alpha\left(\mathbf{u}\otimes\mathbf{v}\right)+\beta\left(\mathbf{u}\otimes\mathbf{w}\right)\ . \tag{2.16b}$$

For the two Cartesian vectors $\mathbf{u}=u_i\widehat{\mathbf{e}}_i$ and $\mathbf{v}=v_j\widehat{\mathbf{e}}_j$ with the standard basis $\{\widehat{\mathbf{e}}_i\}$, the collection

$$\left\{\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j,\ i,j=1,2,3\right\}\xlongequal{\text{def}}\left\{\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right\}\ , \tag{2.17}$$

constitutes a basis for $\mathscr{T}_{\rm so}\left(\mathscr{E}_{\rm r}^{\rm o3}\right)$.[3] Hence, the dimension of this new vector space will be $\dim \mathscr{T}_{\rm so} = \left(\dim \mathscr{E}_{\rm r}^{\rm o3}\right)^2 = 3^2$.

It should be noted that an arbitrary tensor in $\mathscr{T}_{\rm so}\left(\mathscr{E}_{\rm r}^{\rm o3}\right)$ is not necessarily of the single form $\mathbf{u} \otimes \mathbf{v}$ but also a linear combination of dyads such as $\alpha\,(\mathbf{u} \otimes \mathbf{v}) + \beta\,(\mathbf{w} \otimes \bar{\mathbf{w}})$, called *dyadic*. This reveals the fact that a dyadic, in general, may not be expressed in terms of only one tensor product. Analogous to vectors, a general element of this space can be expressed in terms of a (given) basis and collection of 3 by 3 scalar numbers. For instance, a tensor $\mathbf{A}$ can be represented with respect to the **basis tensors** (2.17) by

$$\begin{aligned}\mathbf{A} = & A_{11}\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_1 + A_{12}\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_2 + A_{13}\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_3 \\ & + (A_{21}\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_1 + A_{22}\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_2 + A_{23}\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_3) \\ & + [A_{31}\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_1 + A_{32}\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_2 + A_{33}\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_3] \ ,\end{aligned} \tag{2.18}$$

or by means of the summation convention as

$$\boxed{\mathbf{A} = A_{1j}\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_j + \left(A_{2j}\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_j\right) + \left[A_{3j}\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_j\right] = A_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ .} \tag{2.19}$$

Here, $\mathbf{A}$ is called a **Cartesian** (or **rectangular**) **tensor**[4] constructed from the rectangular components A_{ij} as well as the Cartesian basis $\left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right\}$. The components A_{ij} are obtained via the following expression

$$\boxed{A_{ij} = (\mathbf{A})_{ij} = \widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}\widehat{\mathbf{e}}_j\right] \ ,} \tag{2.20}$$

wherein

$$\boxed{\mathbf{A}\widehat{\mathbf{e}}_j = A_{kj}\widehat{\mathbf{e}}_k \ .} \tag{2.21}$$

One way of showing $(2.20)_2$ will be[5]

[3] The result also holds true for any vector $\mathbf{u}$ in a finite-dimensional vector space and another vector $\mathbf{v}$, in a not necessarily identical space, with any associative basis vectors.

[4] Recall that a tensor depicts a mathematical entity that is independent of any coordinate system. It is frequently seen in the literature that tensors are classified into Cartesian, covariant, contravariant, mixed contra-covariant and co-contravariant. However, this classification may not be true since it is precisely the respective components that have been expressed with respect to a specific coordinate system.

[5] Similar procedure can be used to alternatively derive $(2.20)_2$. This can be shown as follows:
$$\begin{aligned}\widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}\widehat{\mathbf{e}}_j\right] &= \widehat{\mathbf{e}}_i \cdot \left[\left(A_{kl}\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right)\widehat{\mathbf{e}}_j\right] = \widehat{\mathbf{e}}_i \cdot \left[A_{kl}\widehat{\mathbf{e}}_k\left(\widehat{\mathbf{e}}_l \cdot \widehat{\mathbf{e}}_j\right)\right] = \widehat{\mathbf{e}}_i \cdot \left[A_{kl}\widehat{\mathbf{e}}_k\left(\delta_{lj}\right)\right] = \widehat{\mathbf{e}}_i \cdot \left[A_{kj}\widehat{\mathbf{e}}_k\right] \\ &= A_{kj}\left[\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_k\right] = A_{kj}\left[\delta_{ik}\right] = A_{ij} = (\mathbf{A})_{ij} \ .\end{aligned}$$

$$
\begin{aligned}
A_{ij} &\xlongequal[\text{(1.36)}]{\text{from}} A_{kj} \left[\delta_{ik} \right] \\
&\xlongequal[\text{(1.35)}]{\text{from}} A_{kj} \left[\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_k \right] \\
&\xlongequal[\text{(1.9a)-(1.9c)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[A_{kj} \widehat{\mathbf{e}}_k \right] \\
&\xlongequal[\text{(1.36)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[A_{kl} \widehat{\mathbf{e}}_k \left(\delta_{lj} \right) \right] \\
&\xlongequal[\text{(1.35)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[A_{kl} \widehat{\mathbf{e}}_k \left(\widehat{\mathbf{e}}_l \cdot \widehat{\mathbf{e}}_j \right) \right] \\
&\xlongequal[\text{(2.8a), (2.8b) and (2.13)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\left(A_{kl} \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \right) \widehat{\mathbf{e}}_j \right] \\
&\xlongequal[\text{(2.19)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\mathbf{A} \widehat{\mathbf{e}}_j \right] \ .
\end{aligned}
$$

The expressions $(2.19)_2$ and $(2.20)_2$ now help obtain the (Cartesian) coordinate representation of the linear mapping (2.1) as

$$
\boxed{\mathbf{v} = (\mathbf{A})_{ij} \, (\mathbf{u})_j \, \widehat{\mathbf{e}}_i \quad \text{with} \quad (\mathbf{v})_i = (\mathbf{A}\mathbf{u})_i = (\mathbf{A})_{ij} \, (\mathbf{u})_j \ ,} \tag{2.22}
$$

since

$$
\mathbf{v} \xlongequal[\text{(2.1)}]{\text{from}} \mathbf{A}\mathbf{u} \xlongequal[\text{(1.34)}]{\text{from}} \mathbf{A} \left[u_j \widehat{\mathbf{e}}_j \right] \xlongequal[\text{(2.2)}]{\text{from}} u_j \mathbf{A} \widehat{\mathbf{e}}_j \ ,
$$

helps obtain

$$
v_i \xlongequal[\text{(1.33)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \mathbf{v} \xlongequal[\text{the above result}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[u_j \mathbf{A} \widehat{\mathbf{e}}_j \right] \xlongequal[\text{(1.9a)-(1.9c)}]{\text{from}} \left[\widehat{\mathbf{e}}_i \cdot \mathbf{A} \widehat{\mathbf{e}}_j \right] u_j \xlongequal[\text{(2.20)}]{\text{from}} A_{ij} u_j \ .
$$

From now on, every second-order tensor in this text should be realized as a Cartesian tensor, if not otherwise stated. Accordingly, the special unit tensor $\mathbf{I}$ can be represented by

$$
\left.
\begin{aligned}
\mathbf{I}\mathbf{u} &\xlongequal[\text{(2.5)}]{\text{from}} \mathbf{u} \\
&\xlongequal[\text{(1.34)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \cdot \mathbf{u} \right) \widehat{\mathbf{e}}_i \\
&\xlongequal[\text{(2.13)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_i \right) \mathbf{u} \\
&\xlongequal[\text{(1.36)}]{\text{from}} \left(\delta_{ij} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \right) \mathbf{u}
\end{aligned}
\right\}
\xrightarrow[\text{(2.6)}]{\text{from}} \quad \mathbf{I} = \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_i = \delta_{ij} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ . \tag{2.23}
$$

Knowing that generally $A_{ij} \neq u_i v_j$, the map $\boldsymbol{g}^{\text{tp}}_{\text{so}} \left(\mathbf{u}, \mathbf{v} \right) = \mathbf{u} \otimes \mathbf{v}$ now admits the representation

$$
\begin{aligned}
\mathbf{u} \otimes \mathbf{v} &\xlongequal[\text{(1.34)}]{\text{from}} (u_i \widehat{\mathbf{e}}_i) \otimes \left(v_j \widehat{\mathbf{e}}_j\right) \\
&\xlongequal[\text{(2.16a) and (2.16b)}]{\text{from}} u_i v_j \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \\
&= (\mathbf{u})_i \, (\mathbf{v})_j \, \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \quad \text{with} \quad (\mathbf{u} \otimes \mathbf{v})_{ij} = u_i v_j = (\mathbf{u})_i \, (\mathbf{v})_j \ .
\end{aligned} \tag{2.24}
$$

The importance of dyadic product in tensor algebra should now be clear; a powerful tool that enables one to generate tensors from the given vectors. One can finally deduce that:

> Tensors and their relationships naturally carry the mathematical characteristics of vectors.

2.3 Tensor Operations

2.3.1 Composition

The *composition* (or *dot product*) of two tensors **A** and **B**, designated by **AB**, is again a tensor satisfying

$$
\boxed{(\mathbf{AB})\,\mathbf{u} = \mathbf{A}\,(\mathbf{Bu}) \ , \quad \text{for all } \mathbf{A}, \mathbf{B} \in \mathscr{T}_{\text{so}} \text{ and } \mathbf{u} \in \mathscr{E}_{\text{r}}^{\text{o3}} \ .} \tag{2.25}
$$

This operation is extensively utilized in tensor identities and generally does not have the **commutative property**, i.e. $\mathbf{AB} \neq \mathbf{BA}$. In what follows, the goal is to characterize how **A** and **B** interact in a coordinate system. One thus needs to write the above expression in indicial notation. By means of $(2.22)_3$, the direct form **Bu** now becomes $(\mathbf{B})_{kj}\,(\mathbf{u})_j$ and, therefore, the tensorial form $\mathbf{A}\,(\mathbf{Bu})$ takes the coordinate representation $(\mathbf{A})_{ik}\,(\mathbf{B})_{kj}\,(\mathbf{u})_j$. Accordingly, the left hand side of (2.25) renders $((\mathbf{AB})\,\mathbf{u})_i = (\mathbf{AB})_{ij}\,(\mathbf{u})_j$. Comparing the results, by considering (2.6), reveals that

$$
\boxed{(\mathbf{AB})_{ij} = (\mathbf{A})_{ik}\,(\mathbf{B})_{kj} \ .} \tag{2.26}
$$

The interested reader may want to arrive at (2.26) in an alternatively way:

$$
\begin{aligned}
(\mathbf{AB})_{ij} &\xlongequal[\text{(2.20)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[(\mathbf{AB})\,\widehat{\mathbf{e}}_j\right] \\
&\xlongequal[\text{(2.25)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}\left(\mathbf{B}\widehat{\mathbf{e}}_j\right)\right] \\
&\xlongequal[\text{(2.21)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}\left(B_{kj}\widehat{\mathbf{e}}_k\right)\right]
\end{aligned}
$$

$$\overset{\text{from}}{\underset{(2.2)}{=\!=}} \hat{\mathbf{e}}_i \cdot \left[B_{kj} \left(\mathbf{A}\hat{\mathbf{e}}_k \right) \right]$$

$$\overset{\text{from}}{\underset{(1.9a)\text{ to }(1.9c)}{=\!=\!=\!=}} \left[\hat{\mathbf{e}}_i \cdot \left(\mathbf{A}\hat{\mathbf{e}}_k \right) \right] B_{kj}$$

$$\overset{\text{from}}{\underset{(2.20)}{=\!=}} \left[A_{ik} \right] B_{kj} \ .$$

Accordingly,

$$\boxed{\mathbf{AB} = \left(\mathbf{AB} \right)_{ij} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j = \left(\mathbf{A} \right)_{ik} \left(\mathbf{B} \right)_{kj} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \ .} \tag{2.27}$$

In a similar manner, the product of the three tensors **A**, **B** and **C** obeys

$$\underbrace{\mathbf{ABC} = \left(\mathbf{AB} \right) \mathbf{C} = \mathbf{A} \left(\mathbf{BC} \right)}_{\text{in index notation : } (\mathbf{ABC})_{ij} = (\mathbf{A})_{im} (\mathbf{B})_{mn} (\mathbf{C})_{nj}} \quad . \quad \leftarrow \overset{\text{and also}}{(\mathbf{ABCD})_{ij} = (\mathbf{A})_{im} (\mathbf{B})_{mn} (\mathbf{C})_{no} (\mathbf{D})_{oj}} \tag{2.28}$$

It is easy to see that the dot product of two tensors is a **bilinear form** (not the symmetric one owing to $\mathbf{AB} \neq \mathbf{BA}$), that is,

$$\mathbf{A} \left(\alpha \mathbf{B} + \beta \mathbf{C} \right) = \alpha \mathbf{AB} + \beta \mathbf{AC} \quad , \quad \left(\alpha \mathbf{A} + \beta \mathbf{C} \right) \mathbf{B} = \alpha \mathbf{AB} + \beta \mathbf{CB} \ . \tag{2.29}$$

The result (2.27) is in alignment with the following rule

$$\boxed{\left(\mathbf{u} \otimes \mathbf{v} \right) \left(\mathbf{w} \otimes \bar{\mathbf{w}} \right) = \left(\mathbf{v} \cdot \mathbf{w} \right) \mathbf{u} \otimes \bar{\mathbf{w}} \ ,} \tag{2.30}$$

since

$$\begin{aligned} \mathbf{AB} &= \left(A_{ik} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_k \right) \left(B_{lj} \hat{\mathbf{e}}_l \otimes \hat{\mathbf{e}}_j \right) = A_{ik} B_{lj} \left(\hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_l \right) \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \\ &= A_{ik} B_{lj} \left(\delta_{kl} \right) \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j = A_{ik} B_{kj} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \ . \end{aligned} \tag{2.31}$$

having in mind that the dot product is a bilinear form. A similar procedure then reveals that

$$\boxed{\mathbf{A} \left(\mathbf{u} \otimes \mathbf{v} \right) = \left(\mathbf{Au} \right) \otimes \mathbf{v} \ .} \tag{2.32}$$

By means of $A_{ik} \delta_{kj} = A_{ij} = \delta_{ik} A_{kj}$ and (2.23), one immediately obtains

$$\mathbf{AI} = \mathbf{A} \quad , \quad \mathbf{IA} = \mathbf{A} \ , \tag{2.33}$$

and if $\mathbf{A} = \mathbf{I}$, it follows that

$$\mathbf{I}\,\mathbf{I} = \mathbf{I} \ . \tag{2.34}$$

From (2.23), (2.32) and $(2.33)_2$, one will have another useful identity as follows:

$$\mathbf{AI} = \mathbf{A} \left(\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_i \right) = \mathbf{A} \quad \Longrightarrow \quad \boxed{\left(\mathbf{A}\hat{\mathbf{e}}_i \right) \otimes \hat{\mathbf{e}}_i = \mathbf{A} \ .} \tag{2.35}$$

Note that in general $\widehat{\mathbf{e}}_i \otimes (\mathbf{A}\widehat{\mathbf{e}}_i) \neq \mathbf{A}$, see (2.55c).

Hint: In this text, the product $\mathbf{AA}$ will be denoted by $\mathbf{A}^2$ which motivates to define the *powers* (or *monomials*) of tensors by

$$\mathbf{A}^m = \underbrace{\mathbf{AA}\ldots\mathbf{A}}_{m \text{ times}} \,, \quad m = 1, 2, \ldots, \quad \text{and } \mathbf{A}^0 = \mathbf{I} \,. \tag{2.36}$$

For any nonnegative integers m and n, the following useful properties hold

$$\underbrace{\mathbf{A}^m\mathbf{A}^n}_{= \mathbf{A}^n\mathbf{A}^m} = \mathbf{A}^{m+n} \quad , \quad \underbrace{\left(\mathbf{A}^m\right)^n}_{= (\mathbf{A}^n)^m} = \mathbf{A}^{mn} \quad , \quad (\alpha\mathbf{A})^m = \alpha^m\mathbf{A}^m \,. \tag{2.37}$$

A **tensor function** is a function that maps a tensor into another one. An example of which, based on the tensor powers, is the following (general) tensor polynomial

$$\mathbf{H}(\mathbf{A}) = \sum_{i=0}^{m} \alpha_i \mathbf{A}^i \,. \tag{2.38}$$

This can be specialized to various tensor functions such as trigonometric or logarithm functions. Of particular interest is the *exponential tensor function*

$$\boxed{\mathbf{exp}(\mathbf{A}) = \sum_{i=0}^{\infty} \frac{1}{i!}\mathbf{A}^i \,,} \tag{2.39}$$

which is widely used in many branches of science and engineering. The main reason is that this tensor function helps obtain the solution of systems of linear ordinary differential equations, see Exercise 6.7. See also Simo [1], Simo and Hughes [2] and de Souza Neto et al. [3] for an application in **multiplicative plasticity**.

Hint: The operations (2.22) and (2.26) are often called *single contraction* (or *simple contraction*) since the sum of the orders of two tensors in both cases is reduced by two. For instance, a tensor $\mathbf{A}$ and a vector $\mathbf{u}$ with total order of three after single contraction become a tensor $\mathbf{v}$ of order one (recall that it is characterized by only one free index). In this regard, the inner product $\mathbf{u} \cdot \mathbf{v}$ basically manifests a single contraction. In accord with the scalar product $\mathbf{u} \cdot \mathbf{v}$, that is designated by one dot $(\cdot)$ between the variables, some authors still prefer to keep this notation in the present context and write $\mathbf{v} = \mathbf{A} \cdot \mathbf{u}$ (or $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$). This notation has not been utilized in this text for convenience.

The algebraic relations introduced so far are based on the definition of right mapping. By contrast, a **left mapping** is one for which a tensor $\mathbf{A}$ is premultiplied by a vector $\mathbf{v}$. For completeness, some fundamental relations for this mapping are established in the following.

It is also a linear map

$$\boxed{\mathbf{u} = \mathbf{vA} \quad \text{with} \quad u_i = v_j A_{ji} \quad \text{or} \quad (\mathbf{u})_i = (\mathbf{v})_j \, (\mathbf{A})_{ji} \ ,} \quad \longleftarrow \text{see (2.1) and (2.22)} \tag{2.40}$$

that satisfies the linearity condition

$$(\alpha\mathbf{v} + \beta\mathbf{w})\,\mathbf{A} = \alpha\,(\mathbf{vA}) + \beta\,(\mathbf{wA}) \ , \quad \longleftarrow \text{see (2.2)} \tag{2.41}$$

and obeys the rules

$$\mathbf{v}\,(\mathbf{A} + \mathbf{B}) = \mathbf{vA} + \mathbf{vB} \ , \quad \longleftarrow \text{see (2.8a)} \tag{2.42a}$$

$$\mathbf{v}\,(\alpha\mathbf{A}) = \alpha\,(\mathbf{vA}) = (\alpha\mathbf{v})\,\mathbf{A} \ . \quad \longleftarrow \text{see (2.8b)} \tag{2.42b}$$

If the dyad $\mathbf{u} \otimes \mathbf{v}$ is premultiplied by $\mathbf{w}$, the result will be a vector in the direction of $\mathbf{v}$ according to

$$\boxed{\underbrace{\mathbf{w}\,(\mathbf{u} \otimes \mathbf{v}) = (\mathbf{w} \cdot \mathbf{u})\,\mathbf{v}}_{\text{note that } \mathbf{w}\,(\mathbf{u} \otimes \mathbf{v}) \,=\, \mathbf{u}\,(\mathbf{w} \otimes \mathbf{v})} \ ,} \quad \longleftarrow \text{see (2.13)} \tag{2.43}$$

which reveals the fact that the dyad is, in general, not commutative, i.e. $\mathbf{u} \otimes \mathbf{v} \neq \mathbf{v} \otimes \mathbf{u}$. Similarly to (2.14), one can arrive at

$$\begin{aligned}(\alpha\mathbf{w} + \beta\bar{\mathbf{w}})\,(\mathbf{u} \otimes \mathbf{v}) &= \alpha\,(\mathbf{w} \cdot \mathbf{u})\,\mathbf{v} + \beta\,(\bar{\mathbf{w}} \cdot \mathbf{u})\,\mathbf{v} \\ &= \alpha\mathbf{w}\,(\mathbf{u} \otimes \mathbf{v}) + \beta\bar{\mathbf{w}}\,(\mathbf{u} \otimes \mathbf{v}) \ .\end{aligned} \tag{2.44}$$

In accord with (2.21), one can also obtain

$$\widehat{\mathbf{e}}_j \mathbf{A} = A_{jk} \widehat{\mathbf{e}}_k \ . \tag{2.45}$$

The dot product here follows the requirement

$$\boxed{\mathbf{v}\,(\mathbf{AB}) = (\mathbf{vA})\,\mathbf{B} \ .} \quad \longleftarrow \text{see (2.25)} \tag{2.46}$$

It is worthwhile to point out that the left and right mappings are related through the following expression

$$\underbrace{(\mathbf{vA}) \cdot \mathbf{u}}_{= (\mathbf{vA})_j \, (\mathbf{u})_j \,=\, (v_i A_{ij})\, u_j} \quad = \quad \underbrace{\mathbf{v} \cdot (\mathbf{Au})}_{= (\mathbf{v})_i \, (\mathbf{Au})_i \,=\, v_i \,(A_{ij} u_j)} \ . \tag{2.47}$$

2.3.2 Transposition

The *transpose* of a tensor $\mathbf{A}$, denoted by $\mathbf{A}^{\mathrm{T}}$, is defined as

$$\boxed{\mathbf{A}^{\mathrm{T}}\mathbf{u} = \mathbf{u}\mathbf{A}\ ,} \tag{2.48}$$

in which

$$\left(\mathbf{A}^{\mathrm{T}}\mathbf{u}\right)_i = (\mathbf{u}\mathbf{A})_i \Longrightarrow \underbrace{\left(\mathbf{A}^{\mathrm{T}}\right)_{ij}(\mathbf{u})_j}_{\text{see (2.22)}} = \underbrace{(\mathbf{u})_j\,(\mathbf{A})_{ji}}_{\text{see (2.40)}} \xrightarrow[(2.6)]{\text{from}} \left(\mathbf{A}^{\mathrm{T}}\right)_{ij} = (\mathbf{A})_{ji}\ . \tag{2.49}$$

Thus,

$$\boxed{\mathbf{A}^{\mathrm{T}} = A_{ji}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = A_{ij}\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i\ .} \tag{2.50}$$

As a result, the following identities hold

$$\underbrace{\mathbf{A}\mathbf{u} = \mathbf{u}\mathbf{A}^{\mathrm{T}}}_{\text{with } A_{ij}u_j = u_j A^{\mathrm{T}}_{ji}}\ , \tag{2.51a}$$

$$\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{A}\mathbf{v} = \mathbf{v}\cdot\mathbf{A}^{\mathrm{T}}\mathbf{u}}_{\text{or } A^{\mathrm{T}}_{ij}u_j v_i = u_j A_{ji} v_i = v_i A^{\mathrm{T}}_{ij} u_j}\ , \tag{2.51b}$$

$$\underbrace{\mathbf{A}\mathbf{u}\cdot\mathbf{B}\mathbf{v} = \mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{u}\cdot\mathbf{v}}_{\text{or } A_{ij}u_j B_{ik} v_k = B^{\mathrm{T}}_{ki} A_{ij} u_j v_k}\ , \tag{2.51c}$$

$$\underbrace{\mathbf{A}\mathbf{u}\cdot\mathbf{B}\mathbf{v} = \mathbf{u}\cdot\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{v}}_{\text{or } A_{ij}u_j B_{ik} v_k = u_j A^{\mathrm{T}}_{ji} B_{ik} v_k}\ . \tag{2.51d}$$

The linearity of transposition is then implied by

$$\left.\begin{aligned}
(\mathbf{A}+\mathbf{B})^{\mathrm{T}}\mathbf{u} &\overset{\text{from}}{\underset{(2.48)}{=\!=\!=}} \mathbf{u}\,(\mathbf{A}+\mathbf{B}) \\
&\overset{\text{from}}{\underset{(2.42a)}{=\!=\!=}} \mathbf{u}\mathbf{A} + \mathbf{u}\mathbf{B} \\
&\overset{\text{from}}{\underset{(2.48)}{=\!=\!=}} \mathbf{A}^{\mathrm{T}}\mathbf{u} + \mathbf{B}^{\mathrm{T}}\mathbf{u} \\
&\overset{\text{from}}{\underset{(2.8a)}{=\!=\!=}} \left(\mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}\right)\mathbf{u}
\end{aligned}\right\} \xrightarrow[(2.6)]{\text{from}} (\mathbf{A}+\mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}\ . \tag{2.52}$$

and

$$\left.\begin{aligned}(\alpha\mathbf{A})^{\mathrm{T}}\,\mathbf{u} &\overset{\text{from}}{\underset{(2.48)}{=\!=\!=}} \mathbf{u}\,(\alpha\mathbf{A}) \\ &\overset{\text{from}}{\underset{(2.42\text{b})}{=\!=\!=}} \alpha\,(\mathbf{u}\mathbf{A}) \\ &\overset{\text{from}}{\underset{(2.8\text{b}) \text{ and } (2.48)}{=\!=\!=}} \left(\alpha\mathbf{A}^{\mathrm{T}}\right)\mathbf{u}\end{aligned}\right\} \overset{\text{from}}{\underset{(2.6)}{\Longrightarrow}} (\alpha\mathbf{A})^{\mathrm{T}} = \alpha\mathbf{A}^{\mathrm{T}}\,. \tag{2.53}$$

The transpose of a dyad is also implied by

$$\left.\begin{aligned}(\mathbf{v}\otimes\mathbf{w})^{\mathrm{T}}\,\mathbf{u} &\overset{\text{from}}{\underset{(2.48)}{=\!=\!=}} \mathbf{u}\,(\mathbf{v}\otimes\mathbf{w}) \\ &\overset{\text{from}}{\underset{(2.43)}{=\!=\!=}} (\mathbf{u}\cdot\mathbf{v})\,\mathbf{w} \\ &\overset{\text{from}}{\underset{(2.13)}{=\!=\!=}} (\mathbf{w}\otimes\mathbf{v})\,\mathbf{u}\end{aligned}\right\} \overset{\text{from}}{\underset{(2.6)}{\Longrightarrow}} (\mathbf{v}\otimes\mathbf{w})^{\mathrm{T}} = \mathbf{w}\otimes\mathbf{v}\,. \tag{2.54}$$

The following identities trivially hold

$$\mathbf{I}^{\mathrm{T}} = \mathbf{I}\,, \quad \longleftarrow \text{since } \delta_{ij} = \delta_{ji} \tag{2.55a}$$

$$\left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathbf{A}\,, \quad \longleftarrow \begin{array}{c}\text{this means that the transpose of}\\ \text{a transposed tensor is again that tensor}\end{array} \tag{2.55b}$$

$$\mathbf{A}^{\mathrm{T}} = \widehat{\mathbf{e}}_i \otimes \mathbf{A}\widehat{\mathbf{e}}_i\,, \quad \longleftarrow \text{note that } \left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \left(\widehat{\mathbf{e}}_i \otimes \mathbf{A}\widehat{\mathbf{e}}_i\right)^{\mathrm{T}} \text{ yields (2.35)} \tag{2.55c}$$

$$(\mathbf{A}\mathbf{B})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \quad , \quad (\mathbf{A}\mathbf{B}\mathbf{C})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\,, \tag{2.55d}$$

$$\left(\mathbf{A}^{\mathrm{T}}\right)^{k} = \left(\mathbf{A}^{k}\right)^{\mathrm{T}}\,, \quad \text{for any integer } k\,, \tag{2.55e}$$

$$\left[\mathbf{exp}\,(\mathbf{A})\right]^{\mathrm{T}} = \mathbf{exp}\left(\mathbf{A}^{\mathrm{T}}\right)\,. \tag{2.55f}$$

From (2.32), (2.54) and $(2.55\text{d})_1$, one finally obtains

$$\boxed{\underbrace{(\mathbf{u}\otimes\mathbf{v})\,\mathbf{A}^{\mathrm{T}} = \mathbf{u}\otimes(\mathbf{A}\mathbf{v})}_{\text{since } (\mathbf{u}\otimes\mathbf{v})\,\mathbf{A}^{\mathrm{T}} = (\mathbf{A}\,(\mathbf{v}\otimes\mathbf{u}))^{\mathrm{T}} = ((\mathbf{A}\mathbf{v})\otimes\mathbf{u})^{\mathrm{T}} = \mathbf{u}\otimes(\mathbf{A}\mathbf{v})}\,.} \tag{2.56}$$

2.3.3 Decomposition into Symmetric and Skew-Symmetric Parts

A second-order tensor **A** is said to be *symmetric* if

$$\boxed{\underbrace{\mathbf{A}^{\mathrm{T}} = \mathbf{A}}_{\text{in index notation } A_{ji} = A_{ij}} \qquad \text{or, equivalently,} \qquad \underbrace{\mathbf{u}\cdot\mathbf{A}\mathbf{v} = \mathbf{v}\cdot\mathbf{A}\mathbf{u}}_{\text{in index notation } u_i A_{ij} v_j = v_i A_{ij} u_j}\,,} \tag{2.57}$$

and is called *skew-symmetric* (or simply *skew* or *antisymmetric*) when

$$\boxed{\mathbf{A}^{\mathrm{T}} = -\mathbf{A} \quad \text{or, equivalently,} \quad \mathbf{u} \cdot \mathbf{A}\mathbf{v} = -\mathbf{v} \cdot \mathbf{A}\mathbf{u}\ .} \tag{2.58}$$

The symmetric (skew) part of **A** is denoted by sym**A** (skw**A**). The tensors sym**A** and skw**A** are governed by

$$\mathrm{sym}\mathbf{A} = \frac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} \quad , \quad \mathrm{skw}\mathbf{A} = \frac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2}\ . \tag{2.59}$$

The following relations hold

$$\boxed{\left.\begin{aligned} \mathrm{sym}\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{A}\right) &= \mathbf{A}^{\mathrm{T}}\left(\mathrm{sym}\mathbf{B}\right)\mathbf{A} \\ \mathrm{sym}\left(\mathbf{A}\mathbf{B}\mathbf{A}^{\mathrm{T}}\right) &= \mathbf{A}\left(\mathrm{sym}\mathbf{B}\right)\mathbf{A}^{\mathrm{T}} \end{aligned}\right\} \quad , \quad \left.\begin{aligned} \mathrm{skw}\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{A}\right) &= \mathbf{A}^{\mathrm{T}}\left(\mathrm{skw}\mathbf{B}\right)\mathbf{A} \\ \mathrm{skw}\left(\mathbf{A}\mathbf{B}\mathbf{A}^{\mathrm{T}}\right) &= \mathbf{A}\left(\mathrm{skw}\mathbf{B}\right)\mathbf{A}^{\mathrm{T}} \end{aligned}\right\}\ .} \tag{2.60}$$

In this text, for the sake of clarification, **S** (**W**) is used to denote any symmetric (skew) tensor as well as the symmetric (skew) part of an arbitrary tensor. One then trivially has

$$\underbrace{\mathrm{sym}\mathbf{S} = \mathbf{S}}_{\text{since } S_{ij} = S_{ji}} \quad , \quad \underbrace{\mathrm{skw}\mathbf{W} = \mathbf{W}}_{\text{since } W_{ij} = -W_{ji}} \quad , \quad \underbrace{\mathrm{skw}\mathbf{S} = \mathrm{sym}\mathbf{W} = \mathbf{O}}_{\text{since, e.g. } \frac{1}{2}\left(S_{ij} - S_{ji}\right) = \frac{1}{2}\left(S_{ij} - S_{ij}\right) = 0} \quad . \tag{2.61}$$

Any tensor **A** can now be decomposed as

$$\boxed{\mathbf{A} = \frac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} + \frac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2} = \mathbf{S} + \mathbf{W} \quad \text{where} \quad \mathbf{S} = \mathrm{sym}\mathbf{A}\ ,\ \mathbf{W} = \mathrm{skw}\mathbf{A}\ .} \tag{2.62}$$

Note that this **additive decomposition** is unique. The set of all symmetric tensors[6]

$$\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}} = \left\{\mathbf{A} \in \mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right) \mid \mathbf{A}^{\mathrm{T}} = \mathbf{A}\right\}\ , \tag{2.63}$$

is a subspace of all second-order tensors $\mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right)$ and[7]

$$\mathscr{T}_{\mathrm{so}}^{\mathrm{skw}} = \left\{\mathbf{A} \in \mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right) \mid \mathbf{A}^{\mathrm{T}} = -\mathbf{A}\right\}\ , \tag{2.64}$$

indicates the set of all antisymmetric tensors forming another subset of $\mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right)$.

For any skew tensor **W**, there exists a unique vector $\boldsymbol{\omega}$ - called the *axial vector* of **W** - such that

[6] In the literature, the set $\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}$ is often denoted by Sym.

[7] Note that the set $\mathscr{T}_{\mathrm{so}}^{\mathrm{skw}}$ is sometimes denoted by Skw in the literature.

$$\boxed{\mathbf{W}\mathbf{u} = \boldsymbol{\omega} \times \mathbf{u} \,, \quad \text{for all } \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{o3} \,.} \tag{2.65}$$

Therein, the (Cartesian) axial vector $\boldsymbol{\omega}$ corresponding to $\mathbf{W}$ is given by

$$\boldsymbol{\omega} = \omega_m \widehat{\mathbf{e}}_m = (-W_{23})\,\widehat{\mathbf{e}}_1 + (W_{13})\,\widehat{\mathbf{e}}_2 + (-W_{12})\,\widehat{\mathbf{e}}_3 \,. \tag{2.66}$$

In matrix notation, one will have

$$[\,\boldsymbol{\omega}\,] = \left[\,-W_{23} \;\; W_{13} \;\; -W_{12}\,\right]^{\mathrm{T}} \,. \tag{2.67}$$

And its length, using (1.11), renders

$$|\boldsymbol{\omega}| = \sqrt{W_{12}^2 + W_{13}^2 + W_{23}^2} \,. \tag{2.68}$$

As can be seen, the **area vector** $\boldsymbol{\omega} \times \mathbf{u}$ is implied by operating $\mathbf{W}$ on $\mathbf{u}$. It turns out that the skew tensor $\mathbf{W}$ has only three independent components which are all off-diagonal terms. Hence, $\mathbf{W}$ is a tensor with $W_{11} = W_{22} = W_{33} = 0$.

The goal is now to verify (2.65) and (2.66). To do so, consider first

$$\begin{aligned}
W_{ij} &\overset{\text{from}}{\underset{(2.61)}{=\!=}} \frac{1}{2}\left(W_{ij} - W_{ji}\right) \\
&\overset{\text{from}}{\underset{(1.36)}{=\!=}} \frac{1}{2} W_{kl}\left(\delta_{ik}\delta_{jl} - \delta_{jk}\delta_{il}\right) \\
&\overset{\text{from}}{\underset{(1.58a)}{=\!=}} \frac{1}{2} W_{kl}\left(-\varepsilon_{ijm}\varepsilon_{lkm}\right) \\
&\overset{\text{from}}{\underset{(1.54)}{=\!=}} -\left(-\frac{1}{2} W_{kl}\varepsilon_{klm}\right)\varepsilon_{ijm} \,.
\end{aligned} \tag{2.69}$$

Then, by defining

$$\underbrace{\boldsymbol{\omega} = -\frac{1}{2} W_{kl}\varepsilon_{klm}\widehat{\mathbf{e}}_m \quad \text{with} \quad \omega_m = -\frac{1}{2} W_{kl}\varepsilon_{klm}}_{\substack{\text{by expansion: } -2\omega_1 = W_{kl}\varepsilon_{kl1} = W_{23}\varepsilon_{231} + W_{32}\varepsilon_{321} = W_{23} - W_{32} = 2W_{23}\,, \\ -2\omega_2 = W_{13}\varepsilon_{132} + W_{31}\varepsilon_{312} = -W_{13} + W_{31} = -2W_{13} \text{ and } -2\omega_3 = W_{12}\varepsilon_{123} + W_{21}\varepsilon_{213} = 2W_{12}}} \,, \tag{2.70}$$

the expression $(2.69)_4$ takes the form

$$\boxed{W_{ij} = -\varepsilon_{ijm}\omega_m \,.} \tag{2.71}$$

Postmultiplying both sides of (2.71) by an arbitrary vector $\mathbf{u}$ and subsequently using (1.54) and (1.67) finally yields

$$W_{ij}u_j = -\varepsilon_{ijm}\omega_m u_j = \omega_m u_j \varepsilon_{mji} \quad \text{or} \quad (\mathbf{W}\mathbf{u})_i = (\boldsymbol{\omega} \times \mathbf{u})_i \,. \tag{2.72}$$

2.3.4 Contraction and Trace

The *contraction* (or *double contraction* or *scalar product*) of two second-order tensors produces a scalar number, meaning that the sum of the orders of tensors is reduced by four. And this enables one to define scalar-valued functions of tensorial variables. The double contraction between two tensors $\mathbf{u} \otimes \mathbf{v}$ and $\mathbf{w} \otimes \bar{\mathbf{w}}$, characterized by two dots (:), is defined as

$$\boxed{(\mathbf{u} \otimes \mathbf{v}) : (\mathbf{w} \otimes \bar{\mathbf{w}}) = (\mathbf{u} \cdot \mathbf{w}) (\mathbf{v} \cdot \bar{\mathbf{w}}) \ .} \tag{2.73}$$

Recall that the inner product $g^{\text{ip}} (\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ is a **symmetric bilinear form** which takes two vectors and delivers a scalar quantity. Its counterpart here, designated by $g^{\text{c}} (\mathbf{A}, \mathbf{B}) = \mathbf{A} : \mathbf{B}$, will be characterized by the identical properties. For any α , $\beta \in$ R and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} \in \mathscr{T}_{\text{so}}$, they are listed in the following

$$\underbrace{g^{\text{c}} (\mathbf{A}, \mathbf{B}) = g^{\text{c}} (\mathbf{B}, \mathbf{A})}_{\text{or } \mathbf{A} : \mathbf{B} = \mathbf{B} : \mathbf{A} \text{ representing commutative or symmetry property}} \quad , \tag{2.74a}$$

$$\underbrace{g^{\text{c}} (\alpha\mathbf{A} + \beta\mathbf{B}, \mathbf{C}) = \alpha g^{\text{c}} (\mathbf{A}, \mathbf{C}) + \beta g^{\text{c}} (\mathbf{B}, \mathbf{C})}_{\text{or } (\alpha\mathbf{A} + \beta\mathbf{B}) : \mathbf{C} = \alpha\mathbf{A} : \mathbf{C} + \beta\mathbf{B} : \mathbf{C} \text{ rendering linearity in the first argument}} \quad , \tag{2.74b}$$

$$\underbrace{\left.\begin{array}{ll} g^{\text{c}} (\mathbf{A}, \mathbf{A}) > 0 & \text{if } \mathbf{A} \neq \mathbf{O} \\ g^{\text{c}} (\mathbf{A}, \mathbf{A}) = 0 & \text{if } \mathbf{A} = \mathbf{O} \end{array}\right\}}_{\text{or } \mathbf{A} : \mathbf{A} \geq 0, \text{ for which } \mathbf{A} : \mathbf{A} = 0 \Longleftrightarrow \mathbf{A} = \mathbf{O}, \text{ expressing positive-definite property}} \quad . \tag{2.74c}$$

For any two tensors $\mathbf{A}$ and $\mathbf{B}$ of the form $(2.19)_2$, the expressions (2.73), (2.74a) and (2.74b) help obtain

$$\begin{aligned} \mathbf{A} : \mathbf{B} &= \left(A_{ij} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \right) : \left(B_{kl} \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \right) \\ &= A_{ij} B_{kl} \left(\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_k \right) \left(\widehat{\mathbf{e}}_j \cdot \widehat{\mathbf{e}}_l \right) \\ &= A_{ij} B_{kl} \left(\delta_{ik} \right) \left(\delta_{jl} \right) \\ &= \boxed{A_{ij} B_{ij} \ .} \end{aligned} \tag{2.75}$$

This clearly shows that in indicial notation, the double contraction between two second-order tensors has no free index. Only two dummy indices turn out that should be summed over them to obtain the desired result.

The double contraction is also an algebraic operation. Its last property in (2.74c) then helps define the **norm** of a second-order tensor $\mathbf{A}$ via

$$\boxed{|\mathbf{A}| = \sqrt{\mathbf{A} : \mathbf{A}} \ .} \quad \longleftarrow \text{see (1.11)} \tag{2.76}$$

Another useful consequence of (2.74c) is

$$\text{if} \quad \mathbf{A} : \mathbf{B} = 0 \quad \text{for all tensors } \mathbf{B} \text{ then} \quad \mathbf{A} = \mathbf{O} \,, \tag{2.77}$$

which can easily be verified by setting $\mathbf{B} = \mathbf{A}$. The converse relation

$$\boxed{\mathbf{O} : \mathbf{B} = \mathbf{B} : \mathbf{O} = 0 \,,} \tag{2.78}$$

also holds true considering $\mathbf{O}$ as a tensor with zero components in any coordinate system. As a result, for all tensors $\mathbf{B}$, $\mathbf{A} = \mathbf{O}$ if and only if $\mathbf{A} : \mathbf{B} = 0$.

By means of $(1.38)_7$, $(2.22)_{2-3}$, $(2.24)_5$, (2.26), (2.49), $(2.61)_{1-2}$ and $(2.75)_4$, some useful properties can be established:

$$\underbrace{\mathbf{A} : \mathbf{B}}_{= (\mathbf{A})_{ij} (\mathbf{B})_{ij}} = \underbrace{\mathbf{A}^{\mathrm{T}} : \mathbf{B}^{\mathrm{T}}}_{= \left(\mathbf{A}^{\mathrm{T}}\right)_{lk} \left(\mathbf{B}^{\mathrm{T}}\right)_{lk} = (\mathbf{A})_{kl} (\mathbf{B})_{kl} = (\mathbf{A})_{ij} (\mathbf{B})_{ij}} \,, \tag{2.79a}$$

$$\begin{aligned} \underbrace{\mathbf{A} : (\mathbf{B}\mathbf{C})}_{= (\mathbf{A})_{ij} (\mathbf{B})_{ik} (\mathbf{C})_{kj}} &= \underbrace{\left(\mathbf{B}^{\mathrm{T}}\mathbf{A}\right) : \mathbf{C}}_{= \left(\mathbf{B}^{\mathrm{T}}\mathbf{A}\right)_{kj} (\mathbf{C})_{kj} = \left(\mathbf{B}^{\mathrm{T}}\right)_{ki} (\mathbf{A})_{ij} (\mathbf{C})_{kj} = (\mathbf{B})_{ik} (\mathbf{A})_{ij} (\mathbf{C})_{kj}} \\ &= \underbrace{\left(\mathbf{A}\mathbf{C}^{\mathrm{T}}\right) : \mathbf{B}}_{= \left(\mathbf{A}\mathbf{C}^{\mathrm{T}}\right)_{ik} (\mathbf{B})_{ik} = (\mathbf{A})_{ij} \left(\mathbf{C}^{\mathrm{T}}\right)_{jk} (\mathbf{B})_{ik} = (\mathbf{A})_{ij} (\mathbf{C})_{kj} (\mathbf{B})_{ik}} \,, \end{aligned} \tag{2.79b}$$

$$\begin{aligned} \underbrace{(\mathbf{u} \otimes \mathbf{v}) : \mathbf{B}}_{= (\mathbf{u})_i (\mathbf{v})_j (\mathbf{B})_{ij}} &= \underbrace{\mathbf{u} \cdot \mathbf{B}\mathbf{v}}_{= (\mathbf{u})_i (\mathbf{B}\mathbf{v})_i = (\mathbf{u})_i (\mathbf{B})_{ij} (\mathbf{v})_j} \\ &= \underbrace{\mathbf{v} \cdot \mathbf{B}^{\mathrm{T}}\mathbf{u}}_{= (\mathbf{v})_j \left(\mathbf{B}^{\mathrm{T}}\mathbf{u}\right)_j = (\mathbf{v})_j \left(\mathbf{B}^{\mathrm{T}}\right)_{ji} (\mathbf{u})_i = (\mathbf{v})_j (\mathbf{B})_{ij} (\mathbf{u})_i} \,, \end{aligned} \tag{2.79c}$$

$$\begin{aligned} \underbrace{(\mathbf{A}\mathbf{u} \otimes \mathbf{v}) : \mathbf{B}}_{= (\mathbf{A})_{ik} (\mathbf{u})_k (\mathbf{v})_j (\mathbf{B})_{ij}} &= \underbrace{\mathbf{A}\mathbf{u} \cdot \mathbf{B}\mathbf{v}}_{= (\mathbf{A}\mathbf{u})_i (\mathbf{B}\mathbf{v})_i = (\mathbf{A})_{ik} (\mathbf{u})_k (\mathbf{B})_{ij} (\mathbf{v})_j} \\ &= \underbrace{\mathbf{u} \cdot \left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right) \mathbf{v}}_{= (\mathbf{u})_k \left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{v}\right)_k = (\mathbf{u})_k \left(\mathbf{A}^{\mathrm{T}}\right)_{ki} (\mathbf{B})_{ij} (\mathbf{v})_j = (\mathbf{u})_k (\mathbf{A})_{ik} (\mathbf{B})_{ij} (\mathbf{v})_j} \,, \end{aligned} \tag{2.79d}$$

$$\begin{aligned} \underbrace{(\mathbf{u} \otimes \mathbf{A}\mathbf{v}) : \mathbf{B}}_{= (\mathbf{u})_i (\mathbf{A})_{jk} (\mathbf{v})_k (\mathbf{B})_{ij}} &= \underbrace{\mathbf{B}^{\mathrm{T}}\mathbf{u} \cdot \mathbf{A}\mathbf{v}}_{= \left(\mathbf{B}^{\mathrm{T}}\mathbf{u}\right)_j (\mathbf{A}\mathbf{v})_j = \left(\mathbf{B}^{\mathrm{T}}\right)_{ji} (\mathbf{u})_i (\mathbf{A})_{jk} (\mathbf{v})_k = (\mathbf{B})_{ij} (\mathbf{u})_i (\mathbf{A})_{jk} (\mathbf{v})_k} \\ &= \underbrace{\mathbf{u} \cdot \mathbf{B}\mathbf{A}\mathbf{v}}_{= (\mathbf{u})_i (\mathbf{B}\mathbf{A}\mathbf{v})_i = (\mathbf{u})_i (\mathbf{B}\mathbf{A})_{ik} (\mathbf{v})_k = (\mathbf{u})_i (\mathbf{B})_{ij} (\mathbf{A})_{jk} (\mathbf{v})_k} \,. \end{aligned} \tag{2.79e}$$

$$\underbrace{\mathbf{S} : \mathbf{B}}_{= (\mathbf{S})_{ij} (\mathbf{B})_{ij}} = \underbrace{\mathbf{S} : \frac{1}{2}\left(\mathbf{B} + \mathbf{B}^{\mathrm{T}}\right)}_{\text{since } 2 (\mathbf{S})_{ij} (\mathbf{B})_{ij} = (\mathbf{S})_{ij} (\mathbf{B})_{ij} + (\mathbf{S})_{ji} (\mathbf{B})_{ji} = (\mathbf{S})_{ij} (\mathbf{B})_{ij} + (\mathbf{S})_{ij} (\mathbf{B})_{ji}} \,, \tag{2.79f}$$

$$\underbrace{\mathbf{W} : \mathbf{B}}_{= (\mathbf{W})_{ij}\, (\mathbf{B})_{ij}} = \underbrace{\mathbf{W} : \frac{1}{2}\left(\mathbf{B} - \mathbf{B}^{\mathrm{T}}\right)}_{\text{since } 2\, (\mathbf{W})_{ij}\, (\mathbf{B})_{ij} = (\mathbf{W})_{ij}\, (\mathbf{B})_{ij} + (\mathbf{W})_{ji}\, (\mathbf{B})_{ji} = (\mathbf{W})_{ij}\, (\mathbf{B})_{ij} - (\mathbf{W})_{ij}\, (\mathbf{B})_{ji}} , \tag{2.79g}$$

$$\underbrace{\mathbf{S} : \mathbf{W}}_{= (\mathbf{S})_{ij}\, (\mathbf{W})_{ij}} = \underbrace{0}_{\text{since } 2\, (\mathbf{S})_{ij}\, (\mathbf{W})_{ij} = (\mathbf{S})_{ij}\, (\mathbf{W})_{ij} + (\mathbf{S})_{ji}\, (\mathbf{W})_{ji} = (\mathbf{S})_{ij}\, (\mathbf{W})_{ij} - (\mathbf{S})_{ij}\, (\mathbf{W})_{ij}} . \tag{2.79h}$$

The above expressions clearly demonstrate the capability of indicial notation in establishing vector and tensor identities. The last property shows that symmetric and antisymmetric tensors are mutually **orthogonal**. In this regard, the subspaces $\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}$ and $\mathscr{T}_{\mathrm{so}}^{\mathrm{skw}}$, according to (2.63) and (2.64), are said to be orthogonal.

Similarly to (2.79h), the following identities hold true

$$\underbrace{\varepsilon_{ijk} S_{jk} = 0 \quad , \quad S_{ij}\varepsilon_{ijk} = 0 \quad , \quad \varepsilon_{ijk} u_j u_k = 0 \quad , \quad u_i u_j \varepsilon_{ijk} = 0}_{\text{since, e.g., from } \varepsilon_{ijk} = -\varepsilon_{ikj} \text{ and } S_{jk} = S_{kj} \text{ one will arrive at } \varepsilon_{ijk} S_{jk} = \varepsilon_{ikj} S_{kj} = -\varepsilon_{ijk} S_{jk}} . \tag{2.80}$$

Following discussions analogous to those that led to (1.14), one can deduce that[8]

$$\boxed{\mathbf{A} : \mathbf{C} = \mathbf{B} : \mathbf{C} \quad \text{for all tensors } \mathbf{C} \text{ if and only if} \quad \mathbf{A} = \mathbf{B}\ .} \quad \longleftarrow \text{see (2.6)} \tag{2.81}$$

The *trace* of a tensor **A** is also a scalar and denoted here by tr**A**. It is defined by means of the double contraction in the following manner

$$\boxed{\mathrm{tr}\mathbf{A} = \mathbf{I} : \mathbf{A} = g^{\mathrm{c}}\left(\mathbf{I}, \mathbf{A}\right)\ .} \tag{2.82}$$

Then, the properties (2.74a) and (2.74b) immediately imply

$$\mathrm{tr}\mathbf{A} = \mathbf{I} : \mathbf{A} = \mathbf{A} : \mathbf{I} \quad , \quad \mathrm{tr}\left(\alpha\mathbf{A} + \beta\mathbf{B}\right) = \alpha\mathrm{tr}\left(\mathbf{A}\right) + \beta\mathrm{tr}\left(\mathbf{B}\right) . \tag{2.83}$$

Considering the symmetry of double contraction, the following expressions

$$\left.\begin{aligned} \mathbf{A} : \mathbf{B} &\xlongequal[\text{(2.33)}]{\text{from}} \mathbf{A}\mathbf{I} : \mathbf{B} \xlongequal[\text{(2.79b) and (2.83)}]{\text{from}} \mathbf{I} : \mathbf{A}^{\mathrm{T}}\mathbf{B} \\ &\xlongequal[\text{(2.33)}]{\text{from}} \mathbf{I}\mathbf{A} : \mathbf{B} \xlongequal[\text{(2.79b) and (2.83)}]{\text{from}} \mathbf{I} : \mathbf{B}\mathbf{A}^{\mathrm{T}} \end{aligned}\right\} , \tag{2.84}$$

---

[8] The proof is not difficult. If $\mathbf{A} : \mathbf{C} = \mathbf{B} : \mathbf{C}$ holds for all **C**, then one can obtain $(\mathbf{A} - \mathbf{B}) : \mathbf{C} = 0$. Now, by choosing $\mathbf{C} = \mathbf{A} - \mathbf{B}$, one can arrive at $g^{\mathrm{c}}\left(\mathbf{A} - \mathbf{B}, \mathbf{A} - \mathbf{B}\right) = 0$. And this result with the aid of (2.74c) delivers $\mathbf{A} = \mathbf{B}$. The converse is immediate.

The result (2.81) can be shown in an alternative way. Suppose that **C** is a tensor with only one nonzero component. Denoting by C the $(ij)^{\mathrm{th}}$ nonzero component of **C**, the relation $(\mathbf{A} - \mathbf{B}) : \mathbf{C} = 0$ now simplifies to $\left(A_{ij} - B_{ij}\right) C = 0$. And this immediately implies that $A_{ij} = B_{ij},\ i, j = 1, 2, 3$.

help establish

$$\operatorname{tr}\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right) = \operatorname{tr}\left(\mathbf{B}\mathbf{A}^{\mathrm{T}}\right) = \operatorname{tr}\left(\mathbf{B}^{\mathrm{T}}\mathbf{A}\right) = \operatorname{tr}\left(\mathbf{A}\mathbf{B}^{\mathrm{T}}\right) \ . \tag{2.85}$$

Any two second-order tensors **A** and **B** also satisfy

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}) = \operatorname{tr}\left(\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\right) = \operatorname{tr}\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\right) \ . \tag{2.86}$$

Let **A**, **B** and **C** be three arbitrary tensors. Then,

$$\boxed{\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{BCA}) = \operatorname{tr}(\mathbf{CAB}) \ .} \tag{2.87}$$

This is known as the *cyclic* property of the trace operator. Now, let **S**, **T** and **U** be three **symmetric** tensors. Then, one can further establish

$$\boxed{\operatorname{tr}(\mathbf{STU}) = \operatorname{tr}(\mathbf{TSU}) = \operatorname{tr}(\mathbf{UTS}) = \operatorname{tr}(\mathbf{SUT}) \ .} \tag{2.88}$$

By means of the expressions (1.36), (2.23), $(2.75)_4$ and $(2.82)_1$, the trace of the Cartesian tensor $\mathbf{A} = A_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$ in $(2.19)_2$, the dyad $\mathbf{u} \otimes \mathbf{v} = u_i v_j \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$ in $(2.24)_3$, the product $\mathbf{AB} = A_{ik}B_{kj}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$ in $(2.31)_4$, the transposed tensor $\mathbf{A}^{\mathrm{T}} = A_{ji}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$ in $(2.50)_1$ and the skew tensor **W** in $(2.61)_2$ can be computed according to

$$\operatorname{tr}\mathbf{A} = \delta_{ij}A_{ij} = A_{ii} \ , \tag{2.89a}$$

$$\operatorname{tr}(\mathbf{u} \otimes \mathbf{v}) = \delta_{ij}u_i v_j = u_i v_i \ , \tag{2.89b}$$

$$\operatorname{tr}(\mathbf{AB}) = \delta_{ij}A_{ik}B_{kj} = A_{ik}B_{ki} \ , \tag{2.89c}$$

$$\operatorname{tr}\mathbf{A}^{\mathrm{T}} = \delta_{ij}A_{ji} = A_{ii} = \operatorname{tr}\mathbf{A} \ , \tag{2.89d}$$

$$\operatorname{tr}\mathbf{W} = \delta_{ij}W_{ij} = W_{ii} = 0 \ . \tag{2.89e}$$

As a consequence of $(2.89\text{a})_2$, the trace of the unit tensor **I** in (2.23), using $(1.37)_1$, renders

$$\boxed{\operatorname{tr}\mathbf{I} = \mathbf{I} : \mathbf{I} = \delta_{ii} = 3 \ .} \tag{2.90}$$

Guided by (2.78), another special case is

$$\operatorname{tr}\mathbf{O} = \mathbf{I} : \mathbf{O} = \mathbf{O} : \mathbf{I} = 0 \ . \tag{2.91}$$

2.4 Tensor in Matrix Notation

As discussed in the previous chapter, manipulations of vectors and tensors in computer codes will naturally be treated by means of matrix notation. Here, the components of a second-order tensor **A** are collected in a 3 by 3 matrix. It is not difficult to show that all the operations and operators defined so far for tensors are consistently associated with the same operations and operators applied in matrix algebra. Some major relations are presented in the following.

In view of the Cartesian basis vectors (1.42), the dyad $\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$ in this useful notation is written as

$$\left[\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right] = \begin{bmatrix} \delta_{i1} \\ \delta_{i2} \\ \delta_{i3} \end{bmatrix} \begin{bmatrix} \delta_{j1} & \delta_{j2} & \delta_{j3} \end{bmatrix} = \begin{bmatrix} \delta_{i1}\delta_{j1} & \delta_{i1}\delta_{j2} & \delta_{i1}\delta_{j3} \\ \delta_{i2}\delta_{j1} & \delta_{i2}\delta_{j2} & \delta_{i2}\delta_{j3} \\ \delta_{i3}\delta_{j1} & \delta_{i3}\delta_{j2} & \delta_{i3}\delta_{j3} \end{bmatrix} . \tag{2.92}$$

By means of $(1.43)_2$ and $(2.92)_{1-2}$, the tensor product $\mathbf{u} \otimes \mathbf{v}$ takes the form

$$[\,\mathbf{u} \otimes \mathbf{v}\,] = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{bmatrix} . \tag{2.93}$$

Guided by $(2.92)_2$, the tensor $\mathbf{A}$, according to $(2.19)_2$, and its transpose in $(2.50)_1$ will again render $(1.47)_{1-2}$ in a consistent manner. Accordingly, the simple contractions $\mathbf{Au}$ and $\mathbf{AB}$, according to $(2.22)_1$ and $(2.31)_4$, simply follow the ordinary definition of the matrix multiplication, that is,

$$\boxed{[\,\mathbf{Au}\,] = [\,\mathbf{A}\,][\,\mathbf{u}\,] \quad , \quad [\,\mathbf{AB}\,] = [\,\mathbf{A}\,][\,\mathbf{B}\,] \ .} \tag{2.94}$$

A scalar of the form $\mathbf{v} \cdot \mathbf{Au}$, making use of $(1.45)_{1-2}$, $(1.47)_1$ and $(2.94)_1$, is then written as

$$\mathbf{v} \cdot \mathbf{Au} = [\mathbf{v}]^{\mathrm{T}}\,[\mathbf{A}]\,[\mathbf{u}] = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} . \tag{2.95}$$

The symmetric tensor $\mathbf{S}$ as well as the skew tensor $\mathbf{W}$ are represented by

$$[\,\mathbf{S}\,] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{bmatrix} \quad , \quad [\,\mathbf{W}\,] = \begin{bmatrix} 0 & W_{12} & W_{13} \\ -W_{12} & 0 & W_{23} \\ -W_{13} & -W_{23} & 0 \end{bmatrix} . \tag{2.96}$$

Hint: For the so-called general basis $\{\mathbf{g}_i\}$, whose every element can be expressed in terms of the standard basis $\{\widehat{\mathbf{e}}_i\}$, a tensor $\mathbf{A}$ in matrix notation admits some different forms, see Sect. 5.5.2.

2.5 Determinant, Inverse and Cofactor of a Tensor

2.5.1 Determinant of a Tensor

Recall that the determinant of a matrix $[\,\mathbf{A}\,]$ was computed by means of either (1.79) or (1.80). From computational point of view, the determinant of a second-order tensor equals the determinant of its matrix form, i.e.

$$\boxed{\det \mathbf{A} = \det [\,\mathbf{A}\,]\ .} \tag{2.97}$$

But technically, there is a precise definition carrying the concept of tensor. Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a set of three **linearly independent** vectors. The determinant of a tensor $\mathbf{A}$ is then defined by

$$\boxed{\det \mathbf{A} = \frac{\mathbf{A}\mathbf{u} \cdot (\mathbf{A}\mathbf{v} \times \mathbf{A}\mathbf{w})}{\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})}\ .} \quad \longleftarrow \text{see Exercise 4.2} \tag{2.98}$$

In light of Fig. 1.5, the determinant of $\mathbf{A}$ can geometrically be interpreted as the volume of the parallelepiped spanned by $\{\mathbf{A}\mathbf{u}, \mathbf{A}\mathbf{v}, \mathbf{A}\mathbf{w}\}$ over the volume of the parallelepiped constructed from $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$.

Consistent with matrix algebra, some favorable properties of determinant are given in the following:

$$\det \mathbf{A}^{\mathrm{T}} = \det \mathbf{A}\ , \tag{2.99a}$$

$$\det (\mathbf{A}\mathbf{B}) = (\det \mathbf{A})\,(\det \mathbf{B}) = \det (\mathbf{B}\mathbf{A})\ , \tag{2.99b}$$

$$\det (\alpha\mathbf{A}) = \alpha^3 \det \mathbf{A}\ . \tag{2.99c}$$

These properties help establish some identities. First, the expressions $(2.33)_1$ and $(2.99\text{b})_1$ help obtain the determinant of the identity tensor:

$$\underbrace{\det (\mathbf{A}\mathbf{I}) = \det \mathbf{A} \quad \Rightarrow \quad (\det \mathbf{A})\,(\det \mathbf{I}) = \det \mathbf{A}}_{\text{note that these relations hold for all tensors } \mathbf{A}} \quad \Rightarrow \quad \boxed{\det \mathbf{I} = 1\ .} \tag{2.100}$$

Then, by means of (f) in (2.12), $(2.58)_1$, (2.99a) and (2.99c), the determinant of a skew tensor $\mathbf{W}$ renders

$$\left.\begin{aligned} \det \mathbf{W} &= \det \mathbf{W}^{\mathrm{T}} = \det (-\mathbf{W}) = \det ((-1)\,\mathbf{W}) \\ &= (-1)^3 \det \mathbf{W} = -\det \mathbf{W} \end{aligned}\right\} \Rightarrow \quad \boxed{\det \mathbf{W} = 0\ .} \tag{2.101}$$

The following identity also holds true

$$\boxed{\det (\mathbf{u} \otimes \mathbf{v}) = 0\ .} \quad \longleftarrow \text{see (2.180)} \tag{2.102}$$

It can easily be shown that the determinant and trace of a tensor $\mathbf{A}$ are related via the relation

$$\boxed{\det \mathbf{A} = \frac{1}{6}\left[2\text{tr}\mathbf{A}^3 - 3\left(\text{tr}\mathbf{A}\right)\left(\text{tr}\mathbf{A}^2\right) + \left(\text{tr}\mathbf{A}\right)^3\right] .} \tag{2.103}$$

This expression can be extracted from the Cayley-Hamilton equation (4.21), see also (4.15a)–(4.15c) and $(4.17\text{c})_{1-2}$.

Since $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in the definition (2.98) are linearly independent, they form a basis and, therefore, $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) \neq 0$ is guaranteed. As a result, $\det \mathbf{A} \neq 0$ if and only if $\{\mathbf{Au}, \mathbf{Av}, \mathbf{Aw}\}$ forms a basis. In tensor algebra, whether $\det \mathbf{A} = 0$ or $\det \mathbf{A} \neq 0$ is crucially important. A tensor $\mathbf{A}$ is referred to as *singular* if $\det \mathbf{A} = 0$. By contrast, a *nonsingular* tensor is one for which $\det \mathbf{A} \neq 0$. With regard to this, a tensor $\mathbf{A}$ represents an *invertible* tensor if

★ $\det \mathbf{A} \neq 0$,
★ $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) \neq 0$ implies that $\mathbf{Au} \cdot (\mathbf{Av} \times \mathbf{Aw}) \neq 0$,
★ $\mathbf{u} \times \mathbf{v} \neq 0$ implies that $\mathbf{Au} \times \mathbf{Av} \neq 0$, or
★ $\mathbf{Au} = \mathbf{0}$ implies that $\mathbf{u} = \mathbf{0}$.[9]

See Gurtin et al. [4] for more discussions and proof.

2.5.2 *Inverse of a Tensor*

Let $\mathbf{A}$ be an invertible tensor. Then, for the linear map $\mathbf{v} = \mathbf{Au}$, there exists a tensor $\mathbf{A}^{-1} \in \mathscr{T}_{\text{so}}$ such that

$$\mathbf{u} = \mathbf{A}^{-1}\mathbf{v} . \tag{2.104}$$

Here, $\mathbf{A}^{-1}$ is called the *inverse* of $\mathbf{A}$. Then, the identities $\mathbf{u} = \mathbf{A}^{-1}\mathbf{v} = \mathbf{A}^{-1}\left(\mathbf{Au}\right)$ and $\mathbf{v} = \mathbf{A}\left(\mathbf{A}^{-1}\mathbf{v}\right)$ with the aid of (2.5), (2.6) and (2.25), imply the following reciprocal expression[10]

$$\boxed{\mathbf{AA}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A} .} \tag{2.105}$$

[9] Recall from linear algebra that if $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{Au} = \mathbf{0}$ then $\det \mathbf{A} = 0$.

[10] Having in mind the identity matrix $(1.46)_3$, the relations in (2.105) numerically hold true for all invertible tensors. But from the consistency point of view, it may not hold for some special tensors. For instance, consider the simple contraction between a covariant tensor $\mathbf{A}$ and its inverse $\mathbf{A}^{-1}$ which is basically a contravariant tensor. The result $\mathbf{AA}^{-1}$ thus manifests the co-contravariant unit tensor which should be distinguished from the contra-covariant unit tensor $\mathbf{A}^{-1}\mathbf{A}$, see Chap. 5. Another discrepancy can be observed in the context of nonlinear solid mechanics for which different configurations needs to be taken into account for developing its theoretical formulations. Eventually, vast majority of tensors in the finite deformation problems address one specific configuration. But there exists some particular tensors, called two-point tensors, that interact between two different configurations; an example of which will be the deformation gradient $\mathbf{F}$. In can be shown that $\mathbf{F}^{-1}\mathbf{F}$ is the unit tensor in the material description while $\mathbf{FF}^{-1}$ presents the spatial unit tensor, see Exercise 6.16. Thus, the reciprocal expression (2.105) reveals the lack of consistency regarding the two-point tensors within the context of nonlinear continuum mechanics of solids.

By setting $\mathbf{A} = \mathbf{I}$ in (2.105)$_1$ and then using (2.33)$_2$, one then obtains

$$\mathbf{I}^{-1} = \mathbf{I} \ . \tag{2.106}$$

Consistent with (2.19)$_2$, the Cartesian tensor $\mathbf{A}^{-1}$ can be expressed as

$$\boxed{\mathbf{A}^{-1} = A_{ij}^{-1} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \quad \text{where} \quad A_{ij}^{-1} := \left(\mathbf{A}^{-1}\right)_{ij} = \widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}^{-1} \widehat{\mathbf{e}}_j\right] \ .} \tag{2.107}$$

As a result, the component representation of (2.105) will take the form

$$A_{ik} A_{kj}^{-1} = \delta_{ij} = A_{ik}^{-1} A_{kj} \ . \tag{2.108}$$

For any two invertible tensors $\mathbf{A}$ and $\mathbf{B}$, the following properties hold

$$\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1} \ , \ \leftarrow \text{ since } \mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \Rightarrow \det\left(\mathbf{A}\mathbf{A}^{-1}\right) = \det \mathbf{I} \Rightarrow \det \mathbf{A} \det \mathbf{A}^{-1} = 1 \tag{2.109a}$$

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \ , \ \leftarrow \text{ since } (\mathbf{A}\mathbf{B})(\mathbf{A}\mathbf{B})^{-1} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}\mathbf{I}\mathbf{A}^{-1} = \mathbf{A}\mathbf{B}\mathbf{B}^{-1}\mathbf{A}^{-1} \tag{2.109b}$$

$$\left(\mathbf{A}^{-1}\right)^m = \left(\mathbf{A}^m\right)^{-1} \ , \ \leftarrow \text{ where } m \text{ denotes a nonnegative integer} \tag{2.109c}$$

$$(\alpha\mathbf{A})^{-1} = \alpha^{-1}\mathbf{A}^{-1} \ , \ \leftarrow \text{ since } (\alpha\mathbf{A})^{-1}(\alpha\mathbf{A}) = \mathbf{I} = \alpha^{-1}\mathbf{A}^{-1}(\alpha\mathbf{A}) \tag{2.109d}$$

$$\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A} \ , \ \leftarrow \text{ since } \left(\mathbf{A}^{-1}\right)\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A} \tag{2.109e}$$

$$\left(\mathbf{A}^{-1}\right)^{\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)^{-1} \ , \ \leftarrow \text{ since } \left(\mathbf{A}^{-1}\mathbf{A}\right)^{\mathrm{T}} = \mathbf{I}^{\mathrm{T}} \Rightarrow \mathbf{A}^{\mathrm{T}}\left(\mathbf{A}^{-1}\right)^{\mathrm{T}} = \mathbf{I} = \mathbf{A}^{\mathrm{T}}\left(\mathbf{A}^{\mathrm{T}}\right)^{-1} \tag{2.109f}$$

$$\mathrm{tr}\left(\mathbf{A}\mathbf{B}\mathbf{A}^{-1}\right) = \mathrm{tr}\mathbf{B} \ . \ \leftarrow \text{ since } \delta_{ij} A_{ik} B_{kl} A_{lj}^{-1} = A_{ik} B_{kl} A_{li}^{-1} = \delta_{kl} B_{kl} \tag{2.109g}$$

For simplicity, the following notation

$$\mathbf{A}^{-\mathrm{T}} := \left(\mathbf{A}^{-1}\right)^{\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)^{-1} \quad , \quad \mathbf{A}^{-m} := \left(\mathbf{A}^{-1}\right)^m = \left(\mathbf{A}^m\right)^{-1} \ , \tag{2.110}$$

will be adopted. By means of $\mathbf{I} : \mathbf{I} = 3$ and $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, it follows that

$$\mathbf{I} : \mathbf{A}\mathbf{A}^{-1} = 3 \quad \Rightarrow \quad \boxed{\mathbf{A}^{-\mathrm{T}} : \mathbf{A} = 3.} \quad \leftarrow \text{recall from (2.84) that } \mathbf{A} : \mathbf{B} = \mathbf{I} : \mathbf{B}\mathbf{A}^{\mathrm{T}} \tag{2.111}$$

The expression (2.105) only shows that $\mathbf{A}^{-1}$ can uniquely be determined from $\mathbf{A}$ and does not provide any information regarding its computation. With regard to the coordinate representation (2.107)$_1$, the goal is now to obtain A_{ij}^{-1} (or equivalently calculate the inverse of the matrix $[\,\mathbf{A}\,]$). There are various methods in the literature. An effective way with closed-form solution will be discussed in the following.

2.5.3 Cofactor of a Tensor

The *cofactor* of a tensor $\mathbf{A}$, designated by $\mathbf{A}^{\mathrm{c}}$, is defined by

$$\boxed{\mathbf{A}^{\mathrm{c}}\left(\mathbf{u}\times\mathbf{v}\right) = \left(\mathbf{A}\mathbf{u}\right)\times\left(\mathbf{A}\mathbf{v}\right)\ ,} \tag{2.112}$$

for all linearly independent vectors $\mathbf{u}$ and $\mathbf{v}$ for which $\mathsf{A} = |\mathbf{u}\times\mathbf{v}| \neq 0$, see Fig. 1.4. It is a **linear transformation** that operates on the area vector $\mathbf{u}\times\mathbf{v}$ to produce another area vector $\mathbf{Au}\times\mathbf{Av}$. And, it satisfies the following property

$$\underbrace{\left(\mathbf{A}^{\mathrm{c}}\mathbf{B}^{\mathrm{c}}\right)\left(\mathbf{u}\times\mathbf{v}\right) = \left(\mathbf{ABu}\right)\times\left(\mathbf{ABv}\right) = \left(\mathbf{AB}\right)^{\mathrm{c}}\left(\mathbf{u}\times\mathbf{v}\right)}_{\text{since } \mathbf{A}^{\mathrm{c}}\left(\mathbf{B}^{\mathrm{c}}\left(\mathbf{u}\times\mathbf{v}\right)\right) = \mathbf{A}^{\mathrm{c}}\left(\left(\mathbf{Bu}\right)\times\left(\mathbf{Bv}\right)\right) = \mathbf{A}\left(\mathbf{Bu}\right)\times\mathbf{A}\left(\mathbf{Bv}\right) = \left(\mathbf{AB}\right)\mathbf{u}\times\left(\mathbf{AB}\right)\mathbf{v},\ \text{see } (2.25)} \quad . \tag{2.113}$$

Let $\mathbf{A}$ be an invertible tensor. Then, the tensor $\mathbf{A}^{\mathrm{c}}$ in (2.112) is expressible in terms of $\mathbf{A}^{-1}$ as follows:

$$\begin{aligned}
&\xrightarrow[(2.98)]{\text{from}} && (\det\mathbf{A})\,\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = (\mathbf{Au})\cdot(\mathbf{Av}\times\mathbf{Aw}) \\
&\xrightarrow[(1.9\text{c}) \text{ and } (2.112)]{\text{from}} && ((\det\mathbf{A})\,\mathbf{u})\cdot(\mathbf{v}\times\mathbf{w}) = \mathbf{Au}\cdot\mathbf{A}^{\mathrm{c}}(\mathbf{v}\times\mathbf{w}) \\
&\xrightarrow[(2.5) \text{ and } (2.51\text{c})]{\text{from}} && ((\det\mathbf{A})\,\mathbf{Iu})\cdot(\mathbf{v}\times\mathbf{w}) = \left(\mathbf{A}^{\mathrm{c}\,\mathrm{T}}\mathbf{Au}\right)\cdot(\mathbf{v}\times\mathbf{w}) \\
&\xrightarrow[(1.9\text{a}),\ (2.7) \text{ and } (2.8\text{b})]{\text{from}} && (\det\mathbf{A})\,\mathbf{I} = \mathbf{A}^{\mathrm{c}\,\mathrm{T}}\mathbf{A} \quad \text{or} \quad (\det\mathbf{A})\,\delta_{ij} = \left(\mathbf{A}^{\mathrm{c}\,\mathrm{T}}\right)_{ik}(\mathbf{A})_{kj} \\
&\xrightarrow[(2.33) \text{ and } (2.105)]{\text{from}} && \boxed{\mathbf{A}^{\mathrm{c}\,\mathrm{T}} = (\det\mathbf{A})\,\mathbf{A}^{-1}\ .}
\end{aligned} \tag{2.114}$$

In matrix notation and index notation, one will have

$$\left[\mathbf{A}^{-1}\right] = \frac{1}{\det\mathbf{A}}\left[\mathbf{A}^{\mathrm{c}}\right]^{\mathrm{T}} \quad , \quad \left(\mathbf{A}^{-1}\right)_{ij} = \frac{1}{\det\mathbf{A}}\left(\mathbf{A}^{\mathrm{c}}\right)_{ji}\ . \tag{2.115}$$

To compute $\mathbf{A}^{-1}$, one thus needs to have the coordinate representation of $\mathbf{A}^{\mathrm{c}}$. To do so, consider first

$$\begin{aligned}
\mathbf{A}^{\mathrm{c}}\left(\mathbf{u}\times\mathbf{v}\right) &\xlongequal[\text{hand from } (1.67)]{\text{on the one}} \mathbf{A}^{\mathrm{c}}\left(u_m v_n \varepsilon_{mnj}\widehat{\mathbf{e}}_j\right) \\
&\xlongequal[(2.2)]{\text{from}} \varepsilon_{mnj} u_m v_n \left(\mathbf{A}^{\mathrm{c}}\widehat{\mathbf{e}}_j\right) \\
&\xlongequal[(2.21)]{\text{from}} \left(A^{\mathrm{c}}_{ij}\varepsilon_{mnj} u_m v_n\right)\widehat{\mathbf{e}}_i \\
&\xlongequal[\text{hand from } (2.112)]{\text{on the other}} \left(\mathbf{Au}\right)\times\left(\mathbf{Av}\right) \\
&\xlongequal[(2.22)]{\text{from}} \left(A_{km} u_m \widehat{\mathbf{e}}_k\right)\times\left(A_{ln} v_n \widehat{\mathbf{e}}_l\right)
\end{aligned}$$

$$\xlongequal[\text{(1.49a) and (1.49b)}]{\text{from}} (A_{km} A_{ln} u_m v_n) (\widehat{\mathbf{e}}_k \times \widehat{\mathbf{e}}_l)$$
$$\xlongequal[\text{(1.64)}]{\text{from}} (A_{km} A_{ln} \varepsilon_{kli} u_m v_n) \widehat{\mathbf{e}}_i \ ,$$

which results in

$$\left(A^{\mathrm{c}}_{ij} \varepsilon_{mnj} - A_{km} A_{ln} \varepsilon_{kli}\right) u_m v_n = 0 \ .$$

Arbitrariness of **u** and **v** then implies that

$$\boxed{A^{\mathrm{c}}_{io} \varepsilon_{mno} = A_{km} A_{ln} \varepsilon_{kli} \ .} \tag{2.116}$$

Finally, multiplying both sides of the above expression by ε_{mnj}, taking into account (1.54) and $(1.58\text{b})_3$, yields

$$\boxed{\left(\mathbf{A}^{\mathrm{c}}\right)_{ij} = \frac{\varepsilon_{ikl} \varepsilon_{jmn} (\mathbf{A})_{km} (\mathbf{A})_{ln}}{2} \quad \text{and, hence,} \quad \mathbf{A}^{\mathrm{c}} = \frac{\varepsilon_{ikl} \varepsilon_{jmn} A_{km} A_{ln}}{2} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ .} \tag{2.117}$$

The interested reader can arrive at this result in an alternative way. Guided by the expression $(2.20)_2$, one can write

$$\begin{aligned}
\widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}^{\mathrm{c}} \widehat{\mathbf{e}}_j\right] &\xlongequal[\text{(1.66)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\mathbf{A}^{\mathrm{c}} \left(\frac{1}{2} \varepsilon_{jmn} \widehat{\mathbf{e}}_m \times \widehat{\mathbf{e}}_n\right)\right] \\
&\xlongequal[\text{(2.2)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\frac{1}{2} \varepsilon_{jmn} \mathbf{A}^{\mathrm{c}} (\widehat{\mathbf{e}}_m \times \widehat{\mathbf{e}}_n)\right] \\
&\xlongequal[\text{(2.112)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\frac{1}{2} \varepsilon_{jmn} (\mathbf{A} \widehat{\mathbf{e}}_m) \times (\mathbf{A} \widehat{\mathbf{e}}_n)\right] \\
&\xlongequal[\text{(2.21)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\frac{1}{2} \varepsilon_{jmn} (A_{km} \widehat{\mathbf{e}}_k) \times (A_{ln} \widehat{\mathbf{e}}_l)\right] \\
&\xlongequal[\text{(1.49a) and (1.49b)}]{\text{from}} \widehat{\mathbf{e}}_i \cdot \left[\frac{1}{2} \varepsilon_{jmn} A_{km} A_{ln} (\widehat{\mathbf{e}}_k \times \widehat{\mathbf{e}}_l)\right] \\
&\xlongequal[\text{(1.9a) to (1.9c)}]{\text{from}} \frac{1}{2} \varepsilon_{jmn} A_{km} A_{ln} [\widehat{\mathbf{e}}_i \cdot (\widehat{\mathbf{e}}_k \times \widehat{\mathbf{e}}_l)] \\
&\xlongequal[\text{(1.65)}]{\text{from}} \frac{1}{2} \varepsilon_{ikl} \varepsilon_{jmn} A_{km} A_{ln} \ .
\end{aligned} \tag{2.118}$$

It is interesting to point out that the cofactor of the identity tensor is equal to itself. This can be verified as follows:

$$\begin{aligned}
\left(\mathbf{I}^{\mathrm{c}}\right)_{ij} &\xlongequal[\text{(2.117)}]{\text{from}} \frac{1}{2} \varepsilon_{ikl} \varepsilon_{jmn} \delta_{km} \delta_{ln} \\
&\xlongequal[\text{(1.36)}]{\text{from}} \frac{1}{2} \left[\varepsilon_{ikl} \varepsilon_{jkl}\right]
\end{aligned}$$

$$\underset{(1.58b)}{\overset{\text{from}}{=\!=}} \frac{1}{2}\left[2\delta_{ij}\right]$$
$$\underset{(2.23)}{\overset{\text{from}}{=\!=}} (\mathbf{I})_{ij} \ . \tag{2.119}$$

This result can also be deduced from (2.112) taking into account (2.5) and (2.6).

The tensor $\mathbf{A}^{-1}$, with the aid of $(2.107)_1$, $(2.115)_2$ and $(2.117)_1$, can now be computed according to

$$\boxed{\mathbf{A}^{-1} = (2 \det \mathbf{A})^{-1} \, \varepsilon_{imn} \varepsilon_{jkl} A_{km} A_{ln} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ .} \tag{2.120}$$

The matrix form of a cofactor tensor is presented in what follows. First, one needs to define a matrix $[\,\mathbf{M}\,]$ in such a way that any of its components, M_{ij}, is constructed from the determinant of a matrix generated by removing the i th row and j th column of $[\,\mathbf{A}\,]$ in $(1.47)_1$. Then, the components of $\mathbf{A}^{\mathrm{c}}$ read $(\mathbf{A}^{\mathrm{c}})_{ij} = (-1)^{i+j} (\mathbf{M})_{ij}$. The result thus renders

$$\left[\,\mathbf{A}^{\mathrm{c}}\,\right] = \begin{bmatrix} A_{22}A_{33} - A_{32}A_{23} & A_{31}A_{23} - A_{21}A_{33} & A_{21}A_{32} - A_{31}A_{22} \\ A_{32}A_{13} - A_{12}A_{33} & A_{11}A_{33} - A_{31}A_{13} & A_{31}A_{12} - A_{11}A_{32} \\ A_{12}A_{23} - A_{22}A_{13} & A_{21}A_{13} - A_{11}A_{23} & A_{11}A_{22} - A_{21}A_{12} \end{bmatrix} . \tag{2.121}$$

In light of (1.79) or $(2.117)_1$, the above expression admits some alternative representations. One form is given by

$$\left[\,\mathbf{A}^{\mathrm{c}}\,\right] = \begin{bmatrix} \varepsilon_{1jk}A_{j2}A_{k3} & \varepsilon_{1kj}A_{j1}A_{k3} & \varepsilon_{1jk}A_{j1}A_{k2} \\ \varepsilon_{2jk}A_{j2}A_{k3} & \varepsilon_{2kj}A_{j1}A_{k3} & \varepsilon_{2jk}A_{j1}A_{k2} \\ \varepsilon_{3jk}A_{j2}A_{k3} & \varepsilon_{3kj}A_{j1}A_{k3} & \varepsilon_{3jk}A_{j1}A_{k2} \end{bmatrix} . \tag{2.122}$$

2.6 Positive Definite and Negative Definite Tensors

A second-order tensor $\mathbf{A}$ is said to be *positive* (*semi*-) *definite* if it satisfies the following condition

$$\underbrace{\mathbf{u} \cdot \mathbf{A}\mathbf{u}}_{\text{or } u_i A_{ij} u_j \text{ or } [\mathbf{u}]^{\mathrm{T}} [\mathbf{A}] [\mathbf{u}]} > 0 \ \ (\geq 0) \ , \quad \text{for all } \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{o3} \ , \ \mathbf{u} \neq \mathbf{0} \ . \tag{2.123}$$

Conversely, a *negative* (*semi*-) *definite* tensor $\mathbf{A}$ is one for which

$$\mathbf{u} \cdot \mathbf{A}\mathbf{u} < 0 \ \ (\leq 0) \ , \quad \text{for all } \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{o3} \ , \ \mathbf{u} \neq \mathbf{0} \ . \tag{2.124}$$

The positive definite tensors are of great importance in a wide variety of theory- and computation-based research. A necessary and sufficient condition for a tensor $\mathbf{A}$ to be positive definite is that its symmetric part, i.e. sym**A**, be positive definite.[11] Hence, the upcoming discussions are restricted to only symmetric tensors.

There are equivalent conditions for a symmetric tensor to be positive definite. The most common one is that a symmetric tensor **S** is positive definite if and only if all of its eigenvalues are positive, see Chap. 4.

Equivalently, **S** is positive definite if and only if the determinants corresponding to all upper-left submatrices of [**S**] are positive, i.e.

$$q_1 = \det\left[S_{11}\right] > 0 \ , \tag{2.125a}$$

$$q_2 = \det\begin{bmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{bmatrix} > 0 \ , \tag{2.125b}$$

$$q_3 = \det\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{bmatrix} > 0 \ . \tag{2.125c}$$

The above expressions can also be written as[12]

$$\bar{q}_1 = u_1\left[S_{11}\right]u_1 > 0 \ , \tag{2.126a}$$

$$\bar{q}_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix}\begin{bmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} > 0 \ , \tag{2.126b}$$

$$\bar{q}_3 = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} > 0 \ , \tag{2.126c}$$

for all $u_1 \neq 0$, $u_2 \neq 0$ and $u_3 \neq 0$. Practically, there are necessary (but not sufficient) conditions for a symmetric tensor **S** to be positive definite:

---

[11] This stems from the fact that a skew-symmetric tensor **W** cannot be positive definite since $\mathbf{u} \cdot \mathbf{W}\mathbf{u} = \mathbf{u} \cdot \mathbf{W}^{\mathrm{T}}\mathbf{u} = -\mathbf{u} \cdot \mathbf{W}\mathbf{u}$ simply yields $\mathbf{u} \cdot \mathbf{W}\mathbf{u} = 0$.

[12] The equivalence between (2.125a)–(2.125c) and (2.126a)–(2.126c) will be verified here. To begin with, recall from (2.123) that $u_1 \neq 0$, $u_2 \neq 0$ and $u_3 \neq 0$.

First, if $S_{11} > 0$ then $u_1^2 S_{11} > 0$. Conversely, $u_1^2 S_{11} > 0$ implies that $S_{11} > 0$. Hence, $q_1 > 0$ holds if and only if $\bar{q}_1 > 0$.

Then, in a similar manner, $q_1 > 0$ and $q_2 > 0$ are the necessary and sufficient conditions for the *quadratic form*

$$\bar{q}_2 = \left[q_1\right]\left(u_1 + \frac{S_{12}}{S_{11}}u_2\right)^2 + \left[\frac{q_2}{q_1}\right](u_2)^2 \ ,$$

to be positive definite.

At the end, a similar procedure shows that $q_1 > 0$, $q_2 > 0$ and $q_3 > 0$ hold if and only if

$$\bar{q}_3 = \left[q_1\right]\left(u_1 + \frac{S_{12}}{S_{11}}u_2 + \frac{S_{13}}{S_{11}}u_3\right)^2 + \left[\frac{q_2}{q_1}\right]\left(u_2 + \frac{S_{11}S_{23} - S_{13}S_{12}}{S_{11}}u_3\right)^2 + \left[\frac{q_3}{q_2}\right](u_3)^2 > 0 \ .$$

- Its diagonal components S_{11}, S_{22} and S_{33} are positive.
- Its largest component is a diagonal element.
- $S_{ii} + S_{jj} > 2S_{ij}$, for $i \neq j$ (no sum over i, j).
- $\det \mathbf{S} > 0$.

The last condition shows that the positive definite tensors are always **invertible**.

Of interest here is to examine positive-definiteness of the particular tensors $\mathbf{F}^{\mathrm{T}}\mathbf{F}$ and $\mathbf{F}\mathbf{F}^{\mathrm{T}}$ that are extensively used in nonlinear solid mechanics, see Exercise 6.16. First, one needs to check their symmetry:

$$\left.\begin{aligned}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right)^{\mathrm{T}} &= \mathbf{F}^{\mathrm{T}}\left(\mathbf{F}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathbf{F}^{\mathrm{T}}\mathbf{F} \\ \left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right)^{\mathrm{T}} &= \left(\mathbf{F}^{\mathrm{T}}\right)^{\mathrm{T}}\mathbf{F}^{\mathrm{T}} = \mathbf{F}\mathbf{F}^{\mathrm{T}}\end{aligned}\right\} . \tag{2.127}$$

Then, by means of (1.11) and $(2.51\mathrm{b})_1$, one will have

$$\mathbf{u} \cdot \mathbf{F}^{\mathrm{T}}\mathbf{F}\mathbf{u} = \mathbf{F}\mathbf{u} \cdot \mathbf{F}\mathbf{u} = |\mathbf{F}\mathbf{u}|^2 \quad , \quad \mathbf{u} \cdot \mathbf{F}\mathbf{F}^{\mathrm{T}}\mathbf{u} = \mathbf{F}^{\mathrm{T}}\mathbf{u} \cdot \mathbf{F}^{\mathrm{T}}\mathbf{u} = \left|\mathbf{F}^{\mathrm{T}}\mathbf{u}\right|^2 .$$

Knowing that $|\mathbf{F}\mathbf{u}|^2$ and $\left|\mathbf{F}^{\mathrm{T}}\mathbf{u}\right|^2$ are always nonnegative, the symmetric tensors $\mathbf{F}^{\mathrm{T}}\mathbf{F}$ and $\mathbf{F}\mathbf{F}^{\mathrm{T}}$ will be positive semi-definite. And, they are positive definite tensors if $\det \mathbf{F} = \det \mathbf{F}^{\mathrm{T}} \neq 0$.

2.7 Orthogonal Tensor

A second-order tensor $\mathbf{Q}$ is called *orthogonal* if[13]

$$\boxed{\mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{v} \ , \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathscr{E}_{\mathrm{r}}^{o3} \text{ and } \mathbf{Q} \in \mathscr{T}_{\mathrm{so}} \ .} \tag{2.128}$$

The action of $\mathbf{Q}$ in the above condition can be examined geometrically as well as algebraically. As can be seen, the inner product of two vectors remains unchanged if both transform by an orthogonal tensor. Since the inner product contains the concepts of length and angle, the length $|\mathbf{Q}\mathbf{u}|$ and the angle $\theta(\mathbf{Q}\mathbf{u}, \mathbf{Q}\mathbf{v})$ need to be identified as follows:

$$\text{Length}: \ |\mathbf{Q}\mathbf{u}|^2 = \mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{u} = \mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 \quad \Longrightarrow \quad |\mathbf{Q}\mathbf{u}| = |\mathbf{u}| \ , \tag{2.129a}$$

$$\text{Angle}: \ \cos\theta = \frac{\mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{v}}{|\mathbf{Q}\mathbf{u}|\,|\mathbf{Q}\mathbf{v}|} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} \quad \Longrightarrow \quad \theta(\mathbf{Q}\mathbf{u}, \mathbf{Q}\mathbf{v}) = \theta(\mathbf{u}, \mathbf{v}) \ . \tag{2.129b}$$

One can then deduce that:

[13] In this text, an orthogonal tensor is denoted by $\mathbf{Q}$. And a proper orthogonal (or rotation) tensor is denoted by $\mathbf{R}$.

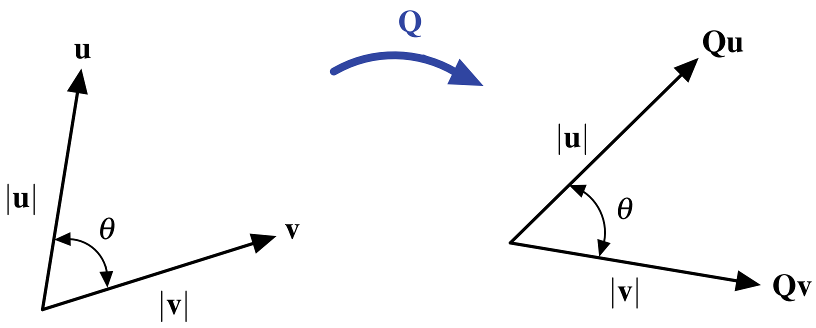


Fig. 2.1 Orthogonal transformation

> The length of any vector as well as the angle between any two vectors are preserved during the orthogonal transformation.

The geometrical interpretation of this linear transformation is depicted in Fig. 2.1.

Next, the algebraic description of an orthogonal tensor is presented. The expression (2.128) holds if and only if

$$\underbrace{\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I} = \mathbf{Q}\mathbf{Q}^{\mathrm{T}}}_{\text{in indicial notation : } Q_{ki}Q_{kj} = \delta_{ij} = Q_{ik}Q_{jk}} , \tag{2.130}$$

or

$$\boxed{\underbrace{\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1}}_{\text{in indicial notation : } Q_{ij} = Q^{-1}_{ji}} .} \tag{2.131}$$

The proof is not difficult. Let $\mathbf{Q}$ be an orthogonal tensor. Then,

$$\mathbf{Qu} \cdot \mathbf{Qv} = \mathbf{u} \cdot \mathbf{v} \xrightarrow[\text{(2.5) and (2.51d)}]{\text{from}} \mathbf{u} \cdot \mathbf{Q}^{\mathrm{T}}\mathbf{Qv} = \mathbf{u} \cdot \mathbf{Iv} \xrightarrow[\text{(2.7)}]{\text{from}} \boxed{\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}}$$

$$\Rightarrow \mathbf{Q}\mathbf{Q}^{\mathrm{T}}\left(\mathbf{Q}\mathbf{Q}^{-1}\right) = \mathbf{QIQ}^{-1} \xrightarrow[\text{(2.28), (2.33) and (2.105)}]{\text{from}} \boxed{\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}\ .}$$

Conversely,

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I} \quad \Rightarrow \quad \mathbf{Q}^{\mathrm{T}}\mathbf{Qv} = \mathbf{Iv} \quad \Rightarrow \quad \mathbf{u} \cdot \mathbf{Q}^{\mathrm{T}}\mathbf{Qv} = \mathbf{u} \cdot \mathbf{Iv} \quad \Rightarrow \quad \boxed{\mathbf{Qu} \cdot \mathbf{Qv} = \mathbf{u} \cdot \mathbf{v}\ .}$$

The set of all orthogonal tensors forms a **group**[14] called *orthogonal group*; indicated by[15]

$$\mathscr{O} = \left\{ \mathbf{Q} \in \mathscr{T}_{so}\left(\mathscr{E}_r^{o3}\right) \mid \mathbf{Q}^T = \mathbf{Q}^{-1} \right\} . \tag{2.132}$$

If $\det \mathbf{Q} = +1\ (-1)$, then $\mathbf{Q}$ is called *proper orthogonal* or *rotation* (*improper orthogonal* or *reflection*). In this regard, the set $\mathscr{O}^+$ also forms a group called *proper orthogonal group*[16]:

$$\mathscr{O}^+ = \{ \mathbf{Q} \in \mathscr{O} \mid \det \mathbf{Q} = +1 \} . \tag{2.133}$$

For any two orthogonal tensors $\mathbf{Q}$ and $\bar{\mathbf{Q}}$, the following identities hold

$$\left(\mathbf{Q}\bar{\mathbf{Q}}\right)^T = \left(\mathbf{Q}\bar{\mathbf{Q}}\right)^{-1} , \quad \leftarrow \text{since } \left(\mathbf{Q}\bar{\mathbf{Q}}\right)^T = \bar{\mathbf{Q}}^T\mathbf{Q}^T = \bar{\mathbf{Q}}^{-1}\mathbf{Q}^{-1} \tag{2.134a}$$

$$\mathbf{Q} : \mathbf{Q} = \mathbf{Q}\bar{\mathbf{Q}} : \mathbf{Q}\bar{\mathbf{Q}} = 3 , \quad \leftarrow \text{since } \mathbf{Q}\bar{\mathbf{Q}} : \mathbf{Q}\bar{\mathbf{Q}} = \mathbf{Q}^T\mathbf{Q} : \bar{\mathbf{Q}}\bar{\mathbf{Q}}^T = \mathbf{I} : \mathbf{I} \tag{2.134b}$$

$$\det \mathbf{Q} = \pm 1 , \quad \leftarrow \text{since } \det\left(\mathbf{Q}^T\mathbf{Q}\right) = \det \mathbf{I} \Rightarrow \left(\det \mathbf{Q}^T\right)\left(\det \mathbf{Q}\right) = 1 \Rightarrow \left(\det \mathbf{Q}\right)^2 = 1 \tag{2.134c}$$

$$\mathbf{Q}^c = \mathbf{Q} , \ \text{if } \det \mathbf{Q} = 1 . \quad \leftarrow \text{since } \mathbf{Q}^{cT} = \left(\det \mathbf{Q}\right)\mathbf{Q}^{-1} \Rightarrow \mathbf{Q}^c = \left(\mathbf{Q}^T\right)^T \tag{2.134d}$$

The last property states that the area of parallelogram defined by the two vectors $\mathbf{u}$ and $\mathbf{v}$ remains unchanged under the proper orthogonal transformation:

$$\boxed{\left|\mathbf{Q}\mathbf{u} \times \mathbf{Q}\mathbf{v}\right| \xlongequal[(2.112)]{\text{from}} \left|\mathbf{Q}^c\left(\mathbf{u} \times \mathbf{v}\right)\right| \xlongequal[(2.134d)]{\text{from}} \left|\mathbf{Q}\left(\mathbf{u} \times \mathbf{v}\right)\right| \xlongequal[(2.129a)]{\text{from}} \left|\mathbf{u} \times \mathbf{v}\right| .} \tag{2.135}$$

Moreover, the rotation tensor preserves the volume spanned by the three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ owing to

$$\boxed{\mathbf{Q}\mathbf{u} \cdot \left(\mathbf{Q}\mathbf{v} \times \mathbf{Q}\mathbf{w}\right) \xlongequal[(2.112) \text{ and } (2.134d)]{\text{from}} \mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\left(\mathbf{v} \times \mathbf{w}\right) \xlongequal[(2.128)]{\text{from}} \mathbf{u} \cdot \left(\mathbf{v} \times \mathbf{w}\right) .} \tag{2.136}$$

Observe that orientation of the basis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is also preserved under the action of rotation tensor. From the previous chapter, recall that $\mathbf{u} \times \mathbf{v}$ obeyed the right-hand screw rule and $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ was declared as positively oriented basis. In this regard, the above result shows that the basis $\{\mathbf{Q}\mathbf{u}, \mathbf{Q}\mathbf{v}, \mathbf{Q}\mathbf{w}\}$ is also positively oriented and a Cartesian coordinate frame defined by the origin $\mathbf{o}$ and a positively oriented orthonormal basis $\{\mathbf{Q}\widehat{\mathbf{e}}_1, \mathbf{Q}\widehat{\mathbf{e}}_2, \mathbf{Q}\widehat{\mathbf{e}}_3\}$ is right-handed.

---

[14] A set $\mathscr{G}$ is said to be a group if it satisfies the following axioms:

1. Closure: For any $A, B \in \mathscr{G}$, $AB \in \mathscr{G}$.
2. Associativity: For any $A, B, C \in \mathscr{G}$, $A(BC) = (AB)C$.
3. Identity: There exists $1 \in \mathscr{G}$ such that $1A = A1 = A$ for any $A \in \mathscr{G}$.
4. Inverse: For any $A \in \mathscr{G}$, there exists $A^{-1} \in \mathscr{G}$ such that $AA^{-1} = A^{-1}A = 1$.

[15] In the literature, the sets $\mathscr{O}$ and $\mathscr{O}^+$ (given in (2.133)) are often denoted by Orth and Orth^+, respectively.

[16] It is worthwhile to mention that the set of all orthogonal tensors $\mathbf{Q}$ with $\det \mathbf{Q} = -1$ does not form a group since the product of two reflection tensors is a rotation tensor.

Let $\mathbf{Q}$ be an orthogonal tensor with the matrix form

$$[\,\mathbf{Q}\,] = \begin{bmatrix} \cdots & \mathbf{u}_1 & \cdots \\ \cdots & \mathbf{u}_2 & \cdots \\ \cdots & \mathbf{u}_3 & \cdots \end{bmatrix} . \tag{2.137}$$

Then, $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ or

$$\begin{bmatrix} \cdots & \mathbf{u}_1 & \cdots \\ \cdots & \mathbf{u}_2 & \cdots \\ \cdots & \mathbf{u}_3 & \cdots \end{bmatrix} \begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} , \tag{2.138}$$

implies that $\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij}$. One can now conclude that:

> The rows (or columns) of an orthogonal matrix form an **orthonormal** basis.

2.8 Projection, Spherical and Deviatoric Tensors

A second-order *projection tensor* $\mathbf{P}$ is a linear transformation satisfying

$$\boxed{\mathbf{P}^2 = \mathbf{P}\ .} \quad \longleftarrow \ \text{it can easily be shown that } \mathbf{P}^m = \mathbf{P} \text{ where } m \text{ denotes an integer number} \tag{2.139}$$

This condition states that when this linear function operates once or twice on an arbitrary vector, the result will be identical, i.e. $\mathbf{P}\,(\mathbf{P}\mathbf{u}) = \mathbf{P}\mathbf{u}$.

Let $\widehat{\mathbf{e}}$ be an arbitrary unit vector. Consistent with Fig. 1.2, any vector $\mathbf{u}$ can be projected onto $\widehat{\mathbf{e}}$ and onto the plane whose normal is that unit vector according to $\mathbf{proj}_{\widehat{\mathbf{e}}}\,\mathbf{u} = (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}}$ and $\mathbf{reje}_{\widehat{\mathbf{e}}}\,\mathbf{u} = \mathbf{u} - (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}}$, respectively. Now, these orthogonal projections help extract two projection tensors via

$$\mathbf{proj}_{\widehat{\mathbf{e}}}\mathbf{u} = (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}} = \underbrace{(\widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}})}_{:=\ \mathbf{P}^{\parallel}_{\widehat{\mathbf{e}}}}\,\mathbf{u} = \mathbf{P}^{\parallel}_{\widehat{\mathbf{e}}}\mathbf{u}\ , \tag{2.140a}$$

$$\mathbf{reje}_{\widehat{\mathbf{e}}}\mathbf{u} = \mathbf{I}\mathbf{u} - (\widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}})\,\mathbf{u} = \underbrace{(\mathbf{I} - \widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}})}_{:=\ \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}}\,\mathbf{u} = \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}\mathbf{u}\ , \tag{2.140b}$$

where (2.5), (2.10) and (2.13) have been used. As can be seen, $\mathbf{P}^{\parallel}_{\widehat{\mathbf{e}}}$ and $\mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}$ represent **symmetric** second-order tensors. They satisfy the following properties

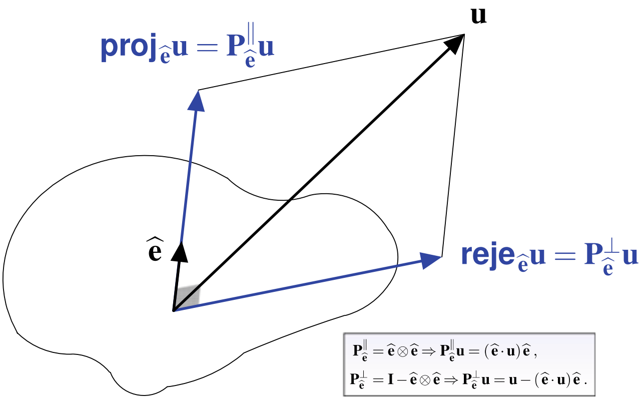


Fig. 2.2 Projection tenors

$$\boxed{\mathbf{P}^{\|}_{\widehat{\mathbf{e}}} + \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}} = \mathbf{I} \quad , \quad \mathbf{P}^{\|}_{\widehat{\mathbf{e}}}\mathbf{P}^{\|}_{\widehat{\mathbf{e}}} = \mathbf{P}^{\|}_{\widehat{\mathbf{e}}} \quad , \quad \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}\mathbf{P}^{\perp}_{\widehat{\mathbf{e}}} = \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}} \quad , \quad \mathbf{P}^{\|}_{\widehat{\mathbf{e}}}\mathbf{P}^{\perp}_{\widehat{\mathbf{e}}} = \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}\mathbf{P}^{\|}_{\widehat{\mathbf{e}}} = \mathbf{O} \ .} \tag{2.141}$$

See Fig. 2.2 for a geometrical interpretation.

Any second-order tensor $\mathbf{A}$ admits the **additive decomposition**

$$\boxed{\mathbf{A} = \mathrm{sph}\mathbf{A} + \mathrm{dev}\mathbf{A} \ ,} \quad \leftarrow \text{see (2.62)} \tag{2.142}$$

into its *spherical* (or *hydrostatic* or *volumetric*) part, sph$\mathbf{A}$, and its *deviatoric* part, dev$\mathbf{A}$, such that[17]

$$\boxed{\mathrm{tr}\,(\mathrm{dev}\mathbf{A}) = 0 \ .} \quad \leftarrow \text{see (2.82)} \tag{2.143}$$

Suppose that[18]

$$\boxed{\mathrm{sph}\mathbf{A} = \frac{1}{3}\,(\mathrm{tr}\mathbf{A})\,\mathbf{I} \ .} \tag{2.144}$$

Then, dev$\mathbf{A}$ can uniquely be determined via

$$\boxed{\mathrm{dev}\mathbf{A} = \mathbf{A} - \frac{1}{3}\,(\mathrm{tr}\mathbf{A})\,\mathbf{I} \ ,} \tag{2.145}$$

and, therefore, the condition (2.143) is trivially satisfied.

The spherical and deviatoric tensors constitute **orthogonal subspaces** of $\mathscr{T}_{\mathrm{so}}\left(\mathscr{E}^{\mathrm{o}3}_{\mathrm{r}}\right)$ owing to

[17] The condition (2.143) is often used to define a deviatoric tensor.

[18] A second-order tensor with identical eigenvalues is called a spherical tensor. Denoting by λ the common value of all eigenvalues, any spherical tensor will be of the form $\lambda\mathbf{I}$. It is thus a symmetric tensor with the quadratic form $\mathbf{u} \cdot (\lambda\mathbf{I})\,\mathbf{u} = \lambda\,|\mathbf{u}|^2$ which can be either positive definite or negative definite, see Beju et al. [5].

$$\text{sph}\mathbf{A} : \text{dev}\mathbf{B} = 0 \ . \quad \leftarrow \text{since } \frac{\text{tr}\mathbf{A}}{3}\mathbf{I} : \left(\mathbf{B} - \frac{\text{tr}\mathbf{B}}{3}\mathbf{I}\right) = \frac{\text{tr}\mathbf{A}\text{tr}\mathbf{B}}{3} - 3\frac{\text{tr}\mathbf{A}\text{tr}\mathbf{B}}{3^2} = 0 \tag{2.146}$$

At the end, one can further establish

$$\text{sph}\,(\text{sph}\mathbf{A}) = \text{sph}\mathbf{A} \quad , \quad \text{sph}\,(\text{dev}\mathbf{A}) = \mathbf{O} \ , \tag{2.147a}$$

$$\text{dev}\,(\text{sph}\mathbf{A}) = \mathbf{O} \quad , \quad \text{dev}\,(\text{dev}\mathbf{A}) = \text{dev}\mathbf{A} \ , \tag{2.147b}$$

and

$$\text{sph}\mathbf{A} : \mathbf{B} = \mathbf{A} : \text{sph}\mathbf{B} \ , \tag{2.148a}$$

$$\text{dev}\mathbf{A} : \mathbf{B} = \mathbf{A} : \text{dev}\mathbf{B} \ , \tag{2.148b}$$

$$\mathbf{A} : \mathbf{B} = \text{sph}\mathbf{A} : \text{sph}\mathbf{B} + \text{dev}\mathbf{A} : \text{dev}\mathbf{B} \ . \tag{2.148c}$$

2.9 Transformation Laws

Recall that a vector space had infinitely many bases and every vector in that space could uniquely be expressed as a linear combination of the elements of a basis. Moreover, recall that a tensor was independent of any coordinate system. But its components depend on an arbitrary chosen coordinate system since they are determined by use of the corresponding basis vectors. That is why the components of a vector or tensor are always declared in advance with respect to a chosen basis. To solve many problems in physics and engineering, it is often desirable to use different coordinate systems. An example of which includes global and local coordinate systems in finite element procedures or spatial multibody dynamics. The components of a vector are known in one coordinate system and it is thus necessary to find its components in the other coordinate system. This motivates to develop relationships between bases of a vector space, known as *transformation laws for basis vectors*. It is also the goal of this section to establish some appropriate transformation laws that enables one to transform the components of a vector (or tensor) with respect to one basis into the components relative to another basis.

Let $\mathscr{G}_{\text{b}}^{o} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be a basis for $\mathscr{E}_{\text{r}}^{o3}$ and $\mathscr{G}_{\text{b}}^{n} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be another basis for the same space. Note that the '*new*' basis $\mathscr{G}_{\text{b}}^{n}$ has been provided upon a linear change in the '*old*' basis $\mathscr{G}_{\text{b}}^{o}$ (in alignment with the linear property of vector space). One can now express each element of the new basis in terms of the elements of the old basis as

$$\left.\begin{aligned} \mathbf{v}_1 &= M_{11}\mathbf{u}_1 + M_{21}\mathbf{u}_2 + M_{31}\mathbf{u}_3 \\ \mathbf{v}_2 &= M_{12}\mathbf{u}_1 + M_{22}\mathbf{u}_2 + M_{32}\mathbf{u}_3 \\ \mathbf{v}_3 &= M_{13}\mathbf{u}_1 + M_{23}\mathbf{u}_2 + M_{33}\mathbf{u}_3 \end{aligned}\right\} \ , \tag{2.149}$$

or

$$\mathbf{v}_i = M_{ji}\mathbf{u}_j \ , \tag{2.150}$$

to introduce the following 3 × 3 **invertible** matrix

$$[\,\mathbf{M}\,] = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_1 \cdot \mathbf{v}_2 & \mathbf{u}_1 \cdot \mathbf{v}_3 \\ \mathbf{u}_2 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_2 & \mathbf{u}_2 \cdot \mathbf{v}_3 \\ \mathbf{u}_3 \cdot \mathbf{v}_1 & \mathbf{u}_3 \cdot \mathbf{v}_2 & \mathbf{u}_3 \cdot \mathbf{v}_3 \end{bmatrix} , \tag{2.151}$$

known as the *change-of-basis matrix* from $\mathscr{G}_{\mathrm{b}}^{o}$ to $\mathscr{G}_{\mathrm{b}}^{n}$. This transformation matrix helps rewrite (2.149) in the convenient form

$$[\,\mathbf{V}\,] = [\,\mathbf{M}\,]^{\mathrm{T}}\,[\,\mathbf{U}\,] \ , \tag{2.152}$$

where

$$[\,\mathbf{V}\,] = \begin{bmatrix} \cdots & \mathbf{v}_1 & \cdots \\ \cdots & \mathbf{v}_2 & \cdots \\ \cdots & \mathbf{v}_3 & \cdots \end{bmatrix} \ , \quad [\,\mathbf{U}\,] = \begin{bmatrix} \cdots & \mathbf{u}_1 & \cdots \\ \cdots & \mathbf{u}_2 & \cdots \\ \cdots & \mathbf{u}_3 & \cdots \end{bmatrix} . \tag{2.153}$$

Vector transformation law. Any vector $\mathbf{w} \in \mathscr{E}_{\mathrm{r}}^{o3}$ with respect to the aforementioned general bases can be represented by

$$\mathbf{w} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + a_3\mathbf{u}_3 = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + b_3\mathbf{v}_3 \ . \tag{2.154}$$

Collecting the components of $\mathbf{w}$ relative to the bases $\mathscr{G}_{\mathrm{b}}^{o}$ and $\mathscr{G}_{\mathrm{b}}^{n}$ in the single-column matrices

$$[\,\mathbf{w}\,]^{o} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \ , \quad [\,\mathbf{w}\,]^{n} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} , \tag{2.155}$$

helps represent (2.154) as

$$[\,\mathbf{w}\,] = [\,\mathbf{U}\,]^{\mathrm{T}}\,[\,\mathbf{w}\,]^{o} = [\,\mathbf{V}\,]^{\mathrm{T}}\,[\,\mathbf{w}\,]^{n} = [\,\mathbf{U}\,]^{\mathrm{T}}\,[\,\mathbf{M}\,]\,[\,\mathbf{w}\,]^{n} \ . \tag{2.156}$$

And this implies that the components $[\,\mathbf{w}\,]^{o}$ and $[\,\mathbf{w}\,]^{n}$ of $\mathbf{w}$ with respect to the different bases $\mathscr{G}_{\mathrm{b}}^{o}$ and $\mathscr{G}_{\mathrm{b}}^{n}$ are related through the transformation matrix $[\,\mathbf{M}\,]$ via

$$[\,\mathbf{w}\,]^{n} = [\,\mathbf{M}\,]^{-1}\,[\,\mathbf{w}\,]^{o} \ . \tag{2.157}$$

Suppose one is given the two orthonormal bases $\{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2, \widehat{\mathbf{u}}_3\}$ and $\{\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \widehat{\mathbf{v}}_3\}$. Then, the change-of-basis matrix $[\,\mathbf{M}\,]$ becomes an orthogonal matrix (because the columns or rows of an orthogonal matrix constitute an orthonormal basis).

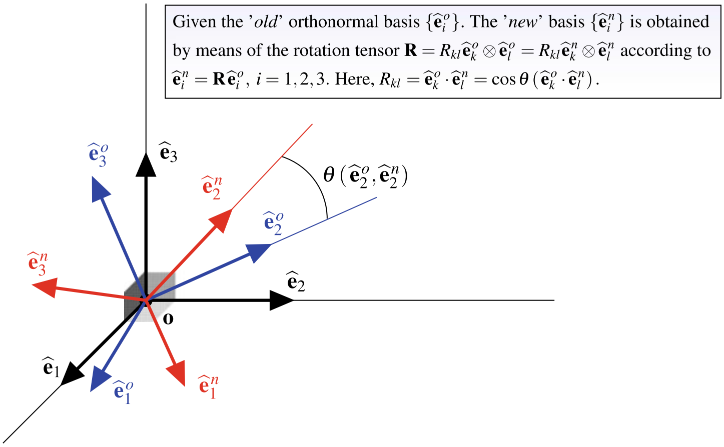


Fig. 2.3 Rotations of a right-handed Cartesian coordinate frame about its origin

2.9.1 Change of Cartesian Basis

The goal is now to specialize the general transformation rule (2.152) to the broadly used Cartesian coordinate system. Let $\left\{\widehat{\mathbf{e}}_i^o\right\}$ be an orthonormal basis of an '*old*' Cartesian coordinate frame. Further, let $\left\{\widehat{\mathbf{e}}_i^n\right\}$ be another set of mutually orthogonal unit vectors corresponding to a '*new*' Cartesian coordinate frame. It is assumed that both frames share the same origin and the new frame has been provided upon a **rotation** of the old one around the origin, see Fig. 2.3. Note that the latter assumption is made to keep the right-handedness of the standard basis.

Under a proper orthogonal transformation of $\left\{\widehat{\mathbf{e}}_i^o\right\}$, one will have

$$\widehat{\mathbf{e}}_i^n = \left(\widehat{\mathbf{e}}_j^o \cdot \widehat{\mathbf{e}}_i^n\right)\widehat{\mathbf{e}}_j^o = \left[\cos\theta\left(\widehat{\mathbf{e}}_j^o, \widehat{\mathbf{e}}_i^n\right)\right]\widehat{\mathbf{e}}_j^o = R_{ji}\widehat{\mathbf{e}}_j^o \ . \tag{2.158}$$

In an alternative way, one can also have

$$\widehat{\mathbf{e}}_j^o \xlongequal[(2.5)]{\text{from}} \mathbf{I}\widehat{\mathbf{e}}_j^o \xlongequal[(2.23)]{\text{from}} \left(\widehat{\mathbf{e}}_i^n \otimes \widehat{\mathbf{e}}_i^n\right)\widehat{\mathbf{e}}_j^o \xlongequal[\text{(2.8a) and (2.13)}]{\text{from}} \left(\widehat{\mathbf{e}}_j^o \cdot \widehat{\mathbf{e}}_i^n\right)\widehat{\mathbf{e}}_i^n = R_{ji}\widehat{\mathbf{e}}_i^n \ . \tag{2.159}$$

The change-of-basis matrix in (2.151) now takes the following form

$$[\,\mathbf{R}\,] = \begin{bmatrix} \cos\theta\left(\widehat{\mathbf{e}}_1^o, \widehat{\mathbf{e}}_1^n\right) & \cos\theta\left(\widehat{\mathbf{e}}_1^o, \widehat{\mathbf{e}}_2^n\right) & \cos\theta\left(\widehat{\mathbf{e}}_1^o, \widehat{\mathbf{e}}_3^n\right) \\ \cos\theta\left(\widehat{\mathbf{e}}_2^o, \widehat{\mathbf{e}}_1^n\right) & \cos\theta\left(\widehat{\mathbf{e}}_2^o, \widehat{\mathbf{e}}_2^n\right) & \cos\theta\left(\widehat{\mathbf{e}}_2^o, \widehat{\mathbf{e}}_3^n\right) \\ \cos\theta\left(\widehat{\mathbf{e}}_3^o, \widehat{\mathbf{e}}_1^n\right) & \cos\theta\left(\widehat{\mathbf{e}}_3^o, \widehat{\mathbf{e}}_2^n\right) & \cos\theta\left(\widehat{\mathbf{e}}_3^o, \widehat{\mathbf{e}}_3^n\right) \end{bmatrix} . \tag{2.160}$$

These nine quantities, that construct a non-symmetric matrix, are known as *direction cosines*. Notice that the columns (or rows) of this matrix are mutually orthogonal and each column is of unit length. This implies that $[\,\mathbf{R}\,]$ is an orthogonal matrix which satisfies $[\,\mathbf{R}\,]^{\mathrm{T}}[\,\mathbf{R}\,] = [\,\mathbf{I}\,] = [\,\mathbf{R}\,][\,\mathbf{R}\,]^{\mathrm{T}}$. And the determinant of this matrix is obviously $+1$.

The proper orthogonal matrix $[\,\mathbf{R}\,]$ translates the old basis $\{\widehat{\mathbf{e}}_i^o\}$ into the new basis $\{\widehat{\mathbf{e}}_i^n\}$. This is indicated by $[\,\widehat{\mathbf{e}}_i^n\,] = [\,\mathbf{R}\,][\,\widehat{\mathbf{e}}_i^o\,]$, $i = 1, 2, 3$. Consistent with this, $[\,\mathbf{R}\,]^{\mathrm{T}}$ rotates $\{\widehat{\mathbf{e}}_i^n\}$ back to $\{\widehat{\mathbf{e}}_i^o\}$, i.e. $[\,\widehat{\mathbf{e}}_i^o\,] = [\,\mathbf{R}\,]^{\mathrm{T}}[\,\widehat{\mathbf{e}}_i^n\,]$, $i = 1, 2, 3$. The proper orthogonal matrix (2.160) is basically the matrix representation of

$$\mathbf{R} = \widehat{\mathbf{e}}_i^n \otimes \widehat{\mathbf{e}}_i^o = R_{ji}\widehat{\mathbf{e}}_j^o \otimes \widehat{\mathbf{e}}_i^o \quad \text{or} \quad \mathbf{R} = \widehat{\mathbf{e}}_j^n \otimes \widehat{\mathbf{e}}_j^o = R_{ji}\widehat{\mathbf{e}}_j^n \otimes \widehat{\mathbf{e}}_i^n \ . \tag{2.161}$$

Observe that the components of the rotation tensor $\mathbf{R}$ are identical in either basis. Guided by Fig. 2.3, one can also introduce the rotation tensors $\mathbf{R}^o = \widehat{\mathbf{e}}_i^o \otimes \widehat{\mathbf{e}}_i$ and $\mathbf{R}^n = \widehat{\mathbf{e}}_i^n \otimes \widehat{\mathbf{e}}_i$ satisfying

$$\widehat{\mathbf{e}}_i^o = \mathbf{R}^o\,\widehat{\mathbf{e}}_i \quad , \quad \widehat{\mathbf{e}}_i^n = \mathbf{R}^n\,\widehat{\mathbf{e}}_i \ , \quad i = 1, 2, 3 \ . \tag{2.162}$$

They help represent the rotation tensor $\mathbf{R} = \widehat{\mathbf{e}}_i^n \otimes \widehat{\mathbf{e}}_i^o$ according to

$$\mathbf{R} \xrightarrow[\text{(2.161) and (2.162)}]{\text{from}} \mathbf{R}^n\,\widehat{\mathbf{e}}_i \otimes \mathbf{R}^o\,\widehat{\mathbf{e}}_i \xrightarrow[\text{(2.32) and (2.55c)}]{\text{from}} \mathbf{R}^n\mathbf{R}^{o\,\mathrm{T}} \ . \tag{2.163}$$

This result basically shows that $\mathbf{R}$ has multiplicatively been decomposed into $\mathbf{R}^n$ and $\mathbf{R}^{o\,\mathrm{T}}$. It is worthwhile to mention that $\mathbf{R}$ is independent of the angles that the old basis vectors make with the standard ones of the frame of reference.[19]

---

[19] Here, the introduced rotation tensors are expressed in the two-dimensional space for the sake of clarity. To have a frame of reference, consider a conventional x_1x_2-Cartesian coordinate system in two dimensions possessing the origin $\mathbf{o}$ and the standard basis $[\,\widehat{\mathbf{e}}_i\,] = \left[\,\delta_{i1}\ \delta_{i2}\,\right]^{\mathrm{T}}$, $i = 1, 2$. Let $\widehat{\mathbf{e}}_1^o$ and $\widehat{\mathbf{e}}_2^o$ be the orthonormal basis vectors of an old $x_1^o x_2^o$-Cartesian coordinate system that is obtained by counterclockwise rotating the x_1 and x_2 axes through an angle α about the origin $\mathbf{o}$. The old basis $\{\widehat{\mathbf{e}}_1^o, \widehat{\mathbf{e}}_2^o\}$ and the rotation tensor $\mathbf{R}^o$ that implies this transformation are then given by

$$\begin{aligned}
&\widehat{\mathbf{e}}_1^o = \cos\alpha\,\widehat{\mathbf{e}}_1 + \sin\alpha\,\widehat{\mathbf{e}}_2 \quad , \quad \widehat{\mathbf{e}}_2^o = -\sin\alpha\,\widehat{\mathbf{e}}_1 + \cos\alpha\,\widehat{\mathbf{e}}_2 \ , \\
&\mathbf{R}^o = \cos\alpha\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_1 - \sin\alpha\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_2 + \sin\alpha\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_1 + \cos\alpha\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_2 \ .
\end{aligned}$$

One can further rotate the x_1^o and x_2^o axes counterclockwise through an angle θ about the origin $\mathbf{o}$ to produce a new $x_1^n x_2^n$-Cartesian coordinate system. The new basis $\{\widehat{\mathbf{e}}_1^n, \widehat{\mathbf{e}}_2^n\}$ then admits the following two forms

$$\begin{aligned}
&\widehat{\mathbf{e}}_1^n = \cos(\alpha+\theta)\,\widehat{\mathbf{e}}_1 + \sin(\alpha+\theta)\,\widehat{\mathbf{e}}_2 \quad , \quad \widehat{\mathbf{e}}_2^n = -\sin(\alpha+\theta)\,\widehat{\mathbf{e}}_1 + \cos(\alpha+\theta)\,\widehat{\mathbf{e}}_2 \ , \\
&\widehat{\mathbf{e}}_1^n = \cos\theta\,\widehat{\mathbf{e}}_1^o + \sin\theta\,\widehat{\mathbf{e}}_2^o \quad , \quad \widehat{\mathbf{e}}_2^n = -\sin\theta\,\widehat{\mathbf{e}}_1^o + \cos\theta\,\widehat{\mathbf{e}}_2^o \ .
\end{aligned}$$

Accordingly, the rotation tensors $\mathbf{R}^n$ (transforming $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2\}$ into $\{\widehat{\mathbf{e}}_1^n, \widehat{\mathbf{e}}_2^n\}$) and $\mathbf{R}$ (transforming $\{\widehat{\mathbf{e}}_1^o, \widehat{\mathbf{e}}_2^o\}$ into $\{\widehat{\mathbf{e}}_1^n, \widehat{\mathbf{e}}_2^n\}$) render

$$\begin{aligned}
\mathbf{R}^n &= \cos(\alpha+\theta)\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_1 - \sin(\alpha+\theta)\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_2 + \sin(\alpha+\theta)\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_1 + \cos(\alpha+\theta)\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_2 \\
&= \cos\theta\,\widehat{\mathbf{e}}_1^o \otimes \widehat{\mathbf{e}}_1 - \sin\theta\,\widehat{\mathbf{e}}_1^o \otimes \widehat{\mathbf{e}}_2 + \sin\theta\,\widehat{\mathbf{e}}_2^o \otimes \widehat{\mathbf{e}}_1 + \cos\theta\,\widehat{\mathbf{e}}_2^o \otimes \widehat{\mathbf{e}}_2 \ , \\
\mathbf{R} &= \cos\theta\,\widehat{\mathbf{e}}_1^o \otimes \widehat{\mathbf{e}}_1^o - \sin\theta\,\widehat{\mathbf{e}}_1^o \otimes \widehat{\mathbf{e}}_2^o + \sin\theta\,\widehat{\mathbf{e}}_2^o \otimes \widehat{\mathbf{e}}_1^o + \cos\theta\,\widehat{\mathbf{e}}_2^o \otimes \widehat{\mathbf{e}}_2^o \ .
\end{aligned}$$

Hint: In many applications the reference frame and the old frame coincide, i.e. $\{\widehat{\mathbf{e}}_i^o\} = \{\widehat{\mathbf{e}}_i\}$. In this case, $\mathbf{R}^o = \mathbf{I}$ and hence (2.163) becomes $\mathbf{R} = \mathbf{R}^n$. In this regard, the components of a vector with respect to these two frames are known as its *global* and *local* coordinates. Note that an arbitrary basis of the local coordinate system is often not orthonormal.

2.9.2 Change of Cartesian Components

Let $u_i^o,\ i = 1, 2, 3$, be the old Cartesian components of an arbitrary vector $\mathbf{u}$ resolved along $\{\widehat{\mathbf{e}}_i^o\}$. Further, let $u_i^n,\ i = 1, 2, 3$, be the new rectangular components of that vector with respect to $\{\widehat{\mathbf{e}}_i^n\}$ in which every element of this new basis obeys (2.158). The **vectorial** transformation law then renders

$$\boxed{u_i^n \xlongequal[\text{(1.33)}]{\text{from}} \mathbf{u} \cdot \widehat{\mathbf{e}}_i^n \xlongequal[\text{(2.158)}]{\text{from}} \mathbf{u} \cdot \left(R_{ji} \widehat{\mathbf{e}}_j^o\right) \xlongequal[\text{(1.9a)-(1.9c)}]{\text{from}} R_{ji} \mathbf{u} \cdot \widehat{\mathbf{e}}_j^o \xlongequal[\text{(1.33)}]{\text{from}} R_{ji} u_j^o \ ,} \tag{2.164}$$

or

$$\boxed{[\,\mathbf{u}\,]^n \xlongequal[\text{(2.157)}]{\text{from}} [\,\mathbf{R}\,]^{-1} [\,\mathbf{u}\,]^o \xlongequal[\text{(2.131)}]{\text{from}} [\,\mathbf{R}\,]^{\mathrm{T}} [\,\mathbf{u}\,]^o \ ,} \tag{2.165}$$

where

$$[\,\mathbf{u}\,]^o = \begin{bmatrix} u_1^o & u_2^o & u_3^o \end{bmatrix}^{\mathrm{T}} \quad , \quad [\,\mathbf{u}\,]^n = \begin{bmatrix} u_1^n & u_2^n & u_3^n \end{bmatrix}^{\mathrm{T}} \ . \tag{2.166}$$

The components of a vector eventually represent that vector in a coordinate system. And if these components transform according to $(2.164)_4$, then they always construct the same vector. With regard to this, some authors prefer to introduce a vector as a set of numbers that obey this transformation rule under a change of coordinates. For the vectorial transformation law in a general curvilinear coordinate system, see (5.59)–(5.60) and (5.105a)–(5.105b).

Hint: Note that the vector $\mathbf{u}$ with two forms $\mathbf{u} = u_k^o \widehat{\mathbf{e}}_k^o = u_k^n \widehat{\mathbf{e}}_k^n$ can itself translate into another vector $\mathbf{u}_{\text{rot}} = \mathbf{R}\mathbf{u}$ with two forms $\mathbf{u}_{\text{rot}} = u_{\text{rot}\,i}^o \widehat{\mathbf{e}}_i^o = u_{\text{rot}\,i}^n \widehat{\mathbf{e}}_i^n$ where $u_{\text{rot}\,i}^o = R_{ij} u_j^o$ and $u_{\text{rot}\,i}^n = R_{ij} u_j^n$, see Fig. 2.4.

A simple algebraic manipulation reveals the fact that the components of $\mathbf{R}$ remain the same with respect to either $\{\widehat{\mathbf{e}}_1^n, \widehat{\mathbf{e}}_2^n\}$ or $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2\}$, that is,

$$\begin{aligned} \mathbf{R} &= \cos\theta\, \widehat{\mathbf{e}}_1^n \otimes \widehat{\mathbf{e}}_1^n - \sin\theta\, \widehat{\mathbf{e}}_1^n \otimes \widehat{\mathbf{e}}_2^n + \sin\theta\, \widehat{\mathbf{e}}_2^n \otimes \widehat{\mathbf{e}}_1^n + \cos\theta\, \widehat{\mathbf{e}}_2^n \otimes \widehat{\mathbf{e}}_2^n \\ &= \cos\theta\, \widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_1 - \sin\theta\, \widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_2 + \sin\theta\, \widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_1 + \cos\theta\, \widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_2 \ . \end{aligned}$$

It is then a simple exercise to show that $\mathbf{R}^n \mathbf{R}^{o\,\mathrm{T}} = \mathbf{R}$ is independent of α. The interested reader may now want to verify this result in the three-dimensional space.

Consider again the old orthonormal basis $\{\widehat{\mathbf{e}}_i^o\}$ that upon a proper orthogonal transformation converts into the new basis $\{\widehat{\mathbf{e}}_i^n\}$ according to (2.158). By requiring $\mathbf{A} = A_{ij}^o \widehat{\mathbf{e}}_i^o \otimes \widehat{\mathbf{e}}_j^o = A_{ij}^n \widehat{\mathbf{e}}_i^n \otimes \widehat{\mathbf{e}}_j^n$, the goal is to determine the relationship between A_{ij}^o and A_{ij}^n. Guided by (2.164) and (2.165), the following relation

$$\boxed{A_{ij}^n \xlongequal[\text{(2.20)}]{\text{from}} \widehat{\mathbf{e}}_i^n \cdot \mathbf{A}\widehat{\mathbf{e}}_j^n \xlongequal[\text{(2.2) and (2.158)}]{\text{from (1.9a)-(1.9c),}} R_{ki} R_{lj} \widehat{\mathbf{e}}_k^o \cdot \mathbf{A}\widehat{\mathbf{e}}_l^o \xlongequal[\text{(2.20)}]{\text{from}} R_{ki} R_{lj} A_{kl}^o \ ,} \tag{2.167}$$

represents the desired **tensorial** transformation law. The result $(2.167)_3$ in matrix representation renders

$$\boxed{[\,\mathbf{A}\,]^n = [\,\mathbf{R}\,]^{\mathrm{T}} [\,\mathbf{A}\,]^o [\,\mathbf{R}\,] \ ,} \tag{2.168}$$

where

$$[\,\mathbf{A}\,]^o = \begin{bmatrix} A_{11}^o & A_{12}^o & A_{13}^o \\ A_{21}^o & A_{22}^o & A_{23}^o \\ A_{31}^o & A_{32}^o & A_{33}^o \end{bmatrix} \ , \quad [\,\mathbf{A}\,]^n = \begin{bmatrix} A_{11}^n & A_{12}^n & A_{13}^n \\ A_{21}^n & A_{22}^n & A_{23}^n \\ A_{31}^n & A_{32}^n & A_{33}^n \end{bmatrix} \ . \tag{2.169}$$

A tensor may now be thought of as a collection of numbers that transform according to $(2.167)_3$ under a change of basis. For the tensorial transformation law in a general curvilinear coordinate system, see (5.59)–(5.60) and (5.106a)–(5.106d).

2.9.3 *Isotropic Vectors and Tensors*

A tensor is said to be *isotropic* if its components remain unchanged under all proper and improper orthogonal transformations (i.e. rotations and reflections). An obvious example of which includes zeroth-order tensors or scalars. A tensor is called *hemitropic* when it is unaffected by all proper orthogonal transformations. The condition of isotropy for vectors and tensors will be discussed in the following.

Isotropic tensors of order one. A first-order tensor $\mathbf{u}$ is isotropic if

$$u_i = Q_{ji} u_j \quad \text{or} \quad [\,\mathbf{u}\,] = [\,\mathbf{Q}\,]^{\mathrm{T}} [\,\mathbf{u}\,] \ . \quad \longleftarrow \text{see (2.164) and (2.165)} \tag{2.170}$$

As a result, one can write $\mathbf{u} = u_i \widehat{\mathbf{e}}_i^o = u_i \widehat{\mathbf{e}}_i^n$. And this holds true when all components of $\mathbf{u}$ are zero, i.e.

$$\boxed{\mathbf{u} = \mathbf{0} \ .} \quad \longleftarrow \text{see (2.173)} \tag{2.171}$$

Proof The fact that $(2.170)_{1-2}$ should hold true for all orthogonal tensors allows one to choose

$$[\mathbf{Q}] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Consequently, $u_1 = -u_1, u_2 = -u_2$ or $u_1 = 0, u_2 = 0$. Now, consider

$$[\mathbf{Q}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} .$$

This finally implies that $u_3 = -u_3$ or $u_3 = 0$. Note that apparently the zero vector is also a hemitropic object. Thus, in this simple case, one could use the rotation tensor to introduce the condition (2.170). This is frequently seen in the literature.

Isotropic tensors of order two. The isotropy condition for a second-order tensor **A** reads

$$A_{ij} = Q_{ki} Q_{lj} A_{kl} \quad \text{or} \quad [\mathbf{A}] = [\mathbf{Q}]^{\mathrm{T}} [\mathbf{A}] [\mathbf{Q}] . \quad \longleftarrow \text{see (2.167) and (2.168)} \tag{2.172}$$

This helps write $\mathbf{A} = A_{ij} \widehat{\mathbf{e}}_i^o \otimes \widehat{\mathbf{e}}_j^o = A_{ij} \widehat{\mathbf{e}}_i^n \otimes \widehat{\mathbf{e}}_j^n$. And this holds true if (Tadmor et al. [6])

$$\boxed{\mathbf{A} = \lambda \mathbf{I} \quad \text{where } \lambda \text{ is a constant} .} \quad \longleftarrow \text{see (3.25)} \tag{2.173}$$

In other words, a second-order tensor is isotropic when it is spherical (note that the second-order zero tensor is trivially an isotropic tensor).

Proof In this case, one may choose

$$[\mathbf{Q}] = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} .$$

It then follows that

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} A_{22} & -A_{23} & A_{21} \\ -A_{32} & A_{33} & -A_{31} \\ A_{12} & -A_{13} & A_{11} \end{bmatrix} .$$

At this stage, consider

$$[\mathbf{Q}] = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} .$$

This leads to

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} A_{22} & -A_{23} & -A_{21} \\ -A_{32} & A_{33} & A_{31} \\ -A_{12} & A_{13} & A_{11} \end{bmatrix} .$$

One can now infer that

$$A_{11} = A_{22} = A_{33} \quad , \quad A_{12} = A_{21} = A_{13} = A_{31} = A_{23} = A_{32} = 0 .$$

This result can be achieved in an alternative way. In the following, this will be verified again for the interested reader (see Hodge [7]).

Let

$$R_{ij} = \delta_{ij} - \varepsilon_{ijm}\omega_m \ , \tag{2.174}$$

be an **infinitesimal rotation** where ω_m denotes the axial vector of a skew tensor W_{ij} with infinitesimal magnitude. Introducing (2.174) into $(2.172)_1$ gives

$$\omega_m \left(\varepsilon_{ljm} A_{il} + \varepsilon_{kim} A_{kj} \right) = 0 \ , \tag{2.175}$$

where the higher-order term has been neglected. The fact that $\omega_i, \ i = 1, 2, 3,$ are arbitrary then implies that

$$\varepsilon_{ljm} A_{il} + \varepsilon_{kim} A_{kj} = 0 \ . \tag{2.176}$$

Multiplying both sides of this expression by ε_{njm}, taking into account (1.58a) and $(1.58b)_3$, results in

$$2A_{in} + A_{ni} = \delta_{ni} A_{kk} \ . \tag{2.177}$$

This can also be written as $2A_{ni} + A_{in} = \delta_{ni} A_{kk}$. It is then easy to see that $A_{in} = A_{ni}$. Thus,

$$\boxed{A_{in} = \frac{A_{kk}}{3} \delta_{in} \ .} \tag{2.178}$$

For representation of an isotropic scalar-, vector- or tensor-valued function of a system of tensorial variables, see Sect. 6.2.

2.10 Exercises

Exercise 2.1

Consider an arbitrary vector **u** and let $\widehat{\mathbf{e}}$ be any vector of unit length. Use (2.140a), (2.140b) and the triple vector product (1.71) to show that **u** can be resolved into components parallel (i.e. $\mathbf{proj}_{\widehat{\mathbf{e}}}\mathbf{u}$) and perpendicular (i.e. $\mathbf{reje}_{\widehat{\mathbf{e}}}\mathbf{u}$) to that unit vector according to

$$\mathbf{u} = \mathsf{proj}_{\widehat{\mathbf{e}}}\mathbf{u} + \mathsf{reje}_{\widehat{\mathbf{e}}}\mathbf{u} = \mathbf{P}^{\parallel}_{\widehat{\mathbf{e}}}\mathbf{u} + \mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}\mathbf{u} = (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}} + \widehat{\mathbf{e}} \times (\mathbf{u} \times \widehat{\mathbf{e}}) \ . \tag{2.179}$$

Solution. Guided by Fig. 2.2, one only needs to show that

$$\mathbf{P}^{\perp}_{\widehat{\mathbf{e}}}\mathbf{u} = \mathbf{u} - (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}} = (\widehat{\mathbf{e}} \cdot \widehat{\mathbf{e}})\,\mathbf{u} - (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}} = \widehat{\mathbf{e}} \times (\mathbf{u} \times \widehat{\mathbf{e}}) \ .$$

Exercise 2.2

Let $\mathbf{A}$ be an invertible tensor and let $\mathbf{u}$ and $\mathbf{v}$ be two arbitrary vectors. Further, let $\mathbf{B}$ be an arbitrary tensor. First, show that for any two scalars α and β, the following relation holds

$$\det(\alpha\mathbf{I} + \beta\mathbf{u} \otimes \mathbf{v}) = \alpha^3 + \alpha^2\beta\mathbf{u} \cdot \mathbf{v} \ . \tag{2.180}$$

Then, make use of this result to obtain

$$\det(\alpha\mathbf{A} + \beta\mathbf{u} \otimes \mathbf{v}) = \alpha^3 \det\mathbf{A} + \alpha^2\beta\mathbf{u} \cdot \mathbf{A}^{\mathrm{c}}\mathbf{v} \ . \tag{2.181}$$

Finally, verify that

$$\det(\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha^3 \det\mathbf{A} + \alpha^2\beta\mathbf{A}^{\mathrm{c}} : \mathbf{B} + \alpha\beta^2\mathbf{A} : \mathbf{B}^{\mathrm{c}} + \beta^3 \det\mathbf{B} \ . \tag{2.182}$$

Solution. By use of (1.36), $(1.38)_7$, (1.52), $(1.79)_1$, (2.23), $(2.24)_4$ and $(2.80)_4$, the first relation can be verified as follows:

$$\begin{aligned}
\det(\alpha\mathbf{I} + \beta\mathbf{u} \otimes \mathbf{v}) &= \varepsilon_{ijk}\,(\alpha\delta_{i1} + \beta u_i v_1)\left(\alpha\delta_{j2} + \beta u_j v_2\right)(\alpha\delta_{k3} + \beta u_k v_3) \\
&= \alpha^3\varepsilon_{ijk}\delta_{i1}\delta_{j2}\delta_{k3} + \alpha^2\beta\varepsilon_{ijk}\delta_{i1}\delta_{j2}u_k v_3 \\
&\quad + \alpha^2\beta\varepsilon_{ijk}\delta_{i1}\delta_{k3}u_j v_2 + \underbrace{\alpha\beta^2\varepsilon_{ijk}\delta_{i1}u_j u_k v_2 v_3}_{=0} \\
&\quad + \alpha^2\beta\varepsilon_{ijk}\delta_{j2}\delta_{k3}u_i v_1 + \underbrace{\alpha\beta^2\varepsilon_{ijk}\delta_{j2}u_i u_k v_1 v_3}_{=0} \\
&\quad + \underbrace{\alpha\beta^2\varepsilon_{ijk}\delta_{k3}u_i u_j v_1 v_2}_{=0} + \beta^3\underbrace{\varepsilon_{ijk}u_i u_j u_k v_1 v_2 v_3}_{=0} \\
&= \alpha^3\varepsilon_{123} + \alpha^2\beta\left(\varepsilon_{i23}u_i v_1 + \varepsilon_{1j3}u_j v_2 + \varepsilon_{12k}u_k v_3\right) \\
&= \alpha^3 + \alpha^2\beta\,(u_1 v_1 + u_2 v_2 + u_3 v_3) = \alpha^3 + \alpha^2\beta\mathbf{u} \cdot \mathbf{v} \ .
\end{aligned}$$

From (2.5), (2.25), (2.32), $(2.33)_1$, (2.51d), (2.55b), $(2.99\text{b})_1$, $(2.105)_1$, $(2.110)_1$, (2.114) and (2.181), the second relation can be shown as

$$\begin{aligned}\det(\alpha\mathbf{A}+\beta\mathbf{u}\otimes\mathbf{v}) &= (\det\mathbf{A})\det\left(\alpha\mathbf{I}+\beta\mathbf{A}^{-1}\mathbf{u}\otimes\mathbf{v}\right)\\ &= (\det\mathbf{A})\left(\alpha^3+\alpha^2\beta\mathbf{A}^{-1}\mathbf{u}\cdot\mathbf{v}\right)\\ &= \alpha^3\det\mathbf{A}+\alpha^2\beta\left(\mathbf{u}\cdot(\det\mathbf{A})\,\mathbf{A}^{-\mathrm{T}}\mathbf{v}\right)\\ &= \alpha^3\det\mathbf{A}+\alpha^2\beta\mathbf{u}\cdot\mathbf{A}^{\mathrm{c}}\mathbf{v}\ .\end{aligned}$$

The above procedures can be combined to arrive at the third relation. To begin with, one should realize that

$$\det(\alpha\mathbf{A}+\beta\mathbf{B}) = (\det\mathbf{A})\det\left(\alpha\mathbf{I}+\beta\mathbf{A}^{-1}\mathbf{B}\right)\ .$$

Then,

$$\begin{aligned}\det\left(\alpha\mathbf{I}+\beta\mathbf{A}^{-1}\mathbf{B}\right) &= \alpha^3\underbrace{\varepsilon_{ijk}\delta_{i1}\delta_{j2}\delta_{k3}}_{=\varepsilon_{123}=1}\\ &\quad+\alpha^2\beta\underbrace{\varepsilon_{ijk}\delta_{i1}\delta_{j2}A^{-1}_{ko}B_{o3}}_{=\varepsilon_{12k}A^{-1}_{ko}B_{o3}=A^{-1}_{3o}B_{o3}}\\ &\quad+\alpha^2\beta\underbrace{\varepsilon_{ijk}\delta_{i1}A^{-1}_{jn}B_{n2}\delta_{k3}}_{=\varepsilon_{1j3}A^{-1}_{jn}B_{n2}=A^{-1}_{2n}B_{n2}}\\ &\quad+\alpha\beta^2\underbrace{\left[\varepsilon_{ijk}A^{-1}_{jn}A^{-1}_{ko}\right]\left[\delta_{i1}B_{n2}B_{o3}\right]}_{=\left[\left(\det\mathbf{A}^{-1}\right)A_{mi}\varepsilon_{mno}\right]\left[\delta_{i1}B_{n2}B_{o3}\right]=\left(\det\mathbf{A}^{-1}\right)A_{m1}B^{\mathrm{c}}_{m1}}\\ &\quad+\alpha^2\beta\underbrace{\varepsilon_{ijk}A^{-1}_{im}B_{m1}\delta_{j2}\delta_{k3}}_{=\varepsilon_{i23}A^{-1}_{im}B_{m1}=A^{-1}_{1m}B_{m1}}\\ &\quad+\alpha\beta^2\underbrace{\left[\varepsilon_{ijk}A^{-1}_{im}A^{-1}_{ko}\right]\left[\delta_{j2}B_{m1}B_{o3}\right]}_{=\left[\left(\det\mathbf{A}^{-1}\right)A_{nj}\varepsilon_{mno}\right]\left[\delta_{j2}B_{m1}B_{o3}\right]=\left(\det\mathbf{A}^{-1}\right)A_{n2}B^{\mathrm{c}}_{n2}}\\ &\quad+\alpha\beta^2\underbrace{\left[\varepsilon_{ijk}A^{-1}_{im}A^{-1}_{jn}\right]\left[\delta_{k3}B_{m1}B_{n2}\right]}_{=\left[\left(\det\mathbf{A}^{-1}\right)A_{ok}\varepsilon_{mno}\right]\left[\delta_{k3}B_{m1}B_{n2}\right]=\left(\det\mathbf{A}^{-1}\right)A_{o3}B^{\mathrm{c}}_{o3}}\\ &\quad+\beta^3\underbrace{\varepsilon_{ijk}A^{-1}_{im}B_{m1}A^{-1}_{jn}B_{n2}A^{-1}_{ko}B_{o3}}_{=\left(\frac{1}{6}\varepsilon_{ijk}\varepsilon_{qrs}A^{-1}_{iq}A^{-1}_{jr}A^{-1}_{ks}\right)\left(\varepsilon_{mno}B_{m1}B_{n2}B_{o3}\right)=\left(\det\mathbf{A}^{-1}\right)(\det\mathbf{B})}\\ &= \alpha^3+\alpha^2\beta\left(A^{-\mathrm{T}}_{m1}B_{m1}+A^{-\mathrm{T}}_{m2}B_{m2}+A^{-\mathrm{T}}_{m3}B_{m3}\right)\\ &\quad+\alpha\beta^2\frac{A_{m1}B^{\mathrm{c}}_{m1}+A_{m2}B^{\mathrm{c}}_{m2}+A_{m3}B^{\mathrm{c}}_{m3}}{\det\mathbf{A}}+\beta^3\frac{\det\mathbf{B}}{\det\mathbf{A}}\\ &= \alpha^3+\frac{\alpha^2\beta A^{\mathrm{c}}_{mn}B_{mn}+\alpha\beta^2A_{mn}B^{\mathrm{c}}_{mn}+\beta^3\det\mathbf{B}}{\det\mathbf{A}}\ .\end{aligned}$$

Multiplying both sides of the above expression by $\det \mathbf{A}$ finally yields the desired result.

Exercise 2.3

Consider a skew tensor $\mathbf{W}^{\mathbf{u}}$ with $W^{\mathbf{u}}_{ij} = -\varepsilon_{ijk} u_k$ and $\mathbf{W}^{\mathbf{v}}$ with $W^{\mathbf{v}}_{ij} = -\varepsilon_{ijk} v_k$ satisfying

$$\mathbf{W}^{\mathbf{u}}\mathbf{w} = \mathbf{u} \times \mathbf{w} \quad , \quad \mathbf{W}^{\mathbf{v}}\mathbf{w} = \mathbf{v} \times \mathbf{w} \, , \quad \text{for all } \mathbf{w} \in \mathscr{E}^{o3}_{\mathrm{r}} \, . \tag{2.183}$$

Verify the following identities

$$\mathbf{W}^{\mathbf{u}}\mathbf{W}^{\mathbf{v}} = \mathbf{v} \otimes \mathbf{u} - [\operatorname{tr}(\mathbf{v} \otimes \mathbf{u})]\, \mathbf{I} \, , \quad \leftarrow \text{note that } \operatorname{tr}(\mathbf{v} \otimes \mathbf{u}) = \mathbf{v} \cdot \mathbf{u} \tag{2.184a}$$

$$\begin{aligned} \mathbf{W}^{\mathbf{u}}\mathbf{A}\mathbf{W}^{\mathbf{v}} &= [\operatorname{tr}(\mathbf{v} \otimes \mathbf{u})]\, \mathbf{A}^{\mathrm{T}} - \mathbf{v}\mathbf{A} \otimes \mathbf{u} - (\operatorname{tr}\mathbf{A})\, [\operatorname{tr}(\mathbf{v} \otimes \mathbf{u})]\, \mathbf{I} \\ &\quad + [\operatorname{tr}(\mathbf{v} \otimes \mathbf{A}\mathbf{u})]\, \mathbf{I} + (\operatorname{tr}\mathbf{A})\, (\mathbf{v} \otimes \mathbf{u}) - \mathbf{v} \otimes \mathbf{A}\mathbf{u} \, . \end{aligned} \tag{2.184b}$$

Solution. This exercise will be solved by means of index notation. With the aid of the expressions (1.36), (1.38)$_7$, (1.54), (1.58a), (2.23) and (2.24)$_4$, the first identity can be shown as follows:

$$\begin{aligned} W^{\mathbf{u}}_{im} W^{\mathbf{v}}_{mj} &= +\varepsilon_{imk} u_k \varepsilon_{mjl} v_l \\ &= \left(-\varepsilon_{mik} \varepsilon_{mjl}\right) u_k v_l \\ &= \left(-\delta_{ij}\delta_{kl} + \delta_{il}\delta_{kj}\right) u_k v_l \\ &= -\delta_{ij} u_k v_k + u_j v_i \\ &= \left(\mathbf{v} \otimes \mathbf{u} - (\mathbf{v} \cdot \mathbf{u})\, \mathbf{I}\right)_{ij} \ . \end{aligned}$$

In a similar fashion, one can establish

$$\begin{aligned} W^{\mathbf{u}}_{im} A_{mn} W^{\mathbf{v}}_{nj} &= (\varepsilon_{imk} u_k)\, A_{mn} \left(\varepsilon_{njl} v_l\right) \\ &= \left(\varepsilon_{imk} \varepsilon_{njl}\right) (u_k A_{mn} v_l) \\ &= \underbrace{\delta_{in} \left(\delta_{mj}\delta_{kl} - \delta_{kj}\delta_{ml}\right) (u_k A_{mn} v_l)}_{= u_k A_{ji} v_k - u_j A_{li} v_l = [\operatorname{tr}(\mathbf{v} \otimes \mathbf{u})] \left(\mathbf{A}^{\mathrm{T}}\right)_{ij} - (\mathbf{v}\mathbf{A} \otimes \mathbf{u})_{ij}} \\ &\quad + \underbrace{\delta_{ij} \left(-\delta_{mn}\delta_{kl} + \delta_{kn}\delta_{ml}\right) (u_k A_{mn} v_l)}_{= \delta_{ij} \left(-u_k A_{mm} v_k + u_k A_{mk} v_m\right) = [-(\operatorname{tr}\mathbf{A})\, (\operatorname{tr}(\mathbf{v} \otimes \mathbf{u})) + \operatorname{tr}(\mathbf{v} \otimes \mathbf{A}\mathbf{u})]\, (\mathbf{I})_{ij}} \\ &\quad + \underbrace{\delta_{il} \left(\delta_{mn}\delta_{kj} - \delta_{kn}\delta_{mj}\right) (u_k A_{mn} v_l)}_{= u_j A_{mm} v_i - u_k A_{jk} v_i = (\operatorname{tr}\mathbf{A})\, (\mathbf{v} \otimes \mathbf{u})_{ij} - (\mathbf{v} \otimes \mathbf{A}\mathbf{u})_{ij}} \ . \end{aligned}$$

Exercise 2.4

Find the axial vector of the skew-symmetric tensor $\mathbf{W} = (1/2)\,(\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u})$.

Solution. In view of (2.65), the goal is to find a vector $\boldsymbol{\omega}$ such that

$$(\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u})\,\mathbf{w} = 2\boldsymbol{\omega} \times \mathbf{w}\ .$$

From (1.72), (2.10) and (2.13), one can arrive at

$$(\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u})\,\mathbf{w} = (\mathbf{v} \cdot \mathbf{w})\,\mathbf{u} - (\mathbf{u} \cdot \mathbf{w})\,\mathbf{v} = -\,(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}\ .$$

By use of (1.49a), one will finally have

$$\mathbf{W} = \operatorname{skw}(\mathbf{u} \otimes \mathbf{v}) \quad \text{for which} \quad \boldsymbol{\omega} = \frac{1}{2}\mathbf{v} \times \mathbf{u}\ . \tag{2.185}$$

Exercise 2.5

Let $\mathbf{W}$ be a skew tensor with the associated axial vector $\boldsymbol{\omega}$. Regarding (2.112), show that

$$\mathbf{W}^{\mathrm{c}} = \boldsymbol{\omega} \otimes \boldsymbol{\omega}\ . \tag{2.186}$$

Solution. First, recall (2.65), i.e. $\mathbf{W}\mathbf{u} = \boldsymbol{\omega} \times \mathbf{u}$. Then, the relation (2.112), with the aid of (1.4a), (1.4d), (1.49a), $(1.49\mathrm{c})_1$, $(1.73)_1$, $(1.78\mathrm{b})_1$ and (2.13), helps obtain

$$\begin{aligned}
\mathbf{W}^{\mathrm{c}}\,(\mathbf{u} \times \mathbf{v}) &= (\mathbf{W}\mathbf{u}) \times (\mathbf{W}\mathbf{v}) \\
&= (\boldsymbol{\omega} \times \mathbf{u}) \times (\boldsymbol{\omega} \times \mathbf{v}) \\
&= \underbrace{[\boldsymbol{\omega} \cdot (\boldsymbol{\omega} \times \mathbf{v})]\,\mathbf{u}}_{=\,[0]\,\mathbf{u}\,=\,\mathbf{0}} - \underbrace{[\mathbf{u} \cdot (\boldsymbol{\omega} \times \mathbf{v})]\,\boldsymbol{\omega}}_{=\,[\boldsymbol{\omega} \cdot (\mathbf{v} \times \mathbf{u})]\,\boldsymbol{\omega}\,=\,-\,[\boldsymbol{\omega} \cdot (\mathbf{u} \times \mathbf{v})]\,\boldsymbol{\omega}} \\
&= \mathbf{0} + [\boldsymbol{\omega} \cdot (\mathbf{u} \times \mathbf{v})]\,\boldsymbol{\omega} \\
&= (\boldsymbol{\omega} \otimes \boldsymbol{\omega})\,(\mathbf{u} \times \mathbf{v})\ ,
\end{aligned}$$

which, by means of (2.6), delivers the desired result.

Exercise 2.6

Suppose one is given a skew tensor $\mathbf{W}$ with the associated axial vector $\boldsymbol{\omega}$. First, show that the determinant of a tensor $\mathbf{I} + \mathbf{W}$ is

$$\det(\mathbf{I}+\mathbf{W}) = 1 + \boldsymbol{\omega}\cdot\boldsymbol{\omega} \ . \tag{2.187}$$

Then, prove that

$$\mathbf{Q} = (\mathbf{I}-\mathbf{W})(\mathbf{I}+\mathbf{W})^{-1} \ , \tag{2.188}$$

is an orthogonal tensor.

Solution. The expression (2.187) reveals the fact that $\mathbf{I}+\mathbf{W}$ is nonsingular and hence its inverse in (2.188) exists. Consequently, $\mathbf{Q}$ is well-defined. Guided by the results of previous exercises, the first part can be verified as follows:

$$\begin{aligned}
\det(\mathbf{I}+\mathbf{W}) \overset{\text{from}}{\underset{(2.182)}{=\!=}} & \underbrace{\det\mathbf{I}}_{=1, \text{ according to } (2.100)} + \underbrace{\mathbf{I}^{c}:\mathbf{W}}_{=\mathbf{I}:\mathbf{W}, \text{ according to } (2.119)} \\
& + \underbrace{\mathbf{I}:\mathbf{W}^{c}}_{=\mathbf{I}:\boldsymbol{\omega}\otimes\boldsymbol{\omega}, \text{ according to } (2.186)} + \underbrace{\det\mathbf{W}}_{=0, \text{ according to } (2.101)} \\
&= 1 + \underbrace{\mathbf{I}:\mathbf{W}}_{=0, \text{ according to } (2.89\text{e})} + \underbrace{\mathbf{I}:\boldsymbol{\omega}\otimes\boldsymbol{\omega}}_{=\omega_i\omega_i, \text{ according to } (2.89\text{b})} \\
&= 1 + \boldsymbol{\omega}\cdot\boldsymbol{\omega} \ .
\end{aligned}$$

It is then easy to see that

$$\det(\mathbf{I}-\mathbf{W}) = 1 + \boldsymbol{\omega}\cdot\boldsymbol{\omega} \ . \tag{2.189}$$

For the second part, one needs to show that $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$. The key point is that the two tensors $\mathbf{I}-\mathbf{W}$ and $\mathbf{I}+\mathbf{W}$ **commute**, that is,

$$\begin{aligned}
(\mathbf{I}-\mathbf{W})(\mathbf{I}+\mathbf{W}) &= \mathbf{I}-\mathbf{W}+\mathbf{W}-\mathbf{W}^2 \\
&= \mathbf{I}-\mathbf{W}^2 \ , \\
(\mathbf{I}+\mathbf{W})(\mathbf{I}-\mathbf{W}) &= \mathbf{I}+\mathbf{W}-\mathbf{W}-\mathbf{W}^2 \\
&= \mathbf{I}-\mathbf{W}^2 \\
&= (\mathbf{I}-\mathbf{W})(\mathbf{I}+\mathbf{W}) \ .
\end{aligned}$$

This result, along with (2.34), (2.55a), $(2.55\text{d})_1$, $(2.58)_1$, $(2.105)_{1-2}$ and (2.109f), helps obtain

$$\begin{aligned}
\mathbf{Q}^{\mathrm{T}}\mathbf{Q} &= \left(\mathbf{I}^{\mathrm{T}}+\mathbf{W}^{\mathrm{T}}\right)^{-1}\left(\mathbf{I}^{\mathrm{T}}-\mathbf{W}^{\mathrm{T}}\right)(\mathbf{I}-\mathbf{W})(\mathbf{I}+\mathbf{W})^{-1} \\
&= (\mathbf{I}-\mathbf{W})^{-1}(\mathbf{I}+\mathbf{W})(\mathbf{I}-\mathbf{W})(\mathbf{I}+\mathbf{W})^{-1} \\
&= (\mathbf{I}-\mathbf{W})^{-1}(\mathbf{I}-\mathbf{W})(\mathbf{I}+\mathbf{W})(\mathbf{I}+\mathbf{W})^{-1} \\
&= \mathbf{I}\,\mathbf{I} \\
&= \mathbf{I} \ .
\end{aligned}$$

Exercise 2.7

The cofactor $\mathbf{A}^{\mathrm{c}}$ of a tensor $\mathbf{A}$ can be represented by[20]

$$\underbrace{\mathbf{A}^{\mathrm{c}} = \left(\mathbf{A}^2\right)^{\mathrm{T}} - (\mathrm{tr}\mathbf{A})\,\mathbf{A}^{\mathrm{T}} + \frac{1}{2}\left[(\mathrm{tr}\mathbf{A})^2 - \left(\mathrm{tr}\mathbf{A}^2\right)\right]\mathbf{I}}_{\text{or } \left(\mathbf{A}^{\mathrm{c}}\right)_{ij} = A^2_{ji} - (\mathrm{tr}\mathbf{A})\,A_{ji} + \frac{1}{2}\left[(\mathrm{tr}\mathbf{A})^2 - \left(\mathrm{tr}\mathbf{A}^2\right)\right]\delta_{ij}} \; . \tag{2.190}$$

The above relation shows that $(\det \mathbf{A})\,\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{c\,T}}$, according to (2.114), can be computed properly even if $\mathbf{A}$ is not invertible. And this means that for any tensor $\mathbf{A}$, the tensorial variable $(\det \mathbf{A})\,\mathbf{A}^{-1}$ exists.

Make use of (2.190) to verify that

$$\begin{aligned}(\alpha\mathbf{A} + \beta\mathbf{B})^{\mathrm{c}} = \alpha^2\mathbf{A}^{\mathrm{c}} + \beta^2\mathbf{B}^{\mathrm{c}} &+ \alpha\beta\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\right) \\ &- \alpha\beta\left[(\mathrm{tr}\mathbf{B})\,\mathbf{A}^{\mathrm{T}} + (\mathrm{tr}\mathbf{A})\,\mathbf{B}^{\mathrm{T}}\right] \\ &+ \alpha\beta\left[(\mathrm{tr}\mathbf{A})\,(\mathrm{tr}\mathbf{B}) - \mathrm{tr}\,(\mathbf{AB})\right]\mathbf{I} \; .\end{aligned} \tag{2.191}$$

Solution. By taking into account the linearity of the trace operator and transposition, one can write

$$\begin{aligned}(\alpha\mathbf{A} + \beta\mathbf{B})^{\mathrm{c}} = &\underbrace{(\alpha\mathbf{A} + \beta\mathbf{B})^{\mathrm{T}}(\alpha\mathbf{A} + \beta\mathbf{B})^{\mathrm{T}}}_{= \alpha^2\mathbf{A}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} + \alpha\beta\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\right) + \beta^2\mathbf{B}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}} \\ &- \underbrace{\left[\mathrm{tr}\,(\alpha\mathbf{A} + \beta\mathbf{B})\right](\alpha\mathbf{A} + \beta\mathbf{B})^{\mathrm{T}}}_{= \alpha^2\,(\mathrm{tr}\mathbf{A})\,\mathbf{A}^{\mathrm{T}} + \alpha\beta\left[(\mathrm{tr}\mathbf{B})\,\mathbf{A}^{\mathrm{T}} + (\mathrm{tr}\mathbf{A})\,\mathbf{B}^{\mathrm{T}}\right] + \beta^2\,(\mathrm{tr}\mathbf{B})\,\mathbf{B}^{\mathrm{T}}} \\ &+ \frac{1}{2}\underbrace{\left(\mathrm{tr}\,(\alpha\mathbf{A} + \beta\mathbf{B})\right)^2}_{= \alpha^2\,(\mathrm{tr}\mathbf{A})^2 + 2\alpha\beta\,(\mathrm{tr}\mathbf{A})\,(\mathrm{tr}\mathbf{B}) + \beta^2\,(\mathrm{tr}\mathbf{B})^2}\mathbf{I} \\ &- \frac{1}{2}\underbrace{\mathrm{tr}\,(\alpha\mathbf{A} + \beta\mathbf{B})^2}_{= \alpha^2\mathrm{tr}\mathbf{A}^2 + 2\alpha\beta\mathrm{tr}\,(\mathbf{AB}) + \beta^2\mathrm{tr}\mathbf{B}^2}\mathbf{I} \; .\end{aligned}$$

which, by simplification, delivers the required result. Some straightforward manipulations now render

$$(\alpha\mathbf{I} + \beta\mathbf{u}\otimes\mathbf{v})^{\mathrm{c}} = \alpha^2\mathbf{I} - \alpha\beta\mathbf{W}^{\mathbf{u}}\mathbf{W}^{\mathbf{v}} \; , \quad \longleftarrow \text{see (2.184a)} \tag{2.192a}$$

$$(\alpha\mathbf{A} + \beta\mathbf{u}\otimes\mathbf{v})^{\mathrm{c}} = \alpha^2\mathbf{A}^{\mathrm{c}} - \alpha\beta\mathbf{W}^{\mathbf{u}}\mathbf{A}\mathbf{W}^{\mathbf{v}} \; , \quad \longleftarrow \text{see (2.184b)} \tag{2.192b}$$

$$(\alpha\mathbf{A} + \beta\mathbf{I})^{\mathrm{c}} = \alpha^2\mathbf{A}^{\mathrm{c}} + \alpha\beta\left[(\mathrm{tr}\mathbf{A})\,\mathbf{I} - \mathbf{A}^{\mathrm{T}}\right] + \beta^2\mathbf{I} \; . \tag{2.192c}$$

Moreover, by (2.114), (2.182) and (2.191),

[20] The expression (2.190) is basically a consequence of the Cayley-Hamilton equation, see (4.21).

$$(\alpha\mathbf{A}+\beta\mathbf{B})^{-1} = \frac{\alpha^2\,(\det\mathbf{A})\,\mathbf{A}^{-1}+\beta^2\,(\det\mathbf{B})\,\mathbf{B}^{-1}}{\alpha^3\det\mathbf{A}+\alpha^2\beta\mathbf{A}^{\mathrm{c}}:\mathbf{B}+\alpha\beta^2\mathbf{A}:\mathbf{B}^{\mathrm{c}}+\beta^3\det\mathbf{B}} + \frac{\alpha\beta\,[\mathbf{AB}+\mathbf{BA}-(\mathrm{tr}\mathbf{B})\,\mathbf{A}-(\mathrm{tr}\mathbf{A})\,\mathbf{B}]}{\alpha^3\det\mathbf{A}+\alpha^2\beta\mathbf{A}^{\mathrm{c}}:\mathbf{B}+\alpha\beta^2\mathbf{A}:\mathbf{B}^{\mathrm{c}}+\beta^3\det\mathbf{B}} + \frac{\alpha\beta\,[(\mathrm{tr}\mathbf{A})\,(\mathrm{tr}\mathbf{B})-\mathrm{tr}\,(\mathbf{AB})]\,\mathbf{I}}{\alpha^3\det\mathbf{A}+\alpha^2\beta\mathbf{A}^{\mathrm{c}}:\mathbf{B}+\alpha\beta^2\mathbf{A}:\mathbf{B}^{\mathrm{c}}+\beta^3\det\mathbf{B}}\;. \tag{2.193}$$

Exercise 2.8

Given the linear transformations

$$\left.\begin{aligned}\mathbf{A}\widehat{\mathbf{e}}_1 &= \cos\theta\widehat{\mathbf{e}}_1+\sin\theta\widehat{\mathbf{e}}_2\\ \mathbf{A}\widehat{\mathbf{e}}_2 &= -\sin\theta\widehat{\mathbf{e}}_1+\cos\theta\widehat{\mathbf{e}}_2\\ \mathbf{A}\widehat{\mathbf{e}}_3 &= \widehat{\mathbf{e}}_3\end{aligned}\right\}\quad\text{and}\quad\left.\begin{aligned}\mathbf{B}\widehat{\mathbf{e}}_1 &= -\widehat{\mathbf{e}}_1\\ \mathbf{B}\widehat{\mathbf{e}}_2 &= \widehat{\mathbf{e}}_2\\ \mathbf{B}\widehat{\mathbf{e}}_3 &= \widehat{\mathbf{e}}_3\end{aligned}\right\}\;. \tag{2.194}$$

First, represent **A** and **B** in matrix notation.

Then, verify that the resulting matrices are orthogonal.

Finally, algebraically as well as geometrically show that [**A**] presents a rotation and [**B**] describes a reflection.

Solution. Guided by (2.21), the matrices are given by

$$[\,\mathbf{A}\,] = \begin{bmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}\quad,\quad [\,\mathbf{B}\,] = \begin{bmatrix}-1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\;. \tag{2.195}$$

To verify that a tensor is orthogonal, one must show that its inverse is equal to its transposes or the simple contraction between the transposed tensor and the tensor itself delivers the unit tensor, see (2.130) and (2.131). Consistent with matrix algebra, these can straightforwardly be shown for the presented matrices.

The matrix [**A**] represents a rotation owing to

$$\det[\,\mathbf{A}\,] = \cos\theta\cos\theta-(-\sin\theta)\sin\theta = \cos^2\theta+\sin^2\theta = 1\;,\qquad \longleftarrow \text{see (2.133)}$$

and [**B**] renders a reflection since $\det[\,\mathbf{B}\,] = (-1)\,(1)\,(1) = -1$.

See Fig. 2.4 for a geometrical interpretation. Therein, [**A**] describes a counterclockwise rotation of a vector $\mathbf{u} = (r\cos\alpha)\widehat{\mathbf{e}}_1 + (r\sin\alpha)\widehat{\mathbf{e}}_2$ with the matrix form $[\,\mathbf{u}\,]^{\mathrm{T}} = \begin{bmatrix} r\cos\alpha & r\sin\alpha & 0\end{bmatrix}$ by an angle θ about the x_3-axis. And [**B**] describes a reflection of the that vector across the x_2-axis. Note that **u** need not necessarily be chosen from the x_1-x_2 plane.

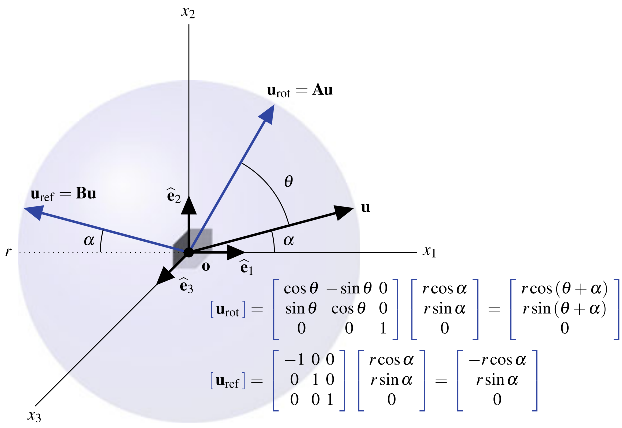


Fig. 2.4 A counterclockwise rotation by an angle θ about the x_3-axis and a reflection across the x_2-axis in a right-handed Cartesian coordinate frame

Exercise 2.9

This exercise aims at characterizing **finite rotation** of an arbitrary vector in the oriented three-dimensional Euclidean vector space. This leads to a well-known formula called the *Rodrigues rotation formula* (see Murray et al. [8]).

Let $\mathbf{u}$ be a vector that is supposed to rotate counterclockwise by an angle θ about a rotation axis defined by the unit vector $\widehat{\mathbf{e}}$. This has been displayed schematically in Fig. 2.5.

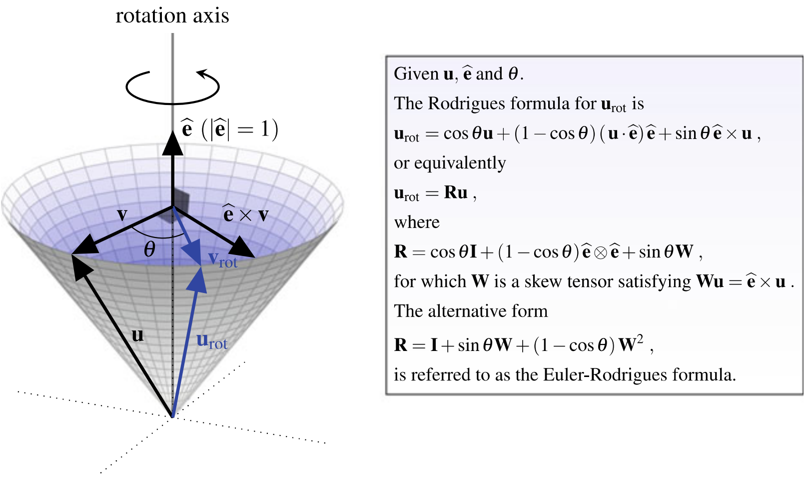


Fig. 2.5 Finite rotation of a vector $\mathbf{u}$ in $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$ based on the Rodrigues rotation formula

First, show that the rotated vector, $\mathbf{u}_{\mathrm{rot}}$, can be written as

$$\underbrace{\mathbf{u}_{\mathrm{rot}}(\theta, \mathbf{u}) = \cos\theta\,\mathbf{u} + (1-\cos\theta)(\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}} + \sin\theta\,\widehat{\mathbf{e}}\times\mathbf{u}}_{\text{this represents the Rodrigues formula}} \ . \tag{2.196}$$

Then, by rewriting (2.196) in the form $\mathbf{u}_{\mathrm{rot}} = \mathbf{R}\mathbf{u}$, show that the rotation tensor $\mathbf{R}$ can be represented by

$$\mathbf{R}(\theta) = \cos\theta\,\mathbf{I} + (1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}} + \sin\theta\,\mathbf{W} \ , \tag{2.197}$$

or

$$\mathbf{R}(\theta) = \mathbf{I} + \sin\theta\,\mathbf{W} + (1-\cos\theta)\,\mathbf{W}^2 \ , \tag{2.198}$$

where $\mathbf{W}$ denotes a skew tensor whose axial vector is $\widehat{\mathbf{e}}$, i.e.

$$\mathbf{W}\mathbf{u} = \widehat{\mathbf{e}}\times\mathbf{u} \ . \tag{2.199}$$

The efficient form (2.198) is known as the *Euler-Rodrigues* formula and frequently utilized in rigid-body dynamics as well as parametrization of shell structures. See Crisfield [9] and Wriggers [10] for more discussions.

Finally, verify that the composition of rotations obeys

$$\mathbf{R}(\theta+\alpha) = \mathbf{R}(\theta)\,\mathbf{R}(\alpha) \ . \tag{2.200}$$

Solution. As can be seen from Fig. 2.5, the unit vector $\widehat{\mathbf{e}}$ is perpendicular to a vector $\mathbf{v}$ whose magnitude is the base radius of the cone. This vector is basically the vector rejection of $\mathbf{u}$ from $\widehat{\mathbf{e}}$, i.e. $\mathbf{v} = \mathbf{u} - (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}}$. If $\mathbf{u}$ rotates by an angle θ about the rotation axis, then $\mathbf{v}$ consistently rotates by the same angle and, therefore, one can write $\mathbf{v}_{\mathrm{rot}} = \mathbf{u}_{\mathrm{rot}} - (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}}$ and subsequently $|\mathbf{v}_{\mathrm{rot}}| = |\mathbf{v}|$. Observe that the cross product $\widehat{\mathbf{e}}\times\mathbf{v}$ with $|\widehat{\mathbf{e}}\times\mathbf{v}| = |\mathbf{v}|$ lies in the plane spanned by the circular base and trivially $(\widehat{\mathbf{e}}\times\mathbf{v})\cdot\mathbf{v} = 0$. This helps project $\mathbf{v}_{\mathrm{rot}}$ along $\mathbf{v}$ and $\widehat{\mathbf{e}}\times\mathbf{v}$ according to

$$\mathbf{v}_{\mathrm{rot}} = |\mathbf{v}_{\mathrm{rot}}|\cos\theta\,\frac{\mathbf{v}}{|\mathbf{v}|} + |\mathbf{v}_{\mathrm{rot}}|\sin\theta\,\frac{\widehat{\mathbf{e}}\times\mathbf{v}}{|\widehat{\mathbf{e}}\times\mathbf{v}|} = \cos\theta\,\mathbf{v} + \sin\theta\,\widehat{\mathbf{e}}\times\mathbf{v} \ . \tag{2.201}$$

It follows that

$$\underbrace{\mathbf{u}_{\mathrm{rot}} = \cos\theta\,\mathbf{v} + \sin\theta\,\widehat{\mathbf{e}}\times\mathbf{v} + (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}}}_{\text{since } \mathbf{v}_{\mathrm{rot}} = \mathbf{u}_{\mathrm{rot}} - (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}} = \cos\theta\,\mathbf{v} + \sin\theta\,\widehat{\mathbf{e}}\times\mathbf{v}} \ . \tag{2.202}$$

By substituting $\mathbf{v} = \mathbf{u} - (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}}$ into the above expression, one then obtains the desired result

$$\begin{aligned}
\mathbf{u}_{\mathrm{rot}} &= \cos\theta\,\mathbf{u} - \cos\theta\,(\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}} + \sin\theta\,\widehat{\mathbf{e}}\times[\mathbf{u} - (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}}] + (\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}} \\
&= \cos\theta\,\mathbf{u} + (1-\cos\theta)\,(\mathbf{u}\cdot\widehat{\mathbf{e}})\,\widehat{\mathbf{e}} + \sin\theta\,\widehat{\mathbf{e}}\times\mathbf{u} \ .
\end{aligned}$$

Next, the goal is to extract a second-order tensor from the above expression. By means of (2.5), (2.13) and (2.65), this can be rewritten in the form $\mathbf{u}_{\mathrm{rot}} = \mathbf{R}\mathbf{u}$ for which the rotation tensor $\mathbf{R}$ is one given in (2.197), i.e.

$$\mathbf{u}_{\mathrm{rot}} = \underbrace{\cos\theta\,\mathbf{I}\mathbf{u} + (1-\cos\theta)\left(\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\right)\mathbf{u} + \sin\theta\,\mathbf{W}\mathbf{u}}_{=\left(\cos\theta\,\mathbf{I} + (1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}} + \sin\theta\,\mathbf{W}\right)\mathbf{u} = \mathbf{R}\mathbf{u}}\,.$$

The interested reader may now want to verify that the introduced second-order tensor $\mathbf{R}$ in (2.197) is orthogonal, i.e. it satisfies (2.130). This can be shown, by means of (1.51), (2.48), (2.52), (2.53), (2.54), (2.55a), $(2.58)_1$, (2.184a) and (2.199), as follows:

$$\begin{aligned}
\mathbf{R}\mathbf{R}^{\mathrm{T}} &= \cos^2\theta\,\mathbf{I} + \cos\theta\,(1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}} \\
&\quad + \cancel{(-\cos\theta\sin\theta)\,\mathbf{W}} + (1-\cos\theta)\cos\theta\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}} \\
&\quad + (1-\cos\theta)^2\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}} - (1-\cos\theta)\sin\theta\,\underbrace{\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\,\mathbf{W}}_{=-\widehat{\mathbf{e}}\otimes\mathbf{W}\widehat{\mathbf{e}} = -\widehat{\mathbf{e}}\otimes(\widehat{\mathbf{e}}\times\widehat{\mathbf{e}}) = \mathbf{O}} \\
&\quad + \cancel{\cos\theta\sin\theta\,\mathbf{W}} + (1-\cos\theta)\sin\theta\,\underbrace{\mathbf{W}\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}}_{=\mathbf{O}} - \sin^2\theta\,\mathbf{W}^2 \\
&= \underbrace{\cos^2\theta\,\mathbf{I} + \sin^2\theta\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}} - \sin^2\theta\,\mathbf{W}^2}_{=\cos^2\theta\,\mathbf{I} + \cancel{\sin^2\theta\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}} - \sin^2\theta\left(\cancel{\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}} - \mathbf{I}\right)} \\
&= \left(\cos^2\theta + \sin^2\theta\right)\mathbf{I} \\
&= \mathbf{I}\,.
\end{aligned} \tag{2.203}$$

Moreover, it truly describes a rotation since

$$\begin{aligned}
\det\mathbf{R} &= \underbrace{\det\left(\cos\theta\,\mathbf{I} + (1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\right)}_{=\cos^3\theta + \cos^2\theta\,(1-\cos\theta)\,\widehat{\mathbf{e}}\cdot\widehat{\mathbf{e}} = \cos^2\theta} \\
&\quad + \sin\theta\,\underbrace{\left(\cos\theta\,\mathbf{I} + (1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\right)^{\mathrm{c}} : \mathbf{W}}_{=\left(\cos^2\theta\,\mathbf{I} - \cos\theta\,(1-\cos\theta)\,\mathbf{W}^2\right) : \mathbf{W} = 0} \quad \swarrow \ \text{note that } \mathbf{I} \text{ and } \mathbf{W}^2 \text{ are symmetric tensors} \\
&\quad + \sin^2\theta\,\underbrace{\left(\cos\theta\,\mathbf{I} + (1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\right) : \mathbf{W}^{\mathrm{c}}}_{=\left(\cos\theta\,\mathbf{I} + (1-\cos\theta)\,\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\right) : \left(\widehat{\mathbf{e}}\otimes\widehat{\mathbf{e}}\right) = \cos\theta\,\widehat{\mathbf{e}}\cdot\widehat{\mathbf{e}} + (1-\cos\theta)\left(\widehat{\mathbf{e}}\cdot\widehat{\mathbf{e}}\right)^2 = 1} \\
&\quad + \sin^3\theta\,\underbrace{\det\mathbf{W}}_{=0} \\
&= \cos^2\theta + \sin^2\theta \\
&= +1\,,
\end{aligned} \tag{2.204}$$

where (2.73), (2.79h), (2.82), $(2.89\mathrm{b})_2$, (2.101), (2.180), (2.182), (2.186), (2.192a) and (2.199) have been used.

The desired expression (2.198) basically exhibits more applicable form of (2.197). To derive it, one first needs to establish the identity

$$\begin{aligned}
\mathbf{W}^2\mathbf{u} &\xlongequal[(2.25)]{\text{from}} \mathbf{W}\,(\mathbf{W}\mathbf{u}) \\
&\xlongequal[(2.199)]{\text{from}} \mathbf{W}\,(\widehat{\mathbf{e}} \times \mathbf{u}) \\
&\xlongequal[(2.199)]{\text{from}} \widehat{\mathbf{e}} \times (\widehat{\mathbf{e}} \times \mathbf{u}) \\
&\xlongequal[(1.71)]{\text{from}} (\widehat{\mathbf{e}} \cdot \mathbf{u})\,\widehat{\mathbf{e}} - (\widehat{\mathbf{e}} \cdot \widehat{\mathbf{e}})\,\mathbf{u} \\
&\xlongequal[\text{to Fig. 2.5}]{\text{with regard}} -\mathbf{v}\,,
\end{aligned} \tag{2.205}$$

and then rephrase (2.196) as

$$\begin{aligned}
\mathbf{u}_{\text{rot}} &= \cos\theta\,\mathbf{u} + (1-\cos\theta)\,(\mathbf{u}-\mathbf{v}) + \sin\theta\,\widehat{\mathbf{e}} \times \mathbf{u} \\
&= \mathbf{u} + \sin\theta\,\widehat{\mathbf{e}} \times \mathbf{u} - (1-\cos\theta)\,\mathbf{v}\,.
\end{aligned} \tag{2.206}$$

Now, making use of (2.5), (2.199) and $(2.205)_5$, the relation $(2.206)_2$ renders

$$\mathbf{u}_{\text{rot}} = \underbrace{\mathbf{I}\mathbf{u} + \sin\theta\,\mathbf{W}\mathbf{u} + (1-\cos\theta)\,\mathbf{W}^2\mathbf{u}}_{=\left(\mathbf{I} + \sin\theta\,\mathbf{W} + (1-\cos\theta)\,\mathbf{W}^2\right)\mathbf{u} = \mathbf{R}\mathbf{u}}\,.$$

The rotation tensor $\mathbf{R}$ in the above expression could be obtained in an alternative way. From (2.184a), one finds that $\widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}} = \mathbf{W}^2 + \mathbf{I}$. Introducing this expression into (2.197) identically delivers the required result.

The composition law (2.200) regarding, e.g., the Euler-Rodrigues formula can be verified in a straightforward manner by having

$$\mathbf{W}^2 = \widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}} - \mathbf{I} \quad \Rightarrow \quad \left.\begin{aligned} \mathbf{W}^3 &= (\widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}} - \mathbf{I})\,\mathbf{W} = -\mathbf{W} \\ \mathbf{W}^4 &= \mathbf{I} - \widehat{\mathbf{e}} \otimes \widehat{\mathbf{e}} = -\mathbf{W}^2 \end{aligned}\right\}\,.$$

Exercise 2.10

Let $\mathbf{exp}\,(\mathbf{A})$ and $\mathbf{exp}\,(\mathbf{B})$ be two exponential tensor functions according to (2.39). Show that the following relations hold

$$\mathbf{exp}\,(\mathbf{A}+\mathbf{B}) = \underbrace{\mathbf{exp}\,(\mathbf{A})\,\mathbf{exp}\,(\mathbf{B})}_{=\,\mathbf{exp}\,(\mathbf{B})\,\mathbf{exp}\,(\mathbf{A})} \quad \text{if} \quad \mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}\,, \tag{2.207a}$$

$$\mathbf{A}^m\left[\mathbf{exp}\left(\mathbf{B}\right)\right] = \left[\mathbf{exp}\left(\mathbf{B}\right)\right]\mathbf{A}^m \ , \quad \text{for any integer } m \text{ if} \quad \mathbf{AB} = \mathbf{BA} \ , \tag{2.207b}$$

$$\mathbf{I} = \underbrace{\mathbf{exp}\left(\mathbf{A}\right)\mathbf{exp}\left(-\mathbf{A}\right)}_{=\,\mathbf{exp}\left(-\mathbf{A}\right)\mathbf{exp}\left(\mathbf{A}\right)} \ , \tag{2.207c}$$

$$\mathbf{I} = \underbrace{\mathbf{exp}\left(\mathbf{W}\right)\left[\mathbf{exp}\left(\mathbf{W}\right)\right]^{\mathrm{T}}}_{=\,\left[\mathbf{exp}\left(\mathbf{W}\right)\right]^{\mathrm{T}}\mathbf{exp}\left(\mathbf{W}\right)} \quad \text{where} \quad \mathbf{W}^{\mathrm{T}} = -\mathbf{W} \ , \tag{2.207d}$$

$$\mathbf{exp}\left(m\mathbf{A}\right) = \left[\mathbf{exp}\left(\mathbf{A}\right)\right]^m \ , \quad \text{for any integer } m \ , \tag{2.207e}$$

$$\mathbf{exp}\left(\mathbf{A} + \mathbf{B}\right) = \mathbf{exp}\left(\mathbf{A}\right) + \mathbf{exp}\left(\mathbf{B}\right) - \mathbf{I} \quad \text{if} \quad \mathbf{AB} = \mathbf{BA} = \mathbf{O} \ , \tag{2.207f}$$

$$\mathbf{exp}\left(\mathbf{QAQ}^{\mathrm{T}}\right) = \mathbf{Q}\left[\mathbf{exp}\left(\mathbf{A}\right)\right]\mathbf{Q}^{\mathrm{T}} \quad \text{where} \quad \mathbf{QQ}^{\mathrm{T}} = \mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I} \ , \tag{2.207g}$$

$$\mathbf{exp}\left(\mathbf{BAB}^{-1}\right) = \mathbf{B}\left[\mathbf{exp}\left(\mathbf{A}\right)\right]\mathbf{B}^{-1} \quad \text{if} \quad \det\mathbf{B} \neq 0 \ . \tag{2.207h}$$

Solution. To begin with, consider the following two special exponential functions

$$\mathbf{exp}\left(\mathbf{O}\right) = \mathbf{I} \quad , \quad \mathbf{exp}\left(\mathbf{I}\right) = \mathrm{e}\mathbf{I} \ , \tag{2.208}$$

where e denotes the Euler's number. The commutative property $\mathbf{AB} = \mathbf{BA}$, along with (2.29) and (2.36), then helps establish

$$\left(\mathbf{A} + \mathbf{B}\right)^2 = \left(\mathbf{A} + \mathbf{B}\right)\left(\mathbf{A} + \mathbf{B}\right) = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2 \ .$$

This result can easily be generalized to the **binomial** formula

$$\left(\mathbf{A} + \mathbf{B}\right)^k = \sum_{i=0}^{k} \frac{k!}{i!\,(k-i)!}\mathbf{A}^{k-i}\mathbf{B}^i \quad \forall\, k \geq 0 \quad \text{if} \quad \mathbf{AB} = \mathbf{BA} \ . \tag{2.209}$$

The expression (2.207a): This relation states that if two arbitrary tensors commute, their exponential tensor functions will consistently commute. Moreover, the single composition of an exponential tensor function, $\mathbf{exp}\left(\mathbf{A}\right)$, with another exponential map, $\mathbf{exp}\left(\mathbf{B}\right)$, equals the exponential of the addition of their arguments provided that $\mathbf{AB} = \mathbf{BA}$. This can be proved directly from the definition of the exponential tensor function. On the one hand, one will have

$$\begin{aligned}
\mathbf{exp}\left(\mathbf{A} + \mathbf{B}\right) &\xlongequal[\text{(2.39)}]{\text{from}} \sum_{k=0}^{\infty} \frac{1}{k!}\left(\mathbf{A} + \mathbf{B}\right)^k \\
&\xlongequal[\text{(2.209)}]{\text{from}} \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^{k} \frac{k!}{i!\,(k-i)!}\mathbf{A}^{k-i}\mathbf{B}^i \\
&\xlongequal[\text{simplification}]{\text{by}} \sum_{k=0}^{\infty} \sum_{i=0}^{k} \frac{\mathbf{A}^{k-i}\mathbf{B}^i}{i!\,(k-i)!}
\end{aligned}$$

$$\xlongequal[\text{the order of summation}]{\text{by changing}} \sum_{i=0}^{\infty} \sum_{k=i}^{\infty} \frac{\mathbf{A}^{k-i}\mathbf{B}^{i}}{i!\,(k-i)!}$$

$$\xlongequal[\text{setting } j=k-i]{\text{by}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\mathbf{A}^{j}\mathbf{B}^{i}}{i!\,j!} \,.$$

And, one the other hand,

$$\mathbf{exp}\,(\mathbf{A})\,\mathbf{exp}\,(\mathbf{B}) \xlongequal[(2.39)]{\text{from}} \left(\sum_{j=0}^{\infty} \frac{1}{j!} \mathbf{A}^{j} \right) \left(\sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{B}^{i} \right)$$

$$\xlongequal[(2.29)]{\text{from}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\mathbf{A}^{j}\mathbf{B}^{i}}{i!\,j!} \,.$$

In a similar manner,

$$\mathbf{exp}\,(\mathbf{B})\,\mathbf{exp}\,(\mathbf{A}) \xlongequal[(2.39)]{\text{from}} \left(\sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{B}^{i} \right) \left(\sum_{j=0}^{\infty} \frac{1}{j!} \mathbf{A}^{j} \right)$$

$$\xlongequal[(2.29)]{\text{from}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\mathbf{B}^{i}\mathbf{A}^{j}}{i!\,j!}$$

$$\xlongequal[\text{considering } \mathbf{AB}=\mathbf{BA}]{\text{by}} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\mathbf{A}^{j}\mathbf{B}^{i}}{i!\,j!} \,.$$

It is worthwhile to point out that the converse of (2.207a), in general, is not true. A lot of examples can be found in the context of matrix algebra to show that $\mathbf{exp}\,(\mathbf{A}+\mathbf{B}) = \mathbf{exp}\,(\mathbf{A})\,\mathbf{exp}\,(\mathbf{B})$ whereas $\mathbf{AB} \neq \mathbf{BA}$. For instance, consider a 4×4 matrix $[\,\mathbf{A}\,]$ with only nonzero elements $A_{34} = 2\pi$ and $A_{43} = -2\pi$. And suppose that $[\,\mathbf{B}\,]$ be a 4×4 matrix with only nonzero elements $B_{31} = 1$, $B_{34} = 2\pi$, $B_{42} = 1$ and $B_{43} = -2\pi$.

The expression (2.207b): If $\mathbf{AB} = \mathbf{BA}$, then $\mathbf{A}^{m}\mathbf{B}^{i} = \mathbf{B}^{i}\mathbf{A}^{m}$ holds for all integers m and i. As a result, the identity (2.207b) is implied by the definition of the exponential tensor function.

The expression (2.207c): Since $\mathbf{A}$ and $-\mathbf{A}$ commute, it follows that

$$\mathbf{exp}\,(\mathbf{A})\,\mathbf{exp}\,(-\mathbf{A}) \xlongequal[(2.207a)]{\text{from}} \mathbf{exp}\,(\mathbf{A}+(-\mathbf{A}))$$

$$\xlongequal[(2.11c)]{\text{from}} \mathbf{exp}\,(\mathbf{O})$$

$$\xlongequal[(2.208)]{\text{from}} \mathbf{I} \,.$$

This result immediately implies that

$$\mathbf{exp}\,(-\mathbf{A}) = \left[\mathbf{exp}\,(\mathbf{A})\right]^{-1} \ . \tag{2.210}$$

The expression (2.207d): Since any skew tensor commutes with its transpose, one can write

$$\begin{aligned}
\mathbf{exp}\,(\mathbf{W})\left[\mathbf{exp}\,(\mathbf{W})\right]^{\mathrm{T}} &\xlongequal[\text{(2.39), (2.52), (2.53) and (2.55e)}]{\text{from}} \mathbf{exp}\,(\mathbf{W})\,\mathbf{exp}\left(\mathbf{W}^{\mathrm{T}}\right) \\
&\xlongequal[\text{(2.58) and (2.207a)}]{\text{from}} \mathbf{exp}\,(\mathbf{W} + (-\mathbf{W})) \\
&\xlongequal[\text{(2.11c) and (2.208)}]{\text{from}} \mathbf{I} \ .
\end{aligned}$$

One can now infer that $\mathbf{exp}\,(\mathbf{W})$ should be an **orthogonal tensor**. By invoking (4.59a) and (4.63a), one can arrive at

$$\det\left[\mathbf{exp}\,(\mathbf{W})\right] = \exp\,(\mathrm{tr}\mathbf{W}) = \exp\,(0) = +1 \ , \tag{2.211}$$

which reveals the fact that $\mathbf{exp}\,(\mathbf{W})$ is a **rotation tensor**.

The expression (2.207e): This identity will be treated by means of **mathematical induction**. First note that when $m = 0$, this expression in the form $\mathbf{exp}\,(0\mathbf{A}) = \left[\mathbf{exp}\,(\mathbf{A})\right]^0$ renders $\mathbf{I} = \mathbf{I}$ using (b) in (2.12), (2.36) and $(2.208)_1$.

For $m = 1$, the relation (2.207e) holds true by the definition of the exponential tensor function along with (2.11e) and (2.36). Now, suppose that it remains valid for $m = k > 0$. It follows that

$$\begin{aligned}
\mathbf{exp}\,[(k+1)\,\mathbf{A}] &\xlongequal[\text{(2.11e) and (2.11g)}]{\text{from}} \mathbf{exp}\,[k\mathbf{A} + \mathbf{A}] \\
&\xlongequal[\text{(2.207a)}]{\text{from}} \mathbf{exp}\,(k\mathbf{A})\,\mathbf{exp}\,(\mathbf{A}) \\
&\xlongequal[\text{assumption}]{\text{by}} \left[\mathbf{exp}\,(\mathbf{A})\right]^k \mathbf{exp}\,(\mathbf{A}) \\
&\xlongequal[\text{(2.37)}]{\text{from}} \left[\mathbf{exp}\,(\mathbf{A})\right]^{k+1} \ .
\end{aligned}$$

If $m = k < 0$, it is not then difficult to see that

$$\mathbf{exp}\,[(k-1)\,\mathbf{A}] = \left[\mathbf{exp}\,(\mathbf{A})\right]^{k-1} \ .$$

The expression (2.207f): It is evident that $(\mathbf{A} + \mathbf{B})^k = \mathbf{A}^k + \mathbf{B}^k$ holds for all positive integers k when $\mathbf{AB} = \mathbf{BA} = \mathbf{O}$. This result along with the useful identities $\mathbf{A} =$

$\mathbf{A} + \mathbf{O} = \mathbf{A} + \mathbf{I} - \mathbf{I}$ and $\mathbf{A}^0 = \mathbf{I}$ yields

$$\begin{aligned}\mathbf{exp}\,(\mathbf{A} + \mathbf{B}) &= \sum_{k=0}^{\infty} \frac{(\mathbf{A} + \mathbf{B})^k}{k!} \\ &= \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!} + \sum_{k=1}^{\infty} \frac{\mathbf{B}^k}{k!} \\ &= \mathbf{exp}\,(\mathbf{A}) + \mathbf{exp}\,(\mathbf{B}) - \mathbf{I} \,.\end{aligned}$$

The expression (2.207g):

$$\begin{aligned}\mathbf{exp}\left(\mathbf{QAQ}^{\mathrm{T}}\right) &\xlongequal[\text{(2.39)}]{\text{from}} \sum_{k=0}^{\infty} \frac{\left(\mathbf{QAQ}^{\mathrm{T}}\right)^k}{k!} \\ &\xlongequal[\text{(2.33), (2.36) and (2.130)}]{\text{from}} \sum_{k=0}^{\infty} \frac{\mathbf{QA}^k\mathbf{Q}^{\mathrm{T}}}{k!} \\ &\xlongequal[\text{(2.29)}]{\text{from}} \mathbf{Q} \left[\sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!} \right] \mathbf{Q}^{\mathrm{T}} \\ &\xlongequal[\text{(2.39)}]{\text{from}} \mathbf{Q} \left[\mathbf{exp}\,(\mathbf{A}) \right] \mathbf{Q}^{\mathrm{T}} \,.\end{aligned}$$

The expression (2.207h):

$$\begin{aligned}\mathbf{exp}\left(\mathbf{BAB}^{-1}\right) &\xlongequal[\text{(2.36) and (2.39)}]{\text{from}} \sum_{k=0}^{\infty} \frac{1}{k!} \underbrace{\mathbf{BAB}^{-1}\mathbf{BAB}^{-1} \cdots \mathbf{BAB}^{-1}}_{k \text{ times}} \\ &\xlongequal[\text{(2.33), (2.36) and (2.105)}]{\text{from}} \sum_{k=0}^{\infty} \frac{\mathbf{BA}^k\mathbf{B}^{-1}}{k!} \\ &\xlongequal[\text{(2.29) and (2.39)}]{\text{from}} \mathbf{B} \left[\mathbf{exp}\,(\mathbf{A}) \right] \mathbf{B}^{-1} \,.\end{aligned}$$

Here, some important properties of the exponential tensor function were investigated. See Exercise 4.6 for more elaboration.

Exercise 2.11

Suppose one is given a tensor $\mathbf{A}$ with the following matrix

$$[\,\mathbf{A}\,] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} .$$

Obtain the spherical and deviatoric parts of $\mathbf{A}$.

Solution. First, one needs to have $\mathrm{tr}\mathbf{A} = 1 + 1 + 1 = 3$. Then, by (2.144) and (2.145), one can obtain

$$\left[\,\mathrm{sph}\mathbf{A}\,\right] = \frac{\mathrm{tr}\mathbf{A}}{3}\,[\,\mathbf{I}\,] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ,$$

and

$$[\,\mathrm{dev}\mathbf{A}\,] = [\,\mathbf{A}\,] - \frac{\mathrm{tr}\mathbf{A}}{3}\,[\,\mathbf{I}\,] = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} .$$

References

1. Simo, J. C. (1998). Numerical analysis and simulation of plasticity. In P. G. Ciarlet & J. L. Lions (Eds.), *Handbook of numerical analysis* (Vol. 6, pp. 183–499. Elsevier Science B.V.
2. Simo, J. C., & Hughes, T. J. R. (1998). *Computational inelasticity*. New York, NY: Springer.
3. de Souza Neto, E. A., Perić, D., & Owen, D. R. J. (2008). *Computational methods for plasticity: Theory and applications*. Wiley.
4. Gurtin, M. E., Fried, E., & Anand, L. (2010). *The mechanics and thermodynamics of continua*. Cambridge: Cambridge University Press.
5. Beju, I., Soós, E., & Teodorescu, P. P. (1983). *Euclidean tensor calculus with applications*. Abacus Press.
6. Tadmor, E. B., Miller, R. E., & Elliott, R. S. (2012). *Continuum mechanics and thermodynamics: From fundamental concepts to governing equations*. Cambridge: Cambridge University Press.
7. Hodge, P. G. (1961). On isotropic Cartesian tensors. *The American Mathematical Monthly, 68*(8), 793–795.

8. Murray, R. M., Li, Z., & Sastry, S. S. (1994). *A mathematical introduction to robotic manipulation*. CRC Press.
9. Crisfield, M. A. (1997). *Non-linear Finite Element Analysis of Solids and Structures, Volume 2: Advanced Topics*. Wiley.
10. Wriggers, P. (2008). *Nonlinear finite element methods*. Berlin, Heidelberg: Springer.

Chapter 3
Algebra of Higher-Order Tensors

Recall from Chap. 1 that scalars were zeroth-order tensors with only one component and vectors or first-order tensors were characterized by three components in three-dimensional spaces. The order of a tensor was consistently increased in Chap. 2 wherein a tensor of order two, characterized by nine components, was declared as a linear mapping that takes a vector as an input and generally delivers another vector as an output. The same idea will be followed in this chapter. Here, higher-order tensors are defined by appealing to the notion of **linear mapping** and their important relationships are characterized. The main focus will be on the fourth-order tensors which are extensively used, for instance, in continuum mechanics of solids. See, e.g., Gurtin [1], Chandrasekharaiah and Debnath [2], Negahban [3], Jog [4] and Irgens [5].

3.1 Tensors of Order Three

A third-order tensor is denoted here by A , B , $\cdots$. It is defined as a linear mapping of first-order tensors in $\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}$ into second-order tensors in $\mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right)$ according to

$$\boxed{\mathbf{A} = \mathsf{A}\mathbf{u} \ , \qquad \mathbf{A} \in \mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right) \ , \ \forall \mathsf{A} \in \mathscr{T}_{\mathrm{to}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right) \ , \ \forall \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}} \ ,} \tag{3.1}$$

where $\mathscr{T}_{\mathrm{to}}$ denotes the set of all third-order tensors. Consistent with (2.1), it renders a right mapping and satisfies the linearity condition

$$\boxed{\mathsf{A}\left(\alpha\mathbf{u} + \beta\mathbf{v}\right) = \alpha\left(\mathsf{A}\mathbf{u}\right) + \beta\left(\mathsf{A}\mathbf{v}\right) \ , \qquad \text{for all } \alpha, \beta \in \mathrm{R} \ ,} \tag{3.2}$$

for which the basic operations **addition** and **scalar multiplication** are defined as follows:

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_3

$$(\mathbf{A} + \mathbf{B})\,\mathbf{u} = \mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{u}\,, \tag{3.3a}$$
$$(\alpha\mathbf{A})\,\mathbf{u} = \alpha\,(\mathbf{A}\mathbf{u}) = \mathbf{A}\,(\alpha\mathbf{u})\;. \tag{3.3b}$$

The *third-order zero tensor* $\mathbf{O}$ is introduced as

$$\boxed{\mathbf{O}\mathbf{u} = \mathbf{O}\,, \qquad \text{for all } \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\;.} \tag{3.4}$$

A dyad $\mathbf{u} \otimes \mathbf{v}$ representing a second-order tensor is now extended to the *triadic product* $\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$ satisfying the property

$$\boxed{(\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w} = \mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w}) = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}\,,} \tag{3.5}$$

which linearly transforms an arbitrary vector $\bar{\mathbf{w}}$ onto a scalar multiple of $\mathbf{u} \otimes \mathbf{v}$ by the rule

$$\boxed{\underbrace{(\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w})\,\bar{\mathbf{w}} = (\mathbf{w} \cdot \bar{\mathbf{w}})\,\mathbf{u} \otimes \mathbf{v}}_{\text{by (1.9a), one can also have } (\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w})\,\bar{\mathbf{w}} = (\mathbf{u} \otimes \mathbf{v} \otimes \bar{\mathbf{w}})\,\mathbf{w}}\;.} \quad \longleftarrow \text{see (2.13)} \tag{3.6}$$

Thus, the linear map $\boldsymbol{g}_{\mathrm{to}}^{\mathrm{tp}}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$ represents a tensor of order three. Consistent with (2.16a)–(2.16b), it is not then difficult to notice that $\boldsymbol{g}_{\mathrm{to}}^{\mathrm{tp}}$ is **trilinear**, i.e.

$$(\alpha\mathbf{u} + \beta\bar{\mathbf{u}}) \otimes \mathbf{v} \otimes \mathbf{w} = \alpha\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} + \beta\bar{\mathbf{u}} \otimes \mathbf{v} \otimes \mathbf{w}\,, \tag{3.7a}$$
$$\mathbf{u} \otimes (\alpha\mathbf{v} + \beta\bar{\mathbf{v}}) \otimes \mathbf{w} = \alpha\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} + \beta\mathbf{u} \otimes \bar{\mathbf{v}} \otimes \mathbf{w}\,, \tag{3.7b}$$
$$\mathbf{u} \otimes \mathbf{v} \otimes (\alpha\mathbf{w} + \beta\bar{\mathbf{w}}) = \alpha\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} + \beta\mathbf{u} \otimes \mathbf{v} \otimes \bar{\mathbf{w}}\;. \tag{3.7c}$$

It may be beneficial to view a tensor of order three as a linear transformation that associates with each tensor in $\mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right)$ a vector in $\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}$. Such a transformation is indicated by

$$\boxed{\mathbf{u} = \mathbf{A} : \mathrm{A}\,, \qquad \mathbf{u} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\,,\ \forall\mathbf{A} \in \mathscr{T}_{\mathrm{to}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right)\,,\ \forall\mathrm{A} \in \mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right),} \tag{3.8}$$

which satisfies

$$\mathbf{A} : (\alpha\mathrm{A} + \beta\mathrm{B}) = \alpha\,(\mathbf{A} : \mathrm{A}) + \beta\,(\mathbf{A} : \mathrm{B})\;, \qquad \text{for all } \alpha, \beta \in \mathrm{R}\,, \tag{3.9}$$

and

$$(\mathbf{A} + \mathbf{B}) : \mathrm{A} = \mathbf{A} : \mathrm{A} + \mathbf{B} : \mathrm{A}\,, \tag{3.10a}$$
$$(\alpha\mathbf{A}) : \mathrm{A} = \alpha\,(\mathbf{A} : \mathrm{A}) = \mathbf{A} : (\alpha\mathrm{A})\;. \tag{3.10b}$$

Now, the double contraction between a triadic product and a dyad obeys

$$\boxed{\mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w}) : \left(\bar{\mathbf{w}} \otimes \hat{\mathbf{w}}\right) = (\mathbf{v} \cdot \bar{\mathbf{w}}) \left(\mathbf{w} \cdot \hat{\mathbf{w}}\right) \mathbf{u} \,.} \tag{3.11}$$

Having in mind (2.74a) and (2.74b), the above operation is a symmetric bilinear form. Guided by (2.17), the collection

$$\left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k,\ i, j, k = 1, 2, 3\right\} \stackrel{\text{def}}{=\!=} \left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k\right\} \,, \tag{3.12}$$

constitutes a basis for $\mathscr{T}_{\text{to}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ and its dimension is $\dim \mathscr{T}_{\text{to}} = \left(\dim \mathscr{E}_{\text{r}}^{\text{o3}}\right)^3 = 3^3$. Subsequently, any element $\mathbf{A}$ of this space can be represented by this basis and a collection of 27 scalar numbers as

$$\boxed{\mathbf{A} = \mathsf{A}_{ijk} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \,.} \quad \longleftarrow \text{see (2.19)} \tag{3.13}$$

Obviously, (3.12) presents a **Cartesian** (or **rectangular**) basis. Thus, $\mathbf{A}$ in (3.13) is called a Cartesian third-order tensor and A_{ijk} will be the respective Cartesian components. They are determined via

$$\boxed{\mathsf{A}_{ijk} = \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\mathbf{A}\widehat{\mathbf{e}}_k\right] \,,} \quad \longleftarrow \text{see (2.20)} \tag{3.14}$$

where

$$\boxed{\mathbf{A}\widehat{\mathbf{e}}_k = \mathsf{A}_{mnk} \widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n \,,} \tag{3.15}$$

since

$$\begin{aligned}
\mathsf{A}_{ijk} &\xrightarrow[\text{(1.36)}]{\text{from}} \mathsf{A}_{mnk} \left[\delta_{im} \delta_{jn}\right] \\
&\xrightarrow[\text{(1.35)}]{\text{from}} \mathsf{A}_{mnk} \left[\left(\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_m\right) \left(\widehat{\mathbf{e}}_j \cdot \widehat{\mathbf{e}}_n\right)\right] \\
&\xrightarrow[\text{(2.73), (2.74a) and (2.74b)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\left(\mathsf{A}_{mnk} \widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n\right)\right] \\
&\xrightarrow[\text{(1.35) and (1.36)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\left(\mathsf{A}_{mno} \widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n\right) \left(\widehat{\mathbf{e}}_o \cdot \widehat{\mathbf{e}}_k\right)\right] \\
&\xrightarrow[\text{(3.3a), (3.3b) and (3.6)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\left(\mathsf{A}_{mno} \widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n \otimes \widehat{\mathbf{e}}_o\right) \widehat{\mathbf{e}}_k\right] \\
&\xrightarrow[\text{(3.13)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\mathbf{A}\widehat{\mathbf{e}}_k\right] \,.
\end{aligned}$$

Let $\mathbf{u}$ and $\mathbf{A}$ are decomposed according to $(1.34)_1$ and $(2.19)_2$, respectively. Further, let $\mathbf{A}$ be a third-order tensor of the form (3.13). Then, using (1.35), (1.36), (3.2)-(3.3b), (3.6) and (3.9)-(3.11), the simple contraction (3.1) and the double contraction (3.8) can be expressed as

$$\begin{aligned}\mathbf{A} = \mathsf{\mathbf{A}}\mathbf{u} &= \left[\mathsf{A}_{ijk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k\right][u_l \widehat{\mathbf{e}}_l] \\ &= \mathsf{A}_{ijk} u_l \delta_{kl} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) \\ &= \mathsf{A}_{ijk} u_k \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) \qquad \text{with } A_{ij} = \mathsf{A}_{ijk} u_k \ , \end{aligned} \tag{3.16a}$$

$$\begin{aligned}\mathbf{u} = \mathsf{\mathbf{A}} : \mathbf{A} &= \left[\mathsf{A}_{ijk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k\right] : [A_{lm} \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_m] \\ &= \mathsf{A}_{ijk} A_{lm} \delta_{jl} \delta_{km} \widehat{\mathbf{e}}_i \\ &= \mathsf{A}_{ijk} A_{jk} \widehat{\mathbf{e}}_i \qquad \text{with } u_i = \mathsf{A}_{ijk} A_{jk} \ . \end{aligned} \tag{3.16b}$$

A well-known example of third-order tensor in continuum mechanics is the **permutation** (or **alternating**) tensor

$$\boxed{\mathsf{\mathbf{E}} = \varepsilon_{ijk} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \ ,} \tag{3.17}$$

where the Cartesian components ε_{ijk} have already been introduced as the permutation symbol in (1.53). Guided by $(3.16a)_4$ and having defined the axial vector $\boldsymbol{\omega}$ of a skew tensor $\mathbf{W}$ according to $(2.70)_1$, operating the permutation tensor $\mathsf{\mathbf{E}}$ on the vector $\boldsymbol{\omega}$ yields

$$\mathsf{\mathbf{E}}\boldsymbol{\omega} = -\mathbf{W} \quad \text{with} \quad \varepsilon_{ijm}\omega_m = -W_{ij} \ . \quad \leftarrow \text{see (2.71)} \tag{3.18}$$

One can also establish

$$\boldsymbol{\omega} = -\frac{1}{2}\mathsf{\mathbf{E}} : \mathbf{W} \quad \text{with} \quad \omega_m = -\frac{1}{2}\varepsilon_{mkl} W_{kl} = -\frac{1}{2} W_{kl} \varepsilon_{klm} \ . \quad \leftarrow \text{see (2.70)} \tag{3.19}$$

Recall from $(2.62)_2$ that any tensor $\mathbf{A}$ can additively be decomposed as $\mathbf{A} = \mathbf{S} + \mathbf{W}$. In light of (2.79h), considering $\varepsilon_{ijk} = -\varepsilon_{ikj}$ and $S_{jk} = S_{kj}$ into $(3.16b)_4$ leads to the useful identity

$$\boxed{\mathsf{\mathbf{E}} : \mathbf{S} = \mathbf{0} \ .} \tag{3.20}$$

This immediately implies that

$$\boldsymbol{\omega} = -\frac{1}{2}\mathsf{\mathbf{E}} : \mathbf{A} \ . \tag{3.21}$$

The double contraction of the permutation tensor $\mathsf{\mathbf{E}}$ and the dyadic product $\mathbf{u} \otimes \mathbf{v}$ results in the cross product of $\mathbf{u}$ and $\mathbf{v}$, that is,

$$\boxed{\mathsf{\mathbf{E}} : (\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \times \mathbf{v} \ ,} \tag{3.22}$$

because

$$\begin{aligned}
\left[\mathbf{E}\right] : [\mathbf{u} \otimes \mathbf{v}] &\xlongequal[\text{(3.17)}]{\text{from}} \left[\varepsilon_{ijk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k\right] : [\mathbf{u} \otimes \mathbf{v}] \\
&\xlongequal[\text{(3.10a)-(3.11)}]{\text{from}} \left[\varepsilon_{ijk}\left(\widehat{\mathbf{e}}_j \cdot \mathbf{u}\right)\left(\widehat{\mathbf{e}}_k \cdot \mathbf{v}\right)\right]\widehat{\mathbf{e}}_i \\
&\xlongequal[\text{(1.33)}]{\text{from}} \left(\varepsilon_{ijk}u_j v_k\right)\widehat{\mathbf{e}}_i \\
&\xlongequal[\text{(1.54)}]{\text{from}} -\left(v_k u_j \varepsilon_{kji}\right)\widehat{\mathbf{e}}_i \\
&\xlongequal[\text{(1.67)}]{\text{from}} -\mathbf{v} \times \mathbf{u} \\
&\xlongequal[\text{(1.49a)}]{\text{from}} \mathbf{u} \times \mathbf{v} \; .
\end{aligned}$$

Isotropic tensors of order three. Under a change of coordinates and consistent with the expression (2.167)$_3$, the old and new Cartesian components of a third-order tensor $\mathbf{A} = \mathsf{A}^o_{ijk}\widehat{\mathbf{e}}^o_i \otimes \widehat{\mathbf{e}}^o_j \otimes \widehat{\mathbf{e}}^o_k = \mathsf{A}^n_{ijk}\widehat{\mathbf{e}}^n_i \otimes \widehat{\mathbf{e}}^n_j \otimes \widehat{\mathbf{e}}^n_k$ are related by

$$\boxed{\mathsf{A}^n_{ijk} = R_{li}R_{mj}R_{nk}\mathsf{A}^o_{lmn} \; .} \tag{3.23}$$

This tensor is said to be **isotropic** when

$$\boxed{\mathsf{A}_{ijk} = Q_{li}Q_{mj}Q_{nk}\mathsf{A}_{lmn} \; .} \tag{3.24}$$

This amounts to writing $\mathbf{A} = \mathsf{A}_{ijk}\widehat{\mathbf{e}}^o_i \otimes \widehat{\mathbf{e}}^o_j \otimes \widehat{\mathbf{e}}^o_k = \mathsf{A}_{ijk}\widehat{\mathbf{e}}^n_i \otimes \widehat{\mathbf{e}}^n_j \otimes \widehat{\mathbf{e}}^n_k$. In the following, it will be verified that a generic isotropic third-order tensor is a scalar multiple of the permutation tensor $\mathbf{E} = \varepsilon_{ijk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k$, i.e.

$$\boxed{\mathsf{A}_{ijk} = \mu\varepsilon_{ijk} \quad \text{where } \mu \text{ is a constant} \; .} \quad \longleftarrow \text{see (3.73)} \tag{3.25}$$

Proof. Consider an infinitesimal rotation of the form (2.174), i.e. $R_{ij} = \delta_{ij} - \varepsilon_{ijm}\omega_m$. Introducing this relation into (3.24) yields

$$\omega_p\left(\varepsilon_{mjp}\mathsf{A}_{imk} + \varepsilon_{lip}\mathsf{A}_{ljk} + \varepsilon_{nkp}\mathsf{A}_{ijn}\right) = 0 \; , \tag{3.26}$$

where the higher-order terms have been neglected. The fact that ω_i, $i = 1, 2, 3$, are arbitrary then implies that

$$\varepsilon_{mjp}\mathsf{A}_{imk} + \varepsilon_{lip}\mathsf{A}_{ljk} + \varepsilon_{nkp}\mathsf{A}_{ijn} = 0 \; . \tag{3.27}$$

By multiplying both sides of this expression by ε_{qip}, ε_{qjp} and ε_{qkp} and then setting $q = i$, $q = j$, and $q = k$ in the resulting expressions, one can arrive at

$$\mathsf{A}_{jik} + 2\mathsf{A}_{ijk} + \mathsf{A}_{kji} = \delta_{ij}\mathsf{A}_{rrk} + \delta_{ik}\mathsf{A}_{rjr} \ , \tag{3.28a}$$

$$2\mathsf{A}_{ijk} + \mathsf{A}_{jik} + \mathsf{A}_{ikj} = \delta_{ji}\mathsf{A}_{rrk} + \delta_{jk}\mathsf{A}_{irr} \ , \tag{3.28b}$$

$$\mathsf{A}_{ikj} + \mathsf{A}_{kji} + 2\mathsf{A}_{ijk} = \delta_{kj}\mathsf{A}_{irr} + \delta_{ki}\mathsf{A}_{rjr} \ . \tag{3.28c}$$

By multiplying both sides of these relations by δ_{jk}, δ_{ik} and δ_{ij} and then setting $i = i$, $j = i$, and $k = i$ in the resulting expressions, one will have

$$\mathsf{A}_{irr} = \mathsf{A}_{rir} = \mathsf{A}_{rri} = 0 \ . \tag{3.29}$$

As a result,

$$2\mathsf{A}_{ijk} + \mathsf{A}_{jik} + \mathsf{A}_{kji} = 0 \ , \tag{3.30a}$$

$$2\mathsf{A}_{ijk} + \mathsf{A}_{jik} + \mathsf{A}_{ikj} = 0 \ , \tag{3.30b}$$

$$2\mathsf{A}_{ijk} + \mathsf{A}_{ikj} + \mathsf{A}_{kji} = 0 \ . \tag{3.30c}$$

It is then easy to see that

$$\mathsf{A}_{ikj} = \mathsf{A}_{kji} = \mathsf{A}_{jik} = -\mathsf{A}_{ijk} \ . \tag{3.31}$$

Considering $(1.52)_3$, $(1.54)_{3-5}$, $(3.29)_{1-3}$ and $(3.31)_{1-3}$, one can finally arrive at the required result.

3.2 Tensors of Order Four

Fourth-order tensors are extensively used in theoretical as well as computational nonlinear solid mechanics and designated in this textbook by $\mathbb{A}$, $\mathbb{a}$, $\cdots$. Denoting by $\mathscr{T}_{\text{fo}}$ the set of all fourth-order tensors, a tensor of order four may be defined as a linear mapping of vectors in $\mathscr{E}_{\text{r}}^{\text{o3}}$ into third-order tensors in $\mathscr{T}_{\text{to}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ according to

$$\mathsf{\mathbf{A}} = \mathbb{A}\mathbf{u} \ , \qquad \mathsf{\mathbf{A}} \in \mathscr{T}_{\text{to}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right) \ , \ \forall\mathbb{A} \in \mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right) \ , \ \forall\mathbf{u} \in \mathscr{E}_{\text{r}}^{\text{o3}} \ . \tag{3.32}$$

Another definition is

$$\boxed{\mathbf{D} = \mathbb{A} : \mathbf{C} \ , \qquad \mathbf{D} \in \mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right) \ , \ \forall\mathbb{A} \in \mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right) \ , \ \forall\mathbf{C} \in \mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right) \ .} \tag{3.33}$$

Of special interest here is to focus on the above useful definition of fourth-order tensors. This enables one to develop a formalism based on the relations established in Chap. 2 for second-order tensors.

In alignment with (2.2) and (3.2), the right mapping (3.33) fulfills the linearity condition

$$\boxed{\mathbb{A} : (\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha\,(\mathbb{A} : \mathbf{A}) + \beta\,(\mathbb{A} : \mathbf{B}) \ , \quad \text{for all } \alpha, \beta \in \mathrm{R} \ .} \tag{3.34}$$

Following discussions analogous to those that led to (2.3), the requirement (3.34) helps establish

$$\boxed{\mathbb{A} : \mathbf{O} = \mathbf{O} \ .} \tag{3.35}$$

One can introduce the *fourth-order zero tensor* $\mathbb{O}$ via

$$\boxed{\mathbb{O} : \mathbf{C} = \mathbf{O} \ , \quad \text{for all } \mathbf{C} \in \mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o}3}\right) \ ,} \tag{3.36}$$

and the *fourth-order unit tensors* $\mathbb{I}$ and $\bar{\mathbb{I}}$ according to

$$\boxed{\mathbb{I} : \mathbf{C} = \mathbf{C} \quad , \quad \bar{\mathbb{I}} : \mathbf{C} = \mathbf{C}^{\mathrm{T}} \ .} \tag{3.37}$$

Two fourth-order tensors $\mathbb{A}$ and $\mathbb{B}$ are said to be *equal* if they identically map all tensors $\mathbf{C}$, that is,

$$\boxed{\mathbb{A} = \mathbb{B} \quad \text{if and only if} \quad \mathbb{A} : \mathbf{C} = \mathbb{B} : \mathbf{C} \ ,} \tag{3.38}$$

or, equivalently,

$$\boxed{\mathbb{A} = \mathbb{B} \quad \text{if and only if} \quad \mathbf{D} : \mathbb{A} : \mathbf{C} = \mathbf{D} : \mathbb{B} : \mathbf{C} \ .} \tag{3.39}$$

Recall that a vector space remained closed with respect to (vector) addition and (scalar) multiplication. And these fundamental operations were indicated in (2.8a) and (2.8b) regarding second-order tensors. Here, they are rewritten as

$$(\mathbb{A} + \mathbb{B}) : \mathbf{C} = \mathbb{A} : \mathbf{C} + \mathbb{B} : \mathbf{C} \ , \tag{3.40a}$$

$$(\alpha\mathbb{A}) : \mathbf{C} = \alpha\,(\mathbb{A} : \mathbf{C}) = \mathbb{A} : (\alpha\mathbf{C}) \ . \tag{3.40b}$$

Having in mind $(-1)\,\mathbf{A} = -\mathbf{A}$ from (2.12), the rule (3.40b) with the aid of (3.38) reveals that

$$\boxed{(-1)\,\mathbb{A} = -\mathbb{A} \ .} \tag{3.41}$$

Similarly to vectors and tensors, the difference between two fourth-order tensors $\mathbb{A}$ and $\mathbb{B}$ are indicated by

$$\mathbb{A} + (-\mathbb{B}) = \mathbb{A} - \mathbb{B} \ . \quad \leftarrow \text{see (2.9)} \tag{3.42}$$

Then, it follows that

$$(\mathbb{A} - \mathbb{B}) : \mathbf{C} = \mathbb{A} : \mathbf{C} - \mathbb{B} : \mathbf{C} \ . \tag{3.43}$$

Consistent with (1.4a) to (1.4h) and also (2.11a) to (2.11h), the following properties holds for any $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C} \in \mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ and α , $\beta \in \text{R}$[1]:

$$\mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A} \ , \tag{3.44a}$$
$$(\mathbb{A} + \mathbb{B}) + \mathbb{C} = \mathbb{A} + (\mathbb{B} + \mathbb{C}) \ , \tag{3.44b}$$
$$\mathbb{A} + (-\mathbb{A}) = \mathbb{O} \ , \tag{3.44c}$$
$$\mathbb{A} + \mathbb{O} = \mathbb{A} \ , \tag{3.44d}$$
$$1\mathbb{A} = \mathbb{A} \ , \tag{3.44e}$$
$$(\alpha\beta)\,\mathbb{A} = \alpha\,(\beta\mathbb{A}) \ , \tag{3.44f}$$
$$(\alpha + \beta)\,\mathbb{A} = \alpha\mathbb{A} + \beta\mathbb{A} \ , \tag{3.44g}$$
$$\alpha\,(\mathbb{A} + \mathbb{B}) = \alpha\mathbb{A} + \alpha\mathbb{B} \ . \tag{3.44h}$$

The set $\mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ is now declared as another **vector space** since not only it remains closed with respect to the operations (3.40a) and (3.40b) but also its elements satisfy the properties (3.44a) to (3.44h).

With regard to the right mapping (3.33), one can also define the **left mapping**

$$\boxed{\mathbf{C} = \mathbf{D} : \mathbb{A} \ ,} \quad \longleftarrow \text{see (2.40)} \tag{3.45}$$

satisfying

$$(\alpha\mathbf{D} + \beta\mathbf{E}) : \mathbb{A} = \alpha\,(\mathbf{D} : \mathbb{A}) + \beta\,(\mathbf{E} : \mathbb{A}) \ , \quad \longleftarrow \text{see (2.41)} \tag{3.46}$$

and

$$\mathbf{D} : (\mathbb{A} + \mathbb{B}) = \mathbf{D} : \mathbb{A} + \mathbf{D} : \mathbb{B} \ , \quad \longleftarrow \text{see (2.42a)} \tag{3.47a}$$
$$\mathbf{D} : (\alpha\mathbb{A}) = \alpha\,(\mathbf{D} : \mathbb{A}) = (\alpha\mathbf{D}) : \mathbb{A} \ . \quad \longleftarrow \text{see (2.42b)} \tag{3.47b}$$

Accordingly, the left and right mappings are related through the following expression

$$\boxed{(\mathbf{D} : \mathbb{A}) : \mathbf{C} = \mathbf{D} : (\mathbb{A} : \mathbf{C}) \ ,} \quad \longleftarrow \text{see (2.47)} \tag{3.48}$$

for any $\mathbf{C}, \mathbf{D} \in \mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$.

---

[1] The proof is not difficult. For instance, the first property can be shown as follows:

$$\left.\begin{aligned} (\mathbb{A} + \mathbb{B}) : \mathbf{C} &\xlongequal[\text{(3.40a)}]{\text{from}} \mathbb{A} : \mathbf{C} + \mathbb{B} : \mathbf{C} \\ &\xlongequal[\text{(2.11a)}]{\text{from}} \mathbb{B} : \mathbf{C} + \mathbb{A} : \mathbf{C} \\ &\xlongequal[\text{(3.40a)}]{\text{from}} (\mathbb{B} + \mathbb{A}) : \mathbf{C} \end{aligned}\right\} \xRightarrow[\text{(3.38)}]{\text{from}} \quad \mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A} \ .$$

And, in a similar manner, one can verify the other properties.

3.2.1 Construction and Representation of Fourth-Order Tensors

Guided by (3.5) and consistent with (3.33), fourth-order tensors here are generated by using second-order tensors. Let $\mathbf{a} \otimes \mathbf{b}$ and $\mathbf{c} \otimes \mathbf{d}$ be two arbitrary tensors that are supposed to deliver a fourth-order tensor. The construction can be done in different ways. Some useful tensor products are listed in the following:

$$(\mathbf{a} \otimes \mathbf{b}) \otimes (\mathbf{c} \otimes \mathbf{d}) = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{d} \ , \tag{3.49a}$$

$$(\mathbf{a} \otimes \mathbf{b}) \boxtimes (\mathbf{c} \otimes \mathbf{d}) = \mathbf{a} \otimes \mathbf{d} \otimes \mathbf{b} \otimes \mathbf{c} \ , \tag{3.49b}$$

$$(\mathbf{a} \otimes \mathbf{b}) \boxplus (\mathbf{c} \otimes \mathbf{d}) = \mathbf{a} \otimes \mathbf{d} \otimes \mathbf{c} \otimes \mathbf{b} \ , \tag{3.49c}$$

$$\begin{aligned}(\mathbf{a} \otimes \mathbf{b}) \odot (\mathbf{c} \otimes \mathbf{d}) &= \frac{1}{2} (\mathbf{a} \otimes \mathbf{b}) \boxtimes (\mathbf{c} \otimes \mathbf{d}) + \frac{1}{2} (\mathbf{a} \otimes \mathbf{b}) \boxplus (\mathbf{c} \otimes \mathbf{d}) \\ &= \frac{1}{2} (\mathbf{a} \otimes \mathbf{d} \otimes \mathbf{b} \otimes \mathbf{c} + \mathbf{a} \otimes \mathbf{d} \otimes \mathbf{c} \otimes \mathbf{b}) \ .\end{aligned} \tag{3.49d}$$

See Itskov [6, 7] and Del Piero [8] for more elaboration on fourth-order tensors. In alignment with the right mapping (3.33), any of the linear transformations introduced in (3.49a) to (3.49d), say $\mathbb{A} = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{d}$, maps any second-order tensor $\mathbf{C} = \mathbf{u} \otimes \mathbf{v}$ onto a scalar multiple of $\mathbf{a} \otimes \mathbf{b}$ via the rule

$$\begin{aligned}\mathbf{A} &= \mathbb{A} : \mathbf{C} \\ &= (\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{d}) : (\mathbf{u} \otimes \mathbf{v}) \\ &= (\mathbf{a} \otimes \mathbf{b}) [(\mathbf{c} \otimes \mathbf{d}) : (\mathbf{u} \otimes \mathbf{v})] \\ &= [(\mathbf{c} \cdot \mathbf{u}) (\mathbf{d} \cdot \mathbf{v})] (\mathbf{a} \otimes \mathbf{b}) \ .\end{aligned} \tag{3.50}$$

A second-order tensor of the form $\mathbb{A} : \mathbf{C}$ naturally arises, for instance, in the linearization of principle of virtual work within the context of mechanics of deformable bodies, see de Souza Neto et al. [9]. Similarly to (3.50), the left mapping (3.45) now obeys

$$\begin{aligned}\mathbf{B} &= \mathbf{C} : \mathbb{A} \\ &= (\mathbf{u} \otimes \mathbf{v}) : (\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{d}) \\ &= [(\mathbf{u} \otimes \mathbf{v}) : (\mathbf{a} \otimes \mathbf{b})] (\mathbf{c} \otimes \mathbf{d}) \\ &= [(\mathbf{u} \cdot \mathbf{a}) (\mathbf{v} \cdot \mathbf{b})] (\mathbf{c} \otimes \mathbf{d}) \ .\end{aligned} \tag{3.51}$$

And obviously $\mathbf{A} \neq \mathbf{B}$ unless $\mathbb{A}$ possesses a special property, see $(3.111)_1$. Guided by these rules, the linearity conditions (3.34) and (3.46) are immediately satisfied owing to

$$
\begin{aligned}
(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}):(\alpha\mathbf{u}\otimes\mathbf{v}+\beta\mathbf{w}\otimes\bar{\mathbf{w}}) &\xlongequal[(3.50)]{\text{from}} \left[(\mathbf{c}\otimes\mathbf{d}):(\alpha\mathbf{u}\otimes\mathbf{v}+\beta\mathbf{w}\otimes\bar{\mathbf{w}})\right](\mathbf{a}\otimes\mathbf{b}) \\
&\xlongequal[(2.74a)\text{ and }(2.74b)]{\text{from}} \left[\alpha\,(\mathbf{c}\otimes\mathbf{d}):(\mathbf{u}\otimes\mathbf{v})\right. \\
&\quad \left.+\beta\,(\mathbf{c}\otimes\mathbf{d}):(\mathbf{w}\otimes\bar{\mathbf{w}})\right](\mathbf{a}\otimes\mathbf{b}) \\
&\xlongequal[(2.11f)\text{–}(2.11g)\text{ and }(2.73)]{\text{from}} \alpha\left[(\mathbf{c}\cdot\mathbf{u})\,(\mathbf{d}\cdot\mathbf{v})\right](\mathbf{a}\otimes\mathbf{b}) \\
&\quad +\beta\left[(\mathbf{c}\cdot\mathbf{w})\,(\mathbf{d}\cdot\bar{\mathbf{w}})\right](\mathbf{a}\otimes\mathbf{b}) \\
&\xlongequal[(3.50)]{\text{from}} \alpha\,(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}):(\mathbf{u}\otimes\mathbf{v}) \\
&\quad +\beta\,(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}):(\mathbf{w}\otimes\bar{\mathbf{w}}) \ , \qquad (3.52a)
\end{aligned}
$$

$$
\begin{aligned}
(\alpha\mathbf{u}\otimes\mathbf{v}+\beta\mathbf{w}\otimes\bar{\mathbf{w}}):(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}) &\xlongequal[(3.51)]{\text{from}} \left[(\alpha\mathbf{u}\otimes\mathbf{v}+\beta\mathbf{w}\otimes\bar{\mathbf{w}}):(\mathbf{a}\otimes\mathbf{b})\right](\mathbf{c}\otimes\mathbf{d}) \\
&\xlongequal[(2.74b)]{\text{from}} \left[\alpha\,(\mathbf{u}\otimes\mathbf{v}):(\mathbf{a}\otimes\mathbf{b})\right. \\
&\quad \left.+\beta\,(\mathbf{w}\otimes\bar{\mathbf{w}}):(\mathbf{a}\otimes\mathbf{b})\right](\mathbf{c}\otimes\mathbf{d}) \\
&\xlongequal[(2.11f)\text{–}(2.11g)\text{ and }(2.73)]{\text{from}} \alpha\left[(\mathbf{u}\cdot\mathbf{a})\,(\mathbf{v}\cdot\mathbf{b})\right](\mathbf{c}\otimes\mathbf{d}) \\
&\quad +\beta\left[(\mathbf{w}\cdot\mathbf{a})\,(\bar{\mathbf{w}}\cdot\mathbf{b})\right](\mathbf{c}\otimes\mathbf{d}) \\
&\xlongequal[(3.51)]{\text{from}} \alpha\,(\mathbf{u}\otimes\mathbf{v}):(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}) \\
&\quad +\beta\,(\mathbf{w}\otimes\bar{\mathbf{w}}):(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}) \ . \qquad (3.52b)
\end{aligned}
$$

Similar procedures then reveal

$$
\begin{aligned}
(\alpha\mathbf{a}\otimes\mathbf{b}+\beta\mathbf{c}\otimes\mathbf{d})\otimes\mathbf{u}\otimes\mathbf{v} &= \alpha\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{u}\otimes\mathbf{v}+\beta\mathbf{c}\otimes\mathbf{d}\otimes\mathbf{u}\otimes\mathbf{v} \ , && (3.53a)\\
\mathbf{a}\otimes\mathbf{b}\otimes(\alpha\mathbf{c}\otimes\mathbf{d}+\beta\mathbf{u}\otimes\mathbf{v}) &= \alpha\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}\otimes\mathbf{d}+\beta\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{u}\otimes\mathbf{v} \ . && (3.53b)
\end{aligned}
$$

The rules (3.49a) to (3.49d) help represent the following identities

$$
\begin{aligned}
&(\mathbf{A}\otimes\mathbf{B}):\mathbf{C}=(\mathbf{B}:\mathbf{C})\,\mathbf{A} \quad , \quad \mathbf{C}:(\mathbf{A}\otimes\mathbf{B})=(\mathbf{C}:\mathbf{A})\,\mathbf{B} \ , && (3.54a)\\
&(\mathbf{A}\boxtimes\mathbf{B}):\mathbf{C}=\mathbf{A}\mathbf{C}\mathbf{B} \quad , \quad \mathbf{C}:(\mathbf{A}\boxtimes\mathbf{B})=\mathbf{A}^{\mathrm{T}}\mathbf{C}\mathbf{B}^{\mathrm{T}} \ , && (3.54b)\\
&(\mathbf{A}\boxplus\mathbf{B}):\mathbf{C}=\mathbf{A}\mathbf{C}^{\mathrm{T}}\mathbf{B} \quad , \quad \mathbf{C}:(\mathbf{A}\boxplus\mathbf{B})=\mathbf{B}\mathbf{C}^{\mathrm{T}}\mathbf{A} \ , && (3.54c)\\
&(\mathbf{A}\odot\mathbf{B}):\mathbf{C}=\mathbf{A}\,(\mathrm{sym}\mathbf{C})\,\mathbf{B} \quad , \quad \mathbf{C}:(\mathbf{A}\odot\mathbf{B})=\mathrm{sym}\left(\mathbf{A}^{\mathrm{T}}\mathbf{C}\mathbf{B}^{\mathrm{T}}\right) \ . && (3.54d)
\end{aligned}
$$

The proof is not difficult. For instance, consider $(3.54c)_2$. To show this result, without loss of generality, suppose that

$$
\mathbf{A}=\mathbf{a}\otimes\mathbf{b} \quad , \quad \mathbf{B}=\mathbf{c}\otimes\mathbf{d} \quad , \quad \mathbf{C}=\mathbf{u}\otimes\mathbf{v} \ .
$$

Then,

$$\mathbf{C} : (\mathbf{A} \boxplus \mathbf{B}) = [\mathbf{u} \otimes \mathbf{v}] : [\mathbf{a} \otimes \mathbf{d} \otimes \mathbf{c} \otimes \mathbf{b}] = (\mathbf{u} \cdot \mathbf{a}) (\mathbf{d} \cdot \mathbf{v}) \, \mathbf{c} \otimes \mathbf{b}$$
$$= (\mathbf{u} \cdot \mathbf{a}) (\mathbf{c} \otimes \mathbf{d}) (\mathbf{v} \otimes \mathbf{b}) = (\mathbf{c} \otimes \mathbf{d}) (\mathbf{v} \otimes \mathbf{u}) (\mathbf{a} \otimes \mathbf{b})$$
$$= (\mathbf{c} \otimes \mathbf{d}) (\mathbf{u} \otimes \mathbf{v})^{\mathrm{T}} (\mathbf{a} \otimes \mathbf{b}) \; . \quad \nwarrow \text{ see also (3.94a)–(3.94i)}$$

Some useful identities can be resulted from (3.54a)–(3.54d). For instance, let $\mathbf{C} = \mathbf{u} \otimes \mathbf{v}$. Then, by means of (2.32), (2.48), (2.54) and (2.79c), one can establish

$$(\mathbf{A} \otimes \mathbf{B}) : (\mathbf{u} \otimes \mathbf{v}) = (\mathbf{u} \cdot \mathbf{B}\mathbf{v}) \, \mathbf{A} \; , \tag{3.55a}$$
$$(\mathbf{u} \otimes \mathbf{v}) : (\mathbf{A} \otimes \mathbf{B}) = (\mathbf{u} \cdot \mathbf{A}\mathbf{v}) \, \mathbf{B} \; , \tag{3.55b}$$
$$(\mathbf{A} \odot \mathbf{B}) : (\mathbf{u} \otimes \mathbf{v}) = \frac{1}{2} \mathbf{A}\mathbf{u} \otimes \mathbf{B}^{\mathrm{T}}\mathbf{v} + \frac{1}{2} \mathbf{A}\mathbf{v} \otimes \mathbf{B}^{\mathrm{T}}\mathbf{u} \; , \tag{3.55c}$$
$$(\mathbf{u} \otimes \mathbf{v}) : (\mathbf{A} \odot \mathbf{B}) = \frac{1}{2} \mathbf{A}^{\mathrm{T}}\mathbf{u} \otimes \mathbf{B}\mathbf{v} + \frac{1}{2} \mathbf{B}\mathbf{v} \otimes \mathbf{A}^{\mathrm{T}}\mathbf{u} \; . \tag{3.55d}$$

Moreover, by choosing $\mathbf{A} = \mathbf{B} = \mathbf{I}$, one immediately obtains

$$(\mathbf{I} \otimes \mathbf{I}) : \mathbf{C} = (\mathrm{tr}\mathbf{C}) \, \mathbf{I} = \mathbf{C} : (\mathbf{I} \otimes \mathbf{I}) \; , \tag{3.56a}$$
$$(\mathbf{I} \boxtimes \mathbf{I}) : \mathbf{C} = \mathbf{C} = \mathbf{C} : (\mathbf{I} \boxtimes \mathbf{I}) \; , \tag{3.56b}$$
$$(\mathbf{I} \boxplus \mathbf{I}) : \mathbf{C} = \mathbf{C}^{\mathrm{T}} = \mathbf{C} : (\mathbf{I} \boxplus \mathbf{I}) \; , \tag{3.56c}$$
$$(\mathbf{I} \odot \mathbf{I}) : \mathbf{C} = \mathrm{sym}\mathbf{C} = \mathbf{C} : (\mathbf{I} \odot \mathbf{I}) \; . \tag{3.56d}$$

Comparing $(3.56b)_1$ and $(3.56c)_1$ with $(3.37)_{1-2}$, taking into account (3.38), now reveals that

$$\boxed{\mathbf{I} \boxtimes \mathbf{I} = \mathbb{I} \quad , \quad \mathbf{I} \boxplus \mathbf{I} = \bar{\mathbb{I}} \, .} \tag{3.57}$$

Having in mind these results, the following special fourth-order tensors are introduced for convenience:

$$\mathbb{P}_{\mathrm{sym}} = \mathbf{I} \odot \mathbf{I} = \frac{1}{2} (\mathbf{I} \boxtimes \mathbf{I} + \mathbf{I} \boxplus \mathbf{I})$$
$$= \frac{1}{2} \left(\mathbb{I} + \bar{\mathbb{I}} \right) \; , \quad \leftarrow \begin{array}{c}\text{super-symmetric projection tensor}\\ \text{(or symmetrizer operator)}\end{array} \tag{3.58a}$$
$$\mathbb{P}_{\mathrm{skw}} = \frac{1}{2} (\mathbf{I} \boxtimes \mathbf{I} - \mathbf{I} \boxplus \mathbf{I}) = \frac{1}{2} \left(\mathbb{I} - \bar{\mathbb{I}} \right) \; , \quad \leftarrow \begin{array}{c}\text{skew-symmetric projection tensor}\\ \text{(or skew-symmetrizer operator)}\end{array} \tag{3.58b}$$
$$\mathbb{P}_{\mathrm{sph}} = \frac{1}{3} \mathbf{I} \otimes \mathbf{I} \; , \quad \leftarrow \begin{array}{c}\text{spherical projection tensor}\\ \text{(or spherical operator)}\end{array} \tag{3.58c}$$
$$\mathbb{P}_{\mathrm{dev}} = \mathbf{I} \boxtimes \mathbf{I} - \frac{1}{3} \mathbf{I} \otimes \mathbf{I} = \mathbb{I} - \mathbb{P}_{\mathrm{sph}} \; . \quad \leftarrow \begin{array}{c}\text{deviatoric projection tensor}\\ \text{(or deviatoric operator)}\end{array} \tag{3.58d}$$

As implied by their names, they are designed to deliver the symmetric, skew-symmetric, spherical and deviatoric parts of an arbitrary second-order tensor:

$$\mathbb{P}_{\text{sym}} : \mathbf{C} = \frac{1}{2}\left(\mathbf{C} + \mathbf{C}^{\text{T}}\right) = \text{sym}\mathbf{C} \ , \tag{3.59a}$$

$$\mathbb{P}_{\text{skw}} : \mathbf{C} = \frac{1}{2}\left(\mathbf{C} - \mathbf{C}^{\text{T}}\right) = \text{skw}\mathbf{C} \ , \tag{3.59b}$$

$$\mathbb{P}_{\text{sph}} : \mathbf{C} = \frac{1}{3}\left(\text{tr}\mathbf{C}\right)\mathbf{I} = \text{sph}\mathbf{C} \ , \tag{3.59c}$$

$$\mathbb{P}_{\text{dev}} : \mathbf{C} = \mathbf{C} - \frac{1}{3}\left(\text{tr}\mathbf{C}\right)\mathbf{I} = \text{dev}\mathbf{C} \ . \tag{3.59d}$$

Regarding symmetric tensors, although these projection tensors remain unchanged but it may be advantageous to rephrase (3.58d) as

$$\mathbb{P}^{\text{s}}_{\text{dev}} = \mathbf{I} \odot \mathbf{I} - \frac{1}{3}\mathbf{I} \otimes \mathbf{I} = \mathbb{P}_{\text{sym}} - \mathbb{P}_{\text{sph}} \ . \longleftarrow \begin{array}{c}\text{super-symmetric deviatoric projection tensor}\\ \text{(or super-symmetric deviatoric operator)}\end{array} \tag{3.60}$$

Hint: The expressions (3.58a) to (3.58d) clearly show that the fourth-order unit tensor $\mathbb{I}$ has additively been decomposed not only to the super-symmetric part $\mathbb{P}_{\text{sym}}$ and skew-symmetric portion $\mathbb{P}_{\text{skw}}$, but also to the spherical part $\mathbb{P}_{\text{sph}}$ and deviatoric portion $\mathbb{P}_{\text{dev}}$.

The Cartesian collection

$$\left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l, \ i, j, k, l = 1, 2, 3\right\} \stackrel{\text{def}}{=\!=} \left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right\} \ , \tag{3.61}$$

is now introduced to denote a basis for $\mathscr{T}_{\text{fo}}\left(\mathscr{E}^{\text{o3}}_{\text{r}}\right)$ with $\dim \mathscr{T}_{\text{fo}} = \left(\dim \mathscr{E}^{\text{o3}}_{\text{r}}\right)^4 = 3^4$. This enables one to express any Cartesian element $\mathbb{A}$ of this new vector space as

$$\boxed{\mathbb{A} = \mathbb{A}_{ijkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \ ,} \quad \longleftarrow \text{see (3.13)} \tag{3.62}$$

where $\mathbb{A}_{ijkl}$ present the **Cartesian components** of $\mathbb{A}$. These 81 scalar numbers are determined by

$$\boxed{\mathbb{A}_{ijkl} = \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\mathbb{A} : \left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right)\right] \ ,} \quad \longleftarrow \text{see (3.14)} \tag{3.63}$$

where

$$\boxed{\mathbb{A} : \left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right) = \mathbb{A}_{mnkl}\widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n \ ,} \quad \longleftarrow \text{see (3.15)} \tag{3.64}$$

owing to

$$\begin{aligned}\mathbb{A}_{ijkl} &\xlongequal[(1.36)]{\text{from}} \mathbb{A}_{mnkl}\left\{\delta_{im}\delta_{jn}\right\}\\ &\xlongequal[(1.35)]{\text{from}} \mathbb{A}_{mnkl}\left\{\left(\widehat{\mathbf{e}}_i\cdot\widehat{\mathbf{e}}_m\right)\left(\widehat{\mathbf{e}}_j\cdot\widehat{\mathbf{e}}_n\right)\right\}\\ &\xlongequal[(2.73),\ (2.74a)\ \text{and}\ (2.74b)]{\text{from}} \left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right):\left\{\mathbb{A}_{mnkl}\left(\widehat{\mathbf{e}}_m\otimes\widehat{\mathbf{e}}_n\right)\right\}\\ &\xlongequal[(1.35)\ \text{and}\ (1.36)]{\text{from}} \left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right):\left\{\left[\mathbb{A}_{mnop}\left(\widehat{\mathbf{e}}_m\otimes\widehat{\mathbf{e}}_n\right)\right]\left(\widehat{\mathbf{e}}_o\cdot\widehat{\mathbf{e}}_k\right)\left(\widehat{\mathbf{e}}_p\cdot\widehat{\mathbf{e}}_l\right)\right\}\\ &\xlongequal[(3.40a),\ (3.40b)\ \text{and}\ (3.50)]{\text{from}} \left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right):\left[\mathbb{A}_{mnop}\left(\widehat{\mathbf{e}}_m\otimes\widehat{\mathbf{e}}_n\otimes\widehat{\mathbf{e}}_o\otimes\widehat{\mathbf{e}}_p\right)\right]:\left(\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right)\\ &\xlongequal[(3.62)]{\text{from}} \left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right):\mathbb{A}:\left(\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right)\ .\end{aligned}$$

Guided by (3.64), the following identity also holds true

$$\boxed{\left(\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right):\mathbb{A}=\mathbb{A}_{klmn}\widehat{\mathbf{e}}_m\otimes\widehat{\mathbf{e}}_n\ .} \tag{3.65}$$

Let $\mathbf{C}$ and $\mathbb{A}$ be two tensorial variables of the forms $(2.19)_2$ and (3.62), respectively. Then,

$$\begin{aligned}\mathbf{A} = \mathbb{A}:\mathbf{C} &= \left[\mathbb{A}_{ijkl}\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right]:\left[C_{mn}\widehat{\mathbf{e}}_m\otimes\widehat{\mathbf{e}}_n\right]\\ &= \mathbb{A}_{ijkl}C_{mn}\delta_{km}\delta_{ln}\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\\ &= \mathbb{A}_{ijkl}C_{kl}\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\quad \text{with } A_{ij}=\mathbb{A}_{ijkl}C_{kl}\ , \end{aligned}\tag{3.66a}$$

$$\begin{aligned}\mathbf{B} = \mathbf{C}:\mathbb{A} &= \left[C_{mn}\widehat{\mathbf{e}}_m\otimes\widehat{\mathbf{e}}_n\right]:\left[\mathbb{A}_{klij}\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\otimes\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right]\\ &= C_{mn}\mathbb{A}_{klij}\delta_{mk}\delta_{nl}\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\\ &= C_{kl}\mathbb{A}_{klij}\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\quad \text{with } B_{ij}=C_{kl}\mathbb{A}_{klij}\ . \end{aligned}\tag{3.66b}$$

Accordingly, the fourth-order unit tensor $\mathbb{I}$ in $(3.37)_1$ takes the form

$$\left.\begin{aligned}\mathbb{I}:\mathbf{C} &\xlongequal[(2.20)\ \text{and}\ (3.37)]{\text{from (2.19),}} \left[\widehat{\mathbf{e}}_i\cdot\mathbf{C}\widehat{\mathbf{e}}_j\right]\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\\ &\xlongequal[(2.79c)]{\text{from}} \left[\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right):\mathbf{C}\right]\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\\ &\xlongequal[(3.50)]{\text{from}} \left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right):\mathbf{C}\end{aligned}\right\}\ \xrightarrow[(3.38)]{\text{from}}\ \boxed{\mathbb{I}=\underbrace{\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j}_{=\,\delta_{ik}\delta_{lj}\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l}\ ,} \tag{3.67}$$

and, in a similar manner, $\bar{\mathbb{I}}$ admits the coordinate representation

$$\left.\begin{aligned}\bar{\mathbb{I}} : \mathbf{C} &\xlongequal[\text{(2.50) and (3.37)}]{\text{from (2.20),}} \left[\widehat{\mathbf{e}}_i \cdot \mathbf{C}\widehat{\mathbf{e}}_j\right]\left(\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i\right) \\ &\xlongequal[\text{(2.79c)}]{\text{from}} \left[\left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \mathbf{C}\right]\left(\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i\right) \\ &\xlongequal[\text{(3.50)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i\right) : \mathbf{C}\end{aligned}\right\} \xrightarrow[\text{(3.38)}]{\text{from}} \boxed{\bar{\mathbb{I}} = \underbrace{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i}_{= \delta_{il}\delta_{kj}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l} .} \tag{3.68}$$

From these results, one can obtain the following identities

$$\mathbf{C} = \mathbf{C} : \mathbb{I} \, , \quad \longleftarrow \text{since } C_{kl} = C_{ij} \left(\mathbb{I}\right)_{ijkl} = C_{ij}\delta_{ik}\delta_{lj} \tag{3.69a}$$

$$\mathbf{C}^{\mathrm{T}} = \mathbf{C} : \bar{\mathbb{I}} \, , \quad \longleftarrow \text{since } \left(\mathbf{C}^{\mathrm{T}}\right)_{kl} = C_{ij} \left(\bar{\mathbb{I}}\right)_{ijkl} = C_{ij}\delta_{il}\delta_{kj} \tag{3.69b}$$

In coordinate representation, the four different fourth-order tensors used in (3.54a)-(3.54d) finally render

$$\mathbf{A} \otimes \mathbf{B} = A_{ij} B_{kl} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \, , \tag{3.70a}$$

$$\mathbf{A} \boxtimes \mathbf{B} = A_{ik} B_{lj} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \, , \tag{3.70b}$$

$$\mathbf{A} \boxplus \mathbf{B} = A_{il} B_{kj} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \, , \tag{3.70c}$$

$$\begin{aligned}\mathbf{A} \odot \mathbf{B} &= \frac{1}{2}\left[\mathbf{A} \boxtimes \mathbf{B} + \mathbf{A} \boxplus \mathbf{B}\right] \\ &= \frac{1}{2}\left[A_{ik} B_{lj} + A_{il} B_{kj}\right] \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \, .\end{aligned} \tag{3.70d}$$

Isotropic tensors of order four. Under a change of coordinates, the old and new Cartesian components of a fourth-order tensor transform as

$$\mathbb{A}^n_{ijkl} = R_{mi} R_{nj} R_{ok} R_{pl} \mathbb{A}^o_{mnop} \, . \tag{3.71}$$

This tensor is said to be **isotropic** if

$$\mathbb{A}_{ijkl} = Q_{mi} Q_{nj} Q_{ok} Q_{pl} \mathbb{A}_{mnop} \, , \tag{3.72}$$

holds true for any arbitrary rotation or reflection of a coordinate frame. This condition helps represent $\mathbb{A} = \mathbb{A}_{ijkl} \widehat{\mathbf{e}}^o_i \otimes \widehat{\mathbf{e}}^o_j \otimes \widehat{\mathbf{e}}^o_k \otimes \widehat{\mathbf{e}}^o_l = \mathbb{A}_{ijkl} \widehat{\mathbf{e}}^n_i \otimes \widehat{\mathbf{e}}^n_j \otimes \widehat{\mathbf{e}}^n_k \otimes \widehat{\mathbf{e}}^n_l$. In the following, it will be shown that this remains true if $\mathbb{A}_{ijkl}$ is a linear combination of $\delta_{ij}\delta_{kl}$, $\delta_{ik}\delta_{jl}$ and $\delta_{il}\delta_{jk}$. Thus, the most general form of an isotropic tensor of rank four represents

$$\boxed{\mathbb{A} = \alpha \mathbf{I} \otimes \mathbf{I} + \beta \mathbf{I} \boxtimes \mathbf{I} + \gamma \mathbf{I} \boxplus \mathbf{I} \quad \text{where } \alpha, \beta, \gamma \text{ are constants} \, .} \quad \longleftarrow \text{see (2.171)} \tag{3.73}$$

Proof. By following similar procedures which led to (3.27), one will have

$$\varepsilon_{mir}\mathbb{A}_{mjkl} + \varepsilon_{njr}\mathbb{A}_{inkl} + \varepsilon_{okr}\mathbb{A}_{ijol} + \varepsilon_{plr}\mathbb{A}_{ijkp} = 0 \ . \tag{3.74}$$

By multiplying both sides of this expression by $\varepsilon_{sir}, \varepsilon_{sjr}, \varepsilon_{skr}$ and ε_{slr} and then setting $s = i, s = j, s = k$ and $s = l$ in the resulting expressions, one can arrive at

$$2\mathbb{A}_{ijkl} + \mathbb{A}_{jikl} + \mathbb{A}_{kjil} + \mathbb{A}_{ljki} = \delta_{ji}\mathbb{A}_{ttkl} + \delta_{ki}\mathbb{A}_{tjtl} + \delta_{li}\mathbb{A}_{tjkt} \ , \tag{3.75a}$$
$$2\mathbb{A}_{ijkl} + \mathbb{A}_{jikl} + \mathbb{A}_{ikjl} + \mathbb{A}_{ilkj} = \delta_{ij}\mathbb{A}_{ttkl} + \delta_{lj}\mathbb{A}_{itkt} + \delta_{kj}\mathbb{A}_{ittl} \ , \tag{3.75b}$$
$$2\mathbb{A}_{ijkl} + \mathbb{A}_{kjil} + \mathbb{A}_{ikjl} + \mathbb{A}_{ijlk} = \delta_{lk}\mathbb{A}_{ijtt} + \delta_{ik}\mathbb{A}_{tjtl} + \delta_{jk}\mathbb{A}_{ittl} \ , \tag{3.75c}$$
$$2\mathbb{A}_{ijkl} + \mathbb{A}_{ljki} + \mathbb{A}_{ilkj} + \mathbb{A}_{ijlk} = \delta_{kl}\mathbb{A}_{ijtt} + \delta_{jl}\mathbb{A}_{itkt} + \delta_{il}\mathbb{A}_{tjkt} \ . \tag{3.75d}$$

Recall that the only isotropic vector was a zero vector. Guided by (2.178) and (3.33), one can then infer that an isotropic fourth-order tensor should be constructed from two identity tensors. Thus, having in mind (3.70a)–(3.70c), it makes sense to assume that

$$\mathbb{A}_{ttij} = \mathbb{A}_{ijtt} \quad , \quad \mathbb{A}_{ittj} = \mathbb{A}_{tijt} \quad , \quad \mathbb{A}_{itjt} = \mathbb{A}_{titj} \ . \tag{3.76}$$

One can further write

$$\mathbb{A}_{ttij} = \bar{\alpha}\delta_{ij} \quad , \quad \mathbb{A}_{itjt} = \bar{\beta}\delta_{ij} \quad , \quad \mathbb{A}_{ittj} = \bar{\gamma}\delta_{ij} \ . \tag{3.77}$$

Consequently,

$$2\mathbb{A}_{ijkl} + \mathbb{A}_{jikl} + \mathbb{A}_{kjil} + \mathbb{A}_{ljki} = \bar{\alpha}\delta_{ji}\delta_{kl} + \bar{\beta}\delta_{ki}\delta_{jl} + \bar{\gamma}\delta_{li}\delta_{jk} \ , \tag{3.78a}$$
$$2\mathbb{A}_{ijkl} + \mathbb{A}_{jikl} + \mathbb{A}_{ikjl} + \mathbb{A}_{ilkj} = \bar{\alpha}\delta_{ij}\delta_{kl} + \bar{\beta}\delta_{lj}\delta_{ik} + \bar{\gamma}\delta_{kj}\delta_{il} \ , \tag{3.78b}$$
$$2\mathbb{A}_{ijkl} + \mathbb{A}_{kjil} + \mathbb{A}_{ikjl} + \mathbb{A}_{ijlk} = \bar{\alpha}\delta_{lk}\delta_{ij} + \bar{\beta}\delta_{ik}\delta_{jl} + \bar{\gamma}\delta_{jk}\delta_{il} \ , \tag{3.78c}$$
$$2\mathbb{A}_{ijkl} + \mathbb{A}_{ljki} + \mathbb{A}_{ilkj} + \mathbb{A}_{ijlk} = \bar{\alpha}\delta_{kl}\delta_{ij} + \bar{\beta}\delta_{jl}\delta_{ik} + \bar{\gamma}\delta_{il}\delta_{jk} \ . \tag{3.78d}$$

By subtracting the sum of one pair of the above expressions from the sum of another pair, one can deduce that

$$\mathbb{A}_{ijkl} = \mathbb{A}_{jilk} = \mathbb{A}_{klij} = \mathbb{A}_{lkji} \ . \tag{3.79}$$

It is then easy to see that

$$2\mathbb{A}_{ijkl} + \mathbb{A}_{ijlk} + \mathbb{A}_{ilkj} + \mathbb{A}_{ikjl} = \bar{\alpha}\delta_{ij}\delta_{kl} + \bar{\beta}\delta_{ik}\delta_{jl} + \bar{\gamma}\delta_{il}\delta_{jk} \ , \tag{3.80a}$$
$$2\mathbb{A}_{iklj} + \mathbb{A}_{ikjl} + \mathbb{A}_{ilkj} + \mathbb{A}_{ijlk} = \bar{\alpha}\delta_{ik}\delta_{jl} + \bar{\beta}\delta_{il}\delta_{jk} + \bar{\gamma}\delta_{ij}\delta_{kl} \ , \tag{3.80b}$$
$$2\mathbb{A}_{iljk} + \mathbb{A}_{ilkj} + \mathbb{A}_{ijlk} + \mathbb{A}_{ikjl} = \bar{\alpha}\delta_{il}\delta_{jk} + \bar{\beta}\delta_{ij}\delta_{kl} + \bar{\gamma}\delta_{ik}\delta_{jl} \ . \tag{3.80c}$$

It follows that

$$2\left(\mathbb{A}_{ijkl}+\mathbb{A}_{iklj}+\mathbb{A}_{iljk}\right)+3\left(\mathbb{A}_{ijlk}+\mathbb{A}_{ilkj}+\mathbb{A}_{ikjl}\right) \\ =\left(\bar{\alpha}+\bar{\beta}+\bar{\gamma}\right)\left(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}\right) \ . \tag{3.81}$$

From this relation, one can obtain

$$\begin{aligned} 6\left(\mathbb{A}_{ijkl}+\mathbb{A}_{iklj}+\mathbb{A}_{iljk}\right)+9\left(\mathbb{A}_{ijlk}+\mathbb{A}_{ilkj}+\mathbb{A}_{ikjl}\right) \\ =3\left(\bar{\alpha}+\bar{\beta}+\bar{\gamma}\right)\left(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}\right) \ , \\ 4\left(\mathbb{A}_{ikjl}+\mathbb{A}_{ijlk}+\mathbb{A}_{ilkj}\right)+6\left(\mathbb{A}_{iklj}+\mathbb{A}_{iljk}+\mathbb{A}_{ijkl}\right) \\ =2\left(\bar{\alpha}+\bar{\beta}+\bar{\gamma}\right)\left(\delta_{ik}\delta_{jl}+\delta_{ij}\delta_{kl}+\delta_{il}\delta_{jk}\right) \ . \end{aligned}$$

Thus,

$$\mathbb{A}_{ijlk}+\mathbb{A}_{ilkj}+\mathbb{A}_{ikjl}=\mathbb{A}_{ijkl}+\mathbb{A}_{iklj}+\mathbb{A}_{iljk} \\ =\frac{1}{5}\left(\bar{\alpha}+\bar{\beta}+\bar{\gamma}\right)\left(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}\right) \ . \tag{3.82}$$

Finally, by substituting (3.82)$_2$ into (3.80a), one can present

$$\mathbb{A}_{ijkl}=\alpha\delta_{ij}\delta_{kl}+\beta\delta_{ik}\delta_{jl}+\gamma\delta_{il}\delta_{jk} \ , \tag{3.83}$$

where

$$\alpha=\frac{4\bar{\alpha}-\bar{\beta}-\bar{\gamma}}{10} \quad , \quad \beta=\frac{4\bar{\beta}-\bar{\alpha}-\bar{\gamma}}{10} \quad , \quad \gamma=\frac{4\bar{\gamma}-\bar{\alpha}-\bar{\beta}}{10} \ . \tag{3.84}$$

3.2.2 Operations with Fourth-Order Tensors

3.2.2.1 Composition

Recall the composition of two second-order tensors introduced in (2.25). The goal is now to extend this operation along with the corresponding relations for the problem at hand. The *composition* (or *double contraction*) of two fourth-order tensors $\mathbb{A}$ and $\mathbb{B}$, designated by $\mathbb{A}:\mathbb{B}$, is again a fourth-order tensor satisfying

$$\boxed{(\mathbb{A}:\mathbb{B}):\mathbf{C}=\mathbb{A}:(\mathbb{B}:\mathbf{C}) \ , \quad \text{for all } \mathbf{C}\in\mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{o3}\right) \ .} \tag{3.85}$$

A general fourth-order tensor of the form $\mathbb{A}:\mathbb{B}$ can be seen, for instance, in material elasticity tensor of compressible hyperelastic materials, see Holzapfel [10]. And, in general, it does not have the commutative property, i.e. $\mathbb{A} : \mathbb{B} \neq \mathbb{B} : \mathbb{A}$. Making use of (3.38), (3.66a)$_5$ and (3.85), its indicial form renders

$$(\mathbb{A} : \mathbb{B})_{ijkl}\, C_{kl} = \mathbb{A}_{ijmn} \underbrace{(\mathbb{B} : \mathbf{C})_{mn}}_{= \mathbb{B}_{mnkl} C_{kl}} \implies \boxed{(\mathbb{A} : \mathbb{B})_{ijkl} = \mathbb{A}_{ijmn} \mathbb{B}_{mnkl}\ .} \tag{3.86}$$

The interested reader may want to arrive at this result in an alternative way. This relies on

$$\begin{aligned}(\mathbb{A} : \mathbb{B})_{ijkl} &\xlongequal[\text{(3.63)}]{\text{from}} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j : [(\mathbb{A} : \mathbb{B}) : \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l] \\ &\xlongequal[\text{(3.85)}]{\text{from}} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j : [\mathbb{A} : (\mathbb{B} : \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l)] \\ &\xlongequal[\text{(3.64)}]{\text{from}} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j : [\mathbb{A} : (\mathbb{B}_{mnkl} \widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n)] \\ &\xlongequal[\text{(3.34) and (3.64)}]{\text{from}} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j : \left[\mathbb{A}_{opmn} \mathbb{B}_{mnkl} \widehat{\mathbf{e}}_o \otimes \widehat{\mathbf{e}}_p\right] \\ &\xlongequal[\text{(2.73), (2.74a) and (2.74b)}]{\text{from}} \mathbb{A}_{opmn} \mathbb{B}_{mnkl} \left(\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_o\right) \left(\widehat{\mathbf{e}}_j \cdot \widehat{\mathbf{e}}_p\right) \\ &\xlongequal[\text{(1.35) and (1.36)}]{\text{from}} \mathbb{A}_{ijmn} \mathbb{B}_{mnkl}\ .\end{aligned}$$

Similar procedures can be followed to arrive at

$$\underbrace{\mathbb{A} : \mathbb{B} : \mathbb{C} = (\mathbb{A} : \mathbb{B}) : \mathbb{C} = \mathbb{A} : (\mathbb{B} : \mathbb{C})}_{\text{in index notation}:\ (\mathbb{A} : \mathbb{B} : \mathbb{C})_{ijkl} = \mathbb{A}_{ijmn} \mathbb{B}_{mnop} \mathbb{C}_{opkl}}\ . \quad \longleftarrow \ \begin{array}{l}\text{and also } (\mathbb{A} : \mathbb{B} : \mathbb{C} : \mathbb{D})_{ijkl} \\ = \mathbb{A}_{ijmn} \mathbb{B}_{mnop} \mathbb{C}_{opqr} \mathbb{D}_{qrkl}\end{array} \tag{3.87}$$

In coordinate representation, the Cartesian fourth-order tensors $\mathbb{A} : \mathbb{B}$ and $\mathbb{A} : \mathbb{B} : \mathbb{C}$ thus render

$$\mathbb{A} : \mathbb{B} = \mathbb{A}_{ijmn} \mathbb{B}_{mnkl}\, \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\ , \tag{3.88a}$$

$$\mathbb{A} : \mathbb{B} : \mathbb{C} = \mathbb{A}_{ijmn} \mathbb{B}_{mnop} \mathbb{C}_{opkl}\, \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\ . \tag{3.88b}$$

In analogy with tensors, its is a simple exercise to show that the composition of two fourth-order tensors is a bilinear form (not the symmetric one since $\mathbb{A} : \mathbb{B} \neq \mathbb{B} : \mathbb{A}$), that is,

$$\mathbb{A} : (\alpha\mathbb{B} + \beta\mathbb{C}) = \alpha\mathbb{A} : \mathbb{B} + \beta\mathbb{A} : \mathbb{C}\ , \tag{3.89a}$$

$$(\alpha\mathbb{A} + \beta\mathbb{B}) : \mathbb{C} = \alpha\mathbb{A} : \mathbb{C} + \beta\mathbb{B} : \mathbb{C}\ . \tag{3.89b}$$

One can now establish the rule

$$\boxed{(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{d}) : (\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w} \otimes \bar{\mathbf{w}}) = (\mathbf{c} \cdot \mathbf{u})\, (\mathbf{d} \cdot \mathbf{v})\, \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{w} \otimes \bar{\mathbf{w}}\ .} \tag{3.90}$$

The result (3.88a) can then be obtained in a more convenient form as follows:

$$\mathbb{A} : \mathbb{B} = \underbrace{\left(\mathbb{A}_{ijmn}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n\right) : \left(\mathbb{B}_{opkl}\widehat{\mathbf{e}}_o \otimes \widehat{\mathbf{e}}_p \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right)}_{= \mathbb{A}_{ijmn}\mathbb{B}_{opkl}\left(\widehat{\mathbf{e}}_m \cdot \widehat{\mathbf{e}}_o\right)\left(\widehat{\mathbf{e}}_n \cdot \widehat{\mathbf{e}}_p\right)\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l}$$
$$= \underbrace{\mathbb{A}_{ijmn}\mathbb{B}_{opkl}\left(\delta_{mo}\right)\left(\delta_{np}\right)\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l}_{= \mathbb{A}_{ijmn}\mathbb{B}_{mnkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l} \ . \tag{3.91}$$

It is then easy to deduce that

$$(\mathbf{C} \otimes \mathbf{D}) : \mathbb{A} = \mathbf{C} \otimes (\mathbf{D} : \mathbb{A}) \quad , \quad \mathbb{A} : (\mathbf{C} \otimes \mathbf{D}) = (\mathbb{A} : \mathbf{C}) \otimes \mathbf{D} \ . \tag{3.92}$$

With the aid of (2.26), (2.49), (2.55d), $(2.59)_1$, (3.53a), (3.53b) and (3.86), the tensor products introduced in (3.49a) to (3.49d) now help represent the following identities

$$(\mathbf{A} \otimes \mathbf{B}) : (\mathbf{C} \otimes \mathbf{D}) = (\mathbf{B} : \mathbf{C})\, \mathbf{A} \otimes \mathbf{D} \ , \quad \swarrow \text{ note that } 2\,(\mathbf{A} \odot \mathbf{B}) = \mathbf{A} \boxtimes \mathbf{B} + \mathbf{A} \boxplus \mathbf{B} \tag{3.93a}$$

$$(\mathbf{A} \odot \mathbf{B}) : (\mathbf{C} \odot \mathbf{D}) = \frac{1}{2}\,(\mathbf{AC}) \odot (\mathbf{DB}) + \frac{1}{2}\left(\mathbf{AD}^{\mathrm{T}}\right) \odot \left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\right) \ , \tag{3.93b}$$

$$(\mathbf{A} \odot \mathbf{B}) : (\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\,(\mathrm{sym}\mathbf{C})\,\mathbf{B}) \otimes \mathbf{D} \ , \tag{3.93c}$$

$$(\mathbf{A} \otimes \mathbf{B}) : (\mathbf{C} \odot \mathbf{D}) = \mathbf{A} \otimes \mathrm{sym}\left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\mathbf{D}^{\mathrm{T}}\right) \ , \tag{3.93d}$$

since

$$((\mathbf{A} \otimes \mathbf{B}) : (\mathbf{C} \otimes \mathbf{D}))_{ijkl} = (\mathbf{A} \otimes \mathbf{B})_{ijmn}\,(\mathbf{C} \otimes \mathbf{D})_{mnkl} = (\mathbf{A})_{ij}\,(\mathbf{B})_{mn}\,(\mathbf{C})_{mn}\,(\mathbf{D})_{kl}$$
$$= [\mathbf{B} : \mathbf{C}]\,(\mathbf{A})_{ij}\,(\mathbf{D})_{kl} = [\mathbf{B} : \mathbf{C}]\,(\mathbf{A} \otimes \mathbf{D})_{ijkl} \ , \tag{3.94a}$$

$$((\mathbf{A} \boxtimes \mathbf{B}) : (\mathbf{C} \boxtimes \mathbf{D}))_{ijkl} = (\mathbf{A} \boxtimes \mathbf{B})_{ijmn}\,(\mathbf{C} \boxtimes \mathbf{D})_{mnkl} = (\mathbf{A})_{im}\,(\mathbf{B})_{nj}\,(\mathbf{C})_{mk}\,(\mathbf{D})_{ln}$$
$$= (\mathbf{AC})_{ik}\,(\mathbf{DB})_{lj} = ((\mathbf{AC}) \boxtimes (\mathbf{DB}))_{ijkl} \ , \tag{3.94b}$$

$$((\mathbf{A} \boxtimes \mathbf{B}) : (\mathbf{C} \boxplus \mathbf{D}))_{ijkl} = (\mathbf{A} \boxtimes \mathbf{B})_{ijmn}\,(\mathbf{C} \boxplus \mathbf{D})_{mnkl} = (\mathbf{A})_{im}\,(\mathbf{B})_{nj}\,(\mathbf{C})_{ml}\,(\mathbf{D})_{kn}$$
$$= (\mathbf{AC})_{il}\,(\mathbf{DB})_{kj} = ((\mathbf{AC}) \boxplus (\mathbf{DB}))_{ijkl} \ , \tag{3.94c}$$

$$((\mathbf{A} \boxplus \mathbf{B}) : (\mathbf{C} \boxtimes \mathbf{D}))_{ijkl} = (\mathbf{A} \boxplus \mathbf{B})_{ijmn}\,(\mathbf{C} \boxtimes \mathbf{D})_{mnkl} = (\mathbf{A})_{in}\,(\mathbf{B})_{mj}\,(\mathbf{C})_{mk}\,(\mathbf{D})_{ln}$$
$$= \left(\mathbf{AD}^{\mathrm{T}}\right)_{il}\left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\right)_{kj} = \left(\left(\mathbf{AD}^{\mathrm{T}}\right) \boxplus \left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\right)\right)_{ijkl} \ , \tag{3.94d}$$

$$((\mathbf{A} \boxplus \mathbf{B}) : (\mathbf{C} \boxplus \mathbf{D}))_{ijkl} = (\mathbf{A} \boxplus \mathbf{B})_{ijmn}\,(\mathbf{C} \boxplus \mathbf{D})_{mnkl} = (\mathbf{A})_{in}\,(\mathbf{B})_{mj}\,(\mathbf{C})_{ml}\,(\mathbf{D})_{kn}$$
$$= \left(\mathbf{AD}^{\mathrm{T}}\right)_{ik}\left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\right)_{lj} = \left(\left(\mathbf{AD}^{\mathrm{T}}\right) \boxtimes \left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\right)\right)_{ijkl} \ , \tag{3.94e}$$

$$((\mathbf{A} \boxtimes \mathbf{B}) : (\mathbf{C} \otimes \mathbf{D}))_{ijkl} = (\mathbf{A} \boxtimes \mathbf{B})_{ijmn}\,(\mathbf{C} \otimes \mathbf{D})_{mnkl} = (\mathbf{A})_{im}\,(\mathbf{B})_{nj}\,(\mathbf{C})_{mn}\,(\mathbf{D})_{kl}$$
$$= (\mathbf{ACB})_{ij}\,(\mathbf{D})_{kl} = ((\mathbf{ACB}) \otimes \mathbf{D})_{ijkl} \ , \tag{3.94f}$$

$$((\mathbf{A} \boxplus \mathbf{B}) : (\mathbf{C} \otimes \mathbf{D}))_{ijkl} = (\mathbf{A} \boxplus \mathbf{B})_{ijmn}\,(\mathbf{C} \otimes \mathbf{D})_{mnkl} = (\mathbf{A})_{in}\,(\mathbf{B})_{mj}\,(\mathbf{C})_{mn}\,(\mathbf{D})_{kl}$$
$$= \left(\mathbf{AC}^{\mathrm{T}}\mathbf{B}\right)_{ij}(\mathbf{D})_{kl} = \left(\left(\mathbf{AC}^{\mathrm{T}}\mathbf{B}\right) \otimes \mathbf{D}\right)_{ijkl} \ , \tag{3.94g}$$

$$((\mathbf{A} \otimes \mathbf{B}) : (\mathbf{C} \boxtimes \mathbf{D}))_{ijkl} = (\mathbf{A} \otimes \mathbf{B})_{ijmn} (\mathbf{C} \boxtimes \mathbf{D})_{mnkl} = (\mathbf{A})_{ij} (\mathbf{B})_{mn} (\mathbf{C})_{mk} (\mathbf{D})_{ln}$$
$$= (\mathbf{A})_{ij} \left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\mathbf{D}^{\mathrm{T}}\right)_{kl} = \left(\mathbf{A} \otimes \left(\mathbf{C}^{\mathrm{T}}\mathbf{B}\mathbf{D}^{\mathrm{T}}\right)\right)_{ijkl} , \tag{3.94h}$$
$$((\mathbf{A} \otimes \mathbf{B}) : (\mathbf{C} \boxplus \mathbf{D}))_{ijkl} = (\mathbf{A} \otimes \mathbf{B})_{ijmn} (\mathbf{C} \boxplus \mathbf{D})_{mnkl} = (\mathbf{A})_{ij} (\mathbf{B})_{mn} (\mathbf{C})_{ml} (\mathbf{D})_{kn}$$
$$= (\mathbf{A})_{ij} \left(\mathbf{D}\mathbf{B}^{\mathrm{T}}\mathbf{C}\right)_{kl} = \left(\mathbf{A} \otimes \left(\mathbf{D}\mathbf{B}^{\mathrm{T}}\mathbf{C}\right)\right)_{ijkl} . \tag{3.94i}$$

By using the identity $\mathbb{A}_{ijmn}\delta_{mk}\delta_{ln} = \delta_{im}\delta_{nj}\mathbb{A}_{mnkl} = \mathbb{A}_{ijkl}$, taking into account (3.67) and (3.86), one can establish

$$\underbrace{\mathbb{A} : \mathbb{I} = \mathbb{I} : \mathbb{A} = \mathbb{A}}_{\text{or } [\mathbb{A} : (\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j)] \otimes (\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j) = (\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j) \otimes [(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j) : \mathbb{A}] = \mathbb{A}} , \tag{3.95}$$

and this immediately yields

$$\boxed{\mathbb{I} : \mathbb{I} = \mathbb{I} \,.} \quad \longleftarrow \text{see (3.153)} \tag{3.96}$$

The result (3.95) basically states that any element $\mathbb{A}$ of the set $\mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o}3}\right)$ remains unchanged under the composition operation with the fourth-order unit tensor $\mathbb{I}$ from the both sides.
Hint: Let's denote $\mathbb{A} : \mathbb{A}$ by $\mathbb{A}^2$ and subsequently define the *powers* (or *monomials*) of fourth-order tensors via

$$\mathbb{A}^m = \underbrace{\mathbb{A} : \mathbb{A} : \ldots : \mathbb{A}}_{m \text{ times}} , \quad m = 1, 2, \ldots , \quad \text{and } \mathbb{A}^0 = \mathbb{I} \,. \tag{3.97}$$

For any two nonnegative integers m and n, the following properties are evident

$$\underbrace{\mathbb{A}^m : \mathbb{A}^n}_{= \mathbb{A}^n : \mathbb{A}^m} = \mathbb{A}^{m+n} \quad , \quad \underbrace{\left(\mathbb{A}^m\right)^n}_{= (\mathbb{A}^n)^m} = \mathbb{A}^{mn} \quad , \quad (\alpha\mathbb{A})^m = \alpha^m \mathbb{A}^m \,. \tag{3.98}$$

The *quadruple contraction* of $\mathbb{A}$ and $\mathbb{B}$ is defined as

$$\boxed{\mathbb{A} \,\vdots\, \mathbb{B} = \mathbb{A}_{ijkl}\mathbb{B}_{ijkl}} . \quad \longleftarrow \text{see (2.75)} \tag{3.99}$$

This helps represent the **norm** of a fourth-order tensor $\mathbb{A}$ according to

$$\boxed{|\mathbb{A}| = \sqrt{\mathbb{A} \,\vdots\, \mathbb{A}} \,.} \quad \longleftarrow \text{see (2.76)} \tag{3.100}$$

3.2.2.2 Simple Composition with Tensors

The interaction between fourth- and second-order tensors is not always of the form (3.33) or (3.45). For instance, a general fourth-order tensor of the form $\mathbf{A}\mathbb{A}\mathbf{B}$ will appear in tensor calculus as a result of differentiating a tensor-valued function of one tensor variable with respect to its argument.[2] The goal here is to define some useful rules for *simple composition* of a fourth-order tensor with two second-order tensors. These rules are based on how a tensor of rank four is constructed from two second-order tensors. Bearing in mind the tensor products introduced in (3.49a)–(3.49d), one can establish

$$\mathbf{A}\,(\mathbf{C}\otimes\mathbf{D})\,\mathbf{B} = (\mathbf{AC})\otimes(\mathbf{DB}) \ , \tag{3.101a}$$

$$\begin{aligned}\mathbf{A}\,(\mathbf{C}\boxtimes\mathbf{D})\,\mathbf{B} &= (\mathbf{A}\boxtimes\mathbf{B}) : (\mathbf{C}\boxtimes\mathbf{D}) \\ &= (\mathbf{AC})\boxtimes(\mathbf{DB}) \ , \quad \longleftarrow \text{see (3.94b)}\end{aligned} \tag{3.101b}$$

$$\begin{aligned}\mathbf{A}\,(\mathbf{C}\boxplus\mathbf{D})\,\mathbf{B} &= (\mathbf{A}\boxtimes\mathbf{B}) : (\mathbf{C}\boxplus\mathbf{D}) \\ &= (\mathbf{AC})\boxplus(\mathbf{DB}) \ , \quad \longleftarrow \text{see (3.94c)}\end{aligned} \tag{3.101c}$$

$$\begin{aligned}\mathbf{A}\,(\mathbf{C}\odot\mathbf{D})\,\mathbf{B} &= (\mathbf{A}\boxtimes\mathbf{B}) : (\mathbf{C}\odot\mathbf{D}) \\ &= (\mathbf{AC})\odot(\mathbf{DB}) \ .\end{aligned} \tag{3.101d}$$

These relations need to be written with respect to the Cartesian basis (3.61). Their right hand sides in indicial notation

$$\begin{aligned}
((\mathbf{AC})\otimes(\mathbf{DB}))_{ijkl} &= (\mathbf{AC})_{ij}\,(\mathbf{DB})_{kl} \\
&= (\mathbf{A})_{im}\left[(\mathbf{C})_{mj}\,(\mathbf{D})_{kn}\right](\mathbf{B})_{nl} \ , \\
((\mathbf{AC})\boxtimes(\mathbf{DB}))_{ijkl} &= (\mathbf{AC})_{ik}\,(\mathbf{DB})_{lj} \\
&= (\mathbf{A})_{im}\left[(\mathbf{C})_{mk}\,(\mathbf{D})_{ln}\right](\mathbf{B})_{nj} \ , \\
((\mathbf{AC})\boxplus(\mathbf{DB}))_{ijkl} &= (\mathbf{AC})_{il}\,(\mathbf{DB})_{kj} \\
&= (\mathbf{A})_{im}\left[(\mathbf{C})_{ml}\,(\mathbf{D})_{kn}\right](\mathbf{B})_{nj} \ , \\
((\mathbf{AC})\odot(\mathbf{DB}))_{ijkl} &= (\mathbf{A})_{im}\left[\frac{1}{2}\,(\mathbf{C})_{mk}\,(\mathbf{D})_{ln} + \frac{1}{2}\,(\mathbf{C})_{ml}\,(\mathbf{D})_{kn}\right](\mathbf{B})_{nj} \ ,
\end{aligned}$$

help represent the compact forms

---

[2] As an example, consider a second-order tensor $\mathbf{A} = \mathbf{B}^2$. The **derivative** of $\mathbf{A}$ with respect to $\mathbf{B}$, denoted by $\partial\mathbf{A}/\partial\mathbf{B}$, is a fourth-order tensor presenting

$$\frac{\partial\mathbf{A}}{\partial\mathbf{B}} = \frac{\partial\mathbf{B}}{\partial\mathbf{B}}\mathbf{B} + \mathbf{B}\frac{\partial\mathbf{B}}{\partial\mathbf{B}} = \mathbf{I}\mathbb{I}\mathbf{B} + \mathbf{B}\mathbb{I}\mathbf{I} \ .$$

See Chap. 6 for more elaborations and applications.

$$\mathbf{A}\mathbb{K}\mathbf{B} = A_{im}\mathbb{K}_{mjkn}B_{nl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \quad \text{if} \quad \mathbb{K} = \mathbf{C} \otimes \mathbf{D}\,, \tag{3.102a}$$

$$\mathbf{A}\mathbb{L}\mathbf{B} = A_{im}\mathbb{L}_{mnkl}B_{nj}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \quad \text{if} \quad \mathbb{L} = \mathbf{C} \boxtimes \mathbf{D}, \mathbf{C} \boxplus \mathbf{D}, \mathbf{C} \odot \mathbf{D}\,. \tag{3.102b}$$

Upon use of the double contraction of these fourth-order tensors with an arbitrary tensor $\mathbf{E}$, one can arrive at the useful rules

$$\mathbf{A}\mathbb{K}\mathbf{B} : \mathbf{E} = \mathbf{A}\left[\mathbb{K} : \left(\mathbf{E}\mathbf{B}^{\mathrm{T}}\right)\right] \quad \text{if} \quad \mathbb{K} = \mathbf{C} \otimes \mathbf{D}\,, \tag{3.103a}$$

$$\mathbf{A}\mathbb{L}\mathbf{B} : \mathbf{E} = \mathbf{A}\left[\mathbb{L} : \mathbf{E}\right]\mathbf{B} \quad \text{if} \quad \mathbb{L} = \mathbf{C} \boxtimes \mathbf{D}, \mathbf{C} \boxplus \mathbf{D}, \mathbf{C} \odot \mathbf{D}\,. \tag{3.103b}$$

3.2.2.3 Transposition

Consistent with (2.48), the *transpose* of a fourth-order tensor $\mathbb{A}$, denoted by $\mathbb{A}^{\mathrm{T}}$, is defined as

$$\boxed{\mathbb{A}^{\mathrm{T}} : \mathbf{C} = \mathbf{C} : \mathbb{A}\,, \quad \text{for all } \mathbf{C} \in \mathscr{T}_{\mathrm{so}}\,.} \tag{3.104}$$

Given the Cartesian basis (3.61), one needs to have

$$\left(\mathbb{A}^{\mathrm{T}} : \mathbf{C}\right)_{ij} = (\mathbf{C} : \mathbb{A})_{ij} \Longrightarrow \underbrace{\mathbb{A}^{\mathrm{T}}_{ijkl}C_{kl}}_{\text{see (3.66a)}} = \underbrace{C_{kl}\mathbb{A}_{klij}}_{\text{see (3.66b)}} \xrightarrow[(3.38)]{\text{from}} \boxed{\mathbb{A}^{\mathrm{T}}_{ijkl} = \mathbb{A}_{klij}\,,} \tag{3.105}$$

in order to represent

$$\boxed{\mathbb{A}^{\mathrm{T}} = \mathbb{A}_{klij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l = \mathbb{A}_{ijkl}\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\,.} \tag{3.106}$$

It is worthwhile to point out that the transpose of a second-order tensor, according to (2.48), is an **unique** operation. But different transpositions can be defined for fourth-order tensors and the definition (3.104) renders the most standard form. There are some specific expressions which cannot be represented in tensorial notation by just using this normal form. This motivates to introduce another useful transposition operation as

$$\boxed{\mathbb{A}^{\hat{\mathrm{T}}} : \mathbf{C} = \mathbb{A} : \mathbf{C}^{\mathrm{T}}\,, \quad \text{for all } \mathbf{C} \in \mathscr{T}_{\mathrm{so}}\,,} \tag{3.107}$$

which delivers

$$\left(\mathbb{A}^{\hat{\mathrm{T}}} : \mathbf{C}\right)_{ij} = \left(\mathbb{A} : \mathbf{C}^{\mathrm{T}}\right)_{ij} \Longrightarrow \underbrace{\mathbb{A}^{\hat{\mathrm{T}}}_{ijkl}C_{kl}}_{\text{see (3.66a)}} = \underbrace{\mathbb{A}_{ijkl}C_{lk}}_{= \mathbb{A}_{ijlk}C_{kl}} \xrightarrow[(3.38)]{\text{from}} \boxed{\mathbb{A}^{\hat{\mathrm{T}}}_{ijkl} = \mathbb{A}_{ijlk}\,.} \tag{3.108}$$

As a result,

$$\boxed{\mathbb{A}^{\hat{\mathrm{T}}} = \mathbb{A}_{ijlk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l = \mathbb{A}_{ijkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_k\,.} \tag{3.109}$$

Note that, in general,

$$\left(\mathbb{A}^{\hat{\mathrm{T}}}\right)^{\mathrm{T}} \neq \left(\mathbb{A}^{\mathrm{T}}\right)^{\hat{\mathrm{T}}} .$$

Now, the required data is available in order to extend the results (2.51a) to (2.56) to the problem at hand (see Exercise 3.1):

$$\left.\begin{array}{l}\mathbb{A} : \mathbf{C} = \mathbf{C} : \mathbb{A}^{\mathrm{T}} \\ \mathbb{A} : \mathbf{C} = \mathbb{A}^{\hat{\mathrm{T}}} : \mathbf{C}^{\mathrm{T}}\end{array}\right\} , \quad \longleftarrow \text{see (2.51a)} \tag{3.110a}$$

$$\left.\begin{array}{r}\mathbf{C} : \mathbb{A}^{\mathrm{T}} : \mathbf{D} = \mathbf{D} : \mathbb{A} : \mathbf{C} \\ \mathbf{C} : \mathbb{A}^{\hat{\mathrm{T}}} : \mathbf{D} = \mathbf{D}^{\mathrm{T}} : \mathbb{A}^{\mathrm{T}} : \mathbf{C} \\ \mathbf{C} : \mathbb{A}^{\hat{\mathrm{T}}} = (\mathbf{C} : \mathbb{A})^{\mathrm{T}}\end{array}\right\} , \tag{3.110b}$$

$$\left.\begin{array}{r}(\mathbb{A} : \mathbf{C}) : (\mathbb{B} : \mathbf{D}) = \left(\left(\mathbb{B}^{\mathrm{T}} : \mathbb{A}\right) : \mathbf{C}\right) : \mathbf{D} \\ \left(\mathbb{A}^{\hat{\mathrm{T}}} : \mathbf{C}\right) : \left(\mathbb{B}^{\hat{\mathrm{T}}} : \mathbf{D}\right) = \left(\left(\mathbb{B}^{\mathrm{T}} : \mathbb{A}\right) : \mathbf{C}^{\mathrm{T}}\right) : \mathbf{D}^{\mathrm{T}}\end{array}\right\} , \tag{3.110c}$$

$$\left.\begin{array}{r}(\mathbb{A} : \mathbf{C}) : (\mathbb{B} : \mathbf{D}) = \mathbf{C} : \left(\left(\mathbb{A}^{\mathrm{T}} : \mathbb{B}\right) : \mathbf{D}\right) \\ \left(\mathbb{A}^{\hat{\mathrm{T}}} : \mathbf{C}\right) : \left(\mathbb{B}^{\hat{\mathrm{T}}} : \mathbf{D}\right) = \mathbf{C}^{\mathrm{T}} : \left(\left(\mathbb{A}^{\mathrm{T}} : \mathbb{B}\right) : \mathbf{D}^{\mathrm{T}}\right)\end{array}\right\} , \tag{3.110d}$$

$$\left.\begin{array}{l}(\mathbb{A} + \mathbb{B})^{\mathrm{T}} = \mathbb{A}^{\mathrm{T}} + \mathbb{B}^{\mathrm{T}} \\ (\mathbb{A} + \mathbb{B})^{\hat{\mathrm{T}}} = \mathbb{A}^{\hat{\mathrm{T}}} + \mathbb{B}^{\hat{\mathrm{T}}}\end{array}\right\} , \quad \longleftarrow \text{see (2.52)} \tag{3.110e}$$

$$\left.\begin{array}{l}(\alpha\mathbb{A})^{\mathrm{T}} = \alpha\mathbb{A}^{\mathrm{T}} \\ (\alpha\mathbb{A})^{\hat{\mathrm{T}}} = \alpha\mathbb{A}^{\hat{\mathrm{T}}}\end{array}\right\} , \tag{3.110f}$$

$$\left.\begin{array}{ll}(\mathbf{A} \otimes \mathbf{B})^{\mathrm{T}} = \mathbf{B} \otimes \mathbf{A} \quad , & \underbrace{2\,(\mathbf{A} \odot \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} \boxtimes \mathbf{B}^{\mathrm{T}} + \mathbf{B} \boxplus \mathbf{A}}_{\text{for which } (\mathbf{A} \boxtimes \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} \boxtimes \mathbf{B}^{\mathrm{T}} ,\ (\mathbf{A} \boxplus \mathbf{B})^{\mathrm{T}} = \mathbf{B} \boxplus \mathbf{A}} \\ (\mathbf{A} \otimes \mathbf{B})^{\hat{\mathrm{T}}} = \mathbf{A} \otimes \mathbf{B}^{\mathrm{T}} \quad , & \underbrace{(\mathbf{A} \odot \mathbf{B})^{\hat{\mathrm{T}}} = \mathbf{A} \odot \mathbf{B}}_{\text{for which } (\mathbf{A} \boxtimes \mathbf{B})^{\hat{\mathrm{T}}} = \mathbf{A} \boxplus \mathbf{B} ,\ (\mathbf{A} \boxplus \mathbf{B})^{\hat{\mathrm{T}}} = \mathbf{A} \boxtimes \mathbf{B}}\end{array}\right\} , \tag{3.110g}$$

$$\left.\begin{array}{ll}\mathbb{I}^{\mathrm{T}} = \mathbb{I} \quad , & \bar{\mathbb{I}}^{\mathrm{T}} = \bar{\mathbb{I}} \\ \mathbb{I}^{\hat{\mathrm{T}}} = \bar{\mathbb{I}} \quad , & \bar{\mathbb{I}}^{\hat{\mathrm{T}}} = \mathbb{I}\end{array}\right\} , \quad \longleftarrow \text{see (2.55a)} \tag{3.110h}$$

$$\left.\begin{array}{ll}\mathbb{A}^{\mathrm{T}\,\mathrm{T}} = \mathbb{A} \quad , & \mathbb{A}^{\mathrm{T}\,\hat{\mathrm{T}}\,\mathrm{T}} = \bar{\mathbb{I}} : \mathbb{A} \\ \mathbb{A}^{\hat{\mathrm{T}}\,\hat{\mathrm{T}}} = \mathbb{A} \quad , & \mathbb{A}^{\hat{\mathrm{T}}\,\mathrm{T}\,\hat{\mathrm{T}}} = \bar{\mathbb{I}} : \mathbb{A}^{\mathrm{T}} : \bar{\mathbb{I}}\end{array}\right\} , \tag{3.110i}$$

$$\left.\begin{array}{l} \mathbb{A}^{\mathrm{T}} = \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) \otimes \left(\mathbb{A} : \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right)\right) \\ \mathbb{A}^{\widehat{\mathrm{T}}} = \left(\left(\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i\right) : \mathbb{A}^{\mathrm{T}}\right) \otimes \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) \\ \mathbb{A}^{\widehat{\mathrm{T}}} = \mathbb{A} : \bar{\mathbb{I}} \end{array}\right\} , \tag{3.110j}$$

$$\left.\begin{array}{ll} (\mathbb{A} : \mathbb{B})^{\mathrm{T}} = \mathbb{B}^{\mathrm{T}} : \mathbb{A}^{\mathrm{T}} \quad , & (\mathbb{A} : \mathbb{B} : \mathbb{C})^{\mathrm{T}} = \mathbb{C}^{\mathrm{T}} : \mathbb{B}^{\mathrm{T}} : \mathbb{A}^{\mathrm{T}} \\ (\mathbb{A} : \mathbb{B})^{\widehat{\mathrm{T}}} = \mathbb{A} : \mathbb{B}^{\widehat{\mathrm{T}}} \quad , & (\mathbb{A} : \mathbb{B} : \mathbb{C})^{\widehat{\mathrm{T}}} = \mathbb{A} : \mathbb{B} : \mathbb{C}^{\widehat{\mathrm{T}}} \end{array}\right\} , \tag{3.110k}$$

$$\left.\begin{array}{l} (\mathbf{C} \otimes \mathbf{D}) : \mathbb{A}^{\mathrm{T}} = \mathbf{C} \otimes (\mathbb{A} : \mathbf{D}) \\ (\mathbf{C} \otimes \mathbf{D}) : \mathbb{A}^{\widehat{\mathrm{T}}} = \mathbf{C} \otimes \left(\mathbb{A}^{\mathrm{T}} : \mathbf{D}\right)^{\mathrm{T}} \end{array}\right\} . \quad \longleftarrow \text{see (2.56)} \tag{3.110l}$$

3.2.3 Major and Minor Symmetries of Fourth-Order Tensors

A fourth-order tensor $\mathbb{A}$ is said to be *major symmetric* (or simply *symmetric*) if

$$\boxed{\mathbb{A}^{\mathrm{T}} = \mathbb{A} \quad \text{or, equivalently,} \quad \mathbf{C} : \mathbb{A} : \mathbf{D} = \mathbf{D} : \mathbb{A} : \mathbf{C} \,,} \tag{3.111}$$

which is in accord with (2.57). One then says that $\mathbb{A}$ possesses the *major symmetries*; indicated by,

$$\boxed{\mathbb{A}_{klij} = \mathbb{A}_{ijkl} \,.} \tag{3.112}$$

Unlike second-order tensors, the condition (3.111) is not the only type of symmetry. For instance, if

$$\boxed{\mathbb{A}^{\widehat{\mathrm{T}}} = \mathbb{A} \quad \text{or, equivalently,} \quad \mathbf{C} : \mathbb{A} : \mathbf{D} = \mathbf{C} : \left(\mathbb{A} : \mathbf{D}^{\mathrm{T}}\right) = (\mathbf{C} : \mathbb{A})^{\mathrm{T}} : \mathbf{D} \,,} \tag{3.113}$$

then $\mathbb{A}$ is called *minor (right) symmetric* and its components will have the *minor (right) symmetries*, that is,

$$\boxed{\mathbb{A}_{ijlk} = \mathbb{A}_{ijkl} \,.} \quad \longleftarrow \text{see (1.97)} \tag{3.114}$$

A fourth-order tensor $\mathbb{A}$ possessing both major and minor symmetries according to

$$\mathbb{A}_{ijkl} = \mathbb{A}_{klij} = \mathbb{A}_{klji} = \mathbb{A}_{ijlk} = \mathbb{A}_{lkij} = \mathbb{A}_{lkji} = \mathbb{A}_{jilk} = \mathbb{A}_{jikl} \,, \tag{3.115}$$

is referred to as *super-symmetric*. The super-symmetric tensors are of great importance in solid mechanics. A well-known example will be the elastic stiffness tensor in linear elasticity (see Exercise 3.4). Another example is the super-symmetric projection tensor $\mathbb{P}_{\text{sym}}$ already introduced in (3.58a). Some important material as well as spatial elasticity tensors for hyperelastic solids have also major and minor symmetries, see Exercise 6.16.

It is easy to see that when a super-symmetric tensor operates on an arbitrary second-order tensor, the result will be a symmetric tensor:

$$\begin{aligned}(\mathbb{A} : \mathbf{C})^{\mathrm{T}} &\xlongequal[\text{(3.110a)}]{\text{from}} \left(\mathbf{C} : \mathbb{A}^{\mathrm{T}}\right)^{\mathrm{T}} \\ &\xlongequal[\text{(3.111)}]{\text{from}} (\mathbf{C} : \mathbb{A})^{\mathrm{T}} \\ &\xlongequal[\text{(3.110b)}]{\text{from}} \mathbf{C} : \mathbb{A}^{\widehat{\mathrm{T}}} \\ &\xlongequal[\text{(3.113)}]{\text{from}} \mathbf{C} : \mathbb{A} \\ &\xlongequal[\text{(3.111)}]{\text{from}} \mathbf{C} : \mathbb{A}^{\mathrm{T}} \\ &\xlongequal[\text{(3.110a)}]{\text{from}} \mathbb{A} : \mathbf{C} \,. \end{aligned} \tag{3.116}$$

The set of all super-symmetric tensors

$$\mathscr{T}_{\text{fo}}^{\text{ss}} = \left\{ \mathbb{A} \in \mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right) \mid \mathbb{A}^{\mathrm{T}} = \mathbb{A} \,,\ \mathbb{A}^{\widehat{\mathrm{T}}} = \mathbb{A} \right\} \,, \tag{3.117}$$

constitutes a subspace of all fourth-order tensors $\mathscr{T}_{\text{fo}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$.

A fourth-order tensor $\mathbb{A}$ is referred to as *major skew-symmetric* (or simply *skew-symmetric*) if

$$\underbrace{\mathbb{A}^{\mathrm{T}} = -\mathbb{A} \quad \text{or, equivalently,} \quad \mathbf{C} : \mathbb{A} : \mathbf{D} = -\mathbf{D} : \mathbb{A} : \mathbf{C}}_{\text{with major skew-symmetries } \mathbb{A}_{klij} = -\mathbb{A}_{ijkl}} \,, \tag{3.118}$$

and *minor (right) skew-symmetric* if

$$\underbrace{\mathbb{A}^{\widehat{\mathrm{T}}} = -\mathbb{A} \text{ or, equivalently, } \mathbf{C} : \mathbb{A} : \mathbf{D} = -\mathbf{C} : \left(\mathbb{A} : \mathbf{D}^{\mathrm{T}}\right) = -\left(\mathbf{C} : \mathbb{A}\right)^{\mathrm{T}} : \mathbf{D}}_{\text{with minor (right) skew-symmetries } \mathbb{A}_{ijlk} = -\mathbb{A}_{ijkl}} \,. \tag{3.119}$$

Consistent with tensors, any fourth-order tensor $\mathbb{A}$ can uniquely be decomposed into its symmetric and skew-symmetric parts

$$\boxed{\mathbb{A} = \underbrace{\frac{1}{2}\left(\mathbb{A} + \mathbb{A}^{\mathrm{T}}\right)}_{:=\,\text{Sym}\mathbb{A}} + \underbrace{\frac{1}{2}\left(\mathbb{A} - \mathbb{A}^{\mathrm{T}}\right)}_{:=\,\text{Skw}\mathbb{A}} = \text{Sym}\mathbb{A} + \text{Skw}\mathbb{A} \,.} \tag{3.120}$$

In a similar manner,

$$\boxed{\mathbb{A} = \underbrace{\frac{1}{2}\left(\mathbb{A} + \mathbb{A}^{\hat{\mathrm{T}}}\right)}_{:=\,\mathrm{sym}\mathbb{A}} + \underbrace{\frac{1}{2}\left(\mathbb{A} - \mathbb{A}^{\hat{\mathrm{T}}}\right)}_{:=\,\mathrm{skw}\mathbb{A}} = \mathrm{sym}\mathbb{A} + \mathrm{skw}\mathbb{A} \ .} \tag{3.121}$$

3.2.4 Fourth-Order Tensor in Matrix Notation

As discussed in the previous chapters, the theoretical side of a physical problem will end up with a tensorial expression (called an equilibrium equation) represented in direct notation. And the derivations of formulations are usually carried out by use of indicial notation. The computational side of the problem then follows that aims at analytically solving the problem, if there is a closed-form solution, or providing an approximate solution by means of numerical procedures. It is exactly at this stage that the role of matrix notation becomes dominant in which the attempt will be made to recast (the components of) tensors into (multi- and/or single-column) matrices. In the following, the focus will only be on second- and fourth-order tensors.

An arbitrary tensor possessing 9 independent components is written here as a single-column matrix:

$$[\,\mathbf{C}\,]_{9\times 1} = \left[\, C_{11}\ C_{22}\ C_{33}\ C_{23}\ C_{13}\ C_{12}\ C_{32}\ C_{31}\ C_{21} \,\right]^{\mathrm{T}} \ . \tag{3.122}$$

Consistent with this structure, an arbitrary fourth-order tensor with 81 independent components will be written as

$$[\,\mathbb{A}\,]_{9\times 9} = \begin{bmatrix} \mathbb{A}_{1111} & \cdot\cdot & \mathbb{A}_{1123} & \mathbb{A}_{1113} & \mathbb{A}_{1112} & \mathbb{A}_{1132} & \mathbb{A}_{1131} & \mathbb{A}_{1121} \\ \mathbb{A}_{2211} & \cdot\cdot & \mathbb{A}_{2223} & \mathbb{A}_{2213} & \mathbb{A}_{2212} & \mathbb{A}_{2232} & \mathbb{A}_{2231} & \mathbb{A}_{2221} \\ \mathbb{A}_{3311} & \cdot\cdot & \mathbb{A}_{3323} & \mathbb{A}_{3313} & \mathbb{A}_{3312} & \mathbb{A}_{3332} & \mathbb{A}_{3331} & \mathbb{A}_{3321} \\ \vdots & \vdots\,\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbb{A}_{2111} & \cdot\cdot & \mathbb{A}_{2123} & \mathbb{A}_{2113} & \mathbb{A}_{2112} & \mathbb{A}_{2132} & \mathbb{A}_{2131} & \mathbb{A}_{2121} \end{bmatrix} . \tag{3.123}$$

The right mapping $(3.66a)_1$ as well as the left mapping $(3.66b)_1$ can then be represented in the convenient forms

$$\boxed{[\,\mathbf{A}\,] = [\,\mathbb{A} : \mathbf{C}\,] = [\,\mathbb{A}\,][\,\mathbf{C}\,] \ ,} \tag{3.124}$$

and

$$\boxed{[\,\mathbf{B}\,] = [\,\mathbf{C} : \mathbb{A}\,] = \left[\,\mathbb{A}^{\mathrm{T}} : \mathbf{C}\,\right] = [\,\mathbb{A}\,]^{\mathrm{T}}[\,\mathbf{C}\,] \ .} \tag{3.125}$$

where $\left[\,\mathbb{A}^{\mathrm{T}}\,\right] = [\,\mathbb{A}\,]^{\mathrm{T}}$ has been used. It follows that

$$\boxed{\mathbf{C} : \mathbf{D} = [\,\mathbf{C}\,]^{\mathrm{T}}\,[\,\mathbf{D}\,]\ ,\ \mathbf{D} : \mathbb{A} : \mathbf{C} = [\,\mathbf{D}\,]^{\mathrm{T}}\,[\,\mathbb{A}\,]\,[\,\mathbf{C}\,]\ ,\ [\,\mathbb{A} : \mathbb{B}\,] = [\,\mathbb{A}\,]\,[\,\mathbb{B}\,]\ .} \tag{3.126}$$

Of particular interest here is to consider symmetric and super-symmetric tensors. Let $\mathbf{S}$ and $\tilde{\mathbf{S}}$ be two symmetric tensors. Further, let $\mathbb{C}$ and $\tilde{\mathbb{C}}$ be two super-symmetric tensors. A symmetric tensor $\mathbf{S}$ with 6 independent components is given by the single-column matrix

$$[\,\mathbf{S}\,]_{6\times 1} = \left[\, S_{11}\ \ S_{22}\ \ S_{33}\ \ \alpha S_{23}\ \ \alpha S_{13}\ \ \alpha S_{12} \,\right]^{\mathrm{T}}\ , \tag{3.127}$$

where α presents a constant which can be either 1 or 2 depending on the operation. Accordingly, 21 independent components of a super-symmetric tensor $\mathbb{C}$ can be collected in the following symmetric matrix

$$[\,\mathbb{C}\,]_{6\times 6} = \begin{bmatrix} \mathbb{C}_{1111} & \mathbb{C}_{1122} & \mathbb{C}_{1133} & \mathbb{C}_{1123} & \mathbb{C}_{1113} & \mathbb{C}_{1112} \\ & \mathbb{C}_{2222} & \mathbb{C}_{2233} & \mathbb{C}_{2223} & \mathbb{C}_{2213} & \mathbb{C}_{2212} \\ & & \mathbb{C}_{3333} & \mathbb{C}_{3323} & \mathbb{C}_{3313} & \mathbb{C}_{3312} \\ & & & \mathbb{C}_{2323} & \mathbb{C}_{2313} & \mathbb{C}_{2312} \\ & sym. & & & \mathbb{C}_{1313} & \mathbb{C}_{1312} \\ & & & & & \mathbb{C}_{1212} \end{bmatrix}\ . \tag{3.128}$$

The explicit form (3.127) (or (3.128)) is referred to as *Voigt* notation. For convenience, one can collect the subscript indices of $\mathbf{S}$ in the following 6×2 matrix:

$$\left[\,\mathbf{V}^{\mathrm{not}}\,\right] = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 2 & 3 \\ 1 & 3 \\ 1 & 2 \end{bmatrix}\ . \tag{3.129}$$

Now, the right and left mappings, according to $(3.66\mathrm{a})_1$ and $(3.66\mathrm{b})_1$, can be written as

$$\boxed{[\,\mathbb{C} : \mathbf{S}\,] = [\,\mathbf{S} : \mathbb{C}\,] = [\,\mathbb{C}\,]\,[\,\mathbf{S}\,]\big|_{\alpha=2}\ .} \tag{3.130}$$

By knowing that

$$\boxed{\tilde{\mathbf{S}} : \mathbf{S} = \left[\,\tilde{\mathbf{S}}\,\right]^{\mathrm{T}}\Big|_{\alpha=1}\ [\,\mathbf{S}\,]\big|_{\alpha=2}\ ,} \tag{3.131}$$

one will have

$$\boxed{\tilde{\mathbf{S}} : \mathbb{C} : \mathbf{S} = \left[\tilde{\mathbf{S}} \right]^{\mathrm{T}} \Big|_{\alpha=2} [\, \mathbb{C} \,] \, [\, \mathbf{S} \,]|_{\alpha=2} \ .} \quad \longleftarrow \text{note that also } \tilde{\mathbf{S}} : \mathbb{C} : \mathbf{S} = \mathbf{S} : \mathbb{C} : \tilde{\mathbf{S}} \tag{3.132}$$

The fourth-order tensor $\tilde{\mathbb{C}} : \mathbb{C}$ possesses the minor symmetries but, in general, it is not super-symmetric due to the lack of major symmetries, i.e. $\tilde{\mathbb{C}} : \mathbb{C} \neq \mathbb{C} : \tilde{\mathbb{C}}$. The matrix from of such a tensorial variable with only minor symmetries is similar to (3.128) but generally may not be symmetric; an example of which with 36 independent components is already given in (1.96b).

Some special symmetric and super-symmetric tensors in Voigt notation are listed in the following.

The unit tensor $\mathbf{I} = \delta_{ij} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$**:**

$$[\, \mathbf{I} \,] = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} . \tag{3.133}$$

The super-symmetric identity tensor $\mathbb{P}_{\mathrm{sym}} = \mathbf{I} \odot \mathbf{I} = \frac{1}{2} \left(\delta_{ik}\delta_{lj} + \delta_{il}\delta_{kj} \right) \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l$**:**

$$\left[\, \mathbb{P}_{\mathrm{sym}} \,\right] = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \tag{3.134}$$

The super-symmetric spherical operator $\mathbb{P}_{\mathrm{sph}} = \frac{1}{3} \mathbf{I} \otimes \mathbf{I} = \frac{1}{3} \delta_{ij} \delta_{kl} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l$**:**

$$\left[\, \mathbb{P}_{\mathrm{sph}} \,\right] = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} . \tag{3.135}$$

The super-symmetric deviatoric operator $\mathbb{P}^{\mathrm{s}}_{\mathrm{dev}} = \mathbb{P}_{\mathrm{sym}} - \mathbb{P}_{\mathrm{sph}} = \mathbf{I} \odot \mathbf{I} - \frac{1}{3} \mathbf{I} \otimes \mathbf{I}$**:**

$$[\,\mathbb{P}^{s}_{dev}\,] = \frac{1}{6}\begin{bmatrix} 4 & -2 & -2 & 0 & 0 & 0 \\ -2 & 4 & -2 & 0 & 0 & 0 \\ -2 & -2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} . \tag{3.136}$$

3.2.5 Determinant and Inverse of Fourth-Order Tensors

The determinant of a fourth-order tensor $\mathbb{A}$, denoted by $\det \mathbb{A}$, is defined as the determinant of its matrix form, i.e. $\det \mathbb{A} = \det [\,\mathbb{A}\,]$, see Steinmann [11]. Guided by (2.99a) to (2.99c) and consistent with matrix algebra, the following useful properties hold true:

$$\det \mathbb{A}^{T} = \det [\,\mathbb{A}\,]^{T} = \det [\,\mathbb{A}\,] = \det \mathbb{A} \ , \tag{3.137a}$$

$$\begin{aligned} \det (\mathbb{A} : \mathbb{B}) &= \det ([\,\mathbb{A} : \mathbb{B}\,]) = \det ([\,\mathbb{A}\,][\,\mathbb{B}\,]) = (\det [\,\mathbb{A}\,]) (\det [\,\mathbb{B}\,]) \\ &= (\det \mathbb{A}) (\det \mathbb{B}) = \det (\mathbb{B} : \mathbb{A}) \ , \end{aligned} \tag{3.137b}$$

$$\det (\alpha \mathbb{A}) = \det [\,\alpha \mathbb{A}\,] = \alpha^{9} \det [\,\mathbb{A}\,] = \alpha^{9} \det \mathbb{A} \ . \tag{3.137c}$$

By means of $(3.95)_2$ and $(3.137b)_4$, one can compute the determinant of the fourth-order unit tensor $\mathbb{I}$ as follows:

$$\det (\mathbb{A} : \mathbb{I}) = \det \mathbb{A} \quad \Rightarrow \quad (\det \mathbb{A}) (\det \mathbb{I}) = \det \mathbb{A} \quad \Rightarrow \quad \det \mathbb{I} = 1 \ . \tag{3.138}$$

The determinant of the fourth-order unit tensor $\bar{\mathbb{I}}$ also shows the same result:

$$\det \bar{\mathbb{I}} = 1 \ . \tag{3.139}$$

From (3.41), (3.118), $(3.137a)_3$ and $(3.137c)_3$, the determinant of a **skew-symmetric** tensor $\mathbb{A}$ becomes zero:

$$\left.\begin{aligned} \det \mathbb{A} &= \det \mathbb{A}^{T} = \det (-\mathbb{A}) = \det ((-1)\, \mathbb{A}) \\ &= (-1)^{9} \det \mathbb{A} = -\det \mathbb{A} \end{aligned}\right\} \Rightarrow \ \det \mathbb{A} = 0 \ . \tag{3.140}$$

By use of (3.41), $(3.58b)_2$, $(3.137b)_4$, $(3.137c)_3$, (3.139) and $(3.154b)_3$, the determinant of the skew-symmetrizer $\mathbb{P}_{skw}$ also renders analogous result:

$$\det \left(\mathbb{P}_{skw} : \bar{\mathbb{I}}\right) = \det (-\mathbb{P}_{skw}) \Rightarrow \det \mathbb{P}_{skw} = -\det \mathbb{P}_{skw} \Rightarrow \det \mathbb{P}_{skw} = 0 \ . \tag{3.141}$$

Analogous to tensors, a fourth-order tensor $\mathbb{A}$ is said to be *invertible* if $\det \mathbb{A} \neq 0$. With regard to the linear mapping $\mathbf{D} = \mathbb{A} : \mathbf{C}$, this means that there exists $\mathbb{A}^{-1} \in \mathscr{T}_{\text{fo}}$, called the *inverse* of $\mathbb{A}$, such that[3]

$$\mathbf{C} = \mathbb{A}^{-1} : \mathbf{D} \,. \tag{3.142}$$

Accordingly, the identities

$$\mathbf{C} = \mathbb{A}^{-1} : \mathbf{D} = \mathbb{A}^{-1} : (\mathbb{A} : \mathbf{C}) \quad , \quad \mathbf{D} = \mathbb{A} : \left(\mathbb{A}^{-1} : \mathbf{D}\right) ,$$

with the aid of $(3.37)_1$, (3.38) and (3.85) imply the reciprocal relation[4]

$$\boxed{\mathbb{A} : \mathbb{A}^{-1} = \mathbb{I} = \mathbb{A}^{-1} : \mathbb{A} \,.} \tag{3.143}$$

As a result,

$$\boxed{\mathbb{I}^{-1} = \mathbb{I} \,.} \tag{3.144}$$

The (Cartesian) coordinate representation of $\mathbb{A}^{-1}$, in alignment with (3.62), renders

$$\mathbb{A}^{-1} = \mathbb{A}^{-1}_{ijkl} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \quad \text{where } \mathbb{A}^{-1}_{ijkl} = \left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right) : \left[\mathbb{A}^{-1} : \left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right)\right] \,. \tag{3.145}$$

This helps present the component from of (3.143) as

$$\mathbb{A}_{ijmn} \mathbb{A}^{-1}_{mnkl} = \delta_{ik} \delta_{lj} = \mathbb{A}^{-1}_{ijmn} \mathbb{A}_{mnkl} \,. \tag{3.146}$$

Any two invertible fourth-order tensors $\mathbb{A}$ and $\mathbb{B}$ satisfy the following properties

$$\det \mathbb{A}^{-1} = (\det \mathbb{A})^{-1} \;, \;\leftarrow \text{ since } \left.\begin{matrix} \mathbb{A} : \mathbb{A}^{-1} = \mathbb{I} \Rightarrow \det\left(\mathbb{A} : \mathbb{A}^{-1}\right) = \det \mathbb{I} \Rightarrow \\ (\det \mathbb{A}) \left(\det \mathbb{A}^{-1}\right) = 1 \Rightarrow \det \mathbb{A}^{-1} = 1/\left(\det \mathbb{A}\right) \end{matrix}\right\} \tag{3.147a}$$

$$(\mathbb{A} : \mathbb{B})^{-1} = \mathbb{B}^{-1} : \mathbb{A}^{-1} \;, \;\leftarrow \text{ since } \left.\begin{matrix} \mathbb{I} = (\mathbb{A} : \mathbb{B}) : (\mathbb{A} : \mathbb{B})^{-1} \\ \mathbb{I} = \mathbb{A} : \mathbb{A}^{-1} = \mathbb{A} : \mathbb{I} : \mathbb{A}^{-1} = \mathbb{A} : \mathbb{B} : \mathbb{B}^{-1} : \mathbb{A}^{-1} \end{matrix}\right\} \tag{3.147b}$$

$$\left(\mathbb{A}^{-1}\right)^m = \left(\mathbb{A}^m\right)^{-1} \;, \;\leftarrow \text{ where } m \text{ denotes a nonnegative integer} \tag{3.147c}$$

$$(\alpha\mathbb{A})^{-1} = \alpha^{-1}\mathbb{A}^{-1} \;, \;\leftarrow \text{ since } (\alpha\mathbb{A})^{-1} : (\alpha\mathbb{A}) = \mathbb{I} = \alpha^{-1}\mathbb{A}^{-1} : (\alpha\mathbb{A}) \tag{3.147d}$$

$$\left(\mathbb{A}^{-1}\right)^{-1} = \mathbb{A} \;, \;\leftarrow \text{ since } \left(\mathbb{A}^{-1}\right) : \left(\mathbb{A}^{-1}\right)^{-1} = \mathbb{I} = \mathbb{A}^{-1} : \mathbb{A} \tag{3.147e}$$

$$\left(\mathbb{A}^{-1}\right)^{\mathrm{T}} = \left(\mathbb{A}^{\mathrm{T}}\right)^{-1} \;, \;\leftarrow \text{ since } \left(\mathbb{A}^{-1} : \mathbb{A}\right)^{\mathrm{T}} = \mathbb{I}^{\mathrm{T}} \Rightarrow \mathbb{A}^{\mathrm{T}} : \left(\mathbb{A}^{-1}\right)^{\mathrm{T}} = \mathbb{I} = \mathbb{A}^{\mathrm{T}} : \left(\mathbb{A}^{\mathrm{T}}\right)^{-1} \tag{3.147f}$$

For subsequent developments, the following notation will be adopted:

---

[3] Note that precisely $\det \mathbb{A} \neq 0$ is the necessary and sufficient condition for $\mathbb{A}^{-1}$ to exist.

[4] The discussions regarding consistency of (2.105) also remain true here. This means that (3.143) may generally not be consistent for all invertible fourth-order tensors, although it always hold true from computational standpoint.

$$\mathbb{A}^{-\mathrm{T}} := \left(\mathbb{A}^{-1}\right)^{\mathrm{T}} = \left(\mathbb{A}^{\mathrm{T}}\right)^{-1} \quad , \quad \mathbb{A}^{-m} := \left(\mathbb{A}^{-1}\right)^{m} = \left(\mathbb{A}^{m}\right)^{-1} . \tag{3.148}$$

Let $\mathbb{C}$ be a super-symmetric tensor. The expression (3.143) then translates to

$$\mathbb{C} : \mathbb{C}^{-1} = \mathbf{I} \odot \mathbf{I} = \mathbb{C}^{-1} : \mathbb{C} , \tag{3.149}$$

since $\mathbf{I} \odot \mathbf{I}$ presents the only fourth-order identity tensor possessing major and minor symmetries (see Exercise 3.4).

3.2.6 Positive Definite and Negative Definite Fourth-Order Tensors

A fourth-order tensor $\mathbb{A}$ is called *positive* (*semi-*) *definite* when its quadratic form satisfies

$$\underbrace{\mathbf{C} : \mathbb{A} : \mathbf{C}}_{\text{or } C_{ij}\,\mathbb{A}_{ijkl}C_{kl} \text{ or } [\mathbf{C}]^{\mathrm{T}}[\mathbb{A}][\mathbf{C}]} > 0 \ (\geq 0) \ , \quad \text{for all } \mathbf{C} \in \mathscr{T}_{\text{so}} \ , \ \mathbf{C} \neq \mathbf{O} \ . \tag{3.150}$$

and referred to as *negative* (*semi-*) *definite* if

$$\mathbf{C} : \mathbb{A} : \mathbf{C} < 0 \ (\leq 0) \ , \quad \text{for all } \mathbf{C} \in \mathscr{T}_{\text{so}} \ , \ \mathbf{C} \neq \mathbf{O} \ . \tag{3.151}$$

Consistent with tensors, positive-definiteness of a fourth-order tensor $\mathbb{A}$ will be guaranteed when its symmetric part, i.e. $\mathrm{Sym}\mathbb{A}$, is positive definite since basically $\mathbf{C} : \mathrm{Skw}\mathbb{A} : \mathbf{C} = \mathbf{C} : (\mathrm{Skw}\mathbb{A})^{\mathrm{T}} : \mathbf{C} = -\mathbf{C} : \mathrm{Skw}\mathbb{A} : \mathbf{C}$ delivers

$$\mathbf{C} : \mathrm{Skw}\mathbb{A} : \mathbf{C} = 0 .$$

3.3 Exercises

Exercise 3.1

Verify $(3.110\text{g})_2$, $(3.110\text{g})_4$, $(3.110\text{i})_2$, $(3.110\text{j})_{1-3}$, $(3.110\text{k})_1$ and $(3.110\text{k})_3$.

Solution. The expressions (3.110a) to (3.110l) present some important relationships of the transposition operations introduced in (3.104) and (3.107). It is recommended that the interested reader proves all these expressions. Here only some major ones are verified. By use of the Cartesian representation of a fourth-order tensor, according to (3.62), the verification for each desired relation will be shown step by step in the following.

The expression (3.110g)$_2$:

$$2\,(\mathbf{A}\odot\mathbf{B})^{\mathrm{T}} \xlongequal[\text{(3.49d) and (3.106)}]{\text{from}} \underbrace{\left(\left(A_{ik}B_{lj}+A_{il}B_{kj}\right)\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right)^{\mathrm{T}}}_{=\left(A_{ik}B_{lj}+A_{il}B_{kj}\right)\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\otimes\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j=\left(A_{ki}B_{jl}+A_{kj}B_{il}\right)\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l}$$

$$\xlongequal[\text{(2.49)}]{\text{from}} \left(A^{\mathrm{T}}_{ik}B^{\mathrm{T}}_{lj}+B_{il}A_{kj}\right)\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l$$

$$\xlongequal[\text{(3.49b) and (3.49c)}]{\text{from}} \mathbf{A}^{\mathrm{T}}\boxtimes\mathbf{B}^{\mathrm{T}}+\mathbf{B}\boxplus\mathbf{A}\ .$$

The expression (3.110g)$_4$:

$$2\,(\mathbf{A}\odot\mathbf{B})^{\widehat{\mathrm{T}}} \xlongequal[\text{(3.49d) and (3.109)}]{\text{from}} \underbrace{\left(\left(A_{ik}B_{lj}+A_{il}B_{kj}\right)\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right)^{\widehat{\mathrm{T}}}}_{=\left(A_{ik}B_{lj}+A_{il}B_{kj}\right)\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_l\otimes\widehat{\mathbf{e}}_k=\left(A_{il}B_{kj}+A_{ik}B_{lj}\right)\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l}$$

$$\xlongequal[\text{(3.49b) and (3.49c)}]{\text{from}} \mathbf{A}\boxplus\mathbf{B}+\mathbf{A}\boxtimes\mathbf{B}$$

$$\xlongequal[\text{(3.49d)}]{\text{from}} 2\,(\mathbf{A}\odot\mathbf{B})\ .$$

The expression (3.110i)$_2$:

$$\mathbb{A}^{\mathrm{T}\widehat{\mathrm{T}}\mathrm{T}} \xlongequal[\text{(3.95)}]{\text{from}} (\mathbb{I}:\mathbb{A})^{\mathrm{T}\widehat{\mathrm{T}}\mathrm{T}}$$

$$\xlongequal[\text{(3.110h) and (3.110k)}]{\text{from}} \left(\mathbb{A}^{\mathrm{T}}:\mathbb{I}\right)^{\widehat{\mathrm{T}}\mathrm{T}}$$

$$\xlongequal[\text{(3.110k)}]{\text{from}} \left(\mathbb{A}^{\mathrm{T}}:\mathbb{I}^{\widehat{\mathrm{T}}}\right)^{\mathrm{T}}$$

$$\xlongequal[\text{(3.110h)}]{\text{from}} \left(\mathbb{A}^{\mathrm{T}}:\bar{\mathbb{I}}\right)^{\mathrm{T}}$$

$$\xlongequal[\text{(3.110i) and (3.110k)}]{\text{from}} \bar{\mathbb{I}}^{\mathrm{T}}:\mathbb{A}$$

$$\xlongequal[\text{(3.110h)}]{\text{from}} \bar{\mathbb{I}}:\mathbb{A}\ .$$

The expression (3.110j)$_1$:

$$\mathbb{A}^{\mathrm{T}} \xlongequal[\text{(3.53b) and (3.106)}]{\text{from}} \widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\left(\mathbb{A}_{klij}\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l\right)$$

$$\xlongequal[\text{(3.64)}]{\text{from}} \left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\otimes\left(\mathbb{A}:\left(\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\right)\right)\ .$$

The expression (3.110j)$_2$:

$$
\begin{aligned}
\mathbb{A}^{\widehat{\mathrm{T}}} &\xlongequal[\text{(3.109)}]{\text{from}} \mathbb{A}_{klji}\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \\
&\xlongequal[\text{(3.53a) and (3.105)}]{\text{from}} \left(\mathbb{A}^{\mathrm{T}}_{jikl}\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l\right) \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \\
&\xlongequal[\text{(3.65)}]{\text{from}} \left(\left(\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i\right) : \mathbb{A}^{\mathrm{T}}\right) \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ .
\end{aligned}
$$

The expression (3.110j)$_3$:

$$
\begin{aligned}
\mathbb{A}^{\widehat{\mathrm{T}}} &\xlongequal[\text{(3.109)}]{\text{from}} \mathbb{A}_{ijlk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&\xlongequal[\text{(1.36)}]{\text{from}} \mathbb{A}_{ijmn}\delta_{ml}\delta_{kn}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&\xlongequal[\text{(3.68)}]{\text{from}} \mathbb{A}_{ijmn}\bar{\mathbb{I}}_{mnkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&\xlongequal[\text{(3.91)}]{\text{from}} \mathbb{A} : \bar{\mathbb{I}} \ .
\end{aligned}
$$

The expression (3.110k)$_1$:

$$
\begin{aligned}
(\mathbb{A} : \mathbb{B})^{\mathrm{T}} &\xlongequal[\text{(3.91) and (3.106)}]{\text{from}} \mathbb{A}_{ijmn}\mathbb{B}_{mnkl}\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \\
&\xlongequal[\text{(3.105)}]{\text{from}} \mathbb{B}^{\mathrm{T}}_{klmn}\mathbb{A}^{\mathrm{T}}_{mnij}\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \\
&\xlongequal[\text{(3.91)}]{\text{from}} \mathbb{B}^{\mathrm{T}} : \mathbb{A}^{\mathrm{T}} \ .
\end{aligned}
$$

The expression (3.110k)$_3$:

$$
\begin{aligned}
(\mathbb{A} : \mathbb{B})^{\widehat{\mathrm{T}}} &\xlongequal[\text{(3.91) and (3.109)}]{\text{from}} \mathbb{A}_{ijmn}\mathbb{B}_{mnlk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&\xlongequal[\text{(3.108)}]{\text{from}} \mathbb{A}_{ijmn}\mathbb{B}^{\widehat{\mathrm{T}}}_{mnkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&\xlongequal[\text{(3.91)}]{\text{from}} \mathbb{A} : \mathbb{B}^{\widehat{\mathrm{T}}} \ .
\end{aligned}
$$

Exercise 3.2

Show that the projection tensors (3.58a) to (3.58d) satisfy the following properties

$$\mathbb{P}_{\text{sym}} : \mathbb{P}_{\text{sym}} = \mathbb{P}_{\text{sym}} \,, \qquad \leftarrow \text{see (2.61)} \tag{3.152a}$$

$$\mathbb{P}_{\text{skw}} : \mathbb{P}_{\text{skw}} = \mathbb{P}_{\text{skw}} \,, \qquad \leftarrow \text{see (2.61)} \tag{3.152b}$$

$$\mathbb{P}_{\text{sym}} : \mathbb{P}_{\text{skw}} = \mathbb{O} = \mathbb{P}_{\text{skw}} : \mathbb{P}_{\text{sym}} \,, \qquad \leftarrow \text{see (2.79h)} \tag{3.152c}$$

$$\mathbb{P}_{\text{sph}} : \mathbb{P}_{\text{sph}} = \mathbb{P}_{\text{sph}} \,, \qquad \leftarrow \text{see } (2.147\text{a})_1 \tag{3.152d}$$

$$\mathbb{P}_{\text{dev}} : \mathbb{P}_{\text{dev}} = \mathbb{P}_{\text{dev}} \,, \qquad \leftarrow \text{see } (2.147\text{b})_2 \tag{3.152e}$$

$$\mathbb{P}_{\text{sph}} : \mathbb{P}_{\text{dev}} = \mathbb{O} = \mathbb{P}_{\text{dev}} : \mathbb{P}_{\text{sph}} \,. \qquad \leftarrow \text{see} (2.147\text{a})_2 \text{ and } (2.147\text{b})_1 \tag{3.152f}$$

It is worthwhile to point out that the results (3.152e) and (3.152f) will be identical if $\mathbb{P}_{\text{dev}}$ is replaced by $\mathbb{P}^{\text{s}}_{\text{dev}}$ in (3.60).

Solution. To begin with, one needs to compute the double contraction between the fourth-order unit tensors. From (3.94b) to (3.94e) along with (3.57), one immediately obtains

$$\mathbb{I} : \mathbb{I} = \mathbb{I} \quad , \quad \mathbb{I} : \bar{\mathbb{I}} = \bar{\mathbb{I}} : \mathbb{I} = \bar{\mathbb{I}} \quad , \quad \bar{\mathbb{I}} : \bar{\mathbb{I}} = \mathbb{I} \,. \tag{3.153}$$

Note that the above results can also be obtained by following the similar procedures shown in the previous exercise. For instance,

$$\begin{aligned}
\mathbb{I} : \bar{\mathbb{I}} &= \mathbb{I}_{ijmn} \bar{\mathbb{I}}_{mnkl} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&= \delta_{im} \delta_{nj} \delta_{ml} \delta_{kn} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&= \delta_{il} \delta_{kj} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&= \bar{\mathbb{I}} \,.
\end{aligned}$$

Let's once more time show another result in (3.153) for the interested reader:

$$\left.\begin{aligned}
(\bar{\mathbb{I}} : \bar{\mathbb{I}}) : \mathbf{C} &\xlongequal[(3.85)]{\text{from}} \bar{\mathbb{I}} : (\bar{\mathbb{I}} : \mathbf{C}) \xlongequal[(3.110\text{a})]{\text{from}} \bar{\mathbb{I}} : \left(\bar{\mathbb{I}}^{\hat{\text{T}}} : \mathbf{C}^{\text{T}}\right) \\
&\xlongequal[(3.110\text{h})]{\text{from}} \bar{\mathbb{I}} : (\mathbb{I} : \mathbf{C}^{\text{T}}) \xlongequal[(3.37)]{\text{from}} \bar{\mathbb{I}} : \mathbf{C}^{\text{T}} \\
&\xlongequal[(3.110\text{a})]{\text{from}} \bar{\mathbb{I}}^{\hat{\text{T}}} : (\mathbf{C}^{\text{T}})^{\text{T}} \xlongequal[(2.55\text{b}) \text{ and } (3.110\text{h})]{\text{from}} \mathbb{I} : \mathbf{C}
\end{aligned}\right\} \xRightarrow[(3.38)]{\text{from}} \bar{\mathbb{I}} : \bar{\mathbb{I}} = \mathbb{I} \,.$$

Making use of the results (3.153), taking into account (3.89a) and (3.89b), the properties (3.152a) to (3.152c) can be shown in a straightforward manner. The results (3.153) can also help understand that the projection tensors $\mathbb{P}_{\text{sym}}$ and $\mathbb{P}_{\text{skw}}$ commute with the unit tensor $\bar{\mathbb{I}}$, that is,

$$\mathbb{P}_{\text{sym}} : \bar{\mathbb{I}} = \frac{1}{2}\left(\mathbb{I} + \bar{\mathbb{I}}\right) : \bar{\mathbb{I}} = \frac{1}{2}\left(\bar{\mathbb{I}} + \mathbb{I}\right) = \mathbb{P}_{\text{sym}}$$
$$= \bar{\mathbb{I}} : \mathbb{P}_{\text{sym}} \,, \tag{3.154a}$$
$$\mathbb{P}_{\text{skw}} : \bar{\mathbb{I}} = \frac{1}{2}\left(\mathbb{I} - \bar{\mathbb{I}}\right) : \bar{\mathbb{I}} = \frac{1}{2}\left(\bar{\mathbb{I}} - \mathbb{I}\right) = -\mathbb{P}_{\text{skw}}$$
$$= \bar{\mathbb{I}} : \mathbb{P}_{\text{skw}} \,. \tag{3.154b}$$

From $(2.90)_1$, $(2.90)_3$, (3.93a) and (3.97), one can verify that the square of the spherical projection tensor is equal to itself:

$$\mathbb{P}^2_{\text{sph}} = \mathbb{P}_{\text{sph}} : \mathbb{P}_{\text{sph}} = \frac{1}{3}\left(\mathbf{I} \otimes \mathbf{I}\right) : \frac{1}{3}\left(\mathbf{I} \otimes \mathbf{I}\right) = \frac{\text{tr}\mathbf{I}}{9}\mathbf{I} \otimes \mathbf{I} = \frac{1}{3}\mathbf{I} \otimes \mathbf{I} = \mathbb{P}_{\text{sph}} \,.$$

By means of (3.89a), (3.89b), $(3.95)_{1-2}$, (3.97), (3.152d) and $(3.153)_1$, one can arrive at the fifth desired relation:

$$\mathbb{P}^2_{\text{dev}} = \mathbb{P}_{\text{dev}} : \mathbb{P}_{\text{dev}} = \underbrace{\mathbb{I} : \mathbb{I}}_{= \mathbb{I}} - \underbrace{\mathbb{I} : \mathbb{P}_{\text{sph}}}_{= \mathbb{P}_{\text{sph}}} - \underbrace{\mathbb{P}_{\text{sph}} : \mathbb{I}}_{= \mathbb{P}_{\text{sph}}} + \underbrace{\mathbb{P}_{\text{sph}} : \mathbb{P}_{\text{sph}}}_{= \mathbb{P}_{\text{sph}}} = \mathbb{I} - \mathbb{P}_{\text{sph}} = \mathbb{P}_{\text{dev}} \,.$$

The expression $(3.152\text{f})_1$ regards the composition of the spherical and deviatoric projection tensors which renders

$$\mathbb{P}_{\text{sph}} : \mathbb{P}_{\text{dev}} = \mathbb{P}_{\text{sph}} : \mathbb{I} - \mathbb{P}_{\text{sph}} : \mathbb{P}_{\text{sph}} = \mathbb{P}_{\text{sph}} - \mathbb{P}_{\text{sph}} = \mathbb{O} \,,$$

and, in a similar manner, $\mathbb{P}_{\text{dev}} : \mathbb{P}_{\text{sph}} = \mathbb{O}$.

Exercise 3.3

Let $\mathbb{A}$ be an arbitrary fourth-order tensor and $\mathbb{C}$ be a super-symmetric tensor. Show that

$$\mathbb{P}_{\text{sym}} : \mathbb{A} : \mathbb{P}_{\text{sym}} = \underbrace{\frac{1}{4}\left(\mathbb{A} + \mathbb{A}^{\hat{\text{T}}} + \mathbb{A}^{\text{T}\hat{\text{T}}\text{T}} + \mathbb{A}^{\text{T}\hat{\text{T}}\text{T}\hat{\text{T}}}\right)}_{\text{or } \frac{1}{4}\left(\mathbb{A}_{ijkl} + \mathbb{A}_{ijlk} + \mathbb{A}_{jikl} + \mathbb{A}_{jilk}\right) = \left(\mathbb{P}_{\text{sym}} : \mathbb{A} : \mathbb{P}_{\text{sym}}\right)_{ijkl}} \,, \tag{3.155a}$$

$$\mathbb{P}_{\text{skw}} : \mathbb{A} : \mathbb{P}_{\text{skw}} = \underbrace{\frac{1}{4}\left(\mathbb{A} - \mathbb{A}^{\hat{\text{T}}} - \mathbb{A}^{\text{T}\hat{\text{T}}\text{T}} + \mathbb{A}^{\text{T}\hat{\text{T}}\text{T}\hat{\text{T}}}\right)}_{\text{or } \frac{1}{4}\left(\mathbb{A}_{ijkl} - \mathbb{A}_{ijlk} - \mathbb{A}_{jikl} + \mathbb{A}_{jilk}\right) = \left(\mathbb{P}_{\text{skw}} : \mathbb{A} : \mathbb{P}_{\text{skw}}\right)_{ijkl}} \,, \tag{3.155b}$$

$$\mathbb{P}_{\text{sym}} : \mathbb{C} : \mathbb{P}_{\text{sym}} = \underbrace{\mathbb{C}}_{\text{or } \mathbb{C} = \mathbb{C} : \mathbb{P}_{\text{sym}} = \mathbb{P}_{\text{sym}} : \mathbb{C}} \,. \tag{3.155c}$$

Make use of (3.155c) to represent any super-symmetric tensor $\mathbb{C}$ according to

$$\mathbb{C} = \frac{1}{2}\mathbb{C}_{ijkl}\left[\left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_k\right) \odot \left(\widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_j\right) + \left(\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k\right) \odot \left(\widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_i\right)\right] . \tag{3.156}$$

Solution. With the aid of $(3.58a)_3$, (3.89a), (3.89b), $(3.95)_{1-2}$, (3.105), (3.108), $(3.110i)_2$, $(3.110j)_3$ and $(3.110k)_3$, one can verify the first desired relation as follows:

$$\begin{aligned}
4\mathbb{P}_{\text{sym}} : \mathbb{A} : \mathbb{P}_{\text{sym}} &= \underbrace{\left(\mathbb{I} + \bar{\mathbb{I}}\right) : \mathbb{A} : \left(\mathbb{I} + \bar{\mathbb{I}}\right)}_{= (\mathbb{I} + \bar{\mathbb{I}}) : (\mathbb{A} : \mathbb{I} + \mathbb{A} : \bar{\mathbb{I}}) = (\mathbb{I} + \bar{\mathbb{I}}) : \left(\mathbb{A} + \mathbb{A}^{\widehat{\mathrm{T}}}\right)} \\
&= \underbrace{\mathbb{I} : \mathbb{A}}_{= \mathbb{A}} + \underbrace{\mathbb{I} : \mathbb{A}^{\widehat{\mathrm{T}}}}_{= \mathbb{A}^{\widehat{\mathrm{T}}}} + \underbrace{\bar{\mathbb{I}} : \mathbb{A}}_{= \mathbb{A}^{\mathrm{T}\widehat{\mathrm{T}}\mathrm{T}}} + \underbrace{\bar{\mathbb{I}} : \mathbb{A}^{\widehat{\mathrm{T}}}}_{= \mathbb{A}^{\mathrm{T}\widehat{\mathrm{T}}\mathrm{T}\widehat{\mathrm{T}}}} \\
&= \underbrace{\mathbb{A}}_{\text{with } \mathbb{A}_{ijkl}} + \underbrace{\mathbb{A}^{\widehat{\mathrm{T}}}}_{\text{with } \mathbb{A}_{ijlk}} + \underbrace{\mathbb{A}^{\mathrm{T}\widehat{\mathrm{T}}\mathrm{T}}}_{\text{with } \mathbb{A}_{jikl}} + \underbrace{\mathbb{A}^{\mathrm{T}\widehat{\mathrm{T}}\mathrm{T}\widehat{\mathrm{T}}}}_{\text{with } \mathbb{A}_{jilk}} .
\end{aligned}$$

By following similar procedures as shown above, one can verify (3.155b) in a straightforward manner.

Having in mind (3.115), the desired relation (3.155c) is basically a consequence of (3.155a). In this regard, (3.155c) in index notation represents

$$\mathbb{C}_{ijkl} = \frac{1}{4}\left(\mathbb{C}_{ijkl} + \mathbb{C}_{ijlk} + \mathbb{C}_{jikl} + \mathbb{C}_{jilk}\right) .$$

Finally, guided by $(3.49d)_2$ and the above result, the super-symmetric tensor $\mathbb{C}$ can be expressed with respect to the Cartesian basis (3.61) as

$$\begin{aligned}
\mathbb{C} &= \frac{1}{4}\left(\mathbb{C}_{ijkl} + \mathbb{C}_{ijlk} + \mathbb{C}_{jikl} + \mathbb{C}_{jilk}\right)\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l \\
&= \underbrace{\frac{1}{4}\left(\mathbb{C}_{ijkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l + \mathbb{C}_{ijkl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_k\right)}_{= \frac{1}{2}\mathbb{C}_{ijkl}\left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_k\right) \odot \left(\widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_j\right)} \\
&\quad + \underbrace{\frac{1}{4}\left(\mathbb{C}_{ijkl}\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l + \mathbb{C}_{ijkl}\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_k\right)}_{= \frac{1}{2}\mathbb{C}_{ijkl}\left(\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k\right) \odot \left(\widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_i\right)} .
\end{aligned}$$

Exercise 3.4

In the linearized theory of elasticity, the stress tensor $\boldsymbol{\sigma}_{\text{ist}}$ for homogeneous isotropic linear elastic materials reads

$$\boxed{\boldsymbol{\sigma}_{\text{ist}} = \lambda \left[\operatorname{tr} \boldsymbol{\varepsilon}_{\text{ist}}\right] \mathbf{I} + 2\mu \boldsymbol{\varepsilon}_{\text{ist}}\ ,} \tag{3.157}$$

where $\boldsymbol{\varepsilon}_{\text{ist}}$ is a **symmetric** second-order tensor known as the *infinitesimal strain* (or *small strain*) *tensor* and the coefficients λ and μ are called the *Lamé constants* (or *Lamé parameters*). These moduli are related to *Young's modulus* E and *Poisson's ratio* ν via the relations

$$\lambda = \frac{E\nu}{(1+\upsilon)(1-2\nu)} \quad , \quad \mu = \frac{E}{2(1+\upsilon)} \ , \tag{3.158}$$

and, conversely,

$$E = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu} \quad , \quad \nu = \frac{\lambda}{2(\lambda+\mu)} \ . \tag{3.159}$$

The stress-strain relation (3.157) is known as the *generalized Hooke's law* for isotropic materials within the realm of *infinitesimal strain theory*. It can also be demonstrated as $\boldsymbol{\sigma}_{\text{ist}} = \mathbb{C}_{\text{ist}} : \boldsymbol{\varepsilon}_{\text{ist}}$ where $\mathbb{C}_{\text{ist}}$ is a fourth-order tensor referred to as the *elasticity tensor* (or *elastic stiffness tensor*). Readers who need an in-depth treatment of theory of elasticity are referred to some standard texts such as Ciarlet [12], Slaughter [13] and Sadd [14].

First, represent $\mathbb{C}_{\text{ist}}$ and then obtain its inverse called the *compliance elasticity tensor* (or *elastic compliance tensor*). Apparently, the fourth-order tensor $\mathbb{C}_{\text{ist}}^{-1}$ can operate on $\boldsymbol{\sigma}_{\text{ist}}$ to deliver $\boldsymbol{\varepsilon}_{\text{ist}}$, i.e. $\boldsymbol{\varepsilon}_{\text{ist}} = \mathbb{C}_{\text{ist}}^{-1} : \boldsymbol{\sigma}_{\text{ist}}$.

Finally, express the resulting expressions $\boldsymbol{\sigma}_{\text{ist}} = \mathbb{C}_{\text{ist}} : \boldsymbol{\varepsilon}_{\text{ist}}$ and $\boldsymbol{\varepsilon}_{\text{ist}} = \mathbb{C}_{\text{ist}}^{-1} : \boldsymbol{\sigma}_{\text{ist}}$ in Voigt notation.

Solution. The elasticity tensor basically presents the sensitivity of the stress field with respect to the strain measure. And it is usually computed by means of the chain rule of differentiation, see Exercise 6.16. See also Exercise 6.17 for its numerical differentiation. Here, there is no need to use differentiation due to the **linear** structure of (3.157). By rewriting (3.157) in indicial form, taking into account (1.36), (2.33), (2.49), $(2.59)_1$, $(2.61)_1$, $(2.89\text{a})_1$, $(3.54\text{a})_1$ and $(3.54\text{d})_1$, the desired elasticity tensor can be extracted from

$$\begin{aligned}
(\boldsymbol{\sigma}_{\text{ist}})_{ij} &= \underbrace{\lambda\left[\delta_{kl}\,(\boldsymbol{\varepsilon}_{\text{ist}})_{kl}\right]\delta_{ij} + \mu\,(\boldsymbol{\varepsilon}_{\text{ist}})_{ij} + \mu\,(\boldsymbol{\varepsilon}_{\text{ist}})_{ji}}_{=\lambda\delta_{kl}(\boldsymbol{\varepsilon}_{\text{ist}})_{kl}\delta_{ij} + \mu\delta_{ik}\delta_{lj}(\boldsymbol{\varepsilon}_{\text{ist}})_{kl} + \mu\delta_{kj}\delta_{il}(\boldsymbol{\varepsilon}_{\text{ist}})_{kl}} \\
&= \underbrace{\left[\lambda\delta_{ij}\delta_{kl} + \mu\left(\delta_{ik}\delta_{lj} + \delta_{il}\delta_{kj}\right)\right](\boldsymbol{\varepsilon}_{\text{ist}})_{kl}}_{=\left[\lambda(\mathbf{I}\otimes\mathbf{I})_{ijkl} + 2\mu(\mathbf{I}\odot\mathbf{I})_{ijkl}\right](\boldsymbol{\varepsilon}_{\text{ist}})_{kl}} \ .
\end{aligned}$$

Thus,

$$\boxed{\mathbb{C}_{\text{ist}} = \lambda\mathbf{I}\otimes\mathbf{I} + 2\mu\mathbf{I}\odot\mathbf{I}\ .} \quad \longleftarrow \text{see (6.182)} \tag{3.160}$$

To proceed, one needs to express $\boldsymbol{\varepsilon}_{\text{ist}}$ in terms of $\boldsymbol{\sigma}_{\text{ist}}$. Having in mind the identity $\text{tr}\mathbf{I} = 3$, the trace of (3.157) gives

$$\text{tr}\,\boldsymbol{\sigma}_{\text{ist}} = 3\lambda\,(\text{tr}\,\boldsymbol{\varepsilon}_{\text{ist}}) + 2\mu\text{tr}\,\boldsymbol{\varepsilon}_{\text{ist}} \quad \Longrightarrow \quad \text{tr}\,\boldsymbol{\varepsilon}_{\text{ist}} = \frac{\text{tr}\,\boldsymbol{\sigma}_{\text{ist}}}{3\lambda+2\mu}\ .$$

It then follows that

$$\begin{aligned}\boldsymbol{\varepsilon}_{\text{ist}} &= \frac{1}{2\mu}\boldsymbol{\sigma}_{\text{ist}} - \frac{\lambda}{2\mu}\left[\text{tr}\,\boldsymbol{\varepsilon}_{\text{ist}}\right]\mathbf{I} \\ &= -\frac{\lambda \text{tr}\,\boldsymbol{\sigma}_{\text{ist}}}{2\mu\left(3\lambda+2\mu\right)}\mathbf{I} + \frac{1}{2\mu}\boldsymbol{\sigma}_{\text{ist}} \\ &\overset{\text{from}}{\underset{(3.158)}{=\!=}} -\frac{\nu \text{tr}\,\boldsymbol{\sigma}_{\text{ist}}}{E}\mathbf{I} + \frac{1+\nu}{E}\boldsymbol{\sigma}_{\text{ist}} \,. \end{aligned} \tag{3.161}$$

Following similar procedures which led to (3.160) then provides

$$\begin{aligned}\mathbb{C}_{\text{ist}}^{-1} &= -\frac{\lambda}{2\mu\left(3\lambda+2\mu\right)}\mathbf{I}\otimes\mathbf{I} + \frac{1}{2\mu}\mathbf{I}\odot\mathbf{I} \\ &= -\frac{\nu}{E}\mathbf{I}\otimes\mathbf{I} + \frac{1+\nu}{E}\mathbf{I}\odot\mathbf{I}\,. \end{aligned} \tag{3.162}$$

Notice that

$$\begin{aligned}\mathbb{C}_{\text{ist}} : \mathbb{C}_{\text{ist}}^{-1} &= \left(\lambda\mathbf{I}\otimes\mathbf{I} + 2\mu\mathbf{I}\odot\mathbf{I}\right) : \left(-\frac{\lambda\mathbf{I}\otimes\mathbf{I}}{2\mu\left(3\lambda+2\mu\right)} + \frac{\mathbf{I}\odot\mathbf{I}}{2\mu}\right) \\ &= -\frac{3\lambda^2\mathbf{I}\otimes\mathbf{I}}{2\mu\left(3\lambda+2\mu\right)} + \frac{\lambda\left(3\lambda+2\mu\right)\mathbf{I}\otimes\mathbf{I}}{2\mu\left(3\lambda+2\mu\right)} - \frac{2\mu\lambda\mathbf{I}\otimes\mathbf{I}}{2\mu\left(3\lambda+2\mu\right)} + \frac{2\mu}{2\mu}\mathbf{I}\odot\mathbf{I} \\ &= \mathbf{I}\odot\mathbf{I}\,. \quad \longleftarrow \text{see (3.149)}\end{aligned}$$

At the end, using (3.127) and $(3.130)_2$, the linear mapping $\boldsymbol{\sigma}_{\text{ist}} = \mathbb{C}_{\text{ist}} : \boldsymbol{\varepsilon}_{\text{ist}}$ in Voigt notation renders

$$\begin{bmatrix} (\boldsymbol{\sigma}_{\text{ist}})_{11} \\ (\boldsymbol{\sigma}_{\text{ist}})_{22} \\ (\boldsymbol{\sigma}_{\text{ist}})_{33} \\ (\boldsymbol{\sigma}_{\text{ist}})_{23} \\ (\boldsymbol{\sigma}_{\text{ist}})_{13} \\ (\boldsymbol{\sigma}_{\text{ist}})_{12} \end{bmatrix} = \begin{bmatrix} 2\mu+\lambda & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & 2\mu+\lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & 2\mu+\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} \begin{bmatrix} (\boldsymbol{\varepsilon}_{\text{ist}})_{11} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{22} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{33} \\ 2\,(\boldsymbol{\varepsilon}_{\text{ist}})_{23} \\ 2\,(\boldsymbol{\varepsilon}_{\text{ist}})_{13} \\ 2\,(\boldsymbol{\varepsilon}_{\text{ist}})_{12} \end{bmatrix} . \tag{3.163}$$

And, in a similar manner, the linear function $\boldsymbol{\varepsilon}_{\text{ist}} = \mathbb{C}_{\text{ist}}^{-1} : \boldsymbol{\sigma}_{\text{ist}}$ takes the form

$$\begin{bmatrix} (\boldsymbol{\varepsilon}_{\text{ist}})_{11} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{22} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{33} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{23} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{13} \\ (\boldsymbol{\varepsilon}_{\text{ist}})_{12} \end{bmatrix} = \frac{1}{E} \begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1+\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1+\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1+\nu}{2} \end{bmatrix} \begin{bmatrix} (\boldsymbol{\sigma}_{\text{ist}})_{11} \\ (\boldsymbol{\sigma}_{\text{ist}})_{22} \\ (\boldsymbol{\sigma}_{\text{ist}})_{33} \\ 2\,(\boldsymbol{\sigma}_{\text{ist}})_{23} \\ 2\,(\boldsymbol{\sigma}_{\text{ist}})_{13} \\ 2\,(\boldsymbol{\sigma}_{\text{ist}})_{12} \end{bmatrix} . \tag{3.164}$$

References

1. Gurtin, M. E. (1981). *An introduction to continuum mechanics*. Academic Press.
2. Chandrasekharaiah, D. S., & Debnath, L. (1994). *Continuum mechanics*. Academic Press.
3. Negahban, M. (2012). *The mechanical and thermodynamical theory of plasticity*. CRC Press.
4. Jog, C. S. (2015). *Continuum mechanics: Foundations and applications of mechanics* (3rd ed., Vol. I). Cambridge: Cambridge University Press.
5. Irgens, F. (2008). *Continuum mechanics*. Berlin, Heidelberg: Springer.
6. Itskov, M. (2000). On the theory of fourth-order tensors and their applications in computational mechanics. *Computer Methods in Applied Mechanics and Engineering, 189*, 419–438.
7. Itskov, M. (2003). Computation of the exponential and other isotropic tensor functions and their derivatives. *Computer Methods in Applied Mechanics and Engineering, 192*, 3985–3999.
8. Del Piero, G. (1979). Some properties of the set of fourth-order tensors, with application to elasticity. *Journal of Elasticity, 9*, 245–261.
9. de Souza Neto, E. A., Perić, D., & Owen, D. R. J. (2008). *Computational methods for plasticity: Theory and applications*. Wiley.
10. Holzapfel, G. A. (2000). *Nonlinear solid mechanics: A continuum approach for engineering*. Wiley.
11. Steinmann, P. (2015). *Geometrical foundations of continuum mechanics: An application to first- and second-order elasticity and elasto-plasticity (Lecture Notes in Applied Mathematics and Mechanics)*. Berlin, Heidelberg: Springer.
12. Ciarlet, P. G. (1988). *Mathematical elasticity, volume I: Three-Dimensional elasticity*. Amsterdam: Elsevier Science.
13. Slaughter, W. S. (2002). *The linearized theory of elasticity*. Boston, MA: Birkhäuser.
14. Sadd, M. H. (2009). *Elasticity: Theory, applications, and numerics*. Academic Press.

Chapter 4
Eigenvalues, Eigenvectors and Spectral Decompositions of Tensors

Eigenvalues (or *characteristic values* or *principal values*) and *eigenvectors* (or *principal axes* or *principal directions*) are extensively used in many branches of physics and engineering; examples of which include quantum mechanics, control theory and stability analysis. They represent mathematical objects that are associated with a linear transformation in the realm of linear algebra. Since the components and basis vectors of a second-order tensor vary from one coordinate system to another, it may be beneficial to seek and find some certain values and directions corresponding to that tensor which remain invariant under the coordinate transformations. These special properties of a tensor that basically illustrate its invariant nature are known as the eigenvalues and eigenvectors.

A tensor can be diagonalized when it is *non-defective* or *diagonalizable*. A representation of a non-defective tensor in terms of its eigenvalues and eigenvectors is known as the *spectral decomposition* of that tensor which is extremely important from computational point of view. The aim of this chapter is thus to characterize the eigenvalues and eigenvectors of second-order tensors due to their great importance in scientific and engineering problems. Spectral decomposition of a symmetric tensor is also studied.

Decent books specifically dedicated to the eigenvalue problem and its applications are, for instance, due to Wilkinson [1], Parlett [2], Saad [3] and Qi et al. [4].

4.1 Eigenvalue Problem

The scalars $\lambda \in \mathrm{C}$ will be the eigenvalues of a tensor $\mathbf{A} \in \mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}\right)$ if there exists corresponding nonzero vectors $\mathbf{n} \in \mathscr{E}_{\mathrm{c}}^{\mathrm{o}3}$ such that

$$\boxed{\mathbf{A}\,\mathbf{n} = \lambda \mathbf{n}}\ , \tag{4.1}$$

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_4

where $\mathbf{n}$ present the eigenvectors of $\mathbf{A}$. An expression of the form (4.1) is called the **eigenvalue problem**. As can be seen, once $\mathbf{n}$ undergo this linear transformation, their directions are preserved (or possibly reversed) while their magnitudes will be changed by the factors $|\lambda|$. Moreover, they are obviously independent of any coordinate system. It is important to note that although all second-order tensors in this text have been declared as real, their eigenvalues and eigenvectors may generally not be real in accord with matrix algebra. For instance, orthogonal as well as skew tensors exhibit complex eigenvalues and eigenvectors, see Exercise 4.9.

Suppose that a tensor $\mathbf{A}$ possesses three **distinct** (or **non-degenerate**) eigenvalues $\lambda_i,\ i = 1, 2, 3$. For each eigenvalue there exists a corresponding eigenvector and, therefore, each of the **eigenpairs** $(\lambda_i, \mathbf{n}_i)\ ,\ i = 1, 2, 3$, satisfies the eigenvalue problem. It is often seen that the eigenvalue problem is represented by the **normalized eigenvectors** $\widehat{\mathbf{n}}$. Indeed, there are infinitely many eigenvectors associated with an eigenvalue since any nonzero scalar multiple of an eigenvector is still an eigenvector. For tensors with distinct eigenvalues, it can be shown that the set of eigenvectors form a basis called **eigenbasis**. But the set of eigenvectors, in general, may not constitute an eigenbasis. See, for instance, Exercise 4.5 wherein the eigenvectors of the so-called *defective* tensors are unable to produce a basis. By use of the so-called *generalized eigenvectors*, however, one can extend the eigenvectors of defective tensors to have an eigenbasis.

The eigenvalues and eigenvectors of a tensor are computationally the eigenvalues and eigenvectors of its matrix form. It is exactly this matrix form served as an input for available computer codes calculating the eigenvalues and eigenvectors. With regard to the eigenvalue problem, one should realize that

- ☛ any nonzero vector constructed from a linear combination of eigenvectors that correspond to the same eigenvalue will again be an eigenvector:

$$\begin{aligned} &\text{if} && \mathbf{A}\,\mathbf{n} = \lambda\mathbf{n}\ ,\ \ \mathbf{A}\,\mathbf{m} = \lambda\mathbf{m}, \\ &\text{then} && \mathbf{A}\,(\alpha\mathbf{n} + \beta\mathbf{m}) = \alpha\mathbf{A}\,\mathbf{n} + \beta\mathbf{A}\,\mathbf{m} = \lambda\,(\alpha\mathbf{n} + \beta\mathbf{m})\,. \end{aligned}$$

- ☛ any two tensors $\mathbf{A}$ and $\mathbf{B}$ that are related, making use of an invertible tensor $\mathbf{M}$, via $\mathbf{B} = \mathbf{M}\mathbf{A}\mathbf{M}^{-1}$ have identical eigenvalues while their eigenvectors will change by that linear mapping[1]:

$$\begin{aligned} &\text{if} \quad \mathbf{A}\,\mathbf{n} = \lambda\mathbf{n}, \\ &\text{then}\ \mathbf{A}\left(\mathbf{M}^{-1}\mathbf{M}\right)\mathbf{n} = \lambda\mathbf{n} \quad \Rightarrow \quad \mathbf{B}\,(\mathbf{M}\mathbf{n}) = \lambda\,(\mathbf{M}\mathbf{n})\,. \end{aligned}$$

The eigenvalue problem $\mathbf{A}\,\mathbf{n} = \lambda\mathbf{n}$ basically renders a **right mapping** and the corresponding eigenvectors $\mathbf{n}$ can accordingly be viewed as the **right eigenvectors**. With regard to this, one can also define the **left eigenvalue problem**

$$\boxed{\mathbf{m}\,\mathbf{A} = \lambda\mathbf{m}\ ,} \tag{4.2}$$

[1] Note that in the context of matrix algebra, two matrices $[\mathbf{A}]$ and $[\mathbf{M}]\,[\mathbf{A}]\,[\mathbf{M}]^{-1}$ are referred to as *similar matrices*.

in which the nonzero vectors **m** denote the **left eigenvectors**. Throughout the developments, it will be seen that $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ have identical eigenvalues. Having this in mind and taking (2.51a) into account, one can deduce that the right eigenvectors of a tensor are the left eigenvectors of its transpose and vice versa. In this text, all eigenvectors should be regarded as right eigenvectors, if not otherwise stated.

From the eigenvalue problem (4.1), one can obtain $\mathbf{A}^2\,\mathbf{n} = \mathbf{A}\,(\lambda\mathbf{n}) = \lambda\mathbf{A}\,\mathbf{n} = \lambda^2\,\mathbf{n}$. Thus,

$$\boxed{\mathbf{A}^k\,\mathbf{n} = \lambda^k\,\mathbf{n}, \quad k = 0, 1, 2, \ldots .} \tag{4.3}$$

This result immediately implies that any eigenvalue of a tensor polynomial of $\mathbf{A}$ according to $\mathbf{H}(\mathbf{A}) = \sum_{k=0}^{n} \alpha_k \mathbf{A}^k$ should be of the form $H(\lambda) = \sum_{k=0}^{n} \alpha_k \lambda^k$ owing to

$$\mathbf{H}(\mathbf{A})\,\mathbf{n} = \sum_{k=0}^{n} \alpha_k \left(\mathbf{A}^k\mathbf{n}\right) = \sum_{k=0}^{n} \alpha_k \left(\lambda^k\mathbf{n}\right) = \left(\sum_{k=0}^{n} \alpha_k\lambda^k\right)\mathbf{n} = H(\lambda)\,\mathbf{n} .$$

4.2 Characteristic Equation and Principal Scalar Invariants

By means of the expression (2.5), the eigenvalue problem (4.1) can be rewritten as $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{n} = 0$. From linear algebra, it is evident that such a homogeneous equation has a nontrivial solution, i.e. $\mathbf{n} \neq 0$, if

$$\boxed{\det(\mathbf{A} - \lambda\mathbf{I}) = 0 ,} \tag{4.4}$$

or, equivalently,

$$\det\begin{bmatrix} A_{11} - \lambda & A_{12} & A_{13} \\ A_{21} & A_{22} - \lambda & A_{23} \\ A_{31} & A_{32} & A_{33} - \lambda \end{bmatrix} = 0 . \tag{4.5}$$

The expression (4.4) is referred to as the *characteristic determinant* of $\mathbf{A}$. It can also be represented in a more convenient form as follows:

$$\boxed{\underbrace{\lambda^3 - I_1(\mathbf{A})\,\lambda^2 + I_2(\mathbf{A})\,\lambda - I_3(\mathbf{A})}_{:=\,p_{\mathbf{A}}(\lambda)} = 0 .} \tag{4.6}$$

This is known as the *characteristic equation* for $\mathbf{A}$ and its left hand side $p_{\mathbf{A}}(\lambda)$ is called the *characteristic polynomial* of $\mathbf{A}$. Here, the scalars $I_k,\ k = 1, 2, 3$, present the *principal scalar invariants* of $\mathbf{A}$. They are given by

$$I_1(\mathbf{A}) = A_{11} + A_{22} + A_{33}$$
$$= \boxed{\frac{1}{2}\varepsilon_{qrn}\varepsilon_{qrs}A_{ns} \xlongequal[\text{(2.89a)}]{\text{from}} A_{nn}\ ,} \tag{4.7a}$$
$$I_2(\mathbf{A}) = A_{11}A_{22} - A_{21}A_{12} + A_{11}A_{33} - A_{31}A_{13} + A_{22}A_{33} - A_{32}A_{23}$$
$$= \boxed{\frac{1}{2}\varepsilon_{qmn}\varepsilon_{qrs}A_{mr}A_{ns}\ ,} \tag{4.7b}$$
$$I_3(\mathbf{A}) = A_{11}(A_{22}A_{33} - A_{32}A_{23}) - A_{21}(A_{12}A_{33} - A_{32}A_{13}) + A_{31}(A_{12}A_{23} - A_{22}A_{13})$$
$$= \boxed{\frac{1}{6}\varepsilon_{lmn}\varepsilon_{qrs}A_{lq}A_{mr}A_{ns} \xlongequal[\text{(1.80)}]{\text{from}} \det \mathbf{A}\ .} \tag{4.7c}$$

For a symmetric tensor $\mathbf{S}$, they render

$$I_1(\mathbf{S}) = S_{11} + S_{22} + S_{33}\ , \tag{4.8a}$$
$$I_2(\mathbf{S}) = S_{11}S_{22} + S_{11}S_{33} + S_{22}S_{33} - S_{12}^2 - S_{13}^2 - S_{23}^2\ , \tag{4.8b}$$
$$I_3(\mathbf{S}) = S_{11}S_{22}S_{33} + 2S_{12}S_{13}S_{23} - S_{23}^2S_{11} - S_{13}^2S_{22} - S_{12}^2S_{33}\ . \tag{4.8c}$$

The characteristic equation (4.6) represents a polynomial of order three in accord with the dimension of the vector space $\mathscr{E}_{\mathrm{r}}^{o3}$. Its roots are basically the principal values which satisfy the eigenvalue problem. It can analytically be solved by a well-known method often attributed to Cardano.[2] It is worthwhile to point out that the

[2] The Cardano's formula is briefly discussed here. Consider a general cubic equation of the form

$$\lambda^3 + a\lambda^2 + b\lambda + c = 0\ ,$$

where the coefficients a, b and c are complex numbers. By use of the substitution $\lambda = -a/3 + \widetilde{\lambda}$, this equation will be reduced to

$$\widetilde{\lambda}^3 + 3Q\widetilde{\lambda} - 2R = 0\ , \quad \text{where}\ \ 9Q = 3b - a^2\ \ ,\ \ 54R = 9ab - 27c - 2a^3\ .$$

It is not then difficult to see that the above useful form can be rewritten according to

$$\left(\widetilde{\lambda} - B\right)\left[\widetilde{\lambda}^2 + B\widetilde{\lambda} + \left(B^2 + 3Q\right)\right] = 0\ , \quad \text{where}\ \ B = \sqrt[3]{R + \sqrt{Q^3 + R^2}} + \sqrt[3]{R - \sqrt{Q^3 + R^2}}\ .$$

The solutions of this equation thus provides the three roots

$$\widetilde{\lambda}_1 = S + T\ \ ,\ \ \widetilde{\lambda}_2 = -\frac{S+T}{2} + \mathrm{i}\frac{\sqrt{3}}{2}(S - T)\ \ ,\ \ \widetilde{\lambda}_3 = -\frac{S+T}{2} - \mathrm{i}\frac{\sqrt{3}}{2}(S - T)\ ,$$

where

$$S = \sqrt[3]{R + \sqrt{D}}\ \ ,\ \ T = \sqrt[3]{R - \sqrt{D}}\ \ ,\ \ D = Q^3 + R^2\ .$$

The roots of the original cubic equation is accordingly represented by $\lambda_k = -a/3 + \widetilde{\lambda}_k,\ k = 1, 2, 3$. Note that in case of real coefficients, i.e. $a\ ,\ b\ ,\ c \in \mathrm{R}$, the sign of the discriminant D determines the nature of roots being real or imaginary. For this case, the cubic equation will have roots $\lambda_k, \ k = 1, 2, 3,$ fundamentally exhibit complex numbers. These roots can be represented in some alternative forms. A popular representation, often seen in the literature based on the trigonometric functions, is given by (Bronshtein et al. [5])

$$\lambda_k = \begin{cases} \frac{1}{3}\left[I_1 + 2\sqrt{I_1^2 - 3I_2}\cos\frac{\theta+2\pi(k-1)}{3}\right] & \text{if } I_1^2 - 3I_2 \neq 0 \\ \frac{1}{3}I_1 + \frac{1}{3}\sqrt[3]{27I_3 - I_1^3}\exp\left(\mathrm{i}\frac{2\pi k}{3}\right) & \text{if } I_1^2 - 3I_2 = 0 \end{cases} , \quad k = 1, 2, 3, \tag{4.9}$$

where

$$\theta = \arccos\frac{2I_1^3 - 9I_1I_2 + 27I_3}{2\sqrt{\left(I_1^2 - 3I_2\right)^3}} \quad , \quad \exp\left(\mathrm{i}\frac{2\pi k}{3}\right) = \cos\frac{2\pi k}{3} + \mathrm{i}\sin\frac{2\pi k}{3} \ , \tag{4.10}$$

and the imaginary unit i satisfies $\mathrm{i}^2 + 1 = 0$.

The cubic equation (4.6) can now be rephrased as

$$\boxed{p_{\mathbf{A}}(\lambda) = \prod_{k=1}^{r}(\lambda - \lambda_k)^{s_k} = 0 \ ,} \tag{4.11}$$

where $r \ (1 \leq r \leq 3)$ is the number of **distinct** eigenvalues and s_k presents the so-called *algebraic multiplicity* of the principal value λ_k. The so-called *geometric multiplicity* t_k of an eigenvalue λ_k, on the other hand, determines the respective number of **linearly independent** eigenvectors. Indicating by

$$\mathscr{N}(\mathbf{A}) = \left\{\mathbf{n} \in \mathscr{E}_{\mathrm{r}}^{o3} \mid (\mathbf{A} - 0\mathbf{I})\,\mathbf{n} = \mathbf{0}\right\} \ , \tag{4.12}$$

the *null space* (or *kernel*) of **A**, the geometric multiplicity of λ_k is basically the dimension of its so-called *eigenspace* (or *characteristic space*) $\mathscr{N}(\mathbf{A} - \lambda_k\mathbf{I})$, that is, $t_k = \dim\mathscr{N}(\mathbf{A} - \lambda_k\mathbf{I})$. The span of eigenvectors corresponding to an eigenvalue λ_k basically defines the eigenspace associated with that eigenvalue, i.e.

$$\mathscr{N}(\mathbf{A} - \lambda_k\mathbf{I}) = \mathrm{Span}\{\mathbf{n}_{k1}, \cdots, \mathbf{n}_{kl}\} \ , \quad k = 1, \cdots, r\,; \ l = t_k \ . \tag{4.13}$$

For sake of convenience, let's agree to write the eigenvectors of a tensor with distinct eigenvalues as $\mathbf{n}_{11} = \mathbf{n}_1$, $\mathbf{n}_{21} = \mathbf{n}_1$ and $\mathbf{n}_{31} = \mathbf{n}_3$.

If $t_k = s_k, \ k = 1, 2, 3,$ there is no deficiency and the eigenvectors of **A** form an eigenbasis, otherwise **A** is **defective** (in this case the geometric multiplicity of

---

☆ three distinct real roots when $D < 0$;
☆ three real roots with repetitions when $D = 0$ (if $Q \neq 0$ and $R \neq 0$, there is one single and one double root; and if $Q = R = 0$, there is simply one triple root) ;
☆ one real root and a pair of complex conjugates when $D > 0$.

See, e.g., Birkhoff and Mac Lane [6] for more details.

an eigenvalue is **smaller** than its algebraic multiplicity). A set of eigenvalues of a tensor **A** in which any eigenvalue is repeated according to its algebraic multiplicity is referred to as the *spectrum* of **A** and denoted here by spec (**A**). These objects are illustrated in the following example. ✌

Let **D**, **E** and **F** be three Cartesian tensors whose matrices are

$$[\mathbf{D}] = \begin{bmatrix} -2 & -4 & 2 \\ -2 & 1 & 2 \\ 4 & 2 & 5 \end{bmatrix} \quad , \quad [\mathbf{E}] = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix} \quad , \quad [\mathbf{F}] = \begin{bmatrix} 1 & 0.3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Their characteristic polynomials are given by

$$p_{\mathbf{D}}(\lambda) = (\lambda - 3)(\lambda + 5)(\lambda - 6) \ ,$$
$$p_{\mathbf{E}}(\lambda) = \lambda^2 (\lambda - 9) \ ,$$
$$p_{\mathbf{F}}(\lambda) = (\lambda - 1)^3 \ .$$

Accordingly, the algebraic multiplicities of the three distinct characteristic values $\lambda = 3, -5, 6$ of **D** are $s = 1, 1, 1$, respectively. If s_k of λ_k is unity, then this principal value is basically **non-multiple**. Such an eigenvalue is thus referred to as *simple eigenvalue*. In this regard, **E** (**F**) is a tensor with one **multiple** eigenvalue for which the algebraic multiplicities of $\lambda = 0, 9$ ($\lambda = 1$) are $s = 2, 1$ ($s = 3$), respectively.

The spectrum of each of these tensor should now be clear:

$$\text{spec}(\mathbf{D}) = \{3, -5, 6\} \quad , \quad \text{spec}(\mathbf{E}) = \{0, 0, 9\} \quad , \quad \text{spec}(\mathbf{F}) = \{1, 1, 1\} \ .$$

By elementary row reduction, one can finally arrive at the eigenspaces

$$\begin{aligned}
\mathcal{N}(\mathbf{D} - 3\mathbf{I}) &= \text{Span}\{-2\widehat{\mathbf{e}}_1 + 3\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3\} &&, \ t_1 = 1 \ , \\
\mathcal{N}(\mathbf{D} + 5\mathbf{I}) &= \text{Span}\{-2\widehat{\mathbf{e}}_1 - \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_1\} &&, \ t_2 = 1 \ , \\
\mathcal{N}(\mathbf{D} - 6\mathbf{I}) &= \text{Span}\left\{\tfrac{1}{16}\widehat{\mathbf{e}}_1 + \tfrac{3}{8}\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3\right\} &&, \ t_3 = 1 \ , \\
\mathcal{N}(\mathbf{E}) &= \text{Span}\{-\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 \ , \ -\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_3\} &&, \ t_1 = 2 \ , \\
\mathcal{N}(\mathbf{E} - 9\mathbf{I}) &= \text{Span}\{\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3\} &&, \ t_2 = 1 \ , \\
\mathcal{N}(\mathbf{F} - \mathbf{I}) &= \text{Span}\{\widehat{\mathbf{e}}_1 \ , \ \widehat{\mathbf{e}}_3\} &&, \ t \ = 2 \ .
\end{aligned}$$

One can now deduce that the eigenvectors of **D** form an eigenbasis. This also holds true for **E**. But, this is not the case for **F** since there does not exist three linearly independent eigenvectors associated with the triple root $\lambda = 1$. In this case, the root $\lambda = 1$ is referred to as the *defective eigenvalue*. ✌

The roots λ_k, $k = 1, 2, 3$, help obtain the coefficients I_k, $k = 1, 2, 3$, in (4.6) according to

$$I_1(\mathbf{A}) = \lambda_1(\mathbf{A}) + \lambda_2(\mathbf{A}) + \lambda_3(\mathbf{A}) \ , \tag{4.14a}$$

$$I_2(\mathbf{A}) = \lambda_1(\mathbf{A})\,\lambda_2(\mathbf{A}) + \lambda_2(\mathbf{A})\,\lambda_3(\mathbf{A}) + \lambda_1(\mathbf{A})\,\lambda_3(\mathbf{A}) \ , \tag{4.14b}$$

$$I_3(\mathbf{A}) = \lambda_1(\mathbf{A})\,\lambda_2(\mathbf{A})\,\lambda_3(\mathbf{A}) \ . \tag{4.14c}$$

They are known as the *Vieta's formulas* (Hazewinkel [7]). Notice that these quantities remain invariant under any permutation of $(1, 2, 3)$. As can be seen from (4.7a)–(4.7c) and (4.14a)–(4.14c), the principal values of a tensor admit some alternative forms. The following decent form, often seen in the literature in advance, is yet another representation of the principal scalar invariants

$$I_1(\mathbf{A}) = \frac{\mathbf{Au}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{Av}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{v}\times\mathbf{Aw})}{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})} \ , \tag{4.15a}$$

$$I_2(\mathbf{A}) = \frac{\mathbf{Au}\cdot(\mathbf{Av}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{Av}\times\mathbf{Aw}) + \mathbf{Au}\cdot(\mathbf{v}\times\mathbf{Aw})}{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})} \ , \tag{4.15b}$$

$$I_3(\mathbf{A}) = \frac{\mathbf{Au}\cdot(\mathbf{Av}\times\mathbf{Aw})}{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})} \overset{\text{from}}{\underset{(2.98)}{=\!=\!=}} \det\mathbf{A} \ , \quad \longleftarrow \ \text{the proof is given in Exercise 4.2} \tag{4.15c}$$

where the three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ constitute an arbitrary basis.[3]

A very practical form of the principal invariants is based on the following recursive formula

$$I_k(\mathbf{A}) = \frac{(-1)^{k+1}}{k}\mathrm{tr}\left(\mathbf{A}^k + \sum_{j=1}^{k-1}(-1)^j\, I_j(\mathbf{A})\,\mathbf{A}^{k-j}\right) \ , \quad k = 1, 2, 3 \ . \tag{4.16}$$

This is known as the *Newton's identity*.[4] By means of (2.36) and $(2.83)_2$, it then provides

$$I_1(\mathbf{A}) = \boxed{\mathrm{tr}\mathbf{A} \ ,} \tag{4.17a}$$

$$I_2(\mathbf{A}) = -\frac{1}{2}\left[\mathrm{tr}\mathbf{A}^2 - I_1\mathrm{tr}\mathbf{A}\right]$$

$$= \boxed{\frac{1}{2}\left[(\mathrm{tr}\mathbf{A})^2 - \mathrm{tr}\mathbf{A}^2\right] \ ,} \tag{4.17b}$$

---

[3] The three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ were naturally assumed to be linearly independent in order to formally define the principal invariants of a tensor $\mathbf{A}$. However, note that, for instance, an expression of the form

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\,I_1(\mathbf{A}) = \mathbf{Au}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{Av}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{v}\times\mathbf{Aw}) \ ,$$

is always true whether these vectors form a basis or not.

[4] Note that the Newton's identity also remains true for finite-dimensional vector spaces.

$$I_3(\mathbf{A}) = \frac{1}{3}\left[\mathrm{tr}\mathbf{A}^3 - I_1\mathrm{tr}\mathbf{A}^2 + I_2\mathrm{tr}\mathbf{A}\right]$$

$$= \boxed{\frac{1}{6}\left[2\mathrm{tr}\mathbf{A}^3 - 3(\mathrm{tr}\mathbf{A})\left(\mathrm{tr}\mathbf{A}^2\right) + (\mathrm{tr}\mathbf{A})^3\right]\ .} \tag{4.17c}$$

With the aid of $(2.86)_3$ and $(2.89\text{d})_3$, one can deduce from (4.17a) to (4.17c) that

$$\boxed{I_k\left(\mathbf{A}^{\mathrm{T}}\right) = I_k(\mathbf{A})\ ,\quad k = 1, 2, 3\ ,} \tag{4.18}$$

which reveals the fact that the tensors $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}$ have identical eigenvalues.

The second invariant of a tensor, according to $(4.17\text{b})_2$, can also be expressed in terms of its cofactor:

$$\boxed{I_2(\mathbf{A}) = \mathrm{tr}\mathbf{A}^{\mathrm{c}}\ ,} \tag{4.19}$$

because

$$[I_2(\mathbf{A}) \xlongequal[\text{(2.26), (2.89a) and (4.17b)}]{\text{from}} \frac{1}{2}A_{mm}A_{nn} - \frac{1}{2}A_{nm}A_{mn}$$
$$\xlongequal[\text{(1.36)}]{\text{from}} \frac{1}{2}\left(\delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm}\right)A_{km}A_{ln}$$
$$\xlongequal[\text{(1.54) and (1.58a)}]{\text{from}} \frac{1}{2}\varepsilon_{ikl}\varepsilon_{imn}A_{km}A_{ln}$$
$$\xlongequal[\text{(1.36) and (2.117)}]{\text{from}} \delta_{ij}\left(\mathbf{A}^{\mathrm{c}}\right)_{ij}$$
$$\xlongequal[\text{(2.89a)}]{\text{from}} \mathrm{tr}\mathbf{A}^{\mathrm{c}}\ .$$

The interested readers may also want to arrive at the result (4.19) in an alternative way:

$$I_2(\mathbf{A}) \xlongequal[\text{(1.73) and (4.15b)}]{\text{from}} \frac{\mathbf{w}\cdot(\mathbf{Au}\times\mathbf{Av}) + \mathbf{u}\cdot(\mathbf{Av}\times\mathbf{Aw}) + \mathbf{v}\cdot(\mathbf{Aw}\times\mathbf{Au})}{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})}$$
$$\xlongequal[\text{(2.51b) and (2.112)}]{\text{from}} \frac{\mathbf{A}^{\mathrm{cT}}\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) + \mathbf{A}^{\mathrm{cT}}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{A}^{\mathrm{cT}}\mathbf{v}\cdot(\mathbf{w}\times\mathbf{u})}{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})}$$
$$\xlongequal[\text{(1.73)}]{\text{from}} \frac{\mathbf{u}\cdot\left(\mathbf{v}\times\mathbf{A}^{\mathrm{cT}}\mathbf{w}\right) + \mathbf{A}^{\mathrm{cT}}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot\left(\mathbf{A}^{\mathrm{cT}}\mathbf{v}\times\mathbf{w}\right)}{\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})}$$
$$\xlongequal[\text{(4.15a)}]{\text{from}} I_1\left(\mathbf{A}^{\mathrm{cT}}\right)$$
$$\xlongequal[\text{(4.17a)}]{\text{from}} \mathrm{tr}\mathbf{A}^{\mathrm{cT}}$$
$$\xlongequal[\text{(2.89d)}]{\text{from}} \mathrm{tr}\mathbf{A}^{\mathrm{c}}\ .$$

At the end, one can establish

$$\boxed{\mathbf{A}^{\mathrm{c\,T}} = I_2(\mathbf{A})\,\mathbf{I} - I_1(\mathbf{A})\,\mathbf{A} + \mathbf{A}^2} \;, \tag{4.20}$$

since

$$\begin{aligned}
[(I_1\mathbf{A})\,\mathbf{u}]\cdot(\mathbf{v}\times\mathbf{w}) &\xlongequal[\text{(1.9c) and (2.8b)}]{\text{from}} \mathbf{A}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\,I_1 \\
&\xlongequal[\text{(2.25) and (4.15a)}]{\text{from}} \mathbf{A}^2\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{A}\mathbf{u}\cdot(\mathbf{A}\mathbf{v}\times\mathbf{w}) + \mathbf{A}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{A}\mathbf{w}) \\
&\xlongequal[\text{(2.5) and (4.15b)}]{\text{from}} \mathbf{A}^2\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{I}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\,I_2 - \mathbf{u}\cdot(\mathbf{A}\mathbf{v}\times\mathbf{A}\mathbf{w}) \\
&\xlongequal[\text{(1.9c) and (2.8b)}]{\text{from}} \mathbf{A}^2\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + (I_2\mathbf{I})\,\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) - \mathbf{u}\cdot(\mathbf{A}\mathbf{v}\times\mathbf{A}\mathbf{w}) \\
&\xlongequal[\text{(2.51b) and (2.112)}]{\text{from}} \mathbf{A}^2\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + (I_2\mathbf{I})\,\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) - \mathbf{A}^{\mathrm{c\,T}}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) \\
&\xlongequal[\text{(1.9b), (2.8a) and (2.10)}]{\text{from}} \left[\left(\mathbf{A}^2 + I_2\mathbf{I} - \mathbf{A}^{\mathrm{c\,T}}\right)\mathbf{u}\right]\cdot(\mathbf{v}\times\mathbf{w}) \;,
\end{aligned}$$

with the aid of (1.9a) and (2.7) will deliver the desired result.

4.3 Cayley-Hamilton Equation

A tensorial analogue of the characteristic equation (4.6) according to

$$\boxed{\underbrace{\mathbf{A}^3 - I_1(\mathbf{A})\,\mathbf{A}^2 + I_2(\mathbf{A})\,\mathbf{A} - I_3(\mathbf{A})\,\mathbf{I}}_{:=\,\mathrm{p}_{\mathbf{A}}(\mathbf{A})} = \mathbf{O}} \;, \tag{4.21}$$

is called the *Cayley-Hamilton equation*. It is basically a theorem stating that any second-order tensor **A** satisfies its own characteristic equation. The proof is not difficult since, on the one hand,

$$\left[\mathbf{A}^{\mathrm{c\,T}}\right]\mathbf{A} \xlongequal[\text{(4.20)}]{\text{from}} \left[I_2\mathbf{I} - I_1\mathbf{A} + \mathbf{A}^2\right]\mathbf{A} \xlongequal[\text{(2.29), (2.33) and (2.36)}]{\text{from}} I_2\mathbf{A} - I_1\mathbf{A}^2 + \mathbf{A}^3 \;,$$

and, on the other hand,

$$\left[\mathbf{A}^{\mathrm{c\,T}}\right]\mathbf{A} \xlongequal[\text{(2.114)}]{\text{from}} \left[(\det\mathbf{A})\,\mathbf{A}^{-1}\right]\mathbf{A} \xlongequal[\text{(2.105)}]{\text{from}} (\det\mathbf{A})\,\mathbf{I} \xlongequal[\text{(4.15c)}]{\text{from}} I_3\mathbf{I} \;.$$

The desired result (4.21) thus follows. In the literature, it is often seen that **A** is assumed to be symmetric in advance to prove the Cayley-Hamilton equation. The reason is that the eigenvectors of a symmetric tensor form an eigenbasis. The proof shown here is irrespective of the nature of tensors in accord with the Cayley-Hamilton theorem that generally holds for any tensor $\mathbf{A}$.

The main application of the Cayley-Hamilton equation is to represent the powers of a tensor $\mathbf{A}$ as a combination of $\mathbf{A}$, $\mathbf{A}^2$ and $\mathbf{A}^3$. As an example, $\mathbf{A}^4$ and $\mathbf{A}^5$ can be computed much easier by the following relations

$$\mathbf{A}^4 = I_1 \mathbf{A}^3 - I_2 \mathbf{A}^2 + I_3 \mathbf{A} \ . \tag{4.22a}$$

$$\mathbf{A}^5 = \left(I_1^2 - I_2\right) \mathbf{A}^3 + \left(I_3 - I_1 I_2\right) \mathbf{A}^2 + I_1 I_3 \mathbf{A} \ . \tag{4.22b}$$

Moreover, $\mathbf{A}^{-1}$ can also be represented in a closed-form expression according to

$$\boxed{\mathbf{A}^{-1} = \frac{\mathbf{A}^2 - I_1 \mathbf{A} + I_2 \mathbf{I}}{I_3} \ .} \tag{4.23}$$

4.4 Spectral Decomposition

The *spectral decomposition* (or *eigenvalue decomposition* or *spectral representation*) is of great importance in matrix as well as tensor algebra and calculus. Computing any polynomial or inverse of second-order tensors more readily and mapping a quadratic function into a simple decoupled quadratic form are just a few applications of this powerful representation. It is also frequently seen in machine learning, mechanics of constitutive modeling and vibration analysis. Considering appropriate conditions upon which a tensor admits this canonical form in terms of its eigenvalues and eigenvectors seems inevitable.

The goal here is thus to represent a tensor in the spectral form or diagonalize its matrix form. And the results will be accompanied by the corresponding theorems.

Theorem A

Any two eigenvectors corresponding to distinct eigenvalues of a tensor are linearly independent. ★

Proof Let $(\lambda_1, \mathbf{n}_1)$ and $(\lambda_2, \mathbf{n}_2)$ be two arbitrary eigenpairs of a tensor $\mathbf{A}$ and suppose that $\lambda_1 \neq \lambda_2$. The goal is then to show that $\mathbf{n}_1$ and $\mathbf{n}_2$ are linearly independent, i.e. the homogeneous equation $a\mathbf{n}_1 + b\mathbf{n}_2 = \mathbf{0}$ has the trivial solution $a = b = 0$. Here, the coefficients a and b can generally be complex numbers. Given that $\mathbf{0} = a\mathbf{n}_1 + b\mathbf{n}_2$, one will have, on the one hand,

$$\mathbf{0} \overset{\text{from}}{\underset{(2.3)}{=\!=}} \mathbf{A}\mathbf{0} \overset{\text{by}}{\underset{\text{assumption}}{=\!=\!=}} \mathbf{A}\left(a\mathbf{n}_1 + b\mathbf{n}_2\right) \overset{\text{from}}{\underset{(2.2)}{=\!=}} a\mathbf{A}\mathbf{n}_1 + b\mathbf{A}\mathbf{n}_2 \overset{\text{from}}{\underset{(4.1)}{=\!=}} a\lambda_1\mathbf{n}_1 + b\lambda_2\mathbf{n}_2 \ ,$$

and, on the other hand,

$$\mathbf{0} \xlongequal[\text{(a) of (1.76)}]{\text{from}} \lambda_1 \mathbf{0} \xlongequal[\text{assumption}]{\text{by}} \lambda_1 \left(a\mathbf{n}_1 + b\mathbf{n}_2\right) \xlongequal[\text{(1.4f) and (1.4h)}]{\text{from}} a\lambda_1\mathbf{n}_1 + b\lambda_1\mathbf{n}_2 \ .$$

As a result, $b(\lambda_1 - \lambda_2)\,\mathbf{n}_2 = \mathbf{0}$ implies that $b = 0$ since $\lambda_1 \neq \lambda_2$ by assumption and $\mathbf{n}_2 \neq \mathbf{0}$ by definition. Following similar procedure then reveals that $a = 0$. Hence, $\mathbf{n}_1$ and $\mathbf{n}_2$ should be linearly independent. This results in the following theorem. ●

Theorem B

A linear combination of two eigenvectors corresponding to distinct eigenvalues of a tensor is not an eigenvector associated with any eigenvalue of that tensor. ★

Proof Consider again two arbitrary eigenpairs $(\lambda_1, \mathbf{n}_1)$ and $(\lambda_2, \mathbf{n}_2)$ of a tensor $\mathbf{A}$ and suppose that $\lambda_1 \neq \lambda_2$. The goal here is to show that $a\mathbf{n}_1 + b\mathbf{n}_2$ is not an eigenvector of $\mathbf{A}$ where a and b are, in general, complex nonzero numbers. The proof is done by contradiction. Suppose that $a\mathbf{n}_1 + b\mathbf{n}_2$ is an eigenvector of $\mathbf{A}$ associated with an eigenvalue λ_3. With the aid of (1.4f), (1.4h), (2.2) and (4.1), one will have on the one hand $\mathbf{A}(a\mathbf{n}_1 + b\mathbf{n}_2) = \lambda_3(a\mathbf{n}_1 + b\mathbf{n}_2) = a\lambda_3\mathbf{n}_1 + b\lambda_3\mathbf{n}_2$ and on the other hand $\mathbf{A}(a\mathbf{n}_1 + b\mathbf{n}_2) = a\mathbf{A}\mathbf{n}_1 + b\mathbf{A}\mathbf{n}_2 = a\lambda_1\mathbf{n}_1 + b\lambda_2\mathbf{n}_2$. Thus,

$$a(\lambda_1 - \lambda_3)\,\mathbf{n}_1 + b(\lambda_2 - \lambda_3)\,\mathbf{n}_2 = \mathbf{0} \ .$$

Now, recall that any two eigenvectors corresponding to distinct eigenvalues are linearly independent. It is then easy to deduce that $\lambda_1 = \lambda_3 = \lambda_2$. And this contradicts the earlier assumption that λ_1 and λ_2 were distinct eigenvalues. ●

Combining the above theorems leads to:

Theorem C

Eigenvectors corresponding to pairwise distinct eigenvalues of a tensor are linearly independent. ★

As a consequence, the eigenvectors of a tensor $\mathbf{A} \in \mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ with distinct eigenvalues always form a basis for $\mathscr{E}_{\text{c}}^{\text{o3}}$. In what follows, the goal is to characterize an important relationship between the left and right eigenvectors associated with distinct eigenvalues. ✚

Theorem D

Left and right eigenvectors corresponding to distinct eigenvalues of a tensor are orthogonal. ★

Proof Denoting by $\mathbf{m}$ and $\mathbf{n}$ the left and right eigenvectors, respectively, let $(\lambda_1, \mathbf{m}_1)$ and $(\lambda_2, \mathbf{n}_2)$ be two arbitrary eigenpairs of a tensor $\mathbf{A}$ and suppose that $\lambda_1 \neq \lambda_2$. The goal is then to show that $\mathbf{m}_1 \cdot \mathbf{n}_2 = 0$. The dot product of $\mathbf{m}_1$ and $\mathbf{A}\mathbf{n}_2$ gives, on the one hand,

$$\mathbf{m}_1 \cdot (\mathbf{A}\mathbf{n}_2) \xlongequal[(4.1)]{\text{from}} \mathbf{m}_1 \cdot (\lambda_2 \mathbf{n}_2) \xlongequal[(1.9a) \text{ and } (1.9c)]{\text{from}} \lambda_2 \mathbf{m}_1 \cdot \mathbf{n}_2 \ ,$$

and, on the other hand,

$$\mathbf{m}_1 \cdot (\mathbf{A}\mathbf{n}_2) \xlongequal[(2.47)]{\text{from}} (\mathbf{m}_1 \mathbf{A}) \cdot \mathbf{n}_2 \xlongequal[(4.2)]{\text{from}} (\lambda_1 \mathbf{m}_1) \cdot \mathbf{n}_2 \xlongequal[(1.9c)]{\text{from}} \lambda_1 \mathbf{m}_1 \cdot \mathbf{n}_2 \ .$$

As a result, $(\lambda_1 - \lambda_2)\, \mathbf{m}_1 \cdot \mathbf{n}_2 = \mathbf{0}$ implies that $\mathbf{m}_1 \cdot \mathbf{n}_2 = 0$ since $\lambda_1 \neq \lambda_2$ by assumption. Note that it is always possible to scale the left and right eigenvectors corresponding to the same eigenvalue in order to provide, for instance, $\mathbf{m}_1 \cdot \mathbf{n}_1 = 1$. The inner product of the left and right eigenvectors can thus be unified as

$$\mathbf{m}_i \cdot \mathbf{n}_j = \mathbf{n}_i \cdot \mathbf{m}_j = \delta_{ij} \ , \quad i, j = 1, 2, 3 \ . \quad : \tag{4.24}$$

Two sets of basis vectors for which (4.24) holds are called *dual bases* and accordingly $\{\mathbf{m}_i\}$ is said to be *dual* to $\{\mathbf{n}_i\}$ or vice versa. Let $[\,\mathbf{M}\,]$ be a matrix whose rows are the left eigenvectors and $[\,\mathbf{N}\,]$ be a matrix whose columns are the right eigenvectors. Then, rewriting $(4.24)_2$ in the convenient form

$$\underbrace{\begin{bmatrix} \cdots & \mathbf{m}_1 & \cdots \\ \cdots & \mathbf{m}_2 & \cdots \\ \cdots & \mathbf{m}_3 & \cdots \end{bmatrix}}_{=[\mathbf{M}]} \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{n}_1 & \mathbf{n}_2 & \mathbf{n}_3 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{=[\mathbf{N}]} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=[\mathbf{I}]} , \tag{4.25}$$

gives the useful result

$$[\,\mathbf{M}\,] = [\,\mathbf{N}\,]^{-1} \ . \quad \longleftarrow \ \text{see (2.105)} \tag{4.26}$$

Attention is now focused on the spectral representation of a tensor. First, consider the case of **non-multiple** eigenvalues. In this case, a tensor $\mathbf{A} \in \mathscr{T}_{\text{so}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ has the triples $(\lambda_k \, , \, \mathbf{m}_k \, , \, \mathbf{n}_k)$, $k = 1, 2, 3$, for which the left and right eigenvectors are dual to each other. Recall from $(2.19)_2$ that a tensor $\mathbf{A}$, with respect to the Cartesian basis $\left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right\}$, was expressed as $\mathbf{A} = A_{ij} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$. This is only one form of representing $\mathbf{A}$. Change of basis from $\left\{\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j\right\}$ to $\left\{\mathbf{m}_i \otimes \mathbf{n}_j\right\}$ and accordingly translating A_{ij} to A^{λ}_{ij} enable one to construct another form of $\mathbf{A}$, known as its spectral representation. Here, A^{λ}_{ij}, $i, j = 1, 2, 3$, are basically the *spectral components* of $\mathbf{A}$. To represent $\mathbf{A} = A^{\lambda}_{ij} \mathbf{n}_i \otimes \mathbf{m}_j$, one needs to have

$$A^{\lambda}_{ij} \xlongequal[(2.20)]{\text{from}} \mathbf{m}_i \cdot \left[\mathbf{A}\mathbf{n}_j\right]$$

$$\overset{\text{from}}{\underset{(4.1)}{=\!=}} \mathbf{m}_i \cdot \underbrace{\left[\lambda_j \mathbf{n}_j\right]}_{\text{no sum}}$$

$$\overset{\text{from (1.9a),}}{\underset{\text{(1.9c) and (4.24)}}{=\!=}} \underbrace{\lambda_j \delta_{ij}}_{\text{no sum}} \ . \tag{4.27}$$

Finally, the spectral decomposition of a tensor **A** with distinct eigenvalues takes the following form

$$\boxed{\mathbf{A} = \sum_{i=1}^{3} \lambda_i \, \mathbf{n}_i \otimes \mathbf{m}_i \ .} \tag{4.28}$$

This expression with the aid of (4.25) and (4.26) can be rewritten as

$$\boxed{[\,\mathbf{A}\,] = [\,\mathbf{N}\,][\,\mathbf{D}\,][\,\mathbf{N}\,]^{-1} \ ,} \tag{4.29}$$

where $[\,\mathbf{D}\,]$ is a **diagonal matrix** of the eigenvalues:

$$[\,\mathbf{D}\,] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \ . \tag{4.30}$$

The expression (4.29) is known as **matrix diagonalization** in the context of matrix algebra. In other words, a matrix $[\,\mathbf{A}\,]$ can be converted into a diagonal matrix $[\,\mathbf{D}\,]$ if there exists an invertible matrix $[\,\mathbf{N}\,]$ such that $[\,\mathbf{D}\,] = [\,\mathbf{N}\,]^{-1}\,[\,\mathbf{A}\,]\,[\,\mathbf{N}\,]$. With regard to this, **A** in (4.28) ($[\,\mathbf{A}\,]$ in (4.29)) is referred to as *diagonalizable tensor* (*diagonalizable matrix*). It is worthwhile to point out that not all second-order tensors take the explicit form (4.28) (or not all square matrices are diagonalizable).

Next, consider non-defective tensors possessing **multiple** eigenvalues which are also diagonalizable. Let **A** be a tensor whose every eigenvalue λ_k has identical algebraic and geometric multiplicities, i.e. $s_k = t_k$. There are again three linearly independent right eigenvectors constituting a basis. With a slight abuse of notation, **A** may be written as

$$\mathbf{A} = \sum_{i=1}^{r} \sum_{j=1}^{r} \sum_{k=1}^{t_i} \sum_{l=1}^{s_j} A^{\lambda}_{ijkl} \mathbf{n}_{ik} \otimes \mathbf{m}_{jl} \ ,$$

where $\mathbf{n}_{ik}$ and $\mathbf{m}_{jl}$ are dual bases which satisfy

$$\begin{aligned} \mathbf{m}_{ik} \cdot \mathbf{n}_{jl} &= \mathbf{n}_{ik} \cdot \mathbf{m}_{jl} \\ &= \delta_{ij} \delta_{kl} \ . \end{aligned} \tag{4.31}$$

Accordingly, the spectral components A^{λ}_{ijkl} are determined by

$$\begin{aligned} A^{\lambda}_{ijkl} &\overset{\text{from}}{\underset{(2.20)}{=\!=}} \mathbf{m}_{ik} \cdot \left[\mathbf{A}\mathbf{n}_{jl}\right] \\ &\overset{\text{from}}{\underset{(4.1)}{=\!=}} \mathbf{m}_{ik} \cdot \underbrace{\left[\lambda_j \mathbf{n}_{jl}\right]}_{\text{no sum}} \\ &\overset{\text{from (1.9a),}}{\underset{\text{(1.9c) and (4.31)}}{=\!=}} \underbrace{\lambda_j \delta_{ij} \delta_{kl}}_{\text{no sum}} \ . \end{aligned} \tag{4.32}$$

At the end, the spectral decomposition in diagonal form of a non-defective tensor $\mathbf{A}$ with possibly multiple eigenvalues is given by

$$\boxed{\mathbf{A} = \sum_{i=1}^{r} \sum_{k=1}^{t_i} \lambda_i \mathbf{n}_{ik} \otimes \mathbf{m}_{ik} \ .} \tag{4.33}$$

Observe that if $r = 3$ (or $t_1 = t_2 = t_3 = 1$), then (4.33) yields (4.28). It is important to note that a defective tensor is not diagonalizable but its eigenspaces, with the aid of generalized eigenvectors, can form an eigenbasis.

As an example, let $\mathbf{A}$ be a Cartesian tensor whose matrix is

$$[\,\mathbf{A}\,] = \begin{bmatrix} 4 & -3 & -3 \\ 3 & -2 & -3 \\ -1 & 1 & 2 \end{bmatrix} \ .$$

By having the characteristic polynomial

$$p_{\mathbf{A}}(\lambda) = -(\lambda - 1)^2 (\lambda - 2) \ ,$$

of this tensor, one finds that the algebraic multiplicities of the two distinct eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$ are $s_1 = 2$ and $s_2 = 1$, respectively. And this renders spec $(\mathbf{A}) = \{1, 1, 2\}$. Making use of elementary row reduction, the eigenspaces are given by

$$\begin{aligned} \mathscr{N}(\mathbf{A} - 1\mathbf{I}) &= \text{Span}\{\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 \, , \, \widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_3\} \quad &, \quad t_1 = 2 \ , \\ \mathscr{N}(\mathbf{A} - 2\mathbf{I}) &= \text{Span}\{-3\widehat{\mathbf{e}}_1 - 3\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3\} \quad &, \quad t_2 = 1 \ . \end{aligned}$$

Since every eigenvalue has equal algebraic and geometric multiplicities, this tensor is not defective and thus admits the spectral decomposition. The three linearly independent vectors

$$\mathbf{n}_{11} = \widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 \quad , \quad \mathbf{n}_{12} = \widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_3 \quad , \quad \mathbf{n}_{21} = -3\widehat{\mathbf{e}}_1 - 3\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3 \ ,$$

now help construct the invertible matrix

$$[\,\mathbf{N}\,] = \begin{bmatrix} 1 & 1 & -3 \\ 1 & 0 & -3 \\ 0 & 1 & 1 \end{bmatrix} .$$

The dual basis of $\{\mathbf{n}_{11}, \mathbf{n}_{12}, \mathbf{n}_{21}\}$, i.e. $\{\mathbf{m}_{11}, \mathbf{m}_{12}, \mathbf{m}_{21}\}$, can be obtained by computing the inverse of this matrix:

$$[\,\mathbf{M}\,] = [\,\mathbf{N}\,]^{-1} = \begin{bmatrix} -3 & 4 & 3 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix} .$$

Hence,

$$\mathbf{m}_{11} = -3\,\widehat{\mathbf{e}}_1 + 4\,\widehat{\mathbf{e}}_2 + 3\,\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{m}_{12} = \widehat{\mathbf{e}}_1 - \widehat{\mathbf{e}}_2 \quad , \quad \mathbf{m}_{21} = -\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3 \ .$$

Note that computing $\{\mathbf{m}_{11}, \mathbf{m}_{12}, \mathbf{m}_{21}\}$ in this way is very favorable from computational standpoint. But, the dual basis of a given basis can also be obtained by means of its *metric tensor*, see Chap. 5.

Finally, by having all required data, the given matrix $[\,\mathbf{A}\,]$ can spectrally be decomposed as

$$[\,\mathbf{A}\,] = \begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{n}_{11} & \mathbf{n}_{12} & \mathbf{n}_{21} \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cdots & \mathbf{m}_{11} & \cdots \\ \cdots & \mathbf{m}_{12} & \cdots \\ \cdots & \mathbf{m}_{21} & \cdots \end{bmatrix} ,$$

which is eventually the matrix form of

$$\mathbf{A} = \sum_{i=1}^{r} \sum_{k=1}^{t_i} \lambda_i \mathbf{n}_{ik} \otimes \mathbf{m}_{ik} = \lambda_1 \left(\mathbf{n}_{11} \otimes \mathbf{m}_{11} + \mathbf{n}_{12} \otimes \mathbf{m}_{12} \right) + \lambda_2 \mathbf{n}_{21} \otimes \mathbf{m}_{21} \ .$$

4.5 Eigenprojections of a Tensor

The spectral representation (4.33) is often demonstrated in the handy form

$$\boxed{\mathbf{A} = \sum_{i=1}^{r} \lambda_i \mathbf{P}_i \ ,} \tag{4.34}$$

where the second-order tensors

$$\boxed{\mathbf{P}_i = \sum_{k=1}^{t_i} \mathbf{n}_{ik} \otimes \mathbf{m}_{ik} \ , \quad i = 1, \cdots, r \ ,} \tag{4.35}$$

are called the *eigenprojections* of $\mathbf{A}$ (their matrix forms $[\,\mathbf{P}_i\,]$, $i = 1, \cdots, r$, are also known as *Frobenius covariants* of $[\,\mathbf{A}\,]$ in the context of matrix algebra). Each eigenprojection $\mathbf{P}_i$ is basically a projection onto the eigenspace associated with the principal value λ_i, as implied by its name. It is then a simple exercise to present some characteristics of eigenprojections:

$$\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_i \ , \qquad i, j = 1, \cdots, r \ ; \text{ no sum} \ , \tag{4.36a}$$

$$\mathbf{P}_i \mathbf{A} = \mathbf{A} \mathbf{P}_i = \lambda_i \mathbf{P}_i \ , \qquad i = 1, \cdots, r \ ; \text{ no sum} \ , \tag{4.36b}$$

$$\sum_{i=1}^{r} \mathbf{P}_i = \mathbf{I} \ . \tag{4.36c}$$

These special properties help compute the following multipurpose tensor functions much easier:

$$\textbf{Tensor powers}: \quad \underbrace{\mathbf{A}^k = \sum_{i=1}^{r} \lambda_i^k \mathbf{P}_i, \quad k = 0, 1, \cdots ,}_{\text{for instance, } \mathbf{A}^2 = \sum_{i=1}^{r} \lambda_i^2 \mathbf{P}_i} \tag{4.37a}$$

$$\textbf{Inverse of tensor powers}: \quad \underbrace{\left(\mathbf{A}^k\right)^{-1} = \sum_{i=1}^{r} \left(\lambda_i^k\right)^{-1} \mathbf{P}_i, \quad k = 0, 1, \cdots ,}_{\text{for instance, } \mathbf{A}^{-1} = \sum_{i=1}^{r} \lambda_i^{-1} \mathbf{P}_i} \tag{4.37b}$$

$$\textbf{Tensor polynomial}: \quad \underbrace{\mathbf{H}(\mathbf{A}) = \sum_{i=1}^{r} H(\lambda_i) \mathbf{P}_i}_{\text{for instance, } \mathbf{exp}(\mathbf{A}) = \sum_{i=1}^{r} \exp(\lambda_i) \mathbf{P}_i} \ . \tag{4.37c}$$

The last expression, known as *Sylvester's matrix theorem* (or *Sylvester's formula*) in matrix theory, has important implications. For instance, it can help represent the eigenprojections of a diagonalizable tensor irrespective of the eigenvalue problem. To show this, consider the following *Lagrange basis polynomials*

$$H_i(\lambda) = \prod_{\substack{j=1 \\ j \neq i}}^{r} \frac{\lambda - \lambda_j}{\lambda_i - \lambda_j} \ , \quad i = 1, \cdots, r > 1 \ , \tag{4.38}$$

satisfying

$$H_i(\lambda_k) = \delta_{ik} \ , \quad i, k = 1, \cdots, r > 1 \ . \tag{4.39}$$

The eigenprojections then take the following **basis-free** form

$$\mathbf{P}_i \overset{\text{from}}{\underset{(1.36)}{=\!=}} \sum_{k=1}^{r} \delta_{ik} \mathbf{P}_k \overset{\text{from}}{\underset{(4.39)}{=\!=}} \sum_{k=1}^{r} H_i\left(\lambda_k\right) \mathbf{P}_k \overset{\text{from}}{\underset{(4.37c) \text{ and } (4.38)}{=\!=\!=\!=}} \prod_{\substack{j=1 \\ j \neq i}}^{r} \frac{\mathbf{A} - \lambda_j \mathbf{I}}{\lambda_i - \lambda_j} . \tag{4.40}$$

Note that for the simple case $r = 1$, the property (4.36c) provides $\mathbf{P}_1 = \mathbf{I}$.

Regarding diagonalizable tensors, the expression (4.37c) can also be utilized to deliver the Cayley-Hamilton equation (4.21). To show this, consider a polynomial function of the form $H(\lambda) = \lambda^3 - I_1\lambda^2 + I_2\lambda - I_3$ which is eventually the characteristic polynomial of a diagonalizable tensor $\mathbf{A}$, see (4.6). Having in mind (4.36c), one can immediately arrive at $\mathbf{H}(\mathbf{A}) = \sum_{i=1}^{r} H(\lambda_i)\mathbf{P}_i = \mathbf{A}^3 - I_1\mathbf{A}^2 + I_2\mathbf{A} - I_3\mathbf{I}$. Finally, the r identities $H(\lambda_i) = 0$ imply the desired result $\mathbf{H}(\mathbf{A}) = \mathbf{O}$.

4.6 Spectral Decomposition of Symmetric Tensors

Of special interest in this section is to consider the spectral form of a symmetric second-order tensor due to its wide application in silence and engineering. There are many theorems (and indeed numerous problems) regarding eigenvalues and eigenvectors of a symmetric tensor. Here only the most suitable ones covering the demand for diagonalization are studied.

To begin with, consider the following theorem determining the nature of eigenvalues and eigenvectors being real or imaginary:

Theorem E

Eigenvalues and eigenvectors of a real symmetric second-order tensor $\mathbf{S} \in \mathscr{T}_{\text{so}}^{\text{s}}\left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)$ belong to R and $\mathscr{E}_{\text{r}}^{\text{o3}}$, respectively. ★

Proof Let $(\lambda_i, \mathbf{n}_i)$ be an arbitrary eigenpair of $\mathbf{S}$ satisfying $\mathbf{S}\,\mathbf{n}_i = \lambda_i \mathbf{n}_i$ according to (4.1). The complex conjugate of this eigenvalue problem is $\overline{\mathbf{S}\,\mathbf{n}_i} = \overline{\lambda_i\,\mathbf{n}_i}$ or simply $\mathbf{S}\,\overline{\mathbf{n}}_i = \overline{\lambda}_i\,\overline{\mathbf{n}}_i$ since $\mathbf{S}$ is real and thus not affected. One will have, on the one hand,

$$\overline{\mathbf{n}}_i \cdot (\mathbf{S}\mathbf{n}_i) = \overline{\mathbf{n}}_i \cdot (\lambda_i \mathbf{n}_i) \quad \Rightarrow \quad \overline{\mathbf{n}}_i \cdot \mathbf{S}\mathbf{n}_i = \lambda_i \overline{\mathbf{n}}_i \cdot \mathbf{n}_i \ , \quad i = 1, 2, 3; \ \text{no sum} \ ,$$

and, on the other hand,

$$(\mathbf{S}\overline{\mathbf{n}}_i) \cdot \mathbf{n}_i = \left(\overline{\lambda}_i \overline{\mathbf{n}}_i\right) \cdot \mathbf{n}_i \quad \Rightarrow \quad \overline{\mathbf{n}}_i \cdot \mathbf{S}\mathbf{n}_i = \overline{\lambda}_i \overline{\mathbf{n}}_i \cdot \mathbf{n}_i \ , \quad i = 1, 2, 3; \ \text{no sum} \ .$$

As a result $\left(\lambda_i - \overline{\lambda}_i\right)\overline{\mathbf{n}}_i \cdot \mathbf{n}_i = 0$ implies that $\lambda_i = \overline{\lambda}_i$ since $\mathbf{n}_i \neq \mathbf{0}$ by definition and thus $\overline{\mathbf{n}}_i \cdot \mathbf{n}_i > 0$ by the positive definite property (1.22c). This reveals the fact that $\overline{\lambda}_i$ has no imaginary part. Indeed, any $c^* \in$ C is real if and only if $\overline{c}^* = c^*$. The result

$\lambda_i = \overline{\lambda}_i$ along with the assumption that $\mathbf{S}$ is real then imply that the eigenspace of λ_i, i.e. $\mathscr{N}(\mathbf{S} - \lambda_i \mathbf{I})$, should not contain any complex vector.

The following theorem regards an important relationship between the eigenvectors associated with distinct eigenvalues:

Theorem F

Eigenvectors of a real symmetric second-order tensor corresponding to distinct eigenvalues are mutually orthogonal.

Proof Let $(\lambda_1, \mathbf{n}_1)$ and $(\lambda_2, \mathbf{n}_2)$ be two arbitrary eigenpairs of a symmetric tensor $\mathbf{S}$ and suppose that $\lambda_1 \neq \lambda_2$. Then,

$$\mathbf{n}_1 \cdot (\mathbf{S}\mathbf{n}_2) \xlongequal[(4.1)]{\text{from}} \mathbf{n}_1 \cdot (\lambda_2 \mathbf{n}_2) \xlongequal[(1.9a) \text{ and } (1.9c)]{\text{from}} \lambda_2 \mathbf{n}_1 \cdot \mathbf{n}_2 \ ,$$

while

$$\mathbf{n}_1 \cdot (\mathbf{S}\mathbf{n}_2) \xlongequal[(2.51b) \text{ and } (2.57)]{\text{from}} (\mathbf{S}\mathbf{n}_1) \cdot \mathbf{n}_2 \xlongequal[(4.1)]{\text{from}} (\lambda_1 \mathbf{n}_1) \cdot \mathbf{n}_2 \xlongequal[(1.9c)]{\text{from}} \lambda_1 \mathbf{n}_1 \cdot \mathbf{n}_2 \ .$$

Thus, $(\lambda_1 - \lambda_2)\, \mathbf{n}_1 \cdot \mathbf{n}_2 = 0$ implies that $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0$ since $\lambda_1 \neq \lambda_2$ by assumption.

The last theorem considered here states that a symmetric tensor has no deficiency and hence it can always be diagonalized (the proof can be found, e.g., in Itskov [8]):

Theorem G

Any eigenvalue of a real symmetric second-order tensor has identical algebraic and geometric multiplicities.

By having all required data, one can now rephrase the previous results regarding the spectral forms of a generic tensor $\mathbf{A} \in \mathscr{T}_{so}$ for the problem at hand.

First, consider a symmetric tensor $\mathbf{S}$ with **distinct** eigenvalues. To begin with, one needs to know that the left and right eigenvectors of a symmetric tensor are identical since $\mathbf{Sn} = \mathbf{nS}$. Hence, (4.24) now takes the form $\mathbf{n}_i \cdot \mathbf{n}_j = \delta_{ij},\ i, j = 1, 2, 3$. And this reveals the fact the eigenvectors of a symmetric tensor to be used for its diagonalization should essentially be unit vectors.

Denoting by $\widehat{\mathbf{n}}_i,\ i = 1, 2, 3$, the normalized eigenvectors of $\mathbf{S}$, the set $\{\widehat{\mathbf{n}}_i\} := \{\widehat{\mathbf{n}}_1, \widehat{\mathbf{n}}_2, \widehat{\mathbf{n}}_3\}$ is an orthonormal basis, see (1.16). Let $[\,\mathbf{Q}\,]$ be a matrix whose columns are these normalized eigenvectors. Then, one can deduce from (4.25) and (4.26) that $[\,\mathbf{Q}\,]^{-1} = [\,\mathbf{Q}\,]^{\mathrm{T}}$. Such a matrix should thus be an orthogonal matrix, see (2.130) and (2.131). With regard to this, an orthogonal matrix must be realized as a matrix whose columns form an orthonormal basis. The rows of such a matrix also constitutes another orthonormal basis.

At the end, any symmetric tensor $\mathbf{S}$ with non-multiple eigenvalues admits the following spectral form

$$\boxed{\mathbf{S} = \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \ .} \quad \longleftarrow \text{see (4.28)} \tag{4.41}$$

In matrix notation, it renders

$$[\,\mathbf{S}\,] = \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ \widehat{\mathbf{n}}_1 & \widehat{\mathbf{n}}_2 & \widehat{\mathbf{n}}_3 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{=[\mathbf{Q}]} \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}}_{=[\mathbf{D}]} \underbrace{\begin{bmatrix} \cdots & \widehat{\mathbf{n}}_1 & \cdots \\ \cdots & \widehat{\mathbf{n}}_2 & \cdots \\ \cdots & \widehat{\mathbf{n}}_3 & \cdots \end{bmatrix}}_{=[\mathbf{Q}]^{-1}} . \longleftarrow \text{see (4.29) and (4.30)} \tag{4.42}$$

These results help introduce a symmetric tensor as an *orthogonally diagonalizable tensor*. Conversely, symmetry of the eigenprojections $\widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i$, $i = 1, 2, 3$, amounts to writing $\mathbf{S}^{\mathrm{T}} = \mathbf{S}$. Consistent with this, a matrix of the form $[\,\mathbf{Q}\,][\,\mathbf{D}\,][\,\mathbf{Q}\,]^{-1}$ with $[\,\mathbf{Q}\,]^{\mathrm{T}} = [\,\mathbf{Q}\,]^{-1}$ and $[\,\mathbf{D}\,]^{\mathrm{T}} = [\,\mathbf{D}\,]$ is always symmetric. As a result, one can state that:

> A tensor is symmetric if and only if it can orthogonally be diagonalized.

Next, consider a symmetric tensor $\mathbf{S}$ with **multiple** eigenvalues. In this case, the expression (4.33) should consistently be modified. Recall from the arguments led to (4.41) that each eigenvector is of unit length and there is no difference between the left and right eigenvectors. The key point here is that the eigenvectors corresponding to a repeated eigenvalue may not be orthogonal although they are linearly independent. A reliable technique such as the *Gram-Schmidt process* is thus required for orthonormalising eigenvectors.[5] Having recorded all required data, the spectral decomposition of a symmetric tensor $\mathbf{S}$ with repeated eigenvalues is given by

---

[5] In an inner product space, there are infinitely many bases and hence any specific basis should be chosen with care. To simplify computations, it is often most convenient to work with an orthogonal basis. Orthogonalization of an arbitrary basis is usually accompanied by normalization. These can be achieved by use of the Gram-Schmidt process which is basically an algorithm that takes a linearly independent set of vectors in an inner product space and delivers an orthonormal set. This process is completed within two steps. Given an arbitrary basis $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ of $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$, *Gram-Schmidt orthogonalization* first generates an orthogonal basis $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ via the relations

$$\mathbf{u} = \mathbf{a} \quad , \quad \mathbf{v} = \mathbf{b} - \frac{\mathbf{b} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a} \quad , \quad \mathbf{w} = \mathbf{c} - \frac{\mathbf{c} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a} - \frac{\mathbf{c} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}}\mathbf{b} \, .$$

Subsequently, *Gram-Schmidt orthonormalization* delivers an orthonormal basis $\{\widehat{\mathbf{u}}, \widehat{\mathbf{v}}, \widehat{\mathbf{w}}\}$ according to

$$\widehat{\mathbf{u}} = \frac{\mathbf{u}}{|\mathbf{u}|} \quad , \quad \widehat{\mathbf{v}} = \frac{\mathbf{v}}{|\mathbf{v}|} \quad , \quad \widehat{\mathbf{w}} = \frac{\mathbf{w}}{|\mathbf{w}|} \, .$$

Note that the Gram–Schmidt process can also be used for any finite-dimensional vector space equipped with an inner product.

$$\mathbf{S} = \sum_{i=1}^{r} \sum_{k=1}^{t_i} \lambda_i \, \widehat{\mathbf{n}}_{ik} \otimes \widehat{\mathbf{n}}_{ik} \ . \tag{4.43}$$

This can also be rephrased as

$$\mathbf{S} = \sum_{i=1}^{r} \lambda_i \widehat{\mathbf{P}}_i \quad \text{where} \quad \widehat{\mathbf{P}}_i = \sum_{k=1}^{t_i} \widehat{\mathbf{n}}_{ik} \otimes \widehat{\mathbf{n}}_{ik} \ , \quad i = 1, \cdots, r \ . \tag{4.44}$$

It is important to note that (4.41) and (4.43) are equivalent if $r = 3$.

For subsequent developments, two special cases are considered. First, suppose that a symmetric tensor $\mathbf{S}$ has two distinct eigenvalues, namely λ_1 with $t_1 = 2$ and λ_2 with $t_2 = 1$. By having the eigenpairs $(\lambda_1, \widehat{\mathbf{n}}_{11})$, $(\lambda_1, \widehat{\mathbf{n}}_{12})$ and $(\lambda_2, \widehat{\mathbf{n}}_{21})$, one can write

$$\mathbf{S} = \lambda_1 \widehat{\mathbf{P}}_1 + \lambda_2 \widehat{\mathbf{P}}_2 = \lambda_1 \left(\widehat{\mathbf{n}}_{11} \otimes \widehat{\mathbf{n}}_{11} + \widehat{\mathbf{n}}_{12} \otimes \widehat{\mathbf{n}}_{12} \right) + \lambda_2 \, \widehat{\mathbf{n}}_{21} \otimes \widehat{\mathbf{n}}_{21} \ .$$

Notice that $\widehat{\mathbf{n}}_{21}$ is perpendicular to a plane spanned by $\widehat{\mathbf{n}}_{11}$ and $\widehat{\mathbf{n}}_{12}$. For notional simplicity, the unit vectors $\widehat{\mathbf{n}}_{11}$, $\widehat{\mathbf{n}}_{12}$ and $\widehat{\mathbf{n}}_{21}$ are denoted by $\widehat{\mathbf{n}}_1$, $\widehat{\mathbf{n}}_2$ and $\widehat{\mathbf{n}}_3$, respectively. Having in mind the identity $\mathbf{I} = \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3$, the projection tensors then render $\widehat{\mathbf{P}}_2 = \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 := \mathbf{P}^{\parallel}_{\widehat{\mathbf{n}}_3}$ and $\widehat{\mathbf{P}}_1 = \mathbf{I} - \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 := \mathbf{P}^{\perp}_{\widehat{\mathbf{n}}_3}$, see (2.140a) and (2.140b). Accordingly,

$$\boxed{\mathbf{S} = \lambda_2 \mathbf{P}^{\parallel}_{\widehat{\mathbf{n}}_3} + \lambda_1 \mathbf{P}^{\perp}_{\widehat{\mathbf{n}}_3} = \lambda_2 \, \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 + \lambda_1 \left(\mathbf{I} - \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \right) \ .} \tag{4.45}$$

Next, consider a symmetric tensor $\mathbf{S}$ having only one distinct eigenvalue; namely, λ with $t = 3$. It is then easy to show that $\mathbf{S}$ becomes the scalar multiplication of the identity tensor by this eigenvalue:

$$\boxed{\mathbf{S} = \lambda \mathbf{I} \ .} \quad \longleftarrow \ \text{this renders a spherical tensor, see (2.144)} \tag{4.46}$$

In this case, any direction is basically a principal direction and any orthonormal basis of $\mathscr{E}^{o3}_{\mathrm{r}}$ can be regarded as a set of eigenvectors.

At the end, consistent with (4.41), (4.45) and (4.46), the eigenprojections of a symmetric tensor $\mathbf{S}$ help additively decompose any vector $\mathbf{w}$ in $\mathscr{E}^{o3}_{\mathrm{r}}$ as

$$\mathbf{w} = \sum_{i=1}^{r} \mathbf{w}_i = \begin{cases} \left(\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \right) \mathbf{w} & \text{if } r = 3 \\ \left(\widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \right) \mathbf{w} + \left(\mathbf{I} - \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \right) \mathbf{w} & \text{if } r = 2 \\ \mathbf{I}\mathbf{w} & \text{if } r = 1 \end{cases} . \tag{4.47}$$

4.7 Exercises

Exercise 4.1

Verify that the quantities in (4.17a) to (4.17c), i.e. I_1, I_2 and I_3, are also the principal invariants of a second-order tensor **A** under an orthogonal transformation, that is,

$$\underbrace{I_i\left(\mathbf{QAQ}^{\mathrm{T}}\right) = I_i\left(\mathbf{A}\right)}_{i\,=\,1,2,3}\ , \quad \text{for all orthogonal tensors } \mathbf{Q} \in \mathscr{T}_{so}\ , \tag{4.48}$$

or, equivalently,

$$\underbrace{\operatorname{tr}\left(\mathbf{QAQ}^{\mathrm{T}}\right)^i = \operatorname{tr}\left(\mathbf{A}^i\right)}_{i\,=\,1,2,3}\ , \quad \text{for all } \mathbf{Q} \in \mathscr{T}_{so}\ . \tag{4.49}$$

Solution. The three relations in (4.49) can be verified as follows:

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{QAQ}^{\mathrm{T}}\right) &\xlongequal[(2.131)]{\text{from}} \operatorname{tr}\left(\mathbf{QAQ}^{-1}\right) \\
&\xlongequal[(2.109\text{g})]{\text{from}} \operatorname{tr}\left(\mathbf{A}\right)\ , \\
\operatorname{tr}\left(\mathbf{QAQ}^{\mathrm{T}}\right)^2 &\xlongequal[(2.33),\ (2.36)\text{ and }(2.130)]{\text{from}} \operatorname{tr}\left(\mathbf{QA}^2\mathbf{Q}^{\mathrm{T}}\right) \\
&\xlongequal[(2.109\text{g})\text{ and }(2.131)]{\text{from}} \operatorname{tr}\left(\mathbf{A}^2\right)\ , \\
\operatorname{tr}\left(\mathbf{QAQ}^{\mathrm{T}}\right)^3 &\xlongequal[(2.33),\ (2.36)\text{ and }(2.130)]{\text{from}} \operatorname{tr}\left(\mathbf{QA}^3\mathbf{Q}^{\mathrm{T}}\right) \\
&\xlongequal[(2.109\text{g})\text{ and }(2.131)]{\text{from}} \operatorname{tr}\left(\mathbf{A}^3\right)\ .
\end{aligned}$$

Exercise 4.2

Show that the principal invariants of a tensor **A** according to (4.15a), (4.15b) and $(4.15\text{c})_1$ also admit the representations $(4.7\text{a})_2$, $(4.7\text{b})_2$ and $(4.7\text{c})_2$, respectively.

Solution. Let $\mathbf{u} = u_i\widehat{\mathbf{e}}_i$, $\mathbf{v} = v_j\widehat{\mathbf{e}}_j$ and $\mathbf{w} = w_k\widehat{\mathbf{e}}_k$ be three vectors which form an arbitrary basis. Having in mind that the inner product of two vectors is a symmetric bilinear form and their cross product has linearity in each argument, this exercise is solved in the following by use of the replacement property of Kronecker delta and some identities regarding the permutation symbol.

The expression (4.15a)*:*

$$
\begin{aligned}
&\mathbf{Au}\cdot(\mathbf{v}\times\mathbf{w})+\mathbf{u}\cdot(\mathbf{Av}\times\mathbf{w})+\mathbf{u}\cdot(\mathbf{v}\times\mathbf{Aw}) \\
&\xlongequal[\text{(2.21)}]{\text{from}} A_{mi}\left[\widehat{\mathbf{e}}_m\cdot\left(\widehat{\mathbf{e}}_j\times\widehat{\mathbf{e}}_k\right)\right]\left[u_i v_j w_k\right]+A_{mj}\left[\widehat{\mathbf{e}}_i\cdot\left(\widehat{\mathbf{e}}_m\times\widehat{\mathbf{e}}_k\right)\right]\left[u_i v_j w_k\right] \\
&\quad + A_{mk}\left[\widehat{\mathbf{e}}_i\cdot\left(\widehat{\mathbf{e}}_j\times\widehat{\mathbf{e}}_m\right)\right]\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.65)}]{\text{from}} \left(\varepsilon_{mjk}A_{mi}+\varepsilon_{imk}A_{mj}+\varepsilon_{ijm}A_{mk}\right)\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.36)}]{\text{from}} A_{mn}\left(\varepsilon_{mjk}\delta_{in}+\varepsilon_{imk}\delta_{jn}+\varepsilon_{ijm}\delta_{kn}\right)\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.59)}]{\text{from}} \delta_{mn}A_{mn}\left[\varepsilon_{ijk}u_i v_j w_k\right] \\
&\xlongequal[\text{with renaming the dummy indices}]{\text{from (1.54) and (1.58b) along}} \frac{1}{2}\varepsilon_{qrn}\varepsilon_{qrs}A_{ns}\left[\varepsilon_{ijk}u_i v_j w_k\right] \\
&\xlongequal[\text{(1.74)}]{\text{from}} \frac{1}{2}\varepsilon_{qrn}\varepsilon_{qrs}A_{ns}\left[\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\right] \ .
\end{aligned}
$$

The expression (4.15b)*:*

$$
\begin{aligned}
&\mathbf{Au}\cdot(\mathbf{Av}\times\mathbf{w})+\mathbf{u}\cdot(\mathbf{Av}\times\mathbf{Aw})+\mathbf{Au}\cdot(\mathbf{v}\times\mathbf{Aw}) \\
&\xlongequal[\text{(2.21)}]{\text{from}} A_{mi}A_{nj}\left[\widehat{\mathbf{e}}_m\cdot\left(\widehat{\mathbf{e}}_n\times\widehat{\mathbf{e}}_k\right)\right]\left[u_i v_j w_k\right]+A_{mj}A_{nk}\left[\widehat{\mathbf{e}}_i\cdot\left(\widehat{\mathbf{e}}_m\times\widehat{\mathbf{e}}_n\right)\right]\left[u_i v_j w_k\right] \\
&\quad + A_{mi}A_{nk}\left[\widehat{\mathbf{e}}_m\cdot\left(\widehat{\mathbf{e}}_j\times\widehat{\mathbf{e}}_n\right)\right]\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.65)}]{\text{from}} \left(\varepsilon_{mnk}A_{mi}A_{nj}+\varepsilon_{imn}A_{mj}A_{nk}+\varepsilon_{mjn}A_{mi}A_{nk}\right)\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.54)}]{\text{from}} \frac{1}{2}\left[\varepsilon_{imn}\left(A_{mj}A_{nk}-A_{nj}A_{mk}\right)-\varepsilon_{jmn}\left(A_{mi}A_{nk}-A_{ni}A_{mk}\right)\right. \\
&\quad \left.+\,\varepsilon_{kmn}\left(A_{mi}A_{nj}-A_{ni}A_{mj}\right)\right]\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.36)}]{\text{from}} \frac{1}{2}\varepsilon_{qmn}A_{mr}A_{ns}\left[\delta_{iq}\left(\delta_{jr}\delta_{ks}-\delta_{kr}\delta_{js}\right)-\delta_{jq}\left(\delta_{ir}\delta_{ks}-\delta_{kr}\delta_{is}\right)\right. \\
&\quad \left.+\,\delta_{kq}\left(\delta_{ir}\delta_{js}-\delta_{jr}\delta_{is}\right)\right]\left[u_i v_j w_k\right] \\
&\xlongequal[\text{(1.57)}]{\text{from}} \frac{1}{2}\varepsilon_{qmn}\varepsilon_{qrs}A_{mr}A_{ns}\left[\varepsilon_{ijk}u_i v_j w_k\right] \\
&\xlongequal[\text{(1.74)}]{\text{from}} \frac{1}{2}\varepsilon_{qmn}\varepsilon_{qrs}A_{mr}A_{ns}\left[\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\right] \ .
\end{aligned}
$$

The expression (4.15c)*:*

$$\begin{aligned}\mathbf{Au}\cdot[\mathbf{Av}\times\mathbf{Aw}] &= \left[\mathbf{A}\widehat{\mathbf{e}}_i\cdot\left(\mathbf{A}\widehat{\mathbf{e}}_j\times\mathbf{A}\widehat{\mathbf{e}}_k\right)\right]\left[u_i v_j w_k\right]\\ &\overset{\text{from}}{\underset{(2.21)}{=\!=}} A_{li}A_{mj}A_{nk}\left[\widehat{\mathbf{e}}_l\cdot\left(\widehat{\mathbf{e}}_m\times\widehat{\mathbf{e}}_n\right)\right]\left[u_i v_j w_k\right]\\ &\overset{\text{from}}{\underset{(1.65)}{=\!=}} \varepsilon_{lmn}A_{li}A_{mj}A_{nk}\left[u_i v_j w_k\right]\\ &\overset{\text{from}}{\underset{(1.81)}{=\!=}} \frac{1}{6}\varepsilon_{lmn}\varepsilon_{qrs}A_{lq}A_{mr}A_{ns}\left[\varepsilon_{ijk}u_i v_j w_k\right]\\ &\overset{\text{from}}{\underset{(1.74)}{=\!=}} \frac{1}{6}\varepsilon_{lmn}\varepsilon_{qrs}A_{lq}A_{mr}A_{ns}\left[\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\right]\ .\end{aligned}$$

Exercise 4.3

Suppose one is given a symmetric tensor $\mathbf{S}$ of the following form

$$\mathbf{S}=\alpha\left(\mathbf{I}-\widehat{\mathbf{e}}_1\otimes\widehat{\mathbf{e}}_1\right)+\beta\left(\widehat{\mathbf{e}}_1\otimes\widehat{\mathbf{e}}_2+\widehat{\mathbf{e}}_2\otimes\widehat{\mathbf{e}}_1\right)\ ,$$

where $\alpha,\ \beta$ are real scalars and $\widehat{\mathbf{e}}_1,\ \widehat{\mathbf{e}}_2$ present orthogonal unit vectors.

1. Show that the principal values $\lambda_i,\ i=1,2,3$, of $\mathbf{S}$ are

$$\lambda_1=\alpha\quad,\quad\lambda_2=\frac{\alpha+\sqrt{\Delta}}{2}\quad,\quad\lambda_3=\frac{\alpha-\sqrt{\Delta}}{2}\ ,$$

where $\Delta=\alpha^2+4\beta^2$.
Solution. Given the nonzero components $S_{22}=S_{33}=\alpha,\ S_{12}=S_{21}=\beta$, the principal invariants (4.7a) to (4.7c) render

$$I_1=2\alpha\quad,\quad I_2=\alpha^2-\beta^2\quad,\quad I_3=-\alpha\beta^2\ ,$$

Accordingly, the characteristic equation for $\mathbf{S}$ takes the form

$$\lambda^3-2\alpha\lambda^2+\left(\alpha^2-\beta^2\right)\lambda+\alpha\beta^2=0\ .$$

Before applying the Cardano's technique to find the roots of this equation, let's intuitively check on whether α or β is a solution or not. Observe that the scalar α presents a root and this helps write

$$\underbrace{(\lambda-\alpha)\left(a\lambda^2+b\lambda+c\right)}_{=a\lambda^3+(b-\alpha a)\lambda^2+(c-\alpha b)\lambda-\alpha c}=0\ .$$

Consequently, the unknown coefficients a, b, c become $a = 1,\ b = -\alpha,\ c = -\beta^2$ by comparison. The problem is now simplified to find the roots of the quadratic equation $\lambda^2 - \alpha\lambda - \beta^2 = 0$. By use of the **quadratic formula**, one can finally arrive at the desired results.

2. Show that the corresponding normalized eigenvectors $\widehat{\mathbf{n}}_i,\ i = 1, 2, 3$ have the following matrix form

$$[\widehat{\mathbf{n}}_1]^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} ,$$

$$[\widehat{\mathbf{n}}_2]^{\mathrm{T}} = \frac{\begin{bmatrix} -\lambda_3\beta^{-1} & 1 & 0 \end{bmatrix}}{\sqrt{1 + \lambda_3^2\beta^{-2}}} \quad , \quad [\widehat{\mathbf{n}}_3]^{\mathrm{T}} = \frac{\begin{bmatrix} -\lambda_2\beta^{-1} & 1 & 0 \end{bmatrix}}{\sqrt{1 + \lambda_2^2\beta^{-2}}} .$$

Solution. The Cartesian tensor **S** is a symmetric tensor with distinct eigenvalues. Thus, each eigenvalue has the identical algebraic and geometric multiplicities $s_i = t_i = 1,\ i = 1, 2, 3$. And this means that it can spectrally decomposed as (4.41) or (4.42). First, the goal is to find all eigenvectors $\mathbf{n}_1 = (\mathbf{n}_1)_i\, \widehat{\mathbf{e}}_i$ corresponding to λ_1. They are nonzero solutions of

$$(\mathbf{S} - \lambda_1\mathbf{I})\, \mathbf{n}_1 = (\mathbf{S} - \alpha\mathbf{I})\, \mathbf{n}_1 = \mathbf{0} .$$

For convenience, by using $(1.43)_1$, $(1.47)_1$ and $(2.94)_1$, this expression can be written as

$$[\,\mathbf{S} - \alpha\mathbf{I}\,]\,[\,\mathbf{n}_1\,] = \begin{bmatrix} -\alpha & \beta & 0 \\ \beta & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (\mathbf{n}_1)_1 \\ (\mathbf{n}_1)_2 \\ (\mathbf{n}_1)_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

It follows that

$$(\mathbf{n}_1)_1 = (\mathbf{n}_1)_2 = 0 \quad , \quad (\mathbf{n}_1)_3 = \gamma \in \mathrm{R} .$$

Subsequently, the corresponding eigenspace is

$$\mathscr{N}\,(\mathbf{S} - \lambda_1\mathbf{I}) = \left\{\mathbf{n}_1 \in \mathscr{E}_{\mathrm{r}}^{o3} \,|\, \mathbf{n}_1 = \gamma\, \widehat{\mathbf{e}}_3\ ,\ \text{for any } \gamma \in \mathrm{R}\right\} ,$$

or

$$\mathscr{N}\,(\mathbf{S} - \lambda_1\mathbf{I}) = \mathrm{Span}\,\{\widehat{\mathbf{e}}_3\} .$$

As a result, $\widehat{\mathbf{e}}_3$ is a basis for $\mathscr{N}\,(\mathbf{S} - \lambda_1\mathbf{I})$ and this delivers $\widehat{\mathbf{n}}_1 = \widehat{\mathbf{e}}_3$.
Next, to find all eigenvectors $\mathbf{n}_2 = (\mathbf{n}_2)_i\, \widehat{\mathbf{e}}_i$ associated with λ_2, one needs to solve

$$(\mathbf{S} - \lambda_2\mathbf{I})\, \mathbf{n}_2 = \left(\mathbf{S} - \frac{\alpha + \sqrt{\varDelta}}{2}\mathbf{I}\right) \mathbf{n}_2 = \mathbf{0} .$$

By means of the elementary row reductions, the above expression in the convenient form

$$-\frac{1}{2}\begin{bmatrix} \sqrt{\Delta}+\alpha & -2\beta & 0 \\ -2\beta & \sqrt{\Delta}-\alpha & 0 \\ 0 & 0 & \sqrt{\Delta}-\alpha \end{bmatrix}\begin{bmatrix} (\mathbf{n}_2)_1 \\ (\mathbf{n}_2)_2 \\ (\mathbf{n}_2)_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} ,$$

will be reduced to

$$\begin{bmatrix} \beta & \lambda_3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}\begin{bmatrix} (\mathbf{n}_2)_1 \\ (\mathbf{n}_2)_2 \\ (\mathbf{n}_2)_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Thus, any solution should satisfy

$$\beta\,(\mathbf{n}_2)_1 + \lambda_3\,(\mathbf{n}_2)_2 = 0 \quad , \quad (\mathbf{n}_2)_3 = 0 \, ,$$

which helps obtain

$$\mathscr{N}\,(\mathbf{S} - \lambda_2\mathbf{I}) = \left\{\mathbf{n}_2 \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}} \mid \mathbf{n}_2 = \gamma\,\left(-\lambda_3\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2\right) \text{ , for any } \gamma \in \mathrm{R}\right\} \, ,$$

or

$$\mathscr{N}\,(\mathbf{S} - \lambda_2\mathbf{I}) = \mathrm{Span}\left\{-\lambda_3\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2\right\} \, .$$

Now, one can see that $-\lambda_3\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2$ is a basis for $\mathscr{N}\,(\mathbf{S} - \lambda_2\mathbf{I})$ and the desired result will be $\widehat{\mathbf{n}}_2 = \left(-\lambda_3\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2\right)/\sqrt{1 + \lambda_3^2\beta^{-2}}$.
Finally, all eigenvectors $\mathbf{n}_3 = (\mathbf{n}_3)_i\,\widehat{\mathbf{e}}_i$ corresponding to λ_3 are nontrivial solutions of the eigenvalue problem

$$(\mathbf{S} - \lambda_3\mathbf{I})\,\mathbf{n}_3 = \left(\mathbf{S} - \frac{\alpha - \sqrt{\Delta}}{2}\mathbf{I}\right)\mathbf{n}_3 = \mathbf{0} \, .$$

By elementary row operations, one can arrive at

$$\begin{bmatrix} \beta & \lambda_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} (\mathbf{n}_3)_1 \\ (\mathbf{n}_3)_2 \\ (\mathbf{n}_3)_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

This holds true if $[\mathbf{n}_3]^{\mathrm{T}} = \left[\,(\mathbf{n}_3)_1\;\;(\mathbf{n}_3)_2\;\;(\mathbf{n}_3)_3\,\right]$ satisfies

$$\beta\,(\mathbf{n}_3)_1 + \lambda_2\,(\mathbf{n}_3)_2 = 0 \quad , \quad (\mathbf{n}_3)_3 = 0 \, ,$$

and, accordingly,

$$\mathscr{N}(\mathbf{S} - \lambda_3\mathbf{I}) = \left\{\mathbf{n}_3 \in \mathscr{E}_{\mathrm{r}}^{o3} \,|\, \mathbf{n}_3 = \gamma \left(-\lambda_2\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2\right) \text{ , for any } \gamma \in \mathrm{R}\right\} ,$$

or

$$\mathscr{N}(\mathbf{S} - \lambda_3\mathbf{I}) = \mathrm{Span}\left\{-\lambda_2\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2\right\} .$$

Therefore, $-\lambda_2\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2$ is a basis for $\mathscr{N}(\mathbf{S} - \lambda_3\mathbf{I})$ and this delivers the desired result $\widehat{\mathbf{n}}_3 = \left(-\lambda_2\beta^{-1}\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2\right)/\sqrt{1 + \lambda_2^2\beta^{-2}}$.
It is evident that $\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2 = \widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_3 = \widehat{\mathbf{n}}_2 \cdot \widehat{\mathbf{n}}_3 = 0$. By use of (1.41) and $(1.43)_1$, one can finally obtain the matrix form of these eigenvectors.

Exercise 4.4

Let **A** be an arbitrary tensor and further let **S** be a symmetric tensor. Then, show that these tensors *commute*, that is,

$$\mathbf{SA} = \mathbf{AS} , \tag{4.51}$$

if and only if **A** leaves all eigenspaces of **S** **invariant**.[6] This is known as the *commutation theorem* (see Liu [9]).

Moreover, show that **A** should be a symmetric tensor when **S** has distinct eigenvalues.

Solution. Suppose that $\mathbf{SA} = \mathbf{AS}$ and let $(\lambda, \mathbf{w})$ be an eigenpair of **S** satisfying $\mathbf{Sw} = \lambda\mathbf{w}$. Then,

$$\begin{aligned}
\mathbf{S}(\mathbf{Aw}) &\xlongequal[(2.25)]{\text{from}} (\mathbf{SA})\,\mathbf{w} \\
&\xlongequal[\text{assumption}]{\text{by}} (\mathbf{AS})\,\mathbf{w} \\
&\xlongequal[(2.25)]{\text{from}} \mathbf{A}(\mathbf{Sw}) \\
&\xlongequal[\text{assumption}]{\text{by}} \mathbf{A}(\lambda\mathbf{w}) \\
&\xlongequal[(2.8\text{b})]{\text{from}} \lambda(\mathbf{Aw}) .
\end{aligned}$$

As can be seen, both **w** and **Aw** belong to the characteristic space $\mathscr{N}(\mathbf{S} - \lambda\mathbf{I})$ of **S**. To prove the converse, recall from (4.47) that any vector **w** in $\mathscr{E}_{\mathrm{r}}^{o3}$ can additively

[6] Let $\mathscr{F}$ be a set of vectors. If **Aw** belongs to $\mathscr{F}$ for any vector **w** in $\mathscr{F}$, one then says that the linear mapping **A** leaves $\mathscr{F}$ invariant.

be decomposed with respect to the eigenspaces of $\mathbf{S}$ as $\mathbf{w} = \sum_{i=1}^{r} \mathbf{w}_i$. Now, if $\mathbf{A}$ preserves all characteristic spaces of $\mathbf{S}$, then $\mathbf{w}_i$, $\mathbf{A}\mathbf{w}_i \in \mathscr{N}(\mathbf{S} - \lambda_i \mathbf{I})$ and

$$\mathbf{S}(\mathbf{A}\mathbf{w}_i) = \lambda_i (\mathbf{A}\mathbf{w}_i) = \mathbf{A}(\lambda_i \mathbf{w}_i) = \mathbf{A}(\mathbf{S}\mathbf{w}_i) \ .$$

This result, along with (2.2) and (2.25), helps obtain

$$(\mathbf{SA})\,\mathbf{w} = \mathbf{S}(\mathbf{A}\mathbf{w}) = \sum_{i=1}^{r} \mathbf{S}(\mathbf{A}\mathbf{w}_i) = \sum_{i=1}^{r} \mathbf{A}(\mathbf{S}\mathbf{w}_i) = \mathbf{A}(\mathbf{S}\mathbf{w}) = (\mathbf{AS})\,\mathbf{w} \ .$$

At the end, $\mathbf{SA} = \mathbf{AS}$ is implied by use of (2.6).

Hint: If $\mathbf{S}$ is spherical, i.e. $\mathbf{S} = \lambda \mathbf{I}$, the expression (4.51) obviously shows that $\mathbf{A}$ need not necessarily be symmetric. But, this relation implies that $\mathbf{A}$ be a **symmetric** tensor when $\mathbf{S}$ has distinct eigenvalues, i.e. $\mathbf{S} = \sum_{k=1}^{3} \lambda_k \, \widehat{\mathbf{n}}_k \otimes \widehat{\mathbf{n}}_k$. This will be shown in the following.

A tensor $\mathbf{A}$ that leaves all characteristic spaces of $\mathbf{S}$ invariant should be of the form $\mathbf{A} = \sum_{i,j=1}^{3} A_{ij} \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_j$. Then,

$$\mathbf{SA} = \sum_{i,j=1}^{3} \lambda_i A_{ij} \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_j \quad , \quad \mathbf{AS} = \sum_{i,j=1}^{3} A_{ij} \lambda_j \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_j \ .$$

Now, $\mathbf{SA} = \mathbf{AS}$ yields

$$\sum_{i,j=1}^{3} A_{ij} \left(\lambda_i - \lambda_j\right) \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_j = \mathbf{O} \ .$$

This expression implies that $A_{ij} = 0$ for $i \neq j$, since $\lambda_i \neq \lambda_j$ when $i \neq j$ by assumption. This reveals the fact that $\mathbf{A}$ should essentially be a symmetric tensor. These considerations help introduce the important notion of *coaxiality* in, for instance, nonlinear continuum mechanics. This is discussed below.

Two symmetric second-order tensors $\mathbf{S}_1$ and $\mathbf{S}_2$ are said to be *coaxial* if their eigenvectors coincide. It is then a simple exercise to show that the two symmetric second-order tensors $\mathbf{S}_1$ and $\mathbf{S}_2$ are coaxial if and only if they commute, that is,

$$\mathbf{S}_1 \, \mathbf{S}_2 = \mathbf{S}_2 \, \mathbf{S}_1 \ . \tag{4.52}$$

Note that if $\mathbf{S}_1$ and $\mathbf{S}_2$ are coaxial, their simple contraction $\mathbf{S}_1\mathbf{S}_2$ renders another symmetric tensor. To show this, consider

$$\mathbf{S}_1\mathbf{S}_2 = \mathbf{S}_2\mathbf{S}_1 \quad \Rightarrow \quad \underbrace{\mathbf{S}_1\mathbf{S}_2 - (\mathbf{S}_1\mathbf{S}_2)^{\mathrm{T}} = \mathbf{O}}_{\text{recall that } 2\mathrm{skw}\mathbf{S} = \mathbf{S} - \mathbf{S}^{\mathrm{T}}} \quad \Rightarrow \quad \underbrace{\mathrm{skw}\,(\mathbf{S}_1\mathbf{S}_2) = \mathbf{O}}_{\text{recall that } \mathbf{S} = \mathrm{sym}\mathbf{S} + \mathrm{skw}\mathbf{S}} \ .$$

Thus,

$$\mathbf{S}_1\mathbf{S}_2 = \operatorname{sym}(\mathbf{S}_1\mathbf{S}_2) \ .$$

A well-known example of coaxial tensors in nonlinear continuum mechanics regards the right Cauchy-Green strain tensor and the second Piola-Kirchhoff stress tensor for isotropic hyperelastic solids. For such solids, the left Cauchy-Green strain tensor and the true stress tensor are also coaxial, see Exercise 6.16.

Exercise 4.5

Not all tensors can be diagonalized. For instance, there are special tensors called **defective tensors** having fewer eigenvectors than the dimension of space to constitute a basis. It is thus the goal of this exercise to resolve the eigenvector issue of defective tensors by use of the so-called *generalized eigenvectors* or *Jordan vectors*. They supplement the eigenvectors of defective tensors to provide a complete set of linearly independent vectors. As will be seen, the eigenvectors introduced so far can be realized as the generalized eigenvectors of rank 1.

Recall from (4.29) that the diagonal matrix $[\,\mathbf{D}\,]$ of a diagonalizable tensor $\mathbf{A}$ is basically the collection of its components with respect to $\{\mathbf{n}_i \otimes \mathbf{m}_i\}$. This diagonal matrix is a special case of the so-called *Jordan normal form* or *Jordan canonical form*. In alignment with (4.29), the components of a defective tensor $\mathbf{A}$ with respect to a complete basis of its extended eigenvectors deliver the Jordan normal form $[\,\mathbf{J}\,]$ of that tensor. This is indicated by

$$[\,\mathbf{J}\,] = [\,\mathbf{N}\,]^{-1}\,[\,\mathbf{A}\,]\,[\,\mathbf{N}\,] \ , \tag{4.53}$$

where $[\,\mathbf{N}\,]$ is an invertible matrix of linearly independent vectors that nearly diagonalizes $[\,\mathbf{A}\,]$ and $[\,\mathbf{J}\,]$ is of the following form

$$[\,\mathbf{J}\,] = \begin{bmatrix} \mathbf{J}_1 & & \\ & \mathbf{J}_2 & \\ & & \ddots \end{bmatrix} . \tag{4.54}$$

In this expression, each **Jordan block** $[\,\mathbf{J}_i\,]$ renders a square matrix of the form

$$[\,\mathbf{J}_i\,] = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} , \tag{4.55}$$

where λ_i is an eigenvalue that can be equal in different blocks. Its geometric multiplicity presents the number of corresponding Jordan blocks. And its algebraic multiplicity is equal to the sum of the orders of all associated Jordan blocks. In the three-dimensional space, the Jordan canonical form can be any of the following matrices

$$\underbrace{\begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}}_{\substack{\text{if } p_{\mathbf{A}} = (\lambda-\lambda_1)^3 \\ \text{and } t_1 = 1}} , \quad \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}}_{\substack{\text{if } p_{\mathbf{A}} = (\lambda-\lambda_1)^3 \\ \text{and } t_1 = 2}} , \quad \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{bmatrix}}_{\substack{\text{if } p_{\mathbf{A}} = (\lambda-\lambda_1)(\lambda-\lambda_2)^2 \\ \text{and } t_1 = t_2 = 1}} . \tag{4.56}$$

Note that the sum of off-diagonal entries (or the number of repeats of 1 on the superdiagonal) presents the number of generalized eigenvectors to be computed. Let λ_i be a defective eigenvalue of $\mathbf{A}$ and $\mathbf{n}_{ij}$, $j = 1, \cdots, t_i$ be its corresponding eigenvectors. Then, any nonzero vector satisfying

$$(\mathbf{A} - \lambda_i \mathbf{I})^{k+1} \mathbf{u}_{ijk} = \mathbf{0} , \quad j = 1, \cdots, t_i \; ; \; k = 1, 2, \cdots , \tag{4.57}$$

is referred to as a *generalized eigenvector*. The chain of generalized eigenvectors in (4.57) is practically computed via

$$\underbrace{(\mathbf{A} - \lambda_i \mathbf{I})\, \mathbf{u}_{ij1} = \mathbf{n}_{ij} \;\; , \;\; (\mathbf{A} - \lambda_i \mathbf{I})\, \mathbf{u}_{ij2} = \mathbf{u}_{ij1} \;\; , \;\; \cdots}_{\substack{\text{note that } \mathbf{n}_{ij} \text{ present the generalized eigenvectors of rank 1} \\ \text{and } \mathbf{u}_{ij1}\, (\mathbf{u}_{ij2}) \text{ are the generalized eigenvectors of rank 2 (3)}}} . \tag{4.58}$$

It is important to point out that not all these non-homogeneous algebraic equations have a solution. In general, a vector on the right hand side may not be within the column space of $\mathbf{A} - \lambda_i \mathbf{I}$ and hence inconsistency happens. But, one can always find the required vectors from these equations to form a basis. This is eventually a special property of defective tensors that always let attain a basis for the vector space. The proof that eigenvectors and generalized eigenvectors of a defective matrix are linearly independent can be found in any decent book on linear algebra. The Jordan normal form of $\mathbf{A}$ can finally be computed by appropriately setting its eigenvectors as well as generalized eigenvectors in the columns of the matrix $[\,\mathbf{N}\,]$.

Let $\mathbf{A}$ be a Cartesian tensor whose matrix form is

$$[\,\mathbf{A}\,] = \begin{bmatrix} 6 & -2 & -1 \\ 3 & 1 & -1 \\ 2 & -1 & 2 \end{bmatrix} .$$

First, show that $\mathbf{A}$ is **defective**. Then, obtain the **Jordan normal form** of $\mathbf{A}$ making use of its eigenvectors as well as **generalized eigenvectors**. Finally, represent that tensor with respect to a complete basis consisting of the resulting vectors.

Solution. The characteristic equation

$$p_{\mathbf{A}}(\lambda) = \lambda^3 - 9\lambda^2 + 27\lambda - 27 = (\lambda - 3)^3 = 0 \ ,$$

renders one distinct eigenvalue $\lambda_1 = 3$ with algebraic multiplicity 3, i.e. $s_1 = 3$. To see whether $\mathbf{A}$ is defective or not, one should have the geometric multiplicity of this multiple eigenvalue. To compute t_1, one needs to find nonzero vectors $\mathbf{n} = n_i \widehat{\mathbf{e}}_i$ satisfying the eigenvalue problem $(\mathbf{A} - 3\mathbf{I})\,\mathbf{n} = \mathbf{0}$. By elementary row operation, the matrix form

$$[\,\mathbf{A} - 3\mathbf{I}\,][\,\mathbf{n}\,] = \begin{bmatrix} 3 & -2 & -1 \\ 3 & -2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \ ,$$

is reduced to

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \ .$$

It follows that $[\,\mathbf{n}\,]^{\mathrm{T}} = \begin{bmatrix} n_1 \ n_2 \ n_3 \end{bmatrix}$ should satisfy $n_1 = n_3$ and $n_2 = n_3$. Thus, the eigenvectors corresponding to $\lambda_1 = 3$ are

$$[\,\mathbf{n}\,] = n_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \ .$$

Consequently,

$$\mathscr{N}(\mathbf{A} - 3\mathbf{I}) = \mathrm{Span}\left\{\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3\right\} \quad \text{delivers} \quad t_1 = 1 \ .$$

This reveals the fact that the deficit of $\lambda_1 = 3$ is $s_1 - t_1 = 2$ and thus $\mathbf{A}$ is defective. Having recorded $\mathbf{n}_{11} = \widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3$, two vectors $\mathbf{u}_{111} = (\mathbf{u}_{111})_i \, \widehat{\mathbf{e}}_i$ and $\mathbf{u}_{112} = (\mathbf{u}_{112})_i \, \widehat{\mathbf{e}}_i$, with the aid of (4.58), should be supplemented for providing a complete basis. It follows that

$$\begin{bmatrix} 3 & -2 & -1 \\ 3 & -2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} (\mathbf{u}_{111})_1 \\ (\mathbf{u}_{111})_2 \\ (\mathbf{u}_{111})_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \left.\begin{array}{l} (\mathbf{u}_{111})_1 = 0 \\ (\mathbf{u}_{111})_2 = 0 \\ (\mathbf{u}_{111})_3 = -1 \end{array}\right\} \ .$$

Thus,

$$\mathbf{u}_{111} = -\widehat{\mathbf{e}}_3 \ .$$

And this helps obtain

$$\begin{bmatrix} 3 & -2 & -1 \\ 3 & -2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} (\mathbf{u}_{112})_1 \\ (\mathbf{u}_{112})_2 \\ (\mathbf{u}_{112})_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad \Rightarrow \quad \left. \begin{array}{l} (\mathbf{u}_{112})_1 = 0 \\ (\mathbf{u}_{112})_2 = -1 \\ (\mathbf{u}_{112})_3 = 2 \end{array} \right\} .$$

Hence,

$$\mathbf{u}_{112} = -\widehat{\mathbf{e}}_2 + 2\,\widehat{\mathbf{e}}_3 \; .$$

Observe that $\mathbf{n}_{11} \cdot (\mathbf{u}_{111} \times \mathbf{u}_{112}) \neq 0$. Thus, the triad $\{\mathbf{n}_{11}, \mathbf{u}_{111}, \mathbf{u}_{112}\}$ forms a basis for $\mathscr{E}_{\mathrm{r}}^{o3}$. This allows one to construct the following nonsingular matrix

$$[\,\mathbf{N}\,] = \begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{n}_{11} & \mathbf{u}_{111} & \mathbf{u}_{112} \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix} ,$$

whose inverse helps provide the dual basis of $\{\mathbf{n}_{11}, \mathbf{u}_{111}, \mathbf{u}_{112}\}$ as follows:

$$[\,\mathbf{N}\,]^{-1} = \begin{bmatrix} \cdots & \mathbf{m}_{11} & \cdots \\ \cdots & \mathbf{v}_{111} & \cdots \\ \cdots & \mathbf{v}_{112} & \cdots \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & -2 & -1 \\ 1 & -1 & 0 \end{bmatrix} .$$

Thus,

$$\mathbf{m}_{11} = \widehat{\mathbf{e}}_1 \quad , \quad \mathbf{v}_{111} = 3\,\widehat{\mathbf{e}}_1 - 2\,\widehat{\mathbf{e}}_2 - \widehat{\mathbf{e}}_3 \quad , \quad \mathbf{v}_{112} = \widehat{\mathbf{e}}_1 - \widehat{\mathbf{e}}_2 \; .$$

By having all required data at hand, one can finally arrive at

$$[\,\mathbf{J}\,] = [\,\mathbf{N}\,]^{-1}\,[\,\mathbf{A}\,]\,[\,\mathbf{N}\,] = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} ,$$

and

$$\mathbf{A} = 3\,(\mathbf{n}_{11} \otimes \mathbf{m}_{11} + \mathbf{u}_{111} \otimes \mathbf{v}_{111} + \mathbf{u}_{112} \otimes \mathbf{v}_{112}) + \mathbf{n}_{11} \otimes \mathbf{v}_{111} + \mathbf{u}_{111} \otimes \mathbf{v}_{112} \; .$$

In this representation, the two extra dyads $\mathbf{n}_{11} \otimes \mathbf{v}_{111}$ and $\mathbf{u}_{111} \otimes \mathbf{v}_{112}$ are resulted from the deficiency of the tensor **A**. As a result, a defective tensor may be thought of as a **nearly diagonalizable tensor**.

Exercise 4.6

Let $\mathbf{A}$ and $\mathbf{B}$ be two second-order tensors with $\mathbf{exp}\,(\mathbf{A}) = \sum_{k=0}^{\infty} (k!)^{-1}\,\mathbf{A}^k$ and $\mathbf{exp}\,(\mathbf{B}) = \sum_{k=0}^{\infty} (k!)^{-1}\,\mathbf{B}^k$, according to (2.39). Then, verify the following identities

$$\det\left[\mathbf{exp}\,(\mathbf{A})\right] = \exp\left(\mathrm{tr}\mathbf{A}\right) \ , \tag{4.59a}$$

$$\det\left[\mathbf{exp}\,(\mathbf{A})\,\mathbf{exp}\,(\mathbf{B})\right] = \underbrace{\exp\left(\mathrm{tr}\left[\mathbf{A}+\mathbf{B}\right]\right)}_{=\det\left[\mathbf{exp}\,(\mathbf{A}+\mathbf{B})\right]} \ . \tag{4.59b}$$

Moreover, prove that

$$\mathbf{exp}\,(\mathbf{W}) = \mathbf{I} + \frac{\sin|\boldsymbol{\omega}|}{|\boldsymbol{\omega}|}\mathbf{W} + \frac{1}{2}\left[\frac{\sin\left(|\boldsymbol{\omega}|/2\right)}{|\boldsymbol{\omega}|/2}\right]^2 \mathbf{W}^2 \ , \tag{4.60}$$

where $\mathbf{W}^{\mathrm{T}} = -\mathbf{W}$ presents a skew tensor and $\boldsymbol{\omega}$ denotes its axial vector satisfying $\mathbf{W}\mathbf{u} = \boldsymbol{\omega} \times \mathbf{u}$ for any generic vector $\mathbf{u}$.

Solution. Guided by (4.3), if $(\lambda_i, \mathbf{n}_i)$ is an eigenpair of $\mathbf{A}$, then $(\exp(\lambda_i), \mathbf{n}_i)$ represents an eigenpair of $\mathbf{exp}\,(\mathbf{A})$, see also (4.37c). By means of (4.14a) and (4.14c), one can write

$$\det\left[\mathbf{exp}\,(\mathbf{A})\right] = \prod_{i=1}^{3} \exp(\lambda_i) = \exp\left(\sum_{i=1}^{3}\lambda_i\right) = \exp\left(\mathrm{tr}\mathbf{A}\right) \ .$$

In the context of matrix algebra, this identity is usually proved taking advantage of the spectral decomposition. This will be carried out in the following for completeness. Let $[\,\mathbf{A}\,]$ be a non-defective matrix that can be diagonalized according to (4.29), i.e. $[\,\mathbf{A}\,] = [\,\mathbf{N}\,][\,\mathbf{D}\,][\,\mathbf{N}\,]^{-1}$. Then, in light of the property (2.207h), one will have

$$\mathbf{exp}\,([\,\mathbf{A}\,]) = [\,\mathbf{N}\,]\,\mathbf{exp}\,([\,\mathbf{D}\,])\left[\,\mathbf{N}^{-1}\,\right] \quad \text{where} \quad \mathbf{exp}\,([\,\mathbf{A}\,]) = \sum_{k=0}^{\infty}\frac{[\,\mathbf{A}\,]^k}{k!} \ . \tag{4.61}$$

It follows that

$$\det\left[\mathbf{exp}\,([\,\mathbf{A}\,])\right] = \det\left[\mathbf{exp}\,([\,\mathbf{D}\,])\right] = \prod_{i=1}^{3}\exp(\lambda_i) = \exp\left(\mathrm{tr}\,[\,\mathbf{A}\,]\right) \ .$$

Now, let $[\,\mathbf{A}\,]$ be a defective matrix. Guided by (4.53), this matrix can be written as $[\,\mathbf{A}\,] = [\,\mathbf{N}\,][\,\mathbf{J}\,][\,\mathbf{N}\,]^{-1}$. Consequently, the expression (4.61) takes the form $\mathbf{exp}\,([\,\mathbf{A}\,]) = [\,\mathbf{N}\,]\,\mathbf{exp}\,([\,\mathbf{J}\,])\left[\,\mathbf{N}^{-1}\,\right]$. The structure of the Jordan canonical form $[\,\mathbf{J}\,]$ is in such a way that its exponential renders an **upper triangular matrix** with e^{λ_i} on the main diagonal. It is evident that the determinant of an upper triangular matrix is independent of its off-diagonal elements (simply the product of the diagonal elements of such a matrix delivers its third invariant). Therefore, the identity (4.59a) is guaranteed for any non-defective as well as defective matrix. Consistent with (4.56), **exp** ([**J**]) can be any of the following matrices in the three-dimensional space

$$\underbrace{\begin{bmatrix} e^{\lambda_1} & e^{\lambda_1} & e^{\lambda_1}/2 \\ 0 & e^{\lambda_1} & e^{\lambda_1} \\ 0 & 0 & e^{\lambda_1} \end{bmatrix}}_{\substack{\text{if } p_{\mathbf{A}} = (\lambda - \lambda_1)^3 \\ \text{and } t_1 = 1}} , \quad \underbrace{\begin{bmatrix} e^{\lambda_1} & 0 & 0 \\ 0 & e^{\lambda_1} & e^{\lambda_1} \\ 0 & 0 & e^{\lambda_1} \end{bmatrix}}_{\substack{\text{if } p_{\mathbf{A}} = (\lambda - \lambda_1)^3 \\ \text{and } t_1 = 2}} , \quad \underbrace{\begin{bmatrix} e^{\lambda_1} & 0 & 0 \\ 0 & e^{\lambda_2} & e^{\lambda_2} \\ 0 & 0 & e^{\lambda_2} \end{bmatrix}}_{\substack{\text{if } p_{\mathbf{A}} = (\lambda - \lambda_1)(\lambda - \lambda_2)^2 \\ \text{and } t_1 = t_2 = 1}} . \tag{4.62}$$

The interested reader may want to use another technique to verify (4.59a). This has been demonstrated in Exercise 6.8.

Having in mind the linearity of the trace operator, the result (4.59a) helps verify (4.59b) as follows:

$$\begin{aligned} \det\left[\mathbf{exp}\,(\mathbf{A})\,\mathbf{exp}\,(\mathbf{B})\right] &= \det\left[\mathbf{exp}\,(\mathbf{A})\right]\det\left[\mathbf{exp}\,(\mathbf{B})\right] \\ &= \exp\,(\text{tr}\mathbf{A})\exp\,(\text{tr}\mathbf{B}) \\ &= \exp\,(\text{tr}\mathbf{A} + \text{tr}\mathbf{B}) \\ &= \exp\,(\text{tr}\,[\mathbf{A} + \mathbf{B}]) \\ &= \det\left[\mathbf{exp}\,(\mathbf{A} + \mathbf{B})\right] . \end{aligned}$$

In the following, the identity (4.60) will be proved making use of the Cayley-Hamilton equation (4.21). Bearing in mind $(2.89\text{a})_2$ and $(2.96)_2$, the principal scalar invariants

$$I_1\,(\mathbf{W}) \xlongequal[\text{(4.17a)}]{\text{from}} \text{tr}\mathbf{W} = 0 , \tag{4.63a}$$

$$I_2\,(\mathbf{W}) \xlongequal[\text{(4.17b) and (4.63a)}]{\text{from}} -\frac{1}{2}\text{tr}\mathbf{W}^2 \xlongequal[\text{(2.68)}]{\text{from}} |\boldsymbol{\omega}|^2 , \tag{4.63b}$$

$$I_3\,(\mathbf{W}) \xlongequal[\text{(4.17c) and (4.63a)}]{\text{from}} \frac{1}{3}\text{tr}\mathbf{W}^3 = 0 , \tag{4.63c}$$

help present some powers of **W** as follows:

$$\left.\begin{aligned} \mathbf{W}^3 &= -\,|\boldsymbol{\omega}|^2\,\mathbf{W} \\ \mathbf{W}^4 &= -\,|\boldsymbol{\omega}|^2\,\mathbf{W}^2 \end{aligned}\right\} , \quad \left.\begin{aligned} \mathbf{W}^5 &= +\,|\boldsymbol{\omega}|^4\,\mathbf{W} \\ \mathbf{W}^6 &= +\,|\boldsymbol{\omega}|^4\,\mathbf{W}^2 \end{aligned}\right\} , \quad \left.\begin{aligned} \mathbf{W}^7 &= -\,|\boldsymbol{\omega}|^6\,\mathbf{W} \\ \mathbf{W}^8 &= -\,|\boldsymbol{\omega}|^6\,\mathbf{W}^2 \end{aligned}\right\} . \tag{4.64}$$

Thus,

$$\mathbf{exp}\,(\mathbf{W}) = \sum_{k=0}^{\infty} \frac{\mathbf{W}^k}{k!} = \mathbf{I} + \left[|\boldsymbol{\omega}| - \frac{|\boldsymbol{\omega}|^3}{3!} + \frac{|\boldsymbol{\omega}|^5}{5!} - \cdots \right] \frac{\mathbf{W}}{|\boldsymbol{\omega}|}$$
$$- \left[-\frac{|\boldsymbol{\omega}|^2}{2!} + \frac{|\boldsymbol{\omega}|^4}{4!} - \frac{|\boldsymbol{\omega}|^6}{6!} + \cdots \right] \frac{\mathbf{W}^2}{|\boldsymbol{\omega}|^2}$$
$$= \mathbf{I} + \frac{\sin |\boldsymbol{\omega}|}{|\boldsymbol{\omega}|}\mathbf{W} - \frac{\cos |\boldsymbol{\omega}| - 1}{|\boldsymbol{\omega}|^2}\mathbf{W}^2 \,. \longleftarrow \text{note that } \sin^2\frac{\theta}{2} = \frac{1-\cos\theta}{2}$$

Exercise 4.7

Let $\text{dev}\mathbf{A}$ be the deviatoric part of a second-order tensor $\mathbf{A}$ according to (2.145), i.e. $\text{dev}\mathbf{A} = \mathbf{A} - (1/3)\,(\text{tr}\mathbf{A})\,\mathbf{I}$. Then, show that its invariants are

$$I_1\,(\text{dev}\mathbf{A}) = 0\,, \tag{4.65a}$$

$$I_2\,(\text{dev}\mathbf{A}) = -\frac{1}{2}\text{tr}\,(\text{dev}\mathbf{A})^2\,, \tag{4.65b}$$

$$I_3\,(\text{dev}\mathbf{A}) = \frac{1}{3}\text{tr}\,(\text{dev}\mathbf{A})^3\,. \tag{4.65c}$$

Solution. Knowing that the trace of a tensor is a symmetric bilinear form, one can write

$$I_1\,(\text{dev}\mathbf{A}) \xlongequal[\text{(4.17a)}]{\text{from}} \text{tr}\,(\text{dev}\mathbf{A})$$
$$\xlongequal[\text{(2.83) and (2.145)}]{\text{from}} \text{tr}\mathbf{A} - \frac{\text{tr}\mathbf{A}}{3}\text{tr}\mathbf{I}$$
$$\xlongequal[\text{(2.90)}]{\text{from}} 0\,,$$
$$I_2\,(\text{dev}\mathbf{A}) \xlongequal[\text{(4.17b)}]{\text{from}} \frac{1}{2}\left[(\text{tr}\,(\text{dev}\mathbf{A}))^2 - \text{tr}\,(\text{dev}\mathbf{A})^2\right]$$
$$\xlongequal[\text{(4.65a)}]{\text{from}} -\frac{1}{2}\text{tr}\,(\text{dev}\mathbf{A})^2\,,$$
$$I_3\,(\text{dev}\mathbf{A}) \xlongequal[\text{(4.17c)}]{\text{from}} \frac{1}{3}\text{tr}\,(\text{dev}\mathbf{A})^3 - \frac{1}{2}\text{tr}\,(\text{dev}\mathbf{A})\,\text{tr}\,(\text{dev}\mathbf{A})^2 + \frac{1}{6}\,(\text{tr}\,(\text{dev}\mathbf{A}))^3$$
$$\xlongequal[\text{(4.65a)}]{\text{from}} \frac{1}{3}\text{tr}\,(\text{dev}\mathbf{A})^3\,.$$

Note that the result (4.65c) can also be obtained from the Cayley-Hamilton equation (4.21) in the following form

$$\underbrace{(\text{dev}\mathbf{A})^3 - \cancelto{0}{I_1(\text{dev}\mathbf{A})}\,(\text{dev}\mathbf{A})^2 + I_2(\text{dev}\mathbf{A})\,\text{dev}\mathbf{A} - I_3(\text{dev}\mathbf{A})\,\mathbf{I} = \mathbf{O}}_{\text{or } I_3(\text{dev}\mathbf{A})\,\mathbf{I} = (\text{dev}\mathbf{A})^3 + I_2(\text{dev}\mathbf{A})\,\text{dev}\mathbf{A}}\,,$$

taking into account $\text{tr}\mathbf{I} = 3$ and $\text{tr}(\text{dev}\mathbf{A}) = 0$.

Exercise 4.8

Use the expressions (4.65a)–(4.65c) to verify that $I_2(\text{dev}\mathbf{A})$ and $I_3(\text{dev}\mathbf{A})$ may be represented, in terms of $I_1(\mathbf{A})$, $I_2(\mathbf{A})$ and $I_3(\mathbf{A})$, by

$$I_2(\text{dev}\mathbf{A}) = I_2(\mathbf{A}) - \frac{1}{3}I_1^2(\mathbf{A}) \;, \tag{4.66a}$$

$$I_3(\text{dev}\mathbf{A}) = I_3(\mathbf{A}) - \frac{1}{3}I_1(\mathbf{A})\,I_2(\mathbf{A}) + \frac{2}{27}I_1^3(\mathbf{A}) \;. \tag{4.66b}$$

That is exactly why $I_1(\mathbf{A})$ to $I_3(\mathbf{A})$ are called the **principal** invariants since any other invariant such as $I_2(\text{dev}\mathbf{A})$ or $I_3(\text{dev}\mathbf{A})$ is expressible in terms of them.

Solution. First, with the aid of (2.29), (2.33), (2.34) and (2.36), one needs to have the useful relations

$$\begin{aligned}(\text{dev}\mathbf{A})^2 &= \left[\mathbf{A} - \frac{1}{3}I_1(\mathbf{A})\,\mathbf{I}\right]\left[\mathbf{A} - \frac{1}{3}I_1(\mathbf{A})\,\mathbf{I}\right] \\ &= \mathbf{A}^2 - \frac{2}{3}I_1(\mathbf{A})\,\mathbf{A} + \frac{1}{9}I_1^2(\mathbf{A})\,\mathbf{I}\,,\end{aligned} \tag{4.67a}$$

$$\begin{aligned}(\text{dev}\mathbf{A})^3 &= \underbrace{\left[\mathbf{A}^2 - \frac{2}{3}I_1(\mathbf{A})\,\mathbf{A} + \frac{1}{9}I_1^2(\mathbf{A})\,\mathbf{I}\right]\left[\mathbf{A} - \frac{1}{3}I_1(\mathbf{A})\,\mathbf{I}\right]}_{= \mathbf{A}^3 - \frac{2}{3}I_1(\mathbf{A})\mathbf{A}^2 + \frac{1}{9}I_1^2(\mathbf{A})\mathbf{A} - \frac{1}{3}I_1(\mathbf{A})\mathbf{A}^2 + \frac{2}{9}I_1^2(\mathbf{A})\mathbf{A} - \frac{1}{27}I_1^3(\mathbf{A})\mathbf{I}} \\ &= \mathbf{A}^3 - I_1(\mathbf{A})\,\mathbf{A}^2 + \frac{1}{3}I_1^2(\mathbf{A})\,\mathbf{A} - \frac{1}{27}I_1^3(\mathbf{A})\,\mathbf{I}\,.\end{aligned} \tag{4.67b}$$

Then, by linearity of the trace operator, $(2.90)_3$ and $(4.17\text{c})_1$, one can obtain

$$\begin{aligned}I_2(\text{dev}\mathbf{A}) &= \underbrace{-\frac{1}{2}\mathbf{I} : \left[\mathbf{A}^2 - \frac{2}{3}I_1(\mathbf{A})\,\mathbf{A} + \frac{1}{9}I_1^2(\mathbf{A})\,\mathbf{I}\right]}_{= -\frac{1}{2}\left[\text{tr}\mathbf{A}^2 - \frac{2}{3}I_1^2(\mathbf{A}) + \frac{1}{3}I_1^2(\mathbf{A})\right]} \\ &= \frac{1}{2}\left[I_1^2(\mathbf{A}) - \text{tr}\mathbf{A}^2 - \frac{2}{3}I_1^2(\mathbf{A})\right] \\ &= I_2(\mathbf{A}) - \frac{1}{3}I_1^2(\mathbf{A}) \;,\end{aligned}$$

$$I_3(\text{dev}\mathbf{A}) = \underbrace{\frac{1}{3}\mathbf{I} : \left[\mathbf{A}^3 - I_1(\mathbf{A})\,\mathbf{A}^2 + \frac{1}{3}I_1^2(\mathbf{A})\,\mathbf{A} - \frac{1}{27}I_1^3(\mathbf{A})\,\mathbf{I}\right]}_{= \frac{1}{3}\text{tr}\mathbf{A}^3 - \frac{1}{3}I_1(\mathbf{A})\,\text{tr}\mathbf{A}^2 + \frac{1}{9}I_1^3(\mathbf{A}) - \frac{1}{27}I_1^3(\mathbf{A})}$$

$$= \underbrace{\left\{\frac{1}{3}\text{tr}\mathbf{A}^3\right\} - \frac{1}{3}I_1(\mathbf{A})\,\text{tr}\mathbf{A}^2 + \frac{2}{27}I_1^3(\mathbf{A})}_{= \left\{I_3(\mathbf{A}) + \cancel{\frac{1}{3}I_1(\mathbf{A})\,\text{tr}\mathbf{A}^2} - \frac{1}{3}I_2(\mathbf{A})\,I_1(\mathbf{A})\right\} \cancel{- \frac{1}{3}I_1(\mathbf{A})\,\text{tr}\mathbf{A}^2} + \frac{2}{27}I_1^3(\mathbf{A})}$$

$$= I_3(\mathbf{A}) - \frac{1}{3}I_1(\mathbf{A})\,I_2(\mathbf{A}) + \frac{2}{27}I_1^3(\mathbf{A}) \ .$$

Exercise 4.9

Let $\mathbf{Q}$ be a real orthogonal tensor and further let $\mathbf{W}$ be a real skew tensor. Then, find the eigenvalues of $\mathbf{Q}$ and show that the eigenvalues of $\mathbf{W}$ are either 0 or purely imaginary numbers.

Solution. Each eigenvalue λ and the corresponding eigenvector $\mathbf{n}$ of $\mathbf{Q}$ satisfy $\mathbf{Q}\,\mathbf{n} = \lambda\mathbf{n}$ and since $\overline{\mathbf{Q}} = \mathbf{Q}$ by assumption, one will have $\mathbf{Q}\,\overline{\mathbf{n}} = \overline{\lambda}\,\overline{\mathbf{n}}$. Pre-multiplying the eigenvalue problem $\mathbf{Q}\,\mathbf{n} = \lambda\mathbf{n}$ by $\mathbf{Q}^{\text{T}}$ yields

$$\mathbf{Q}^{\text{T}}\mathbf{Q}\,\mathbf{n} = \mathbf{Q}^{\text{T}}(\lambda\,\mathbf{n}) \xrightarrow[\text{(2.8b) and (2.130)}]{\text{from}} \mathbf{I}\,\mathbf{n} = \lambda\mathbf{Q}^{\text{T}}\,\mathbf{n} \xrightarrow[\text{(2.5) and (2.48)}]{\text{from}} \mathbf{n}\mathbf{Q} = \lambda^{-1}\mathbf{n}\ ,$$

and, in a similar fashion,

$$\overline{\mathbf{n}}\mathbf{Q} = \overline{\lambda}^{-1}\overline{\mathbf{n}}\ .$$

On the one hand, $\overline{\mathbf{n}}\cdot(\mathbf{Q}\,\mathbf{n}) = \overline{\mathbf{n}}\cdot(\lambda\mathbf{n})$ yields $(\overline{\mathbf{n}}\mathbf{Q})\cdot\mathbf{n} = \lambda\,\overline{\mathbf{n}}\cdot\mathbf{n}$. And, on the other hand, $(\overline{\mathbf{n}}\mathbf{Q})\cdot\mathbf{n} = \left(\overline{\lambda}^{-1}\overline{\mathbf{n}}\right)\cdot\mathbf{n}$ gives $(\overline{\mathbf{n}}\mathbf{Q})\cdot\mathbf{n} = \overline{\lambda}^{-1}\overline{\mathbf{n}}\cdot\mathbf{n}$. Knowing that $\mathbf{n} \neq \mathbf{0}$ by definition, the identity (1.22c) will provide $\overline{\mathbf{n}}\cdot\mathbf{n} > 0$. Thus,

$$\lambda\overline{\lambda} = 1\ .$$

And this implies that

$$\lambda_1 = \pm 1 \quad , \quad \lambda_2 = \underbrace{\exp(\text{i}\theta)}_{= \cos\theta + \text{i}\sin\theta} \quad , \quad \lambda_3 = \underbrace{\exp(-\text{i}\theta)}_{= \cos\theta - \text{i}\sin\theta}\ . \tag{4.68}$$

Guided by these relations, for instance, the eigenvalues of the rotation matrix $(2.195)_1$ will be $\lambda = +1, \exp(\text{i}\theta), \exp(-\text{i}\theta)$ and the eigenvalues of the reflection matrix $(2.195)_2$ are simply $\lambda = -1, 1, 1$.

For a real skew-symmetric tensor $\mathbf{W}$ satisfying $\mathbf{W}^{\text{T}} = -\mathbf{W}$, the eigenvalue problem $\mathbf{W}\,\mathbf{n} = \lambda\mathbf{n}$ or $\mathbf{W}^{\text{T}}\,\mathbf{n} = -\lambda\mathbf{n}$ with the aid of (2.48) delivers $\mathbf{n}\mathbf{W} = -\lambda\mathbf{n}$. The complex conjugate of the resulting expression $\mathbf{n}\mathbf{W} = -\lambda\mathbf{n}$ is $\overline{\mathbf{n}}\,\mathbf{W} = -\overline{\lambda}\,\overline{\mathbf{n}}$. It is then easy to deduce from $\overline{\mathbf{n}} \cdot (\mathbf{W}\,\mathbf{n}) = \overline{\mathbf{n}} \cdot (\lambda \mathbf{n})$ or $(\overline{\mathbf{n}}\,\mathbf{W}) \cdot \mathbf{n} = \lambda \overline{\mathbf{n}} \cdot \mathbf{n}$ and $(\overline{\mathbf{n}}\,\mathbf{W}) \cdot \mathbf{n} = \left(-\overline{\lambda}\overline{\mathbf{n}}\right) \cdot \mathbf{n}$ or $(\overline{\mathbf{n}}\,\mathbf{W}) \cdot \mathbf{n} = -\overline{\lambda}\overline{\mathbf{n}} \cdot \mathbf{n}$ that

$$\lambda = -\overline{\lambda} \,.$$

Finally, one can infer that

$$\lambda_1 = 0 \quad , \quad \lambda_2 = \theta \mathrm{i} \quad , \quad \lambda_3 = -\theta \mathrm{i} \,. \tag{4.69}$$

References

1. Wilkinson, J. H. (1988). *The algebraic eigenvalue problem (Numerical Mathematics and Scientific Computation): Revised Edition*. Oxford Science Publications.
2. Parlett, B. N. (1998). *The symmetric eigenvalue problem*. Society for Industrial and Applied Mathematics.
3. Saad, Y. (2011). *Numerical methods for large eigenvalue Problems: Revised Edition*. Society for Industrial and Applied Mathematics.
4. Qi, L., Chen, H., & Chen, Y. (2018). *Tensor eigenvalues and their applications (Advances in Mechanics and Mathematics 39)*. Singapore: Springer.
5. Bronshtein, I. N., Semendyayev, K. A., Musiol, G., & Mühlig, H. (2015). *Handbook of mathematics* (6th ed.). Berlin, Heidelberg: Springer.
6. Birkhoff, G., & Mac Lane, S. (1969). *A survey of modern algebra* (3rd ed.). Macmillan Company.
7. Hazewinkel, M. (Ed.) (1994). *Encyclopedia of mathematics* (Vol. 10) (Index). Kluwer.
8. Itskov, M. (2015). *Tensor algebra and tensor analysis for engineers: With applications to continuum mechanics* (4th ed.). Cham: Springer.
9. Liu, I.-S. (2002). *Continuum mechanics*. Berlin, Heidelberg: Springer.

Chapter 5
Representation of Tensorial Variables in Curvilinear Coordinates

So far tensorial variables have been expressed with respect to a Cartesian coordinate frame defined by an origin and the standard basis vectors. Sometimes, the symmetry of a problem demands another set of coordinates. For instance, **cylindrical coordinates** are usually used when there is symmetry about the cylindrical axis or **spherical coordinates** are beneficial when geometry of the problem has symmetry about the corresponding origin. These commonly used coordinate systems are examples of what is known as the *curvilinear coordinate system*. This general class of coordinate systems is used in many branches of physics and engineering such as general relativity and structural mechanics. Regarding nonlinear continuum mechanics, there are lots of pioneering articles on material modeling of solids that are written in curvilinear coordinates. Moreover, (mathematical) foundations of elasticity as well as (geometrical) foundations of continuum mechanics are often explained in the literature by means of this powerful coordinate system, see Marsden and Hughes [1] and Steinmann [2]. See also Synge and Schild [3], Başar and Weichert [4] and Hashiguchi [5]. The main goal of this chapter is thus to work with the old art of curvilinear coordinates in order to represent general forms of tensorial variables and study their fundamental properties in a general framework. For more discussions, see Brannon [6] and references therein.

5.1 Basis Vectors

Consider a general curvilinear coordinate system as shown in Fig. 5.1. This can be realized as **local** coordinates embedded in a Cartesian frame of reference as **global** coordinates. As can be seen, the coordinate lines can be curved. This implies that the tangent vectors to these curves change their directions continuously from point to point. For an arbitrary point $\mathbf{x}$ in the space, there is a one-to-one map transforming the curvilinear coordinates $\left(\Theta^1, \Theta^2, \Theta^3\right)$ into the Cartesian coordinates (x_1, x_2, x_3). This is indicated by

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_5

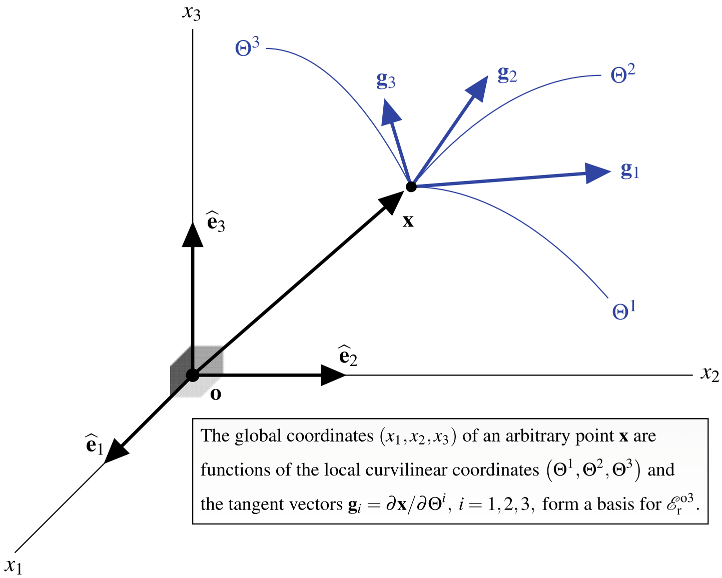


Fig. 5.1 Curvilinear coordinate system

$$\underbrace{x_1 = \hat{x}_1\left(\Theta^1, \Theta^2, \Theta^3\right) \;,\; x_2 = \hat{x}_2\left(\Theta^1, \Theta^2, \Theta^3\right) \;,\; x_3 = \hat{x}_3\left(\Theta^1, \Theta^2, \Theta^3\right)}_{\text{or, simply, } \mathbf{x} = \hat{\mathbf{x}}\left(\Theta^i\right)} . \tag{5.1}$$

A point whose components are functions of one or more variables is called a *point function*. The general transformation (5.1) is assumed to be invertible which allows one to write

$$\underbrace{\Theta^1 = \hat{\Theta}^1\left(x_1, x_2, x_3\right) \;,\; \Theta^2 = \hat{\Theta}^2\left(x_1, x_2, x_3\right) \;,\; \Theta^3 = \hat{\Theta}^3\left(x_1, x_2, x_3\right)}_{\text{or, simply, } \Theta^i = \hat{\Theta}^i(\mathbf{x})} . \tag{5.2}$$

5.1.1 Covariant Basis Vectors

The tangent vectors $\mathbf{g}_i$, $i = 1, 2, 3$, for the curvilinear coordinate system are defined by

$$\mathbf{g}_i = \frac{\partial \mathbf{x}}{\partial \Theta^i} = \lim_{h \to 0} \frac{\hat{\mathbf{x}}\left(\Theta^1 + h\delta_{i1}, \Theta^2 + h\delta_{i2}, \Theta^3 + h\delta_{i3}\right) - \hat{\mathbf{x}}\left(\Theta^1, \Theta^2, \Theta^3\right)}{h} . \tag{5.3}$$

It is important to note that in alignment with the rule (1.23a), for instance, the difference $\hat{\mathbf{x}}\left(\Theta^1 + h, \Theta^2, \Theta^3\right) - \hat{\mathbf{x}}\left(\Theta^1, \Theta^2, \Theta^3\right)$ represents a vector while $\mathbf{x}$ itself is a point. The tangent vectors

$$\{\mathbf{g}_i\} := \{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\} \ , \tag{5.4}$$

basically constitute a basis for $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$. They are referred to as *general basis* or *general basis vectors* in this text. They are eventually general in a way that they can consistently deliver the basis vectors of commonly used coordinate systems. In the following, the correspondence between the global and some well-known local coordinates as well as their basis vectors relationships are presented. ✕

To begin with, for convenience, the global coordinates and basis vectors are denoted by $x_1 = x$, $x_2 = y$, $x_3 = z$ and $\widehat{\mathbf{e}}_1 = \widehat{\mathbf{e}}_x$, $\widehat{\mathbf{e}}_2 = \widehat{\mathbf{e}}_y$, $\widehat{\mathbf{e}}_3 = \widehat{\mathbf{e}}_z$, respectively. First, consider a cylindrical coordinate system for which $\Theta^1 = r$, $\Theta^2 = \theta$ and $\Theta^3 = z$. The **Cartesian coordinates** (x, y, z) of a point can be obtained from its **cylindrical coordinates** (r, θ, z) according to

$$\boxed{x = r\cos\theta \quad , \quad y = r\sin\theta \quad , \quad z = z \ .} \tag{5.5}$$

Conversely, the local coordinates (r, θ, z) can be expressed in terms of the global coordinates (x, y, z) in some alternative forms. One simple form is

$$r = \sqrt{x^2 + y^2} \quad , \quad \theta = \begin{cases} +\cos^{-1}(x/r) & \text{if } y \geq 0 \ , \ r \neq 0 \\ -\cos^{-1}(x/r) & \text{if } y < 0 \ , \ r \neq 0 \\ \text{undefined} & \text{if } r = 0 \end{cases} \quad , \quad z = z \ . \tag{5.6}$$

With $dx = \cos\theta dr - r\sin\theta d\theta$, $dy = \sin\theta dr + r\cos\theta d\theta$ and $dz = dz$, the general basis vectors render

$$\mathbf{g}_1 = \frac{\partial \mathbf{x}}{\partial r} = \widehat{\mathbf{e}}_r \qquad \text{where} \qquad \widehat{\mathbf{e}}_r = \cos\theta\,\widehat{\mathbf{e}}_x + \sin\theta\,\widehat{\mathbf{e}}_y \ , \tag{5.7a}$$

$$\mathbf{g}_2 = \frac{\partial \mathbf{x}}{\partial \theta} = r\,\widehat{\mathbf{e}}_\theta \qquad \text{where} \qquad \widehat{\mathbf{e}}_\theta = -\sin\theta\,\widehat{\mathbf{e}}_x + \cos\theta\,\widehat{\mathbf{e}}_y \ , \tag{5.7b}$$

$$\mathbf{g}_3 = \frac{\partial \mathbf{x}}{\partial z} = \widehat{\mathbf{e}}_z \ , \tag{5.7c}$$

noting that $\{\widehat{\mathbf{e}}_r, \widehat{\mathbf{e}}_\theta, \widehat{\mathbf{e}}_z\}$ represent an orthonormal basis.

Next, consider a spherical coordinate system for which $\Theta^1 = r$, $\Theta^2 = \theta$ and $\Theta^3 = \phi$. The **Cartesian coordinates** (x, y, z) and the **spherical coordinates** (r, θ, ϕ) are related through the following relation

$$\boxed{x = r\sin\theta\cos\phi \quad , \quad y = r\sin\theta\sin\phi \quad , \quad z = r\cos\theta \ ,} \tag{5.8}$$

or, conversely,

$$r = \sqrt{x^2 + y^2 + z^2} \quad , \quad \theta = \cos^{-1} \frac{z}{\sqrt{x^2 + y^2 + z^2}} \quad , \quad \phi = \tan^{-1} \frac{y}{x} \ . \tag{5.9}$$

Now, the line element

$$\begin{aligned} d\mathbf{x} = & (\sin\theta \cos\phi dr + r \cos\theta \cos\phi d\theta - r \sin\theta \sin\phi d\phi) \widehat{\mathbf{e}}_x \\ & + (\sin\theta \sin\phi dr + r \cos\theta \sin\phi d\theta + r \sin\theta \cos\phi d\phi) \widehat{\mathbf{e}}_y \\ & + (\cos\theta dr - r \sin\theta d\theta) \widehat{\mathbf{e}}_z \ , \end{aligned} \tag{5.10}$$

helps provide

$$\mathbf{g}_1 = \frac{\partial \mathbf{x}}{\partial r} = \widehat{\mathbf{e}}_r \quad \text{where} \quad \widehat{\mathbf{e}}_r = \sin\theta \cos\phi \, \widehat{\mathbf{e}}_x + \sin\theta \sin\phi \, \widehat{\mathbf{e}}_y + \cos\theta \, \widehat{\mathbf{e}}_z \ , \tag{5.11a}$$

$$\mathbf{g}_2 = \frac{\partial \mathbf{x}}{\partial \theta} = r \, \widehat{\mathbf{e}}_\theta \quad \text{where} \quad \widehat{\mathbf{e}}_\theta = \cos\theta \cos\phi \, \widehat{\mathbf{e}}_x + \cos\theta \sin\phi \, \widehat{\mathbf{e}}_y - \sin\theta \, \widehat{\mathbf{e}}_z \ , \tag{5.11b}$$

$$\mathbf{g}_3 = \frac{\partial \mathbf{x}}{\partial \phi} = r \sin\theta \, \widehat{\mathbf{e}}_\phi \quad \text{where} \quad \widehat{\mathbf{e}}_\phi = - \sin\phi \, \widehat{\mathbf{e}}_x + \cos\phi \, \widehat{\mathbf{e}}_y \ , \tag{5.11c}$$

noting that $\left\{ \widehat{\mathbf{e}}_r, \widehat{\mathbf{e}}_\theta, \widehat{\mathbf{e}}_\phi \right\}$ renders an orthonormal basis.

Finally, consider a Cartesian coordinate system for which $\Theta^1 = x$, $\Theta^2 = y$ and $\Theta^3 = z$ help simply write $d\mathbf{x} = dx \, \widehat{\mathbf{e}}_x + dy \, \widehat{\mathbf{e}}_y + dz \, \widehat{\mathbf{e}}_z$. Accordingly, one will trivially have

$$\mathbf{g}_1 = \widehat{\mathbf{e}}_x \quad , \quad \mathbf{g}_2 = \widehat{\mathbf{e}}_y \quad , \quad \mathbf{g}_3 = \widehat{\mathbf{e}}_z \ . \quad \times \tag{5.12}$$

The elements of $\{\mathbf{g}_i\}$ are called *covariant basis vectors* and represented by **subscripts**. Sometimes they are referred to as *irregular basis vectors* because they need not necessarily be normalized or perpendicular to each other. They may even be non-right-handed (but absolutely $\mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3) \neq 0$). In this regard, the Cartesian basis $\{\widehat{\mathbf{e}}_i\}$ represents a *regular basis*.

Some useful relations, frequently utilized in this chapter, are introduced in what follows. ✗

Consider first a slightly different form of δ_{ij}, given in (1.35), as

$$\delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \text{for which} \quad \left[\delta^i_j \right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{5.13}$$

Analogous to δ_{ij}, this symbol has the replacement property. Some examples include

$$\left.\begin{array}{l}\delta^i_j\,\widehat{\mathbf{e}}^j = \widehat{\mathbf{e}}^i \\ \delta^j_i\,\widehat{\mathbf{e}}_j = \widehat{\mathbf{e}}_i\end{array}\right\} \;,\; \left.\begin{array}{l}\delta^i_j\,\mathbf{g}^j = \mathbf{g}^i \\ \delta^j_i\,\mathbf{g}_j = \mathbf{g}_i\end{array}\right\} \;,\; \left.\begin{array}{l}\delta^i_j\,\underline{u}^j = \underline{u}^i \\ \delta^j_i\,\underline{u}_j = \underline{u}_i\end{array}\right\} \;,\; \left.\begin{array}{l}\delta^i_j\,\underline{A}^{jk} = \underline{A}^{ik} \\ \delta^j_i\,\underline{A}_{jk} = \underline{A}_{ik}\end{array}\right\} . \tag{5.14}$$

It is known as the *mixed Kronecker delta*. Note that δ^i_j has been defined in alignment with the superscripts (contravariant) and subscripts (covariant) indices of components or basis vectors.

By definition, the standard (as well as any orthonormal) basis is **self-dual** which means that there is no specific distinction between $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{e}}_3\}$ and $\{\widehat{\mathbf{e}}^1, \widehat{\mathbf{e}}^2, \widehat{\mathbf{e}}^3\}$. This is indicated by

$$\widehat{\mathbf{e}}^i = \widehat{\mathbf{e}}_i \quad , \quad \widehat{\mathbf{e}}^i \cdot \widehat{\mathbf{e}}_j = \delta^i_j \overset{\text{by}}{\underset{\text{definition}}{=\!=\!=}} \delta^j_i = \widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}^j \;, \qquad i, j = 1, 2, 3 \,. \tag{5.15}$$

However, for the sake of consistency, it is sometimes beneficial to distinguish between the standard basis vectors with upper indices and ones with lower indices, see (5.20) and (5.26). In this regard, the identity tensor $\mathbf{I} = \delta_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$, according to (2.23), admits the following representations

$$\mathbf{I} = \delta^{ij}\,\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \underbrace{\delta^i_j\,\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}^j}_{=\,\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}^i} = \underbrace{\delta^j_i\,\widehat{\mathbf{e}}^i \otimes \widehat{\mathbf{e}}_j}_{=\,\widehat{\mathbf{e}}^i \otimes \widehat{\mathbf{e}}_i} = \delta_{ij}\,\widehat{\mathbf{e}}^i \otimes \widehat{\mathbf{e}}^j \;, \quad \longleftarrow \text{see (5.78)} \tag{5.16}$$

where

$$\delta^{ij} = \widehat{\mathbf{e}}^i \cdot \widehat{\mathbf{e}}^j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} . \tag{5.17}$$

As a result of the self-duality of $\{\widehat{\mathbf{e}}_i\}$, the permutation symbol in (1.65) renders

$$\varepsilon_{ijk} = \widehat{\mathbf{e}}_i \cdot \left(\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k\right) = \widehat{\mathbf{e}}^i \cdot \left(\widehat{\mathbf{e}}^j \times \widehat{\mathbf{e}}^k\right) = \varepsilon^{ijk} \,. \;\; ✗ \tag{5.18}$$

Regarding the free and dummy indices in this chapter, the following rules must be obeyed:

- ▶ A free index can be either superscript (contravariant) or subscript (covariant) but its specific form should be preserved in all terms of an equation. For instance, an expression of the form $\underline{u}_i = \underline{v}_i + \underline{w}^i$ is not true. The correct form is either $\underline{u}_i = \underline{v}_i + \underline{w}_i$ or $\underline{u}^i = \underline{v}^i + \underline{w}^i$.
- ▶ If there is a dummy index in one term of an expression, one index should appear at the upper level while the other should appear at the lower level. For instance, the last term in $\underline{u}_i\,\underline{A}^{ij}\,\underline{B}_{jk} - \underline{C}_{ki}\,\underline{v}^i + \underline{w}^i\,\underline{D}_{ij}\,\underline{E}_{jk} = 0$ is not consistent. The appropriate form, for instance, is $\underline{w}^i\,\underline{D}_{ij}\,\underline{E}^{\,j}_{\;.k}$.

5.1.2 Contravariant Basis Vectors

Guided by Fig. 5.1, each general basis vector itself can be seen as a Cartesian vector. Therefore, one can write

$$\mathbf{g}_i = (\mathbf{g}_i)^j \,\widehat{\mathbf{e}}_j \ , \qquad i = 1, 2, 3 \ . \quad \longleftarrow \text{see (1.34)} \tag{5.19}$$

This motivates to construct a linear transformation $\mathbf{F}$, relating the local basis $\{\mathbf{g}_i\}$ to the global basis $\{\widehat{\mathbf{e}}_i\}$, according to

$$\boxed{\mathbf{g}_i = \mathbf{F}\,\widehat{\mathbf{e}}_i = F_i^{\,j}\,\widehat{\mathbf{e}}_j \ , \qquad i = 1, 2, 3 \ ,} \quad \longleftarrow \text{see (2.21)} \tag{5.20}$$

where $F^i_{\ j} = \left(\mathbf{g}_j\right)^i$, $i, j = 1, 2, 3$, are the Cartesian components of $\mathbf{F}$. This linear mapping may be represented by

$$\mathbf{F} = \mathbf{g}_i \otimes \widehat{\mathbf{e}}^i \qquad \text{or} \qquad [\,\mathbf{F}\,] = \begin{bmatrix} F^1_{\ 1} & F^1_{\ 2} & F^1_{\ 3} \\ F^2_{\ 1} & F^2_{\ 2} & F^2_{\ 3} \\ F^3_{\ 1} & F^3_{\ 2} & F^3_{\ 3} \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 \\ \vdots & \vdots & \vdots \end{bmatrix} . \tag{5.21}$$

Note that $\mathbf{F}$ is always **invertible** because its columns are linearly independent. Consequently, the inverse of $\mathbf{F}$ according to

$$\mathbf{F}^{-1} = \widehat{\mathbf{e}}_i \otimes \mathbf{g}^i \qquad \text{or} \qquad \left[\,\mathbf{F}^{-1}\,\right] = \begin{bmatrix} \cdots & \mathbf{g}^1 & \cdots \\ \cdots & \mathbf{g}^2 & \cdots \\ \cdots & \mathbf{g}^3 & \cdots \end{bmatrix} , \tag{5.22}$$

delivers another set of irregular basis vectors.[1] This companion triad of vectors

$$\left\{\mathbf{g}^i\right\} := \left\{\mathbf{g}^1, \mathbf{g}^2, \mathbf{g}^3\right\} \ , \tag{5.23}$$

---

[1] The inverse of a tensor will naturally raise or lower the indices. This holds true for the components as well as the bases of a tensor. Given two bases $\{\mathbf{g}_i\}$ and $\left\{\hat{\mathbf{g}}_i\right\}$, some examples include

$$\begin{aligned} \mathbf{A} = \mathbf{g}_i \otimes \hat{\mathbf{g}}_i \quad &\Longleftrightarrow \quad \mathbf{A}^{-1} = \hat{\mathbf{g}}^i \otimes \mathbf{g}^i \ , \\ \mathbf{B} = \mathbf{g}_i \otimes \hat{\mathbf{g}}^i \quad &\Longleftrightarrow \quad \mathbf{B}^{-1} = \hat{\mathbf{g}}_i \otimes \mathbf{g}^i \ , \\ \mathbf{C} = \mathbf{g}^i \otimes \hat{\mathbf{g}}_i \quad &\Longleftrightarrow \quad \mathbf{C}^{-1} = \hat{\mathbf{g}}^i \otimes \mathbf{g}_i \\ \mathbf{D} = \mathbf{g}^i \otimes \hat{\mathbf{g}}^i \quad &\Longleftrightarrow \quad \mathbf{D}^{-1} = \hat{\mathbf{g}}_i \otimes \mathbf{g}_i \ . \end{aligned}$$

These tensors consistently satisfy the reciprocal expression (2.105). Note that the identity tensor here is either $\mathbf{g}_i \otimes \mathbf{g}^i$ or $\mathbf{g}^i \otimes \mathbf{g}_i$, see (5.78).

is known as the *dual basis* of $\{\mathbf{g}_i\}$. They are called *contravariant basis vectors* and shown by **superscripts**. Each contravariant basis vector can be resolved along the standard basis $\{\widehat{\mathbf{e}}^i\}$ as

$$\mathbf{g}^i = \left(\mathbf{g}^i\right)_j \widehat{\mathbf{e}}^j \ , \qquad i = 1, 2, 3 \ . \quad \longleftarrow \text{ see (5.19)} \tag{5.24}$$

Knowing that any dyad has bilinearity property according to (2.15a)$_4$ and (2.15b)$_5$, substituting $\mathbf{g}^i = \left(\mathbf{g}^i\right)_j \widehat{\mathbf{e}}^j$ into $\mathbf{F}^{-1} = \widehat{\mathbf{e}}_i \otimes \mathbf{g}^i$ provides

$$\mathbf{F}^{-1} = \left(\mathbf{F}^{-1}\right)^i_{\ j} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}^j \quad \text{where} \quad \left(\mathbf{F}^{-1}\right)^i_{\ j} = \left(\mathbf{g}^i\right)_j \ . \tag{5.25}$$

Basically, the linear transformation $\mathbf{F}^{-\mathrm{T}}$ translates $\{\widehat{\mathbf{e}}^i\}$ to $\{\mathbf{g}^i\}$:

$$\boxed{\mathbf{g}^i = \mathbf{F}^{-\mathrm{T}} \widehat{\mathbf{e}}^i \ , \qquad i = 1, 2, 3 \ ,} \tag{5.26}$$

since

$$\mathbf{F}^{-1} \xlongequal[(5.22)]{\text{from}} \widehat{\mathbf{e}}_i \otimes \mathbf{g}^i \Rightarrow \mathbf{F}^{-\mathrm{T}} \xlongequal[(2.52) \text{ and } (2.54)]{\text{from}} \mathbf{g}^i \otimes \widehat{\mathbf{e}}_i \ \Rightarrow \ \mathbf{F}^{-\mathrm{T}} \widehat{\mathbf{e}}^j \xlongequal[(2.13),\ (5.14) \text{ and } (5.15)]{\text{from (2.8a)}} \mathbf{g}^j \ .$$

The identities $\mathbf{I} = \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}^i$ and $\mathbf{I} = \mathbf{F}^{-1}\mathbf{F}$ now help establish an important relation between the covariant and contravariant bases:

$$\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}^i = \mathbf{F}^{-1}\mathbf{F} \xlongequal[(5.21) \text{ and } (5.22)]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \mathbf{g}^i\right)\left(\mathbf{g}_j \otimes \widehat{\mathbf{e}}^j\right) \xlongequal[(2.29) \text{ and } (2.30)]{\text{from}} \left(\mathbf{g}^i \cdot \mathbf{g}_j\right)\left(\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}^j\right) \ .$$

It is not then difficult to conclude that

$$\boxed{\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j \quad , \quad \mathbf{g}_i \cdot \mathbf{g}^j = \delta^j_i \ .} \quad \longleftarrow \text{ see (5.15) and (9.33)} \tag{5.27}$$

Geometrically, the condition (5.27) states that each vector of a covariant (contravariant) basis is orthogonal to - a plane defined by - the two vectors of the corresponding contravariant (covariant) basis with different indices, see Fig. 5.2.

The result (5.27)$_1$ may help find the contravariant counterpart of (5.3)$_1$. Since the three curvilinear coordinates Θ^1, Θ^2 and Θ^3 are independent, they trivially satisfy $\partial\Theta^i/\partial\Theta^j = \delta^i_j$. By means of the chain rule of differentiation, one then arrives at

$$\frac{\partial\Theta^i}{\partial\mathbf{x}} \cdot \frac{\partial\mathbf{x}}{\partial\Theta^j} = \delta^i_j \quad \text{or} \quad \frac{\partial\Theta^i}{\partial\mathbf{x}} \cdot \mathbf{g}_j = \delta^i_j \ ,$$

or

$$\boxed{\mathbf{g}^i = \frac{\partial\Theta^i}{\partial\mathbf{x}} \ , \qquad i = 1, 2, 3 \ .} \tag{5.28}$$

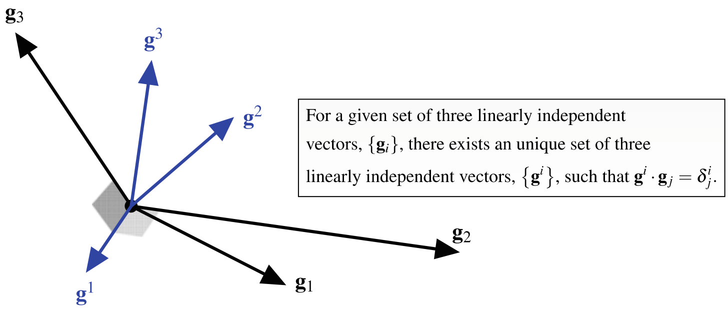


Fig. 5.2 Covariant basis $\{\mathbf{g}_i\}$ with its companion contravariant basis $\{\mathbf{g}^i\}$

The linear transformation $\mathbf{F}$ and its inverse transpose in matrix form can finally be written as

$$[\,\mathbf{F}\,] = \begin{bmatrix} \vdots & \vdots & \vdots \\ \dfrac{\partial \mathbf{x}}{\partial \Theta^1} & \dfrac{\partial \mathbf{x}}{\partial \Theta^2} & \dfrac{\partial \mathbf{x}}{\partial \Theta^3} \\ \vdots & \vdots & \vdots \end{bmatrix} \quad , \quad \mathbf{F}^{-\mathrm{T}} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \dfrac{\partial \Theta^1}{\partial \mathbf{x}} & \dfrac{\partial \Theta^2}{\partial \mathbf{x}} & \dfrac{\partial \Theta^3}{\partial \mathbf{x}} \\ \vdots & \vdots & \vdots \end{bmatrix} . \tag{5.29}$$

The determinant of the linear mapping $\mathbf{F} = \mathbf{g}_i \otimes \widehat{\mathbf{e}}^i$, in (5.21)$_1$, is denoted by

$$\boxed{J := \det[\,\mathbf{F}\,] = \det\left[\,\mathbf{g}_1 \;\; \mathbf{g}_2 \;\; \mathbf{g}_3\,\right] = \mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3)\ ,} \tag{5.30}$$

which, in light of (4.15c), can be computed via

$$\underbrace{J\,\widehat{\mathbf{e}}_i \cdot \left(\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k\right)}_{= J\varepsilon_{ijk}\ ,\ \text{according to (5.18)}} = \underbrace{\mathbf{F}\widehat{\mathbf{e}}_i \cdot \left(\mathbf{F}\widehat{\mathbf{e}}_j \times \mathbf{F}\widehat{\mathbf{e}}_j\right)}_{= \mathbf{g}_i \cdot (\mathbf{g}_j \times \mathbf{g}_k)\ ,\ \text{according to (5.20)}} .$$

Thus,

$$\boxed{J\varepsilon_{ijk} = \mathbf{g}_i \cdot \left(\mathbf{g}_j \times \mathbf{g}_k\right)\ .} \tag{5.31}$$

This scalar variable appears in many relations in this text and referred to as *Jacobian*. The general basis $\{\mathbf{g}_i\}$ is **right-handed** if $J > 0$ and **left-handed** when $J < 0$. Note that the absolute value of Jacobian depicts the volume of a parallelepiped constructed by the three vectors $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{g}_3$.

It is then a simple exercise to see that

$$\boxed{J^{-1}\varepsilon^{ijk} = \mathbf{g}^i \cdot \left(\mathbf{g}^j \times \mathbf{g}^k\right)\ .} \tag{5.32}$$

In what follows, the goal is to represent the contravariant basis vectors in terms of the covariant ones and the Jacobian. Knowing that $\mathbf{g}^i$ is perpendicular to both $\mathbf{g}_j$ and $\mathbf{g}_k$ (when $i \neq j, k$), one may write $\alpha \varepsilon_{ijk} \mathbf{g}^i = \mathbf{g}_j \times \mathbf{g}_k$ where α is an unknown parameter. In light of (5.27)$_1$, this scalar can be determined by requiring $\mathbf{g}^i \cdot \mathbf{g}_l = \delta^i_l$. As a result, $\alpha = J$. Then,

$$\mathbf{g}_j \times \mathbf{g}_k = J \mathbf{g}^i \varepsilon_{ijk} \quad \text{or, using (1.54),} \quad \mathbf{g}_j \times \mathbf{g}_k = J \varepsilon_{jki} \mathbf{g}^i \ . \tag{5.33}$$

This result with the aid of (1.58b)$_3$, i.e. $\varepsilon_{ijk} \varepsilon^{ljk} = 2\delta^l_i$ in this context, and the identity $\mathbf{g}^i \delta^l_i = \mathbf{g}^l$ can be rewritten as

$$\boxed{\mathbf{g}^i = \frac{\varepsilon^{ijk}}{2J} \mathbf{g}_j \times \mathbf{g}_k \ .} \quad \longleftarrow \text{see (5.49)} \tag{5.34}$$

In a similar manner,

$$\mathbf{g}^j \times \mathbf{g}^k = \frac{1}{J} \mathbf{g}_i \varepsilon^{ijk} \quad \text{or, using (1.54),} \quad \mathbf{g}^j \times \mathbf{g}^k = \frac{1}{J} \varepsilon^{jki} \mathbf{g}_i \ . \tag{5.35}$$

The contravariant basis vectors and the Jacobian now help compute the covariant ones:

$$\boxed{\mathbf{g}_i = \frac{J \varepsilon_{ijk}}{2} \mathbf{g}^j \times \mathbf{g}^k \ .} \quad \longleftarrow \text{see (5.43)} \tag{5.36}$$

Hint: Note that any general basis $\{\mathbf{g}_i\}$ in this text belongs to $\mathscr{E}^{o3}_{r}$ and, therefore, the Jacobian is always **positive**.

It is also important to mention that the tangent vectors $\mathbf{g}_i, \ i = 1, 2, 3,$ together with their companion dual vectors $\mathbf{g}^i, \ i = 1, 2, 3,$ generally change from point to point in space and, therefore, they should be regarded as functions of the curvilinear coordinates $\left(\Theta^1, \Theta^2, \Theta^3\right)$.

5.2 Metric Coefficients

The subscript indices of the covariant basis vectors $\mathbf{g}_i, \ i = 1, 2, 3,$ can be raised. In a consistent manner, the superscript indices of the contravariant basis vectors $\mathbf{g}^i, \ i = 1, 2, 3,$ can be lowered. Raising and lowering the indices of components or basis vectors in curvilinear coordinates are enabled by the so-called *metric coefficients*. They carry geometrical characteristics of basis vectors and help change the type of a basis. The operations of raising and lowering indices are basically known as *index juggling*. There are two metric coefficients; namely, the *covariant metric* g_{ij} and the *contravariant metric* g^{ij}. These two quantities, which tensorially transform under a change of coordinates, are the Cartesian components of some special tensors built upon the linear transformation $\mathbf{F}$ in (5.20)–(5.21). This is demonstrated in the following.

5.2.1 Covariant Metric Coefficients

First, the components of the symmetric tensor

$$\begin{aligned}\mathbf{C} &= \mathbf{F}^{\mathrm{T}}\mathbf{F}\\ &\xlongequal[\text{(2.52), (2.54) and (5.21)}]{\text{from}} \left(\widehat{\mathbf{e}}^{i} \otimes \mathbf{g}_i\right)\left(\mathbf{g}_j \otimes \widehat{\mathbf{e}}^{j}\right)\\ &\xlongequal[\text{(2.29) and (2.30)}]{\text{from}} \left(\mathbf{g}_i \cdot \mathbf{g}_j\right)\widehat{\mathbf{e}}^{i} \otimes \widehat{\mathbf{e}}^{j} ,\end{aligned} \tag{5.37}$$

with respect to $\left\{\widehat{\mathbf{e}}^{i}\right\}$ deliver the **covariant metric** g_{ij}. It is basically defined as

$$\boxed{g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j} \quad \text{with} \quad \left[\, g_{ij} \,\right] = [\,\mathbf{C}\,] = \begin{bmatrix} \mathbf{g}_1 \cdot \mathbf{g}_1 & \mathbf{g}_1 \cdot \mathbf{g}_2 & \mathbf{g}_1 \cdot \mathbf{g}_3 \\ \mathbf{g}_2 \cdot \mathbf{g}_1 & \mathbf{g}_2 \cdot \mathbf{g}_2 & \mathbf{g}_2 \cdot \mathbf{g}_3 \\ \mathbf{g}_3 \cdot \mathbf{g}_1 & \mathbf{g}_3 \cdot \mathbf{g}_2 & \mathbf{g}_3 \cdot \mathbf{g}_3 \end{bmatrix} . \tag{5.38}$$

The commutative property of the dot product, according to (1.9a), then implies that

$$\boxed{g_{ij} = g_{ji} \ .} \tag{5.39}$$

This is in alignment with the symmetry of $\mathbf{C}$. As can be seen, the diagonal entries of $\left[\, g_{ij} \,\right]$ help measure the lengths of $\mathbf{g}_i$, $i = 1, 2, 3$, while its off-diagonal elements help determine the angles between these covariant basis vectors, that is,

$$|\mathbf{g}_i| = \sqrt{g_{ii}} \quad , \quad \cos\theta\left(\mathbf{g}_i, \mathbf{g}_j\right) = \frac{g_{ij}}{\sqrt{g_{ii}\, g_{jj}}} \ , \ i, j = 1, 2, 3 \ ; \ \text{no summation} \ . \tag{5.40}$$

Making use of the relation $\det\left[\, g_{ij} \,\right] = \det[\,\mathbf{C}\,] = \det\left[\,\mathbf{F}^{\mathrm{T}}\,\right]\det[\,\mathbf{F}\,]$ along with the Jacobian $J = \det[\,\mathbf{F}\,]$ and the identity $\det\left[\,\mathbf{F}^{\mathrm{T}}\,\right] = \det[\,\mathbf{F}\,]$, one will have

$$\boxed{\det\left[\, g_{ij} \,\right] = J^2 \ .} \tag{5.41}$$

This helps obtain

$$\boxed{\underbrace{J^2 \varepsilon_{ijk} = \varepsilon^{lmn} g_{li}\, g_{mj}\, g_{nk}}_{\text{or } J^2 \varepsilon_{ijk}\, g^{jm} = \varepsilon^{lmn} g_{li}\, g_{nk}} \ .} \quad \longleftarrow \ \text{see (1.85)} \tag{5.42}$$

Having in mind $\mathbf{F} = \mathbf{g}_i \otimes \widehat{\mathbf{e}}^{i}$, $\mathbf{F}^{\mathrm{T}} = \widehat{\mathbf{e}}^{i} \otimes \mathbf{g}_i$, $\mathbf{F}^{-1} = \widehat{\mathbf{e}}_j \otimes \mathbf{g}^j$ and $\mathbf{C} = g_{ik}\widehat{\mathbf{e}}^{i} \otimes \widehat{\mathbf{e}}^{k}$, the identity $\mathbf{F}^{\mathrm{T}} = \mathbf{C}\mathbf{F}^{-1}$ delivers

$$\widehat{\mathbf{e}}^{i} \otimes \mathbf{g}_i = \left(g_{ik}\widehat{\mathbf{e}}^{i} \otimes \widehat{\mathbf{e}}^{k}\right)\left(\widehat{\mathbf{e}}_j \otimes \mathbf{g}^j\right) = g_{ik}\delta^k_j \widehat{\mathbf{e}}^{i} \otimes \mathbf{g}^j = g_{ij}\widehat{\mathbf{e}}^{i} \otimes \mathbf{g}^j \ .$$

And this implies that

$$\boxed{\mathbf{g}_i = g_{ij}\,\mathbf{g}^j \ , \qquad i = 1, 2, 3 \ .} \tag{5.43}$$

The importance of this result in tensor algebra and calculus is that one can favorably use the covariant metric to lower the superscript index of a contravariant basis vector (provided that the objects g_{ij} and $\mathbf{g}^j$ are known). The identity $\mathbf{F}^{\mathrm{T}} = \mathbf{C}\mathbf{F}^{-1}$ in matrix form $\left[\, \mathbf{F}^{\mathrm{T}} \,\right] = \left[\, g_{ij} \,\right] \left[\, \mathbf{F}^{-1} \,\right]$ renders

$$\begin{bmatrix} \cdots & \mathbf{g}_1 & \cdots \\ \cdots & \mathbf{g}_2 & \cdots \\ \cdots & \mathbf{g}_3 & \cdots \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{22} & g_{23} \\ g_{13} & g_{23} & g_{33} \end{bmatrix} \begin{bmatrix} \cdots & \mathbf{g}^1 & \cdots \\ \cdots & \mathbf{g}^2 & \cdots \\ \cdots & \mathbf{g}^3 & \cdots \end{bmatrix} , \tag{5.44}$$

which is basically the matrix representation of (5.43). Guided by (2.152), $\left[\, g_{ij} \,\right]$ in (5.44) can be thought of as a transformation matrix that translates $\left\{ \mathbf{g}^i \right\}$ into $\{\mathbf{g}_i\}$.

Hint: If $g_{ij} = \delta_{ij}$, then $\mathbf{g}_i \cdot \mathbf{g}_j = \delta_{ij}$ implies that each basis vector has unit length and should be perpendicular to any other basis vector whose index is different. This means that $\{\mathbf{g}_i\}$ now forms an orthonormal basis. For such a basis, (5.43) asserts that the covariant and contravariant basis vectors are identical, i.e. $\mathbf{g}_i = \mathbf{g}^i$, $i = 1, 2, 3$. Note that all orthonormal bases are self-dual.

5.2.2 Contravariant Metric Coefficients

Next, the inverse of $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = g_{ij}\,\widehat{\mathbf{e}}^i \otimes \widehat{\mathbf{e}}^j$ according to

$$\begin{aligned} \mathbf{C}^{-1} &= \mathbf{F}^{-1}\mathbf{F}^{-\mathrm{T}} \\ &\xlongequal[\text{(2.52), (2.54) and (5.22)}]{\text{from}} \left(\widehat{\mathbf{e}}_i \otimes \mathbf{g}^i \right) \left(\mathbf{g}^j \otimes \widehat{\mathbf{e}}_j \right) \\ &\xlongequal[\text{(2.29) and (2.30)}]{\text{from}} \left(\mathbf{g}^i \cdot \mathbf{g}^j \right) \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ , \end{aligned} \tag{5.45}$$

helps introduce

$$\boxed{g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j} \quad \text{with} \quad \left[\, g^{ij} \,\right] = \left[\, \mathbf{C}^{-1} \,\right] = \begin{bmatrix} \mathbf{g}^1 \cdot \mathbf{g}^1 & \mathbf{g}^1 \cdot \mathbf{g}^2 & \mathbf{g}^1 \cdot \mathbf{g}^3 \\ \mathbf{g}^2 \cdot \mathbf{g}^1 & \mathbf{g}^2 \cdot \mathbf{g}^2 & \mathbf{g}^2 \cdot \mathbf{g}^3 \\ \mathbf{g}^3 \cdot \mathbf{g}^1 & \mathbf{g}^3 \cdot \mathbf{g}^2 & \mathbf{g}^3 \cdot \mathbf{g}^3 \end{bmatrix} . \tag{5.46}$$

The symmetry of the **contravariant metric**

$$\boxed{g^{ij} = g^{ji} \ ,} \tag{5.47}$$

emanates from the commutative property of the dot product or the symmetry of $\mathbf{C}^{-1}$. Now, the lengths of contravariant basis vectors and the angles between them can be determined by means of the contravariant metric as implied by its name:

$$\left|\mathbf{g}^i\right| = \sqrt{g^{ii}} \quad , \quad \cos\theta\left(\mathbf{g}^i, \mathbf{g}^j\right) = \frac{g^{ij}}{\sqrt{g^{ii}\,g^{jj}}} \; , \; i, j = 1, 2, 3 \; ; \; \text{no summation} \; . \tag{5.48}$$

Substituting $\mathbf{F}^{-1} = \widehat{\mathbf{e}}_i \otimes \mathbf{g}^i$, $\mathbf{C}^{-1} = g^{ik}\,\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_k$ and $\mathbf{F}^{\mathrm{T}} = \widehat{\mathbf{e}}^j \otimes \mathbf{g}_j$ into $\mathbf{F}^{-1} = \mathbf{C}^{-1}\mathbf{F}^{\mathrm{T}}$ yields

$$\widehat{\mathbf{e}}_i \otimes \mathbf{g}^i = \left(g^{ik}\,\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_k\right)\left(\widehat{\mathbf{e}}^j \otimes \mathbf{g}_j\right) = g^{ik}\delta_k^j\,\widehat{\mathbf{e}}_i \otimes \mathbf{g}_j = g^{ij}\,\widehat{\mathbf{e}}_i \otimes \mathbf{g}_j \; .$$

It is then easy to see that

$$\boxed{\mathbf{g}^i = g^{ij}\,\mathbf{g}_j \; , \qquad i = 1, 2, 3 \; .} \tag{5.49}$$

As expected, the contravariant metric helps raise the subscript index of a covariant basis vector (provided that g^{ij} and $\mathbf{g}_j$ are given). Consistent with (5.44), the matrix form of (5.49) may be viewed as a transformation from $\{\mathbf{g}_i\}$ to $\{\mathbf{g}^i\}$ by using $\left[\, g^{ij} \,\right]$, that is,

$$\begin{bmatrix} \cdots & \mathbf{g}^1 & \cdots \\ \cdots & \mathbf{g}^2 & \cdots \\ \cdots & \mathbf{g}^3 & \cdots \end{bmatrix} = \begin{bmatrix} g^{11} & g^{12} & g^{13} \\ g^{12} & g^{22} & g^{23} \\ g^{13} & g^{23} & g^{33} \end{bmatrix} \begin{bmatrix} \cdots & \mathbf{g}_1 & \cdots \\ \cdots & \mathbf{g}_2 & \cdots \\ \cdots & \mathbf{g}_3 & \cdots \end{bmatrix} . \tag{5.50}$$

The relations $\mathbf{C}\mathbf{C}^{-1} = \mathbf{I}$ and $\mathbf{C}^{-1}\mathbf{C} = \mathbf{I}$ then help establish the relationship between the covariant and contravariant metrics:

$$\boxed{\underbrace{g_{ik}\,g^{kj} = \delta_i^j}_{\text{or } g^{ik} g_{kj} = \delta^i_j} \qquad \text{or} \qquad \underbrace{\left[\, g_{ij} \,\right] = \left[\, g^{ij} \,\right]^{-1}}_{\text{or } \left[g^{ij}\right] = \left[g_{ij}\right]^{-1}} \; .} \tag{5.51}$$

From (2.109a), (5.41) and $(5.51)_4$, one will have

$$\boxed{\det\left[\, g^{ij} \,\right] = J^{-2} \; .} \tag{5.52}$$

Using (5.42) and (5.52), one can arrive at

$$\boxed{\underbrace{J^{-2}\varepsilon^{ijk} = \varepsilon_{lmn}\,g^{li}\,g^{mj}\,g^{nk}}_{\text{or } J^{-2}\varepsilon^{ijk} g_{jm} = \varepsilon_{lmn} g^{li} g^{nk}} \; .} \tag{5.53}$$

Now, it should be clear that raising any covariant index of g_{ij} or lowering any contravariant index of g^{ij} will produce the mixed Kronecker delta. In this regard, δ^i_j may be referred to as the *mixed metric coefficients*. As discussed, these quantities may be utilized in an expression to take advantage of their replacement property knowing that they do not change the type of components or basis vectors.

Hint: Given the covariant basis $\{\mathbf{g}_i\}$, there are three procedures for calculating the contravariant basis $\{\mathbf{g}^i\}$. They are summarized in the following.

- The first procedure simply includes two steps. First, by constructing the matrix $[\,\mathbf{F}\,] = [\,\mathbf{g}_1 \;\; \mathbf{g}_2 \;\; \mathbf{g}_3\,]$ and then by calculating the transpose of its inverse according to $[\,\mathbf{F}\,]^{-\mathrm{T}} = [\,\mathbf{g}^1 \;\; \mathbf{g}^2 \;\; \mathbf{g}^3\,]$, one can compute $\{\mathbf{g}^i\}$. This procedure is very favorable from the computational point of view and does not require any information regarding the metrics.
- The second procedure makes use of the cross products of covariant basis vectors and their Jacobian. Here, one will have $\mathbf{g}^1 = J^{-1}\mathbf{g}_2 \times \mathbf{g}_3$, $\mathbf{g}^2 = J^{-1}\mathbf{g}_3 \times \mathbf{g}_1$ and $\mathbf{g}^3 = J^{-1}\mathbf{g}_1 \times \mathbf{g}_2$ where $J = \mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3)$.
- The third procedure eventually represents the desired vectors in a more rigorous way taking advantage of the metric coefficients. Here, one needs to first compute $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$ and then obtain its inverse via $[\,g^{ij}\,] = [\,g_{ij}\,]^{-1}$. This procedure will be completed by calculating $\mathbf{g}^i = g^{ij}\mathbf{g}_j$.

5.3 Tensor Property of Basis Vectors and Metric Coefficients

Consider an old basis $\{\mathbf{g}_i\}$ and a new basis $\{\bar{\mathbf{g}}_i\}$ with the corresponding metric coefficients g_{ij} and $\bar{g}_{ij}$. Indeed, they stem from a change of coordinates from $\left(\Theta^1, \Theta^2, \Theta^3\right)$ to $\left(\bar{\Theta}^1, \bar{\Theta}^2, \bar{\Theta}^3\right)$ which are connected by the following mutually inverse relationships

$$\bar{\Theta}^i = \bar{\Theta}^i\left(\Theta^1, \Theta^2, \Theta^3\right) \quad , \quad \Theta^i = \Theta^i\left(\bar{\Theta}^1, \bar{\Theta}^2, \bar{\Theta}^3\right) . \tag{5.54}$$

This helps write

$$\mathbf{x} = \bar{\hat{\mathbf{x}}}\left(\bar{\Theta}^1, \bar{\Theta}^2, \bar{\Theta}^3\right) = \hat{\mathbf{x}}\left(\Theta^1, \Theta^2, \Theta^3\right) \quad \text{with} \quad \left.\begin{aligned} \bar{\mathbf{g}}_i &= \frac{\partial \mathbf{x}}{\partial \bar{\Theta}^i} \\ \mathbf{g}_i &= \frac{\partial \mathbf{x}}{\partial \Theta^i} \end{aligned}\right\} \, , \quad \left.\begin{aligned} \bar{\mathbf{g}}^i &= \frac{\partial \bar{\Theta}^i}{\partial \mathbf{x}} \\ \mathbf{g}^i &= \frac{\partial \Theta^i}{\partial \mathbf{x}} \end{aligned}\right\} . \tag{5.55}$$

Moreover, by the chain rule of differentiation, the following identities are immediately implied:

$$\delta^i_j = \frac{\partial \bar{\Theta}^i}{\partial \bar{\Theta}^j} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k}\frac{\partial \Theta^k}{\partial \bar{\Theta}^j} \quad , \quad \delta^i_j = \frac{\partial \Theta^i}{\partial \Theta^j} = \frac{\partial \Theta^i}{\partial \bar{\Theta}^k}\frac{\partial \bar{\Theta}^k}{\partial \Theta^j} . \tag{5.56}$$

As a result,

$$\frac{\partial^2 \bar{\Theta}^i}{\partial \Theta^j \partial \Theta^k} = -\frac{\partial \bar{\Theta}^i}{\partial \Theta^l}\frac{\partial \bar{\Theta}^m}{\partial \Theta^j}\frac{\partial \bar{\Theta}^n}{\partial \Theta^k}\frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^m \partial \bar{\Theta}^n} , \tag{5.57a}$$

$$\frac{\partial^2 \Theta^i}{\partial \bar{\Theta}^j \partial \bar{\Theta}^k} = -\frac{\partial \Theta^i}{\partial \bar{\Theta}^l} \frac{\partial \Theta^m}{\partial \bar{\Theta}^j} \frac{\partial \Theta^n}{\partial \bar{\Theta}^k} \frac{\partial^2 \bar{\Theta}^l}{\partial \Theta^m \partial \Theta^n} . \tag{5.57b}$$

Using $(5.55)_{3-6}$ along with the chain rule of differentiation, one can establish the identities

$$\bar{\mathbf{g}}^i \cdot \mathbf{g}_j = \frac{\partial \bar{\Theta}^i}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \Theta^j} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j} , \tag{5.58a}$$

$$\mathbf{g}^i \cdot \bar{\mathbf{g}}_j = \frac{\partial \Theta^i}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \bar{\Theta}^j} = \frac{\partial \Theta^i}{\partial \bar{\Theta}^j} . \tag{5.58b}$$

From (1.9a), (2.5), (2.8a), (2.13), $(5.58a)_2$, $(5.58b)_2$ and invoking $(5.78)_{2-3}$, one can arrive at the **transformation laws**

$$\bar{\mathbf{g}}^i = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j} \mathbf{g}^j \quad , \quad \mathbf{g}^i = \frac{\partial \Theta^i}{\partial \bar{\Theta}^j} \bar{\mathbf{g}}^j \quad , \tag{5.59}$$

note that $\bar{\mathbf{g}}^i = (\mathbf{I})\, \bar{\mathbf{g}}^i = \left(\mathbf{g}^j \otimes \mathbf{g}_j\right) \bar{\mathbf{g}}^i = \left(\bar{\mathbf{g}}^i \cdot \mathbf{g}_j\right) \mathbf{g}^j$ and $\mathbf{g}^i = (\mathbf{I})\, \mathbf{g}^i = \left(\bar{\mathbf{g}}^j \otimes \bar{\mathbf{g}}_j\right) \mathbf{g}^i = \left(\mathbf{g}^i \cdot \bar{\mathbf{g}}_j\right) \bar{\mathbf{g}}^j$

and

$$\bar{\mathbf{g}}_i = \frac{\partial \Theta^j}{\partial \bar{\Theta}^i} \mathbf{g}_j \quad , \quad \mathbf{g}_i = \frac{\partial \bar{\Theta}^j}{\partial \Theta^i} \bar{\mathbf{g}}_j \quad . \tag{5.60}$$

note that $\bar{\mathbf{g}}_i = (\mathbf{I})\, \bar{\mathbf{g}}_i = \left(\mathbf{g}_j \otimes \mathbf{g}^j\right) \bar{\mathbf{g}}_i = \left(\bar{\mathbf{g}}_i \cdot \mathbf{g}^j\right) \mathbf{g}_j$ and $\mathbf{g}_i = (\mathbf{I})\, \mathbf{g}_i = \left(\bar{\mathbf{g}}_j \otimes \bar{\mathbf{g}}^j\right) \mathbf{g}_i = \left(\mathbf{g}_i \cdot \bar{\mathbf{g}}^j\right) \bar{\mathbf{g}}_j$

As a result, the covariant metric coefficients $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$ will **tensorially** transform according to

$$\bar{g}_{ij} = \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} g_{kl} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} . \tag{5.61}$$

And this means that the object g_{ij} is deserved to be called the *covariant metric tensor* (note that the definition of tensor in the literature is not unique). One should now notice that the mixed Kronecker delta δ^i_j is also a tensor because its values in the new coordinate system are governed by the tensor transformation law. This is indicated by

$$\bar{\delta}^i_j = \frac{\partial \bar{\Theta}^i}{\partial \bar{\Theta}^j} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^k}{\partial \bar{\Theta}^j} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \delta^k_l \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} . \tag{5.62}$$

In a similar manner, the contravariant metric coefficients $g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j$ will preserve the tensor property:

$$\bar{g}^{ij} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} g^{kl} \frac{\partial \bar{\Theta}^j}{\partial \Theta^l} . \tag{5.63}$$

In this regard, the variant g^{ij} may also be referred to as the *contravariant metric tensor*.

5.4 Contravariant and Covariant Components of First-Order Tensors

Two general forms of basis vectors were introduced. As a result, an arbitrary vector **u** in curvilinear coordinates has two forms:

$$\mathbf{u} = \underbrace{\underline{u}^i \mathbf{g}_i}_{= \underline{u}^1 \mathbf{g}_1 + \underline{u}^2 \mathbf{g}_2 + \underline{u}^3 \mathbf{g}_3} \ , \tag{5.64a}$$

$$\mathbf{u} = \underbrace{\underline{u}_i \mathbf{g}^i}_{= \underline{u}_1 \mathbf{g}^1 + \underline{u}_2 \mathbf{g}^2 + \underline{u}_3 \mathbf{g}^3} \ , \tag{5.64b}$$

where $\left(\underline{u}^1, \underline{u}^2, \underline{u}^3\right)$ denotes the **contravariant components** of **u** and $\left(\underline{u}_1, \underline{u}_2, \underline{u}_3\right)$ presents its **covariant components**.

The contravariant (covariant) components of a vector basically represent its projection onto the covariant (contravariant) basis. The first form, written in (5.64a), is the natural way of representing a vector. The reason for calling $\underline{u}^i$, $i = 1, 2, 3$, the contravariant components is that they increase when the lengths of covariant basis vectors decrease and vice versa. Consistent with this, $\underline{u}_i$, $i = 1, 2, 3$, are called the covariant components because they increase (decrease) when the lengths of covariant basis vectors increase (decrease). In differential geometry, it is frequently seen that $\underline{u}^i$ is called *vector* and $\underline{u}_i$ is referred to as *covector*. For the sake of clarification, a vector with contravariant components is called a *contravariant vector* in this text and a vector with covariant components is referred to as a *covariant vector*.

It is important to note that in representing a vector, its contravariant (covariant) components should always be written with respect to the covariant (contravariant) basis vectors. Here, the summation convention over a dummy index is only applied when one index is a superscript and the other is a subscript, as shown in (5.64a) and (5.64b).

Having in mind the linearity property (1.10), the dot products of **u** with the covariant and contravariant basis vectors, using $(5.27)_{1-2}$, $(5.38)_1$ and $(5.46)_1$, yield the following useful relations

$$\begin{aligned} \mathbf{u} \cdot \mathbf{g}_i &= \underline{u}^j \mathbf{g}_j \cdot \mathbf{g}_i \\ &= \underline{u}^j g_{ij} \ , \end{aligned} \tag{5.65a}$$

$$\begin{aligned} \mathbf{u} \cdot \mathbf{g}_i &= \underline{u}_j \mathbf{g}^j \cdot \mathbf{g}_i = \underline{u}_j \delta_i^j \\ &= \underline{u}_i \ , \end{aligned} \tag{5.65b}$$

$$\mathbf{u} \cdot \mathbf{g}^i = \underline{u}^j \mathbf{g}_j \cdot \mathbf{g}^i$$

$$= \underline{u}^j \delta^i_j = \underline{u}^i \ , \tag{5.65c}$$

$$\mathbf{u} \cdot \mathbf{g}^i = \underline{u}_j \mathbf{g}^j \cdot \mathbf{g}^i$$
$$= \underline{u}_j g^{ij} \ . \tag{5.65d}$$

As a result,

$$\boxed{\underline{u}_i = g_{ij} \, \underline{u}^j \ ,} \tag{5.66a}$$

$$\boxed{\underline{u}^i = g^{ij} \, \underline{u}_j \ .} \tag{5.66b}$$

These expressions determine the relationships between the covariant and contravariant components of a vector and again show the crucially important role of metric coefficients in raising and lowering indices, see the similar results in (5.43) and (5.49).

5.4.1 Dot Product and Cross Product Between Two Vectors

Given two vectors **u** and **v**, each of which possibly has two forms according to (5.64a) and (5.64b). The goal is now to provide their scalar product as well as cross product having in mind that they are bilinear operators. Making use of $(5.27)_{1-2}$, $(5.38)_1$ and $(5.46)_1$, the inner product of **u** and **v** admits the following four forms

$$\mathbf{u} \cdot \mathbf{v} = \underline{u}^i g_{ij} \, \underline{v}^j \ , \tag{5.67a}$$

$$\mathbf{u} \cdot \mathbf{v} = \underline{u}^i \, \underline{v}_i \ , \tag{5.67b}$$

$$\mathbf{u} \cdot \mathbf{v} = \underline{u}_i \, \underline{v}^i \ . \tag{5.67c}$$

$$\mathbf{u} \cdot \mathbf{v} = \underline{u}_i g^{ij} \, \underline{v}_j \ . \tag{5.67d}$$

Accordingly, the length $|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$ of **u** can be computed according to

$$|\mathbf{u}| = \sqrt{\underline{u}^i g_{ij} \, \underline{u}^j} = \sqrt{\underline{u}^i \, \underline{u}_i} = \sqrt{\underline{u}_i g^{ij} \, \underline{u}_j} \ . \tag{5.68}$$

With the aid of $(5.33)_2$, $(5.35)_2$, (5.66a) and (5.66b), the cross product of **u** and **v** allows the following representations

$$\mathbf{u} \times \mathbf{v} = J \, \underline{u}^i \, \underline{v}^j \varepsilon_{ijk} \mathbf{g}^k = J \, \underline{u}^i g^{jm} \, \underline{v}_m \varepsilon_{ijk} \mathbf{g}^k$$
$$= J g^{im} \, \underline{u}_m \, \underline{v}^j \varepsilon_{ijk} \mathbf{g}^k = J g^{im} \, \underline{u}_m g^{jn} \, \underline{v}_n \varepsilon_{ijk} \mathbf{g}^k \ , \tag{5.69a}$$

$$\mathbf{u} \times \mathbf{v} = \frac{1}{J} g_{im} \, \underline{u}^i g_{jn} \, \underline{v}^n \varepsilon^{ijk} \mathbf{g}_k = \frac{1}{J} g_{im} \, \underline{u}^i \, \underline{v}_j \varepsilon^{ijk} \mathbf{g}_k$$

$$= \frac{1}{J} \underline{u}_i g_{jm} \underline{v}^m \varepsilon^{ijk} \mathbf{g}_k = \frac{1}{J} \underline{u}_i \underline{v}_j \varepsilon^{ijk} \mathbf{g}_k \ . \tag{5.69b}$$

5.4.2 Matrix Notation

Consistent with (1.34)$_1$, (5.64a) and (5.64b), an arbitrary vector $\mathbf{u}$ in this text can have three various forms in matrix notation. Having in mind (1.44), one needs to distinguish between the contravariant single-column matrix

$$[\,\mathbf{u}\,]^{\text{con}} = \left[\, \underline{u}^1 \quad \underline{u}^2 \quad \underline{u}^3 \,\right]^{\text{T}} , \tag{5.70}$$

and its covariant counterpart; written by,

$$[\,\mathbf{u}\,]_{\text{cov}} = \left[\, \underline{u}_1 \quad \underline{u}_2 \quad \underline{u}_3 \,\right]^{\text{T}} . \tag{5.71}$$

As an example, consider the Cartesian vector $\mathbf{u} = 2\,\widehat{\mathbf{e}}_1 - 2\,\widehat{\mathbf{e}}_2 - 4\,\widehat{\mathbf{e}}_3$ as well as the covariant basis vectors $\mathbf{g}_1 = \widehat{\mathbf{e}}_2 + 3\,\widehat{\mathbf{e}}_3$, $\mathbf{g}_2 = -\widehat{\mathbf{e}}_1 + 2\,\widehat{\mathbf{e}}_2 + 4\,\widehat{\mathbf{e}}_3$ and $\mathbf{g}_3 = \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3$, that is,

$$[\,\mathbf{u}\,] = \begin{bmatrix} 2 \\ -2 \\ -4 \end{bmatrix} \ , \quad [\,\mathbf{g}_1\,] = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \ , \quad [\,\mathbf{g}_2\,] = \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix} \ , \quad [\,\mathbf{g}_3\,] = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} .$$

The goal is to represent the single-column matrices $[\,\mathbf{u}\,]^{\text{con}}$ and $[\,\mathbf{u}\,]_{\text{cov}}$. One first needs to provide the inverse transpose of $\mathbf{F} = \mathbf{g}_i \otimes \widehat{\mathbf{e}}^i$ or $[\,\mathbf{F}\,] = \left[\,\mathbf{g}_1 \ \mathbf{g}_2 \ \mathbf{g}_3\,\right]$ in (5.21). Using (2.50)$_2$ and (2.120), it is given by

$$\left[\,\mathbf{F}^{-\text{T}}\,\right] = \begin{bmatrix} 1 & -1 & 1 \\ -0.5 & 0 & 1.5 \\ 0.5 & 0 & -0.5 \end{bmatrix} .$$

Knowing that $\left[\,\mathbf{F}^{-\text{T}}\,\right] = \left[\,\mathbf{g}^1 \ \mathbf{g}^2 \ \mathbf{g}^3\,\right]$, one then obtains the contravariant basis vectors:

$$\left[\,\mathbf{g}^1\,\right] = \left[\,1 \ \ -0.5 \ \ 0.5\,\right]^{\text{T}} \ , \quad \left[\,\mathbf{g}^2\,\right] = \left[\,-1 \ \ 0 \ \ 0\,\right]^{\text{T}} \ , \quad \left[\,\mathbf{g}^3\,\right] = \left[\,1 \ \ 1.5 \ \ -0.5\,\right]^{\text{T}} .$$

or $\mathbf{g}^1 = \widehat{\mathbf{e}}_1 - 0.5\,\widehat{\mathbf{e}}_2 + 0.5\,\widehat{\mathbf{e}}_3$, $\mathbf{g}^2 = -\widehat{\mathbf{e}}_1$, $\mathbf{g}^3 = \widehat{\mathbf{e}}_1 + 1.5\,\widehat{\mathbf{e}}_2 - 0.5\,\widehat{\mathbf{e}}_3$

Consequently, by means of the expressions (1.45)$_1$, (5.38) and (5.46), the metric coefficients $g_{ij} = \left[\,\mathbf{g}_i\,\right]^{\text{T}} \left[\,\mathbf{g}_j\,\right]$ and $g^{ij} = \left[\,\mathbf{g}^i\,\right]^{\text{T}} \left[\,\mathbf{g}^j\,\right]$ become

$$\left[\, g_{ij} \,\right] = \begin{bmatrix} 10 & 14 & 4 \\ 14 & 21 & 6 \\ 4 & 6 & 2 \end{bmatrix} \quad , \quad \left[\, g^{ij} \,\right] = \begin{bmatrix} 1.5 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 3.5 \end{bmatrix} .$$

By (1.45)$_1$, (5.65b)$_3$ and (5.65c)$_3$, the required components are

$$\begin{aligned} \underline{u}^i &= \mathbf{u} \cdot \mathbf{g}^i = [\,\mathbf{u}\,]^{\mathrm{T}} \left[\, \mathbf{g}^i \,\right] \quad &\Rightarrow \quad &\left(\underline{u}^1, \underline{u}^2, \underline{u}^3 \right) = (1, -2, 1) \ , \\ \underline{u}_i &= \mathbf{u} \cdot \mathbf{g}_i = [\,\mathbf{u}\,]^{\mathrm{T}} \left[\, \mathbf{g}_i \,\right] \quad &\Rightarrow \quad &\left(\underline{u}_1, \underline{u}_2, \underline{u}_3 \right) = (-14, -22, -6) \ . \end{aligned}$$

One can now verify the relations $\underline{u}_i = g_{ij}\, \underline{u}^j$ and $\underline{u}^i = g^{ij}\, \underline{u}_j$ given in (5.66a) and (5.66b), respectively. Guided by (5.64a) and (5.64b), these components help express the vector $\mathbf{u}$ as

$$\begin{aligned} \mathbf{u} &= 2\widehat{\mathbf{e}}_1 - 2\widehat{\mathbf{e}}_2 - 4\widehat{\mathbf{e}}_3 \\ &= \mathbf{g}_1 - 2\mathbf{g}_2 + \mathbf{g}_3 \\ &= -14\mathbf{g}^1 - 22\mathbf{g}^2 - 6\mathbf{g}^3 \ . \end{aligned}$$

Finally, this vector in matrix notation admits the following three forms

$$[\,\mathbf{u}\,] = \begin{bmatrix} 2 \\ -2 \\ -4 \end{bmatrix} \quad , \quad [\,\mathbf{u}\,]^{\mathrm{con}} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad , \quad [\,\mathbf{u}\,]_{\mathrm{cov}} = \begin{bmatrix} -14 \\ -22 \\ -6 \end{bmatrix} .$$

5.5 Contravariant, Mixed and Covariant Components of Second-Order Tensors

Recall from (2.17) that the basis $\left\{ \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \right\}$ helped construct a Cartesian tensor according to (2.19)$_2$, i.e. $\mathbf{A} = A_{ij} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$. This reveals the fact that a Cartesian tensor has only one form. Given the covariant basis $\{\mathbf{g}_i\}$ along with its companion contravariant basis $\left\{\mathbf{g}^i\right\}$, one can then construct the following four forms of basis tensors

$$\left\{ \mathbf{g}_i \otimes \mathbf{g}_j \right\} \quad , \quad \left\{ \mathbf{g}_i \otimes \mathbf{g}^j \right\} \quad , \quad \left\{ \mathbf{g}^i \otimes \mathbf{g}_j \right\} \quad , \quad \left\{ \mathbf{g}^i \otimes \mathbf{g}^j \right\} . \tag{5.72}$$

Guided by (5.64a) and (5.64b), any second-order tensor $\mathbf{A} \in \mathscr{T}_{\mathrm{so}} \left(\mathscr{E}_{\mathrm{r}}^{o3} \right)$ can now be decomposed with respect to these basis tensors as

$$\mathbf{A} = \underline{A}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j \ , \tag{5.73a}$$

$$\mathbf{A} = \underline{A}^{i}_{\ .j} \mathbf{g}_i \otimes \mathbf{g}^j \ , \tag{5.73b}$$

$$\mathbf{A} = \underline{A}_{i}^{\ .j} \mathbf{g}^i \otimes \mathbf{g}_j \ , \tag{5.73c}$$

$$\mathbf{A} = \underline{A}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j \ . \tag{5.73d}$$

Here, $\underline{A}^{ij}$ and $\underline{A}_{ij}$ are the **contravariant** and **covariant components** of **A**, respectively. Whereas $\underline{A}^{i}_{.j}$ and $\underline{A}_{i}^{.j}$ present its **mixed contra-covariant** and **mixed co-contravariant components**, respectively. Note that the 'dot' used in front of any index of the mixed components is a placeholder declaring other index as the first index. Dictated by the first index of the mixed tensor bases in (5.73b) and (5.73c), i occurs first in both $\underline{A}^{i}_{.j}$ and $\underline{A}_{i}^{.j}$. Similarly to vectors, the first form according to (5.73a) is the natural way of representing a tensor. For the sake of clarification, a tensor with contravariant (covariant) components is called a *contravariant* (*covariant*) *tensor* in this text. And a *contra-covariant* (*co-contravariant*) *tensor* is one whose components are contra-covariant (co-contravariant).

Considering the four various forms of the components as well as bases of a tensor, by use of (2.73), (2.74a), (2.74b), (2.79c), $(5.27)_{1-2}$, $(5.38)_1$ and $(5.46)_1$, one can write

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}_j = \underline{A}^{mn}\mathbf{g}_m \otimes \mathbf{g}_n : \mathbf{g}_i \otimes \mathbf{g}_j = g_{im}\,\underline{A}^{mn}g_{nj} \;, \tag{5.74a}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}_j = \underline{A}^{m}_{.n}\mathbf{g}_m \otimes \mathbf{g}^n : \mathbf{g}_i \otimes \mathbf{g}_j = g_{im}\,\underline{A}^{m}_{.n}\delta^n_j = g_{im}\,\underline{A}^{m}_{.j} \;, \tag{5.74b}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}_j = \underline{A}_{m}^{.n}\mathbf{g}^m \otimes \mathbf{g}_n : \mathbf{g}_i \otimes \mathbf{g}_j = \delta^m_i\,\underline{A}_{m}^{.n}g_{nj} = \underline{A}_{i}^{.m}g_{mj} \;, \tag{5.74c}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}_j = \underline{A}_{mn}\mathbf{g}^m \otimes \mathbf{g}^n : \mathbf{g}_i \otimes \mathbf{g}_j = \delta^m_i\,\underline{A}_{mn}\delta^n_j = \underline{A}_{ij} \;, \tag{5.74d}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}^j = \underline{A}^{mn}\mathbf{g}_m \otimes \mathbf{g}_n : \mathbf{g}_i \otimes \mathbf{g}^j = g_{im}\,\underline{A}^{mn}\delta^j_n = g_{im}\,\underline{A}^{mj} \;, \tag{5.74e}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}^j = \underline{A}^{m}_{.n}\mathbf{g}_m \otimes \mathbf{g}^n : \mathbf{g}_i \otimes \mathbf{g}^j = g_{im}\,\underline{A}^{m}_{.n}g^{nj} \;, \tag{5.74f}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}^j = \underline{A}_{m}^{.n}\mathbf{g}^m \otimes \mathbf{g}_n : \mathbf{g}_i \otimes \mathbf{g}^j = \delta^m_i\,\underline{A}_{m}^{.n}\delta^j_n = \underline{A}_{i}^{.j} \;, \tag{5.74g}$$

$$\mathbf{g}_i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}_i \otimes \mathbf{g}^j = \underline{A}_{mn}\mathbf{g}^m \otimes \mathbf{g}^n : \mathbf{g}_i \otimes \mathbf{g}^j = \delta^m_i\,\underline{A}_{mn}g^{nj} = \underline{A}_{im}g^{mj} \;, \tag{5.74h}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}_j = \underline{A}^{mn}\mathbf{g}_m \otimes \mathbf{g}_n : \mathbf{g}^i \otimes \mathbf{g}_j = \delta^i_m\,\underline{A}^{mn}g_{nj} = \underline{A}^{im}g_{mj} \;, \tag{5.74i}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}_j = \underline{A}^{m}_{.n}\mathbf{g}_m \otimes \mathbf{g}^n : \mathbf{g}^i \otimes \mathbf{g}_j = \delta^i_m\,\underline{A}^{m}_{.n}\delta^n_j = \underline{A}^{i}_{.j} \;, \tag{5.74j}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}_j = \underline{A}_{m}^{.n}\mathbf{g}^m \otimes \mathbf{g}_n : \mathbf{g}^i \otimes \mathbf{g}_j = g^{im}\,\underline{A}_{m}^{.n}g_{nj} \;, \tag{5.74k}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}_j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}_j = \underline{A}_{mn}\mathbf{g}^m \otimes \mathbf{g}^n : \mathbf{g}^i \otimes \mathbf{g}_j = g^{im}\,\underline{A}_{mn}\delta^n_j = g^{im}\,\underline{A}_{mj} \;, \tag{5.74l}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}^j = \underline{A}^{mn}\mathbf{g}_m \otimes \mathbf{g}_n : \mathbf{g}^i \otimes \mathbf{g}^j = \delta^i_m\,\underline{A}^{mn}\delta^j_n = \underline{A}^{ij} \;, \tag{5.74m}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}^j = \underline{A}^{m}_{.n}\mathbf{g}_m \otimes \mathbf{g}^n : \mathbf{g}^i \otimes \mathbf{g}^j = \delta^i_m\,\underline{A}^{m}_{.n}g^{nj} = \underline{A}^{i}_{.m}g^{mj} \;, \tag{5.74n}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}^j = \underline{A}_{m}^{.n}\mathbf{g}^m \otimes \mathbf{g}_n : \mathbf{g}^i \otimes \mathbf{g}^j = g^{im}\,\underline{A}_{m}^{.n}\delta^j_n = g^{im}\,\underline{A}_{m}^{.j} \;, \tag{5.74o}$$

$$\mathbf{g}^i \cdot \mathbf{A}\mathbf{g}^j = \mathbf{A} : \mathbf{g}^i \otimes \mathbf{g}^j = \underline{A}_{mn}\mathbf{g}^m \otimes \mathbf{g}^n : \mathbf{g}^i \otimes \mathbf{g}^j = g^{im}\,\underline{A}_{mn}g^{nj} \;, \tag{5.74p}$$

wherein

$$\mathbf{A}\mathbf{g}_j = \underline{A}^{mn} g_{nj} \mathbf{g}_m \ , \tag{5.75a}$$
$$\mathbf{A}\mathbf{g}_j = \underline{A}^{m}_{\ .j} \mathbf{g}_m \ , \tag{5.75b}$$
$$\mathbf{A}\mathbf{g}_j = \underline{A}_m^{\ .n} g_{nj} \mathbf{g}^m \ , \tag{5.75c}$$
$$\mathbf{A}\mathbf{g}_j = \underline{A}_{mj} \mathbf{g}^m \ , \tag{5.75d}$$
$$\mathbf{A}\mathbf{g}^j = \underline{A}^{mj} \mathbf{g}_m \ , \tag{5.75e}$$
$$\mathbf{A}\mathbf{g}^j = \underline{A}^{m}_{\ .n} g^{nj} \mathbf{g}_m \ , \tag{5.75f}$$
$$\mathbf{A}\mathbf{g}^j = \underline{A}_m^{\ .j} \mathbf{g}^m \ , \tag{5.75g}$$
$$\mathbf{A}\mathbf{g}^j = \underline{A}_{mn} g^{nj} \mathbf{g}^m \ . \tag{5.75h}$$

The expressions (5.74a)–(5.74p) now deliver the following relationships between the components:

$$\underline{A}_{ij} = g_{ik}\, \underline{A}^{k}_{\ .j} = \underline{A}_i^{\ .k} g_{kj} = g_{ik}\, \underline{A}^{kl} g_{lj} \ , \tag{5.76a}$$
$$\underline{A}_i^{\ .j} = g_{ik}\, \underline{A}^{kj} = \underline{A}_{ik} g^{kj} = g_{ik}\, \underline{A}^{k}_{\ .l} g^{lj} \ , \tag{5.76b}$$
$$\underline{A}^{i}_{\ .j} = g^{ik}\, \underline{A}_{kj} = \underline{A}^{ik} g_{kj} = g^{ik}\, \underline{A}_k^{\ .l} g_{lj} \ , \tag{5.76c}$$
$$\underline{A}^{ij} = g^{ik}\, \underline{A}_k^{\ .j} = \underline{A}^{i}_{\ .k} g^{kj} = g^{ik}\, \underline{A}_{kl} g^{lj} \ . \tag{5.76d}$$

Consistent with (5.64a)–(5.64b) and (5.73a)–(5.73d), the tensor product of two vectors according to $(2.24)_3$ now takes the following forms

$$\mathbf{u} \otimes \mathbf{v} = \underline{u}^i\, \underline{v}^j \mathbf{g}_i \otimes \mathbf{g}_j \ , \tag{5.77a}$$
$$\mathbf{u} \otimes \mathbf{v} = \underline{u}^i\, \underline{v}_j \mathbf{g}_i \otimes \mathbf{g}^j \ , \tag{5.77b}$$
$$\mathbf{u} \otimes \mathbf{v} = \underline{u}_i\, \underline{v}^j \mathbf{g}^i \otimes \mathbf{g}_j \ , \tag{5.77c}$$
$$\mathbf{u} \otimes \mathbf{v} = \underline{u}_i\, \underline{v}_j \mathbf{g}^i \otimes \mathbf{g}^j \ . \tag{5.77d}$$

The identity tensor in curvilinear coordinates, referred to as *metric tensor*, has the following representations

$$\mathbf{I} = g^{ij} \mathbf{g}_i \otimes \mathbf{g}_j = \mathbf{g}_i \otimes \mathbf{g}^i = \mathbf{g}^i \otimes \mathbf{g}_i = g_{ij} \mathbf{g}^i \otimes \mathbf{g}^j \ , \quad \longleftarrow \ \text{see (2.23) and (5.16)} \tag{5.78}$$

owing to

$$(\mathbf{I})^{ij} \overset{\text{from}}{\underset{(5.74\text{m})}{=\!=\!=}} \mathbf{g}^i \cdot \mathbf{I}\mathbf{g}^j \overset{\text{from}}{\underset{(2.5)}{=\!=\!=}} \mathbf{g}^i \cdot \mathbf{g}^j \overset{\text{from}}{\underset{(5.46)}{=\!=\!=}} g^{ij} \ , \tag{5.79a}$$
$$(\mathbf{I})^{i}_{\ .j} \overset{\text{from}}{\underset{(5.74\text{j})}{=\!=\!=}} \mathbf{g}^i \cdot \mathbf{I}\mathbf{g}_j \overset{\text{from}}{\underset{(2.5)}{=\!=\!=}} \mathbf{g}^i \cdot \mathbf{g}_j \overset{\text{from}}{\underset{(5.27)}{=\!=\!=}} \delta^i_j \ , \tag{5.79b}$$
$$(\mathbf{I})_i^{\ .j} \overset{\text{from}}{\underset{(5.74\text{g})}{=\!=\!=}} \mathbf{g}_i \cdot \mathbf{I}\mathbf{g}^j \overset{\text{from}}{\underset{(2.5)}{=\!=\!=}} \mathbf{g}_i \cdot \mathbf{g}^j \overset{\text{from}}{\underset{(5.27)}{=\!=\!=}} \delta_i^j \ , \tag{5.79c}$$
$$(\mathbf{I})_{ij} \overset{\text{from}}{\underset{(5.74\text{d})}{=\!=\!=}} \mathbf{g}_i \cdot \mathbf{I}\mathbf{g}_j \overset{\text{from}}{\underset{(2.5)}{=\!=\!=}} \mathbf{g}_i \cdot \mathbf{g}_j \overset{\text{from}}{\underset{(5.38)}{=\!=\!=}} g_{ij} \ . \tag{5.79d}$$

For the sake of clarification, the **covariant metric tensor** is specifically denoted by

$$\mathbf{g} = g_{ij} \mathbf{g}^i \otimes \mathbf{g}^j \; . \tag{5.80}$$

Its inverse basically renders the **contravariant metric tensor**, that is,

$$\mathbf{g}^{-1} = g^{ij} \mathbf{g}_i \otimes \mathbf{g}_j \; . \tag{5.81}$$

These metric tensors are frequently used in nonlinear continuum mechanics, see Simo and Pister [7], Simo et al. [8] and Ibrahimbegović [9] among many others. Then, the mixed **contra-covariant metric tensor** is denoted by

$$\mathbf{g}_{\text{mix}} = \mathbf{g}_i \otimes \mathbf{g}^i \; . \tag{5.82}$$

The transpose of $\mathbf{g}_{\text{mix}}$ represents the mixed **co-contravariant metric tensor**:

$$\mathbf{g}_{\text{mix}}^{\text{T}} = \mathbf{g}^i \otimes \mathbf{g}_i \; . \tag{5.83}$$

5.5.1 Basic Tensor Relationships

Recall from Chap. 2 that the tensor relationships were expressed in terms of the Cartesian components of tensors. In the following, the introduced relations will be represented in terms of the generalized components of vectors and tensors (see Exercise 5.2).

Mapping of a vector by a tensor: ↙ recall from (2.22) that $\mathbf{v} = \mathbf{A}\mathbf{u} = A_{ij} u_j \widehat{\mathbf{e}}_i$

$$\mathbf{v} = \mathbf{A}\mathbf{u} = \underline{v}^i \mathbf{g}_i \quad \text{where} \quad \underline{v}^i = \underline{A}^{ij} \underline{u}_j = \underline{A}^{im} g_{mj} \underline{u}^j \; , \tag{5.84a}$$

$$\mathbf{v} = \mathbf{A}\mathbf{u} = \underline{v}^i \mathbf{g}_i \quad \text{where} \quad \underline{v}^i = \underline{A}^{i}{}_{.j} \underline{u}^j = \underline{A}^{i}{}_{.m} g^{mj} \underline{u}_j \; , \tag{5.84b}$$

$$\mathbf{v} = \mathbf{A}\mathbf{u} = \underline{v}_i \mathbf{g}^i \quad \text{where} \quad \underline{v}_i = \underline{A}_{ij} \underline{u}^j = \underline{A}_{im} g^{mj} \underline{u}_j \; , \tag{5.84c}$$

$$\mathbf{v} = \mathbf{A}\mathbf{u} = \underline{v}_i \mathbf{g}^i \quad \text{where} \quad \underline{v}_i = \underline{A}_{i}{}^{.j} \underline{u}_j = \underline{A}_{i}{}^{.m} g_{mj} \underline{u}^j \; . \tag{5.84d}$$

Composition: ↙ recall from (2.25) that $(\mathbf{AB})\,\mathbf{u} = \mathbf{A}\,(\mathbf{Bu})$ and this led to $(\mathbf{AB})_{ij} = (\mathbf{A})_{im}\,(\mathbf{B})_{mj}$

$$\begin{aligned}
\mathbf{AB} &= \underline{A}^{i}{}_{.m} g^{mn} \underline{B}_{n}{}^{.j} \mathbf{g}_i \otimes \mathbf{g}_j = \underline{A}^{im} g_{mn} \underline{B}^{nj} \mathbf{g}_i \otimes \mathbf{g}_j \\
&= \underline{A}^{im} \underline{B}_{m}{}^{.j} \mathbf{g}_i \otimes \mathbf{g}_j = \underline{A}^{i}{}_{.m} \underline{B}^{mj} \mathbf{g}_i \otimes \mathbf{g}_j \; ,
\end{aligned} \tag{5.85a}$$

$$\begin{aligned}
\mathbf{AB} &= \underline{A}^{i}{}_{.m} g^{mn} \underline{B}_{nj} \mathbf{g}_i \otimes \mathbf{g}^j = \underline{A}^{im} g_{mn} \underline{B}^{n}{}_{.j} \mathbf{g}_i \otimes \mathbf{g}^j \\
&= \underline{A}^{im} \underline{B}_{mj} \mathbf{g}_i \otimes \mathbf{g}^j = \underline{A}^{i}{}_{.m} \underline{B}^{m}{}_{.j} \mathbf{g}_i \otimes \mathbf{g}^j \; ,
\end{aligned} \tag{5.85b}$$

$$\begin{aligned}
\mathbf{AB} &= \underline{A}_{im} g^{mn} \underline{B}_{n}{}^{.j} \mathbf{g}^i \otimes \mathbf{g}_j = \underline{A}_{i}{}^{.m} g_{mn} \underline{B}^{nj} \mathbf{g}^i \otimes \mathbf{g}_j \\
&= \underline{A}_{i}{}^{.m} \underline{B}_{m}{}^{.j} \mathbf{g}^i \otimes \mathbf{g}_j = \underline{A}_{im} \underline{B}^{mj} \mathbf{g}^i \otimes \mathbf{g}_j \; ,
\end{aligned} \tag{5.85c}$$

$$\begin{aligned}\mathbf{AB} &= \underline{A}_{im} g^{mn}\, \underline{B}_{nj} \mathbf{g}^i \otimes \mathbf{g}^j = \underline{A}_i^{\cdot m} g_{mn}\, \underline{B}^n_{\cdot j} \mathbf{g}^i \otimes \mathbf{g}^j \\ &= \underline{A}_i^{\cdot m}\, \underline{B}_{mj} \mathbf{g}^i \otimes \mathbf{g}^j = \underline{A}_{im}\, \underline{B}^m_{\cdot j} \mathbf{g}^i \otimes \mathbf{g}^j \ .\end{aligned} \tag{5.85d}$$

Transposed tensor: ↙ recall from (2.48) that $\mathbf{A}^T\mathbf{u} = \mathbf{uA}$

$$[\mathbf{A}^T = \underbrace{\left(\mathbf{A}^T\right)^{ij} \mathbf{g}_i \otimes \mathbf{g}_j = (\mathbf{A})^{ji}\, \mathbf{g}_i \otimes \mathbf{g}_j}_{\text{for a symmetric tensor } \mathbf{S}, \text{ one then has } \underline{S}^{ij} = \underline{S}^{ji}} \ , \tag{5.86a}$$

$$\mathbf{A}^T = \underbrace{\left(\mathbf{A}^T\right)^i_{\cdot j} \mathbf{g}_i \otimes \mathbf{g}^j = (\mathbf{A})_j^{\cdot i}\, \mathbf{g}_i \otimes \mathbf{g}^j}_{\text{for a symmetric tensor } \mathbf{S}, \text{ one then has } \underline{S}^i_{\cdot j} = \underline{S}_j^{\cdot i}} \ , \tag{5.86b}$$

$$\mathbf{A}^T = \underbrace{\left(\mathbf{A}^T\right)_i^{\cdot j} \mathbf{g}^i \otimes \mathbf{g}_j = (\mathbf{A})^j_{\cdot i}\, \mathbf{g}^i \otimes \mathbf{g}_j}_{\text{for a symmetric tensor } \mathbf{S}, \text{ one then has } \underline{S}_i^{\cdot j} = \underline{S}^j_{\cdot i}} \ , \tag{5.86c}$$

$$\mathbf{A}^T = \underbrace{\left(\mathbf{A}^T\right)_{ij} \mathbf{g}^i \otimes \mathbf{g}^j = (\mathbf{A})_{ji}\, \mathbf{g}^i \otimes \mathbf{g}^j}_{\text{for a symmetric tensor } \mathbf{S}, \text{ one then has } \underline{S}_{ij} = \underline{S}_{ji}} \ . \tag{5.86d}$$

Contraction: ↙ recall from (2.75) that $\mathbf{A} : \mathbf{B} = A_{ij} B_{ij}$

$$\begin{aligned}\mathbf{A} : \mathbf{B} &= \underline{A}^{ij} g_{im} g_{jn}\, \underline{B}^{mn} = \underline{A}^{ij} g_{im}\, \underline{B}^m_{\cdot j} = \underline{A}^{ij} g_{jm}\, \underline{B}_i^{\cdot m} = \underline{A}^{ij}\, \underline{B}_{ij} \\ &= \underline{A}^i_{\cdot j} g_{im} g^{jn}\, \underline{B}^m_{\cdot n} = \underline{A}^i_{\cdot j} g_{im}\, \underline{B}^{mj} = \underline{A}^i_{\cdot j} g^{jm}\, \underline{B}_{im} = \underline{A}^i_{\cdot j}\, \underline{B}_i^{\cdot j} \\ &= \underline{A}_i^{\cdot j} g^{im} g_{jn}\, \underline{B}_m^{\cdot n} = \underline{A}_i^{\cdot j} g^{im}\, \underline{B}_{mj} = \underline{A}_i^{\cdot j} g_{jm}\, \underline{B}^{im} = \underline{A}_i^{\cdot j}\, \underline{B}^i_{\cdot j} \\ &= \underline{A}_{ij} g^{im} g^{jn}\, \underline{B}_{mn} = \underline{A}_{ij} g^{im}\, \underline{B}_m^{\cdot j} = \underline{A}_{ij} g^{jm}\, \underline{B}^i_{\cdot m} = \underline{A}_{ij}\, \underline{B}^{ij} \ .\end{aligned} \tag{5.87}$$

Trace: ↙ recall from (2.89a) that $\mathrm{tr}\mathbf{A} = \mathbf{I} : \mathbf{A} = A_{ii}$

$$\mathrm{tr}\mathbf{A} = g_{ij}\, \underline{A}^{ij} = \underline{A}^i_{\cdot i} = \underline{A}_i^{\cdot i} = g^{ij}\, \underline{A}_{ij} \ . \tag{5.88}$$

Determinant: ↙ recall from (2.98) that $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) \det \mathbf{A} = \mathbf{Au} \cdot (\mathbf{Av} \times \mathbf{Aw})$

$$\det[\mathbf{A}] = \frac{J^2}{6} \varepsilon_{ijk} \varepsilon_{lmn}\, \underline{A}^{il}\, \underline{A}^{jm}\, \underline{A}^{kn} = J^2 \det \begin{bmatrix} \underline{A}^{11} & \underline{A}^{12} & \underline{A}^{13} \\ \underline{A}^{21} & \underline{A}^{22} & \underline{A}^{23} \\ \underline{A}^{31} & \underline{A}^{32} & \underline{A}^{33} \end{bmatrix} \ , \tag{5.89a}$$

$$\det[\mathbf{A}] = \frac{1}{6} \varepsilon_{ijk} \varepsilon^{lmn}\, \underline{A}^i_{\cdot l}\, \underline{A}^j_{\cdot m}\, \underline{A}^k_{\cdot n} = \det \begin{bmatrix} \underline{A}^1_{\cdot 1} & \underline{A}^1_{\cdot 2} & \underline{A}^1_{\cdot 3} \\ \underline{A}^2_{\cdot 1} & \underline{A}^2_{\cdot 2} & \underline{A}^2_{\cdot 3} \\ \underline{A}^3_{\cdot 1} & \underline{A}^3_{\cdot 2} & \underline{A}^3_{\cdot 3} \end{bmatrix} \ , \tag{5.89b}$$

$$\det[\mathbf{A}] = \frac{1}{6} \varepsilon^{ijk} \varepsilon_{lmn}\, \underline{A}_i^{\cdot l}\, \underline{A}_j^{\cdot m}\, \underline{A}_k^{\cdot n} = \det \begin{bmatrix} \underline{A}_1^{\cdot 1} & \underline{A}_1^{\cdot 2} & \underline{A}_1^{\cdot 3} \\ \underline{A}_2^{\cdot 1} & \underline{A}_2^{\cdot 2} & \underline{A}_2^{\cdot 3} \\ \underline{A}_3^{\cdot 1} & \underline{A}_3^{\cdot 2} & \underline{A}_3^{\cdot 3} \end{bmatrix} \ , \tag{5.89c}$$

$$\det[\mathbf{A}] = \frac{J^{-2}}{6}\varepsilon^{ijk}\varepsilon^{lmn} \underline{A}_{il} \underline{A}_{jm} \underline{A}_{kn} = J^{-2} \det \begin{bmatrix} \underline{A}_{11} & \underline{A}_{12} & \underline{A}_{13} \\ \underline{A}_{21} & \underline{A}_{22} & \underline{A}_{23} \\ \underline{A}_{31} & \underline{A}_{32} & \underline{A}_{33} \end{bmatrix} . \tag{5.89d}$$

Cofactor: ↙ recall from (2.112) that $\mathbf{A}^c(\mathbf{u} \times \mathbf{v}) = (\mathbf{Au}) \times (\mathbf{Av})$

$$\mathbf{A}^c = (\mathbf{A}^c)^{ij} \mathbf{g}_i \otimes \mathbf{g}_j = \frac{1}{2J^2}\varepsilon^{ikl}\varepsilon^{jmn} \underline{A}_{km} \underline{A}_{ln}\mathbf{g}_i \otimes \mathbf{g}_j , \tag{5.90a}$$

$$\mathbf{A}^c = (\mathbf{A}^c)^{i}_{.j} \mathbf{g}_i \otimes \mathbf{g}^j = \frac{1}{2}\varepsilon^{ikl}\varepsilon_{jmn} \underline{A}_k^{.m} \underline{A}_l^{.n}\mathbf{g}_i \otimes \mathbf{g}^j , \tag{5.90b}$$

$$\mathbf{A}^c = (\mathbf{A}^c)_{i}^{.j} \mathbf{g}^i \otimes \mathbf{g}_j = \frac{1}{2}\varepsilon_{ikl}\varepsilon^{jmn} \underline{A}^k_{.m} \underline{A}^l_{.n}\mathbf{g}^i \otimes \mathbf{g}_j , \tag{5.90c}$$

$$\mathbf{A}^c = (\mathbf{A}^c)_{ij} \mathbf{g}^i \otimes \mathbf{g}^j = \frac{J^2}{2}\varepsilon_{ikl}\varepsilon_{jmn} \underline{A}^{km} \underline{A}^{ln}\mathbf{g}^i \otimes \mathbf{g}^j . \tag{5.90d}$$

Inverse tensor: ↙ recall from (2.105) and (2.120) that $\mathbf{AA}^{-1} = \mathbf{A}\,(2\det\mathbf{A})^{-1}\,\varepsilon_{ikl}\varepsilon_{jmn}A_{mk}A_{nl}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \mathbf{I}$

$$\mathbf{A}^{-1} = \left(\mathbf{A}^{-1}\right)^{ij} \mathbf{g}_i \otimes \mathbf{g}_j = \frac{1}{2J^2\det\mathbf{A}}\varepsilon^{imn}\varepsilon^{jkl} \underline{A}_{km} \underline{A}_{ln}\mathbf{g}_i \otimes \mathbf{g}_j , \tag{5.91a}$$

$$\mathbf{A}^{-1} = \left(\mathbf{A}^{-1}\right)^{i}_{.j} \mathbf{g}_i \otimes \mathbf{g}^j = \frac{1}{2\det\mathbf{A}}\varepsilon^{imn}\varepsilon_{jkl} \underline{A}^k_{.m} \underline{A}^l_{.n}\mathbf{g}_i \otimes \mathbf{g}^j , \tag{5.91b}$$

$$\mathbf{A}^{-1} = \left(\mathbf{A}^{-1}\right)_{i}^{.j} \mathbf{g}^i \otimes \mathbf{g}_j = \frac{1}{2\det\mathbf{A}}\varepsilon_{imn}\varepsilon^{jkl} \underline{A}_k^{.m} \underline{A}_l^{.n}\mathbf{g}^i \otimes \mathbf{g}_j , \tag{5.91c}$$

$$\mathbf{A}^{-1} = \left(\mathbf{A}^{-1}\right)_{ij} \mathbf{g}^i \otimes \mathbf{g}^j = \frac{J^2}{2\det\mathbf{A}}\varepsilon_{imn}\varepsilon_{jkl} \underline{A}^{km} \underline{A}^{ln}\mathbf{g}^i \otimes \mathbf{g}^j . \tag{5.91d}$$

It is not then difficult to deduce from $\mathbf{I} = \mathbf{AA}^{-1}$ that

$$g^{rj} = \frac{1}{2\det\mathbf{A}}\varepsilon_{imn}\varepsilon^{jkl} \underline{A}^{ri} \underline{A}_k^{.m} \underline{A}_l^{.n} = \frac{1}{2J^2\det\mathbf{A}}\varepsilon^{imn}\varepsilon^{jkl} \underline{A}^r_{.i} \underline{A}_{km} \underline{A}_{ln} , \tag{5.92a}$$

$$\delta^r_j = \frac{J^2}{2\det\mathbf{A}}\varepsilon_{imn}\varepsilon_{jkl} \underline{A}^{ri} \underline{A}^{km} \underline{A}^{ln} = \frac{1}{2\det\mathbf{A}}\varepsilon^{imn}\varepsilon_{jkl} \underline{A}^r_{.i} \underline{A}^k_{.m} \underline{A}^l_{.n} , \tag{5.92b}$$

$$\delta^j_r = \frac{1}{2\det\mathbf{A}}\varepsilon_{imn}\varepsilon^{jkl} \underline{A}_r^{.i} \underline{A}_k^{.m} \underline{A}_l^{.n} = \frac{1}{2J^2\det\mathbf{A}}\varepsilon^{imn}\varepsilon^{jkl} \underline{A}_{ri} \underline{A}_{km} \underline{A}_{ln} , \tag{5.92c}$$

$$g_{rj} = \frac{J^2}{2\det\mathbf{A}}\varepsilon_{imn}\varepsilon_{jkl} \underline{A}_r^{.i} \underline{A}^{km} \underline{A}^{ln} = \frac{1}{2\det\mathbf{A}}\varepsilon^{imn}\varepsilon_{jkl} \underline{A}_{ri} \underline{A}^k_{.m} \underline{A}^l_{.n} . \tag{5.92d}$$

5.5.2 Matrix Notation

In alignment with $(2.19)_2$ and (5.73a)–(5.73d), any tensor **A** in this text admits five various forms in matrix notation. Besides collecting the Cartesian components of **A** in a matrix according to $(1.47)_1$, one can similarly construct a

- matrix $[\,\mathbf{A}\,]^{\text{con}}$ consisting of its contravariant components.

▼ matrix $[\,\mathbf{A}\,]^{\text{con}}_{\cdot\,\text{cov}}$ consisting of its contra-covariant components.
▼ matrix $[\,\mathbf{A}\,]_{\text{cov}}^{\cdot\,\text{con}}$ consisting of its co-contravariant components.
▼ matrix $[\,\mathbf{A}\,]_{\text{cov}}$ consisting of its covariant components.

As an example, consider a tensor $\mathbf{A}$ with the following **covariant** matrix

$$[\,\mathbf{A}\,]_{\text{cov}} = \begin{bmatrix} 0 & 2 & 1 \\ -3 & 4 & 2 \\ -1 & 0 & 3 \end{bmatrix} ,$$

relative to the contravariant basis vectors

$$\underbrace{\mathbf{g}^1 = 7\,\widehat{\mathbf{e}}_1 - 3\,\widehat{\mathbf{e}}_2 - 3\,\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}^2 = -\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 \quad , \quad \mathbf{g}^3 = -\widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_3}_{\text{or } \mathbf{g}^1 = [\,7 \;\; -3 \;\; -3\,]^{\text{T}} \;\; , \;\; \mathbf{g}^2 = [\,-1 \;\; 1 \;\; 0\,]^{\text{T}} \;\; , \;\; \mathbf{g}^3 = [\,-1 \;\; 0 \;\; 1\,]^{\text{T}}} .$$

The given components and basis vectors help compute the Cartesian components of $\mathbf{A}$. In matrix notation, they render

$$[\,\mathbf{A}\,] = \begin{bmatrix} 16 & -2 & -10 \\ -18 & 7 & 8 \\ -1 & -3 & 3 \end{bmatrix} .$$

To construct other matrices expressing the given tensor, one needs to calculate the covariant basis vectors

$$\underbrace{\mathbf{g}_1 = \widehat{\mathbf{e}}_1 + \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}_2 = 3\,\widehat{\mathbf{e}}_1 + 4\,\widehat{\mathbf{e}}_2 + 3\,\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}_3 = 3\,\widehat{\mathbf{e}}_1 + 3\,\widehat{\mathbf{e}}_2 + 4\,\widehat{\mathbf{e}}_3}_{\text{or } \mathbf{g}_1 = [\,1 \;\; 1 \;\; 1\,]^{\text{T}} \;\; , \;\; \mathbf{g}_2 = [\,3 \;\; 4 \;\; 3\,]^{\text{T}} \;\; , \;\; \mathbf{g}_3 = [\,3 \;\; 3 \;\; 4\,]^{\text{T}}} .$$

It is then easy to see that

$$[\,\mathbf{A}\,]^{\text{con}} = \begin{bmatrix} 1570 & -195 & -272 \\ -319 & 43 & 52 \\ -155 & 16 & 30 \end{bmatrix} ,$$

and

$$[\,\mathbf{A}\,]^{\text{con}}_{\cdot\,\text{cov}} = \begin{bmatrix} 40 & 94 & 17 \\ -7 & -12 & -3 \\ -5 & -16 & -2 \end{bmatrix} \quad , \quad [\,\mathbf{A}\,]_{\text{cov}}^{\cdot\,\text{con}} = \begin{bmatrix} -30 & 5 & 4 \\ -261 & 40 & 38 \\ -97 & 13 & 16 \end{bmatrix} .$$

The simple contraction $\mathbf{v} = \mathbf{A}\mathbf{u}$, according to (5.84a)–(5.84d), in matrix notation takes the following forms

$$[\,\mathbf{v}\,]^{\text{con}} = [\,\mathbf{A}\,]^{\text{con}}\,[\,\mathbf{u}\,]_{\text{cov}} = [\,\mathbf{A}\,]^{\text{con}}_{\cdot\,\text{cov}}\,[\,\mathbf{u}\,]^{\text{con}} , \tag{5.93}$$

and

$$[\,\mathbf{v}\,]_{\text{cov}} = [\,\mathbf{A}\,]_{\text{cov}}\,[\,\mathbf{u}\,]^{\text{con}} = [\,\mathbf{A}\,]_{\text{cov}}^{\cdot\,\text{con}}\,[\,\mathbf{u}\,]_{\text{cov}}\ . \tag{5.94}$$

At the end, matrix forms of the simple contraction **AB** in, for instance, (5.85a) can be written as

$$\begin{aligned}[\,\mathbf{AB}\,]^{\text{con}} &= [\,\mathbf{A}\,]^{\text{con}}_{\cdot\,\text{cov}}\left[\,g^{ij}\,\right][\,\mathbf{B}\,]_{\text{cov}}^{\cdot\,\text{con}} = [\,\mathbf{A}\,]^{\text{con}}\left[\,g_{ij}\,\right][\,\mathbf{B}\,]^{\text{con}} \\ &= [\,\mathbf{A}\,]^{\text{con}}\,[\,\mathbf{B}\,]_{\text{cov}}^{\cdot\,\text{con}} = [\,\mathbf{A}\,]^{\text{con}}_{\cdot\,\text{cov}}\,[\,\mathbf{B}\,]^{\text{con}}\ .\end{aligned} \tag{5.95}$$

5.6 Contravariant, Mixed and Covariant Components of Higher-Order Tensors

Recall from (5.64a)–(5.64b) that a first-order tensor, characterized by one index, had 2 forms. As an extension, a second-order tensor possessing two indices then represented by 2^2 forms according to (5.73a)–(5.73d). Now, it should be clear that a tensor of order three admits 2^3 forms. And a tensor of order four can generally have 2^4 representations. But only some of them are important from the application point of view. This motivates to introduce the widely used forms of higher-order tensors in the literature.

A third-order tensor A, written with respect to the standard basis in (3.13), is now expressed as

$$\mathsf{A} = \underline{\mathsf{A}}^{ijk}\,\mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k\ , \tag{5.96a}$$

$$\mathsf{A} = \underline{\mathsf{A}}^{i\ \cdot\ k}_{\ \cdot\ j\ \cdot}\,\mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}_k\ , \tag{5.96b}$$

$$\mathsf{A} = \underline{\mathsf{A}}^{i\ \cdot\ \cdot}_{\ \cdot\ jk}\,\mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k\ , \tag{5.96c}$$

$$\mathsf{A} = \underline{\mathsf{A}}_{ijk}\,\mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k\ . \tag{5.96d}$$

By analogy with the procedure which led to (5.76a)–(5.76d), the introduced curvilinear components of A are related by

$$\underline{\mathsf{A}}^{ijk} = g^{jm}\,\underline{\mathsf{A}}^{i\ \cdot\ k}_{\ \cdot\ m\ \cdot} = g^{jm}g^{kn}\,\underline{\mathsf{A}}^{i\ \cdot\ \cdot}_{\ \cdot\ mn} = g^{il}g^{jm}g^{kn}\,\underline{\mathsf{A}}_{lmn}\ , \tag{5.97a}$$

$$\underline{\mathsf{A}}^{i\ \cdot\ k}_{\ \cdot\ j\ \cdot} = g_{jm}\,\underline{\mathsf{A}}^{imk} = g^{kn}\,\underline{\mathsf{A}}^{i\ \cdot\ \cdot}_{\ \cdot\ jn} = g^{il}g^{kn}\,\underline{\mathsf{A}}_{ljn}\ , \tag{5.97b}$$

$$\underline{\mathsf{A}}^{i\ \cdot\ \cdot}_{\ \cdot\ jk} = g_{jm}g_{kn}\,\underline{\mathsf{A}}^{imn} = g_{kn}\,\underline{\mathsf{A}}^{i\ \cdot\ n}_{\ \cdot\ j\ \cdot} = g^{il}\,\underline{\mathsf{A}}_{ljk}\ , \tag{5.97c}$$

$$\underline{\mathsf{A}}_{ijk} = g_{il}g_{jm}g_{kn}\,\underline{\mathsf{A}}^{lmn} = g_{il}g_{kn}\,\underline{\mathsf{A}}^{l\ \cdot\ n}_{\ \cdot\ j\ \cdot} = g_{il}\,\underline{\mathsf{A}}^{l\ \cdot\ \cdot}_{\ \cdot\ jk}\ . \tag{5.97d}$$

For subsequent developments, the permutation tensor (3.17) is expressed with respect to the curvilinear basis vectors by the following forms

$$\mathsf{E} = J^{-1}\varepsilon^{ijk}\mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k$$

$$= J \varepsilon_{ijk} \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \ . \quad \longleftarrow \text{ see (5.31) and (5.32)} \tag{5.98}$$

In a similar manner, any fourth-order tensor $\mathbb{A}$ with the Cartesian form (3.62) may be represented with respect to a basis consisting of the covariant and contravariant basis vectors. Therefore,[2]

$$\mathbb{A} = \underline{\mathbb{A}}^{ijkl} \, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k \otimes \mathbf{g}_l \ , \tag{5.99a}$$

$$\mathbb{A} = \underline{\mathbb{A}}^{ij\,.\,l}_{\ .\,.\,k\,.} \, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \otimes \mathbf{g}_l \ , \tag{5.99b}$$

$$\mathbb{A} = \underline{\mathbb{A}}^{ij\,.\,.}_{\ .\,.\,kl} \, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \ , \tag{5.99c}$$

$$\mathbb{A} = \underline{\mathbb{A}}^{i\,.\,.\,l}_{\ .\,jk\,.} \, \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}_l \ , \tag{5.99d}$$

$$\mathbb{A} = \underline{\mathbb{A}}_{ij\,.\,.}^{\ .\,.\,kl} \, \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}_k \otimes \mathbf{g}_l \ , \tag{5.99e}$$

$$\mathbb{A} = \underline{\mathbb{A}}_{ijk\,.}^{\ .\,.\,.\,l} \, \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}_l \ , \tag{5.99f}$$

$$\mathbb{A} = \underline{\mathbb{A}}_{ijkl} \, \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \ , \tag{5.99g}$$

$$\mathbb{A} = \underline{\mathbb{A}}^{i\,.\,.\,.}_{\ .\,jkl} \, \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \ . \tag{5.99h}$$

Guided by (5.76d) and (5.97a), the fully contravariant components $\underline{\mathbb{A}}^{ijkl}$ can be expressed in terms of the other curvilinear components as

$$\begin{aligned} \underline{\mathbb{A}}^{ijkl} &= g^{ko} \, \underline{\mathbb{A}}^{ij\,.\,l}_{\ .\,.\,o\,.} = g^{ko} g^{lp} \, \underline{\mathbb{A}}^{ij\,.\,.}_{\ .\,.\,op} = g^{jo} g^{kp} \, \underline{\mathbb{A}}^{i\,.\,.\,l}_{\ .\,op\,.} \\ &= g^{im} g^{jn} \, \underline{\mathbb{A}}_{mn\,.\,.}^{\ .\,.\,kl} = g^{im} g^{jn} g^{ko} \, \underline{\mathbb{A}}_{mno\,.}^{\ .\,.\,.\,l} \\ &= g^{im} g^{jn} g^{ko} g^{lp} \, \underline{\mathbb{A}}_{mnop} = g^{jn} g^{ko} g^{lp} \, \underline{\mathbb{A}}^{i\,.\,.\,.}_{\ .\,nop} \ . \end{aligned} \tag{5.100}$$

First, let $\mathbb{A}$ be a fully contravariant fourth-order tensor according to (5.99a). Further, let $\mathbf{C}$ be a tensor with various forms in (5.73a)–(5.73d). The right mapping $(3.66a)_4$ and subsequently the left mapping $(3.66b)_4$ can now be rewritten as

---

[2] As an application in nonlinear continuum mechanics, consider the deformation gradient tensor $\mathbf{F} = \underline{F}^{i}_{\ .A} \mathbf{g}_i \otimes \mathbf{G}^A$ which has been defined in this text by only one form, see Exercise 6.16. Accordingly, the right and left Cauchy–Green strain tensors will be $\mathbf{C} = \underline{C}_{AB} \mathbf{G}^A \otimes \mathbf{G}^B$ and $\mathbf{b} = \underline{b}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j$, respectively. Some appropriate derivatives of these strain tensors with respect to each other play an important role in characterizing the stress response of hyperelastic materials. The various forms of a fourth-order tensor introduced in this chapter are mainly in alignment with these partial derivatives. It follows that the sensitivity of

- ♦ $\mathbf{C}^{-1}$ with respect to $\mathbf{C}$, i.e. $\partial \mathbf{C}^{-1} / \partial \mathbf{C}$, is of the form (5.99a),
- ♦ $\mathbf{C}^{-1}$ with respect to $\mathbf{F}$, i.e. $\partial \mathbf{C}^{-1} / \partial \mathbf{F}$, is of the form (5.99b),
- ♦ $\mathbf{b}$ with respect to $\mathbf{b}$, i.e. $\partial \mathbf{b} / \partial \mathbf{b}$, is of the form (5.99c),
- ♦ $\mathbf{F}$ with respect to $\mathbf{F}$, i.e. $\partial \mathbf{F} / \partial \mathbf{F}$, is of the form (5.99d),
- ♦ $\mathbf{C}$ with respect to $\mathbf{C}$, i.e. $\partial \mathbf{C} / \partial \mathbf{C}$, is of the form (5.99e),
- ♦ $\mathbf{C}$ with respect to $\mathbf{F}$, i.e. $\partial \mathbf{C} / \partial \mathbf{F}$, is of the form (5.99f) and
- ♦ $\mathbf{b}^{-1}$ with respect to $\mathbf{b}$, i.e. $\partial \mathbf{b}^{-1} / \partial \mathbf{b}$, is of the form (5.99g).

Note that a tensor of the form (5.99h) is greatly used in differential geometry, see (7.50) and (9.199).

$$\begin{aligned}
\mathbb{A} : \mathbf{C} = \ & \underline{\mathbb{A}}^{ijkl} g_{km} g_{ln}\, \underline{C}^{mn}\, \mathbf{g}_i \otimes \mathbf{g}_j = \underline{\mathbb{A}}^{ijkl} g_{km}\, \underline{C}^{m}_{\ .l}\, \mathbf{g}_i \otimes \mathbf{g}_j \\
= \ & \underline{\mathbb{A}}^{ijkl} g_{ln}\, \underline{C}_{k}^{\ .n}\, \mathbf{g}_i \otimes \mathbf{g}_j = \underline{\mathbb{A}}^{ijkl}\, \underline{C}_{kl}\, \mathbf{g}_i \otimes \mathbf{g}_j \ ,
\end{aligned} \tag{5.101a}$$

$$\begin{aligned}
\mathbf{C} : \mathbb{A} = \ & \underline{C}^{mn} g_{mk} g_{nl}\, \underline{\mathbb{A}}^{klij}\, \mathbf{g}_i \otimes \mathbf{g}_j = \underline{C}^{m}_{\ .l} g_{mk}\, \underline{\mathbb{A}}^{klij}\, \mathbf{g}_i \otimes \mathbf{g}_j \\
= \ & \underline{C}_{k}^{\ .n} g_{nl}\, \underline{\mathbb{A}}^{klij}\, \mathbf{g}_i \otimes \mathbf{g}_j = \underline{C}_{kl}\, \underline{\mathbb{A}}^{klij}\, \mathbf{g}_i \otimes \mathbf{g}_j \ .
\end{aligned} \tag{5.101b}$$

Then, let $\mathbb{A}$ be a fully contravariant fourth-order tensor with $\underline{\mathbb{A}}^{ijmn}$ and $\mathbb{B}$ be a fully covariant one with $\underline{\mathbb{B}}_{opkl}$. Accordingly, the composition $\mathbb{A} : \mathbb{B}$, in light of $(3.91)_4$, renders

$$\mathbb{A} : \mathbb{B} = \underline{\mathbb{A}}^{ijmn}\, \underline{\mathbb{A}}_{mnkl}\, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \ . \tag{5.102}$$

Finally, let $\mathbf{D}$ be a contra-covariant tensor and $\mathbf{E}$ be a co-contravariant tensor. Further, let $\mathbb{A}$ be a tensor of rank four according to (5.99a)–(5.99h). Then, the scalar $\mathbf{D} : \mathbb{A} : \mathbf{E}$ with the Cartesian form $D_{ij} \mathbb{A}_{ijkl} E_{kl}$ can now be demonstrated as

$$\begin{aligned}
\mathbf{D} : \mathbb{A} : \mathbf{E} = \ & \underline{D}^{m}_{\ .j} g_{mi}\, \underline{\mathbb{A}}^{ijkl}\, g_{lp}\, \underline{E}_{k}^{\ .p} = \underline{D}^{m}_{\ .j} g_{mi}\, \underline{\mathbb{A}}^{ij\ .\ l}_{\ .\ .\ k\ .}\, g^{ko} g_{lp}\, \underline{E}_{o}^{\ .p} \\
= \ & \underline{D}^{m}_{\ .j} g_{mi}\, \underline{\mathbb{A}}^{ij\ .\ .}_{\ .\ .\ kl}\, g^{ko}\, \underline{E}_{o}^{\ .l} = \underline{D}^{m}_{\ .n} g_{mi} g^{nj}\, \underline{\mathbb{A}}^{i\ .\ .\ l}_{\ .\ jk\ .}\, g^{ko} g_{lp}\, \underline{E}_{o}^{\ .p} \\
= \ & \underline{D}^{i}_{\ .n} g^{nj}\, \underline{\mathbb{A}}_{ij\ .\ .}^{\ .\ .\ kl}\, g_{lp}\, \underline{E}_{k}^{\ .p} = \underline{D}^{i}_{\ .n} g^{nj}\, \underline{\mathbb{A}}_{ijk\ .}^{\ .\ .\ .\ l}\, g^{ko} g_{lp}\, \underline{E}_{o}^{\ .p} \\
= \ & \underline{D}^{i}_{\ .n} g^{nj}\, \underline{\mathbb{A}}_{ijkl}\, g^{ko}\, \underline{E}_{o}^{\ .l} = \underline{D}^{m}_{\ .n} g_{mi} g^{nj}\, \underline{\mathbb{A}}^{i\ .\ .\ .}_{\ .\ jkl}\, g^{ko}\, \underline{E}_{o}^{\ .l} \ .
\end{aligned} \tag{5.103}$$

5.7 Tensor Property of Components

To begin with, consider a vector $\mathbf{u}$ decomposed according to (5.64a)–(5.64b). Similarly to scalar variables, this object is basically an **invariant**. And this means that it remains unchanged under an arbitrary change of coordinates from $\left(\Theta^1, \Theta^2, \Theta^3\right)$ to $\left(\bar{\Theta}^1, \bar{\Theta}^2, \bar{\Theta}^3\right)$, that is,

$$\begin{aligned}
\mathbf{u} = \ & \underline{\bar{u}}^{i}\, \bar{\mathbf{g}}_i = \underline{\bar{u}}_{i}\, \bar{\mathbf{g}}^i \\
= \ & \underline{u}^{i}\, \mathbf{g}_i = \underline{u}_{i}\, \mathbf{g}^i \ .
\end{aligned} \tag{5.104}$$

From $(5.59)_{1-2}$, $(5.60)_{1-2}$ and $(5.104)_{1-4}$, considering the fact that the decomposition of a tensor with respect to a basis is unique, one can establish the following **transformation laws** for the vector components

$$\underline{\bar{u}}^{i} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\, \underline{u}^{j} \ , \tag{5.105a}$$

$$\underline{\bar{u}}_{i} = \frac{\partial \Theta^j}{\partial \bar{\Theta}^i}\, \underline{u}_{j} \ . \tag{5.105b}$$

The variant $\underline{u}^i$ ($\underline{u}_i$) thus represents a first-order contravariant (covariant) tensor. Next, consider a tensor $\mathbf{A}$ which admits the representations (5.73a)–(5.73d). After some algebraic manipulations, one can infer that the old and new components of this invariant object are **tensorially** related by

$$\bar{\underline{A}}^{ij} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \underline{A}^{kl} \frac{\partial \bar{\Theta}^j}{\partial \Theta^l} , \tag{5.106a}$$

$$\bar{\underline{A}}^i_{\ .j} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \underline{A}^k_{\ .l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} , \tag{5.106b}$$

$$\bar{\underline{A}}_i^{\ .j} = \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \underline{A}_k^{\ .l} \frac{\partial \bar{\Theta}^j}{\partial \Theta^l} , \tag{5.106c}$$

$$\bar{\underline{A}}_{ij} = \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \underline{A}_{kl} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} . \tag{5.106d}$$

Here, the object $\underline{A}^{ij}$ ($\underline{A}^i_{\ .j}$) represents a second-order contravariant (contra-covariant) tensor. And the variant $\underline{A}_i^{\ .j}$ ($\underline{A}_{ij}$) demonstrates a second-order co-contravariant (covariant) tensor. The procedure to extend the rules (5.106a)–(5.106d) to tensors of higher ranks should be clear now. This is left as an exercise to be undertaken by the interested reader.

5.8 Line, Surface and Volume Elements

Consider an arbitrary curvilinear coordinate system embedded in a Cartesian coordinate frame as illustrated in Fig. 5.3. Consider also two infinitesimally close points $\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$ corresponding to $\left(\Theta^1, \Theta^2, \Theta^3\right)$ and $\left(\Theta^1 + d\Theta^1, \Theta^2 + d\Theta^2, \Theta^3 + d\Theta^3\right)$, respectively. Here, the projection of the position increment vector $d\mathbf{x}$ along the tangent vector $\mathbf{g}_i$ is denoted by $d\mathbf{x}^{(i)}$. As can be seen, the linearly independent vectors $d\mathbf{x}^{(i)}$, $i = 1, 2, 3$, form a parallelepiped in space. In the following, the goal is to represent the differential line element $d\mathbf{x}$, the infinitesimal surface elements $d\mathbf{A}^{(i)}$, $i = 1, 2, 3$, and the differential volume element dV of the parallelepiped in terms of the curvilinear coordinate increments and basis vectors. They are essentially needed for the evaluation of line, surface and volume integrals in a curvilinear coordinate system.

First, with the aid of $(5.3)_1$, (5.14), $(5.27)_2$ and (5.64a), the infinitesimal line element $d\mathbf{x}$ can be expressed as

$$\boxed{\underbrace{d\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \Theta^i} d\Theta^i = d\Theta^i \mathbf{g}_i}_{\text{indeed, } d\mathbf{x}^{(1)} = d\Theta^1 \mathbf{g}_1 \text{ , } d\mathbf{x}^{(2)} = d\Theta^2 \mathbf{g}_2 \text{ , } d\mathbf{x}^{(3)} = d\Theta^3 \mathbf{g}_3} \quad \text{with} \quad d\Theta^i = d\mathbf{x} \cdot \mathbf{g}^i} \ . \tag{5.107}$$

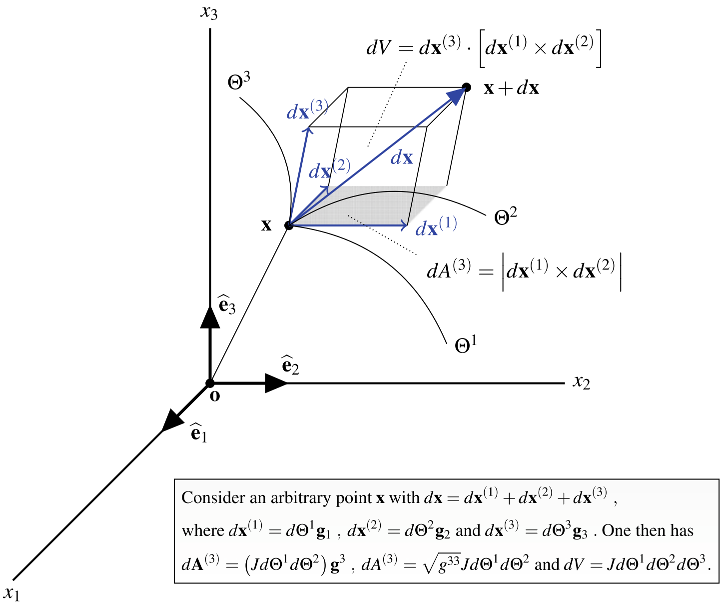


Fig. 5.3 Projection of a position increment vector in an arbitrary curvilinear coordinate system

Note that the curvilinear coordinate increments are the contravariant components of the differential position vector, i.e. $(d\mathbf{x})^i = d\Theta^i$. But, it is not possible to have $\mathbf{x} = \Theta^i \mathbf{g}_i$. It is also important to note that the covariant components of $d\mathbf{x}$ are well-defined and obtained by

$$(d\mathbf{x})_i = g_{ij} \, (d\mathbf{x})^j \quad . \tag{5.108}$$

But,

$$(d\mathbf{x})_i \neq d\Theta_i \quad \text{or} \quad d\Theta_i \neq g_{ij} d\Theta^j \ ,$$

since, it can be shown that, dual coordinates do not exist.

The square of the magnitude of $d\mathbf{x}$ according to

$$ds^2 = d\mathbf{x} \cdot d\mathbf{x} \xrightarrow[\text{(5.107)}]{\text{from}} \left(d\Theta^i \mathbf{g}_i \right) \cdot \left(d\Theta^j \mathbf{g}_j \right) \xrightarrow[\text{(1.9a)-(1.9c) and (5.38)}]{\text{from}} g_{ij} d\Theta^i d\Theta^j \ , \tag{5.109}$$

is referred to as *fundamental differential quadratic form*.

Next, the area elements $d\mathbf{A}^{(i)}, \ i = 1, 2, 3$, can be decomposed as

$$d\mathbf{A}^{(1)} = \underbrace{d\mathbf{x}^{(2)} \times d\mathbf{x}^{(3)}}_{= J d\Theta^2 d\Theta^3 \mathbf{g}^1} \ , \ d\mathbf{A}^{(2)} = \underbrace{d\mathbf{x}^{(3)} \times d\mathbf{x}^{(1)}}_{= J d\Theta^3 d\Theta^1 \mathbf{g}^2} \ , \ d\mathbf{A}^{(3)} = \underbrace{d\mathbf{x}^{(1)} \times d\mathbf{x}^{(2)}}_{= J d\Theta^1 d\Theta^2 \mathbf{g}^3} \ , \tag{5.110}$$

or, collectively as,

$$\boxed{d\mathbf{A}^{(i)} = d\mathbf{x}^{(j)} \times d\mathbf{x}^{(k)} = J d\Theta^j d\Theta^k \mathbf{g}^i \quad \text{where } ijk = 123, 231, 312 \, .} \tag{5.111}$$

The length of $d\mathbf{A}^{(i)}$ then becomes

$$dA^{(i)} = \sqrt{g^{ii}} J d\Theta^j d\Theta^k \quad \text{where } ijk = 123, 231, 312 \ \ (\text{no sum on } i) \, . \tag{5.112}$$

At the end, the infinitesimal volume element dV takes the following form

$$\boxed{dV = d\mathbf{x}^{(i)} \cdot \left[d\mathbf{x}^{(j)} \times d\mathbf{x}^{(k)} \right] = J d\Theta^1 d\Theta^2 d\Theta^3 \quad \text{where } ijk = 123, 231, 312 \, .} \tag{5.113}$$

Consider a parallelepiped defined by the three vectors **u**, **v** and **w**. The so-called *volume form* is sometimes defined to calculate the volume of that object:

$$\boxed{\omega \left(\mathbf{u}, \mathbf{v}, \mathbf{w} \right) = J \varepsilon_{ijk} \, \underline{u}^i \, \underline{v}^j \, \underline{w}^k \, .} \tag{5.114}$$

For the sake of completeness, the Cartesian form of the introduced differential elements are represented in the following:

$$d\mathbf{x} = \underbrace{\frac{\partial \mathbf{x}}{\partial x_i} dx_i = dx_i \widehat{\mathbf{e}}_i \quad \text{with} \quad dx_i = d\mathbf{x} \cdot \widehat{\mathbf{e}}_i}_{\text{indeed, } d\mathbf{x}^{(1)} = dx_1 \widehat{\mathbf{e}}_1 \, , \ d\mathbf{x}^{(2)} = dx_2 \widehat{\mathbf{e}}_2 \, , \ d\mathbf{x}^{(3)} = dx_3 \widehat{\mathbf{e}}_3} \quad , \tag{5.115a}$$

$$d\mathbf{A}^{(i)} = d\mathbf{x}^{(j)} \times d\mathbf{x}^{(k)} = dx_j dx_k \widehat{\mathbf{e}}_i \quad \text{where } ijk = 123, 231, 312 \, , \tag{5.115b}$$

$$dV = d\mathbf{x}^{(i)} \cdot \left[d\mathbf{x}^{(j)} \times d\mathbf{x}^{(k)} \right] = dx_1 dx_2 dx_3 \quad \text{where } ijk = 123, 231, 312 \, . \tag{5.115c}$$

Hint: An arbitrary surface element $d\mathbf{A}$ in three-dimensional space is characterized by its magnitude dA and its unit normal vector $\widehat{\mathbf{n}}$. And it will be referred to as the *surface vector*, i.e. $d\mathbf{A} = \widehat{\mathbf{n}} \, dA$. Such an expression is widely used in **integral theorems**. But, technically, it should be modified to $d\mathbf{A} = \pm \widehat{\mathbf{n}} \, dA$ since a surface element can geometrically admit either the unit normal vector $+\widehat{\mathbf{n}}$ or its additive inverse $-\widehat{\mathbf{n}}$. Accordingly, the appropriate sign is chosen by a **convention** distinguishing between a *closed surface* (used in the divergence theorem) and an *open surface* (utilized in the Stokes' theorem). In Chap. 8, it will be shown that how either $d\mathbf{A} = +\widehat{\mathbf{n}} \, dA$ or $d\mathbf{A} = -\widehat{\mathbf{n}} \, dA$ is determined for such surfaces.

5.9 Exercises

Exercise 5.1

The **covariant** basis vectors of the cylindrical, spherical and Cartesian coordinate systems are already given in (5.7a)–(5.7c), (5.11a)–(5.11c) and (5.12), respectively. Obtain the corresponding **contravariant** basis vectors for these commonly used coordinate systems.

Solution. Recall that the given covariant basis vectors were expressed in terms of the three mutually orthogonal unit vectors $\widehat{\mathbf{e}}_x$, $\widehat{\mathbf{e}}_y$ and $\widehat{\mathbf{e}}_z$. The procedure to attain the companion contravariant basis vectors mainly relies on the scalar product (1.8) which is a symmetric bilinear form according to (1.9a)–(1.9c).

Consider first the cylindrical coordinates (r, θ, z). The covariant metric coefficients in $(5.38)_3$, i.e. $\left[\, g_{ij} \,\right] = \left[\, \mathbf{g}_i \cdot \mathbf{g}_j \,\right]$, then render

$$\underbrace{\left[\, g_{ij} \,\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{with } J = \sqrt{\det g_{ij}} = r} , \quad \longleftarrow \text{cylindrical coordinates} \tag{5.116}$$

Accordingly, the contravariant metric coefficients $(5.51)_4$, i.e. $\left[\, g^{ij} \,\right] = \left[\, g_{ij} \,\right]^{-1}$, take the form

$$\left[\, g^{ij} \,\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & 1 \end{bmatrix} . \quad \longleftarrow \text{cylindrical coordinates} \tag{5.117}$$

Now, the contravariant basis vectors $\mathbf{g}^i = g^{ij}\mathbf{g}_j$, according to (5.49), will be

$$\mathbf{g}^1 = \mathbf{g}_1 = \widehat{\mathbf{e}}_r \qquad \text{where} \qquad \widehat{\mathbf{e}}_r = \cos\theta\,\widehat{\mathbf{e}}_x + \sin\theta\,\widehat{\mathbf{e}}_y \ , \tag{5.118a}$$

$$\mathbf{g}^2 = \frac{\mathbf{g}_2}{r^2} = \frac{\widehat{\mathbf{e}}_\theta}{r} \qquad \text{where} \qquad \widehat{\mathbf{e}}_\theta = -\sin\theta\,\widehat{\mathbf{e}}_x + \cos\theta\,\widehat{\mathbf{e}}_y \ , \tag{5.118b}$$

$$\mathbf{g}^3 = \mathbf{g}_3 = \widehat{\mathbf{e}}_z \ . \tag{5.118c}$$

Next, the spherical coordinates (r, θ, ϕ) are considered. In this case, the covariant metric coefficients become

$$\underbrace{\left[\, g_{ij} \,\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{bmatrix}}_{\text{with } J = \sqrt{\det g_{ij}} = r^2 \sin\theta} . \quad \longleftarrow \text{spherical coordinates} \tag{5.119}$$

Subsequently, the contravariant metric coefficients

$$\left[g^{ij} \right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta} \end{bmatrix} , \quad \longleftarrow \text{spherical coordinates} \tag{5.120}$$

help provide the dual basis vectors

$$\begin{aligned} \mathbf{g}^1 &= \mathbf{g}_1 \\ &= \widehat{\mathbf{e}}_r \qquad \text{where} \qquad \widehat{\mathbf{e}}_r = \sin\theta\cos\phi\,\widehat{\mathbf{e}}_x + \sin\theta\sin\phi\,\widehat{\mathbf{e}}_y + \cos\theta\,\widehat{\mathbf{e}}_z \ , \end{aligned} \tag{5.121a}$$

$$\begin{aligned} \mathbf{g}^2 &= \frac{\mathbf{g}_2}{r^2} \\ &= \frac{\widehat{\mathbf{e}}_\theta}{r} \qquad \text{where} \qquad \widehat{\mathbf{e}}_\theta = \cos\theta\cos\phi\,\widehat{\mathbf{e}}_x + \cos\theta\sin\phi\,\widehat{\mathbf{e}}_y - \sin\theta\,\widehat{\mathbf{e}}_z \ , \end{aligned} \tag{5.121b}$$

$$\begin{aligned} \mathbf{g}^3 &= \frac{\mathbf{g}_3}{r^2 \sin^2\theta} \\ &= \frac{\widehat{\mathbf{e}}_\phi}{r\sin\theta} \qquad \text{where} \qquad \widehat{\mathbf{e}}_\phi = -\sin\phi\,\widehat{\mathbf{e}}_x + \cos\phi\,\widehat{\mathbf{e}}_y \ . \end{aligned} \tag{5.121c}$$

Finally, the reciprocal basis $\left\{ \mathbf{g}^i \right\}$ will be represented for the Cartesian coordinates (x, y, z). The contravariant metric coefficients for such coordinates trivially render $\left[g^{ij} \right] = \left[\delta^{ij} \right]$. Thus,

$$\mathbf{g}^1 = \mathbf{g}_1 = \widehat{\mathbf{e}}_x \quad , \quad \mathbf{g}^2 = \mathbf{g}_2 = \widehat{\mathbf{e}}_y \quad , \quad \mathbf{g}^3 = \mathbf{g}_3 = \widehat{\mathbf{e}}_z \ . \tag{5.122}$$

Exercise 5.2

Verify (5.84a)$_2$, (5.85b)$_2$, (5.86c), (5.87)$_4$, (5.88)$_4$, (5.89a)$_1$, (5.90b)$_2$ and (5.91b)$_2$.

Solution. It is strongly recommended that the interested reader prove all the relations presented in (5.84a)–(5.91d). Here, only some important expressions are chosen for verification. By use of the curvilinear representations of vectors and tensors, according to (5.64a)–(5.64b) and (5.73a)–(5.73d), the derivation of each desired relation will be shown step by step in the following.

The expression (5.84a)$_2$*:* Let **u** be a covariant vector and **A** be a contravariant tensor. Further, let $\mathbf{v} = \mathbf{Au}$. Then,

$$\mathbf{v} \xrightarrow[\text{(5.64b) and (5.73a)}]{\text{from}} \left(\underline{A}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j \right) \left(\underline{u}_k \mathbf{g}^k \right)$$

$$\underset{(2.2),\,(2.8a),\,(2.8b)\text{ and }(2.13)}{\overset{\text{from}}{=\!=\!=}} \underline{A}^{ij}\,\underline{u}_k\left(\mathbf{g}_j\cdot\mathbf{g}^k\right)\mathbf{g}_i$$
$$\underset{(5.27)}{\overset{\text{from}}{=\!=\!=}} \underline{A}^{ij}\,\underline{u}_k\delta^k_j\mathbf{g}_i$$
$$\underset{(5.14)}{\overset{\text{from}}{=\!=\!=}} \underline{A}^{ij}\,\underline{u}_j\mathbf{g}_i\ .$$

Consistent with this result, the vector $\mathbf{v}$, by means of (5.64a), can be decomposed as $\mathbf{v} = \underline{v}^i\mathbf{g}_i$. Since any arbitrary vector has an **unique** representation with respect to a given basis, it follows that $\underline{v}^i = \underline{A}^{ij}\,\underline{u}_j$.

The expression $(5.85\text{b})_2$*:* Let $\mathbf{A}$ be a tensor that is known in its contravariant components. Further, let $\mathbf{B}$ be a contra-covariant tensor. Then,

$$(\mathbf{AB})^i_{.j} \underset{(5.74j)}{\overset{\text{from}}{=\!=\!=}} \mathbf{g}^i\cdot\left[(\mathbf{AB})\,\mathbf{g}_j\right]$$
$$\underset{(2.25)}{\overset{\text{from}}{=\!=\!=}} \mathbf{g}^i\cdot\left[\mathbf{A}\left(\mathbf{B}\mathbf{g}_j\right)\right]$$
$$\underset{(5.75b)}{\overset{\text{from}}{=\!=\!=}} \mathbf{g}^i\cdot\left[\mathbf{A}\left(\underline{B}^n_{.j}\mathbf{g}_n\right)\right]$$
$$\underset{(2.2)}{\overset{\text{from}}{=\!=\!=}} \mathbf{g}^i\cdot\left[\underline{B}^n_{.j}\left(\mathbf{A}\mathbf{g}_n\right)\right]$$
$$\underset{(1.9a)\text{ to }(1.9c)}{\overset{\text{from}}{=\!=\!=}} \mathbf{g}^i\cdot\left[(\mathbf{A}\mathbf{g}_n)\right]\underline{B}^n_{.j}$$
$$\underset{(5.74i)}{\overset{\text{from}}{=\!=\!=}} \underline{A}^{im}g_{mn}\,\underline{B}^n_{.j}\ .$$

The expression (5.86c)*:* Suppose one is given a contra-covariant tensor $\mathbf{A}$. Then,

$$\mathbf{A}^{\mathrm{T}} \underset{(5.73b)}{\overset{\text{from}}{=\!=\!=}} \left(\underline{A}^i_{.j}\mathbf{g}_i\otimes\mathbf{g}^j\right)^{\mathrm{T}}$$
$$\underset{(2.52)\text{ to }(2.54)}{\overset{\text{from}}{=\!=\!=}} \underline{A}^i_{.j}\mathbf{g}^j\otimes\mathbf{g}_i$$
$$\underset{\text{names of the indices}}{\overset{\text{by interchanging the}}{=\!=\!=}} \underline{A}^j_{.i}\mathbf{g}^i\otimes\mathbf{g}_j\ .$$

Consistent with this result, consider $\mathbf{A}^{\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)_i^{.j}\mathbf{g}^i\otimes\mathbf{g}_j$ according to (5.73c). One can thus deduce the desired relation $\left(\mathbf{A}^{\mathrm{T}}\right)_i^{.j} = (\mathbf{A})^j_{.i}$.

The expression $(5.87)_4$*:* Let $\mathbf{A}$ be a tensor that is known in its contravariant components. Further, let $\mathbf{B}$ be a tensor that is known in its covariant components. Then,

$$\mathbf{A}:\mathbf{B} \underset{(5.73a)\text{ and }(5.73d)}{\overset{\text{from}}{=\!=\!=}} \left(\underline{A}^{ij}\mathbf{g}_i\otimes\mathbf{g}_j\right):\left(\underline{B}_{kl}\mathbf{g}^k\otimes\mathbf{g}^l\right)$$
$$\underset{(2.73),\,(2.74a)\text{ and }(2.74b)}{\overset{\text{from}}{=\!=\!=}} \underline{A}^{ij}\,\underline{B}_{kl}\left(\mathbf{g}_i\cdot\mathbf{g}^k\right)\left(\mathbf{g}_j\cdot\mathbf{g}^l\right)$$

$$\xrightarrow[(5.27)]{\text{from}} \underline{A}^{ij}\, \underline{B}_{kl} \delta_i^k \delta_j^l$$
$$\xrightarrow[(5.14)]{\text{from}} \underline{A}^{ij}\, \underline{B}_{ij} \ .$$

The expression (5.88)$_4$*:* Suppose one is given a covariant tensor $\mathbf{A}$. Then,

$$\begin{aligned}
\text{tr}\mathbf{A} &\xrightarrow[(2.83)]{\text{from}} \mathbf{I} : \mathbf{A} \\
&\xrightarrow[(5.73\text{d})]{\text{from}} \mathbf{I} : \left(\underline{A}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j \right) \\
&\xrightarrow[(2.74\text{a}),\ (2.74\text{b})\text{ and }(2.79\text{c})]{\text{from}} \left(\mathbf{g}^i \cdot \mathbf{I} \mathbf{g}^j \right) \underline{A}_{ij} \\
&\xrightarrow[(2.5)\text{ and }(5.46)]{\text{from}} g^{ij}\, \underline{A}_{ij} \ .
\end{aligned}$$

The expression (5.89a)$_1$*:* The goal here is to derive the determinant of a tensor $\mathbf{A}$ assuming that $\mathbf{A}$ is known in its contravariant components. To do so, let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be three arbitrary covariant vectors according to (5.64b), i.e. $\mathbf{u} = \underline{u}_l \mathbf{g}^l$, $\mathbf{v} = \underline{v}_m \mathbf{g}^m$ and $\mathbf{w} = \underline{w}_n \mathbf{g}^n$. Consistent with this, let $\mathbf{Au}$, $\mathbf{Av}$ and $\mathbf{Aw}$ be three linear transformations of the form (5.84a), i.e. $\mathbf{Au} = \underline{A}^{il}\, \underline{u}_l \mathbf{g}_i$, $\mathbf{Av} = \underline{A}^{jm}\, \underline{v}_m \mathbf{g}_j$ and $\mathbf{Aw} = \underline{A}^{kn}\, \underline{w}_n \mathbf{g}_k$. Having in mind that the scalar product and cross product are bilinear operators, it follows that

$$\begin{aligned}
&\xrightarrow[(2.98)]{\text{from}} \underline{u}_l\, \underline{v}_m\, \underline{w}_n \mathbf{g}^l \cdot \left(\mathbf{g}^m \times \mathbf{g}^n \right) \det \mathbf{A} = \underline{u}_l\, \underline{v}_m\, \underline{w}_n \mathbf{g}_i \cdot \left(\mathbf{g}_j \times \mathbf{g}_k \right) \underline{A}^{il}\, \underline{A}^{jm}\, \underline{A}^{kn} \\
&\xrightarrow[\text{of vectors}]{\text{arbitrariness}} \mathbf{g}^l \cdot \left(\mathbf{g}^m \times \mathbf{g}^n \right) \det \mathbf{A} = \mathbf{g}_i \cdot \left(\mathbf{g}_j \times \mathbf{g}_k \right) \underline{A}^{il}\, \underline{A}^{jm}\, \underline{A}^{kn} \\
&\xrightarrow[(5.31)\text{ and }(5.32)]{\text{from}} J^{-1} \varepsilon^{lmn} \det \mathbf{A} = J \varepsilon_{ijk}\, \underline{A}^{il}\, \underline{A}^{jm}\, \underline{A}^{kn} \\
&\xrightarrow[(1.58\text{c})]{\text{from}} 6 \det \mathbf{A} = J^2 \varepsilon_{ijk} \varepsilon_{lmn}\, \underline{A}^{il}\, \underline{A}^{jm}\, \underline{A}^{kn} \ .
\end{aligned}$$

The expression (5.90b)$_2$*:* Suppose one is given a co-contravariant tensor $\mathbf{A}$. The goal here is to derive the desired relation using the basic definition (2.112), i.e. $\mathbf{A}^{\text{c}} (\mathbf{u} \times \mathbf{v}) = (\mathbf{Au}) \times (\mathbf{Av})$.

Let $\mathbf{u}$ and $\mathbf{v}$ be two covariant vectors and consider two linear transformations of the form $\mathbf{Au} = \underline{A}_k^{\ .m}\, \underline{u}_m \mathbf{g}^k$ and $\mathbf{Av} = \underline{A}_l^{\ .n}\, \underline{v}_n \mathbf{g}^l$. Then,

$$\begin{aligned}
(\mathbf{Au}) \times (\mathbf{Av}) &= \left(\underline{A}_k^{\ .m}\, \underline{u}_m \mathbf{g}^k \right) \times \left(\underline{A}_l^{\ .n}\, \underline{v}_n \mathbf{g}^l \right) \\
&\xrightarrow[(1.49\text{a})\text{-}(1.49\text{b})]{\text{from}} \underline{A}_k^{\ .m}\, \underline{A}_l^{\ .n}\, \underline{u}_m\, \underline{v}_n \mathbf{g}^k \times \mathbf{g}^l \\
&\xrightarrow[(5.35)]{\text{from}} J^{-1}\, \underline{A}_k^{\ .m}\, \underline{A}_l^{\ .n}\, \underline{u}_m\, \underline{v}_n \varepsilon^{klp} \mathbf{g}_p \ .
\end{aligned}$$

On the other hand, by using (5.69b)$_4$, one can write $\mathbf{u} \times \mathbf{v} = J^{-1} \underline{u}_m \underline{v}_n \varepsilon^{mno} \mathbf{g}_o$. Consistent with the above result, mapping of this vector by $\mathbf{A}^{\mathrm{c}}$, using (5.84b)$_{2-3}$, then gives

$$\mathbf{A}^{\mathrm{c}} (\mathbf{u} \times \mathbf{v}) = J^{-1} \underline{u}_m \underline{v}_n \varepsilon^{mno} (\mathbf{A}^{\mathrm{c}})^{p}_{.o} \mathbf{g}_p \ .$$

Considering the fact that $\mathbf{u}$ and $\mathbf{v}$ are arbitrary vectors now implies that

$$\varepsilon^{mno} (\mathbf{A}^{\mathrm{c}})^{p}_{.o} \mathbf{g}_p = \underline{A}_k^{\ .m} \underline{A}_l^{\ .n} \varepsilon^{klp} \mathbf{g}_p \ .$$

By multiplying both side of this equation by $\mathbf{g}^i$, taking into account (5.14) and (5.27)$_2$, one will have

$$\varepsilon^{mno} (\mathbf{A}^{\mathrm{c}})^{i}_{.o} = \underline{A}_k^{\ .m} \underline{A}_l^{\ .n} \varepsilon^{kli} \quad \text{or, using (1.54),} \quad \varepsilon^{omn} (\mathbf{A}^{\mathrm{c}})^{i}_{.o} = \underline{A}_k^{\ .m} \underline{A}_l^{\ .n} \varepsilon^{ikl} \ .$$

Multiplying both sides of the above result with ε_{jmn}, using (1.58b)$_3$ and (5.14), finally yields

$$2\delta^o_j (\mathbf{A}^{\mathrm{c}})^{i}_{.o} = \varepsilon^{ikl} \varepsilon_{jmn} \underline{A}_k^{\ .m} \underline{A}_l^{\ .n} \quad \text{or} \quad \boxed{2 (\mathbf{A}^{\mathrm{c}})^{i}_{.j} = \varepsilon^{ikl} \varepsilon_{jmn} \underline{A}_k^{\ .m} \underline{A}_l^{\ .n} \ .}$$

The expression (5.91b)$_2$*:* Suppose one is given a contra-covariant tensor $\mathbf{A}$. Then,

$$\begin{aligned}
\mathbf{A}^{-1} &\xlongequal[(2.114)]{\text{from}} \frac{1}{\det \mathbf{A}} \mathbf{A}^{\mathrm{c}\,\mathrm{T}} \\
&\xlongequal[(5.73b)]{\text{from}} \frac{1}{\det \mathbf{A}} \left(\mathbf{A}^{\mathrm{c}\,\mathrm{T}}\right)^{i}_{.j} \mathbf{g}_i \otimes \mathbf{g}^j \\
&\xlongequal[(5.86b)]{\text{from}} \frac{1}{\det \mathbf{A}} (\mathbf{A}^{\mathrm{c}})_j^{\ .i} \mathbf{g}_i \otimes \mathbf{g}^j \\
&\xlongequal[(5.90c)]{\text{from}} \frac{1}{2 \det \mathbf{A}} \varepsilon^{imn} \varepsilon_{jkl} \underline{A}^{k}_{\ .m} \underline{A}^{l}_{\ .n} \mathbf{g}_i \otimes \mathbf{g}^j \ .
\end{aligned}$$

This result eventually represents the contra-covariant form of the tensor $\mathbf{A}^{-1}$ with $\left(\mathbf{A}^{-1}\right)^{i}_{.j} = (2 \det \mathbf{A})^{-1} \varepsilon^{imn} \varepsilon_{jkl} \underline{A}^{k}_{\ .m} \underline{A}^{l}_{\ .n}$.

Exercise 5.3

Consider an orthogonal tensor $\mathbf{Q}$ with the following two forms

$$\mathbf{Q} = \underline{Q}^{i}_{\ .j} \mathbf{g}_i \otimes \mathbf{g}^j = \underline{Q}_i^{\ .j} \mathbf{g}^i \otimes \mathbf{g}_j \ .$$

1. Demonstrate that the orthogonality condition $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ in index notation can be written as

$$\underline{Q}^{i}_{.k}\, \underline{Q}^{.k}_{j} = \delta^i_j \quad , \quad \underline{Q}^{.k}_{i}\, g_{kl}\, \underline{Q}^{.l}_{j} = g_{ij} \ .$$

Solution. Having in mind (5.73b) and $(5.78)_2$, consider first the contra-covariant form of the three tensors $\mathbf{Q}$, $\mathbf{Q}^{\mathrm{T}}$ and $\mathbf{I}$:

$$\mathbf{Q} = (\mathbf{Q})^{m}_{.n}\, \mathbf{g}_m \otimes \mathbf{g}^n \quad , \quad \mathbf{Q}^{\mathrm{T}} = \left(\mathbf{Q}^{\mathrm{T}}\right)^{k}_{.l} \mathbf{g}_k \otimes \mathbf{g}^l \quad , \quad \mathbf{I} = \mathbf{g}_o \otimes \mathbf{g}^o \ .$$

Then, in light of $(5.85\text{b})_4$ and $(5.86\text{b})_2$, the orthogonality condition $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ renders

$$(\mathbf{Q})^{m}_{.k} \left(\mathbf{Q}^{\mathrm{T}}\right)^{k}_{.l} \mathbf{g}_m \otimes \mathbf{g}^l = \mathbf{g}_o \otimes \mathbf{g}^o \quad \Longrightarrow \quad (\mathbf{Q})^{m}_{.k}\, (\mathbf{Q})^{.k}_{l}\, \mathbf{g}_m \otimes \mathbf{g}^l = \mathbf{g}_o \otimes \mathbf{g}^o \ .$$

Post-multiplying both sides of the above result with $\mathbf{g}_j$ and subsequently multiplying the resulting vectors with $\mathbf{g}^i$, leads to the desired result:

$$(\mathbf{Q})^{m}_{.k}\, (\mathbf{Q})^{.k}_{j}\, \mathbf{g}_m = \mathbf{g}_j \quad \Longrightarrow \quad (\mathbf{Q})^{i}_{.k}\, (\mathbf{Q})^{.k}_{j} = \delta^i_j \ .$$

One can follow the same procedure to verify the second desired relation. Here, one needs to use the following appropriate forms of tensors

$$\mathbf{Q} = (\mathbf{Q})^{.k}_{m}\, \mathbf{g}^m \otimes \mathbf{g}_k \quad , \quad \mathbf{Q}^{\mathrm{T}} = \left(\mathbf{Q}^{\mathrm{T}}\right)^{l}_{.n} \mathbf{g}_l \otimes \mathbf{g}^n \quad , \quad \mathbf{I} = g_{mn} \mathbf{g}^m \otimes \mathbf{g}^n \ .$$

Then, the orthogonality condition $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ delivers

$$\underbrace{(\mathbf{Q})^{.k}_{m}\, g_{kl} \left(\mathbf{Q}^{\mathrm{T}}\right)^{l}_{.n} \mathbf{g}^m \otimes \mathbf{g}^n = g_{mn} \mathbf{g}^m \otimes \mathbf{g}^n}_{\text{or } (\mathbf{Q})^{.k}_{m}\, g_{kl}\, (\mathbf{Q})^{.l}_{n}\, \mathbf{g}^m \otimes \mathbf{g}^n = g_{mn} \mathbf{g}^m \otimes \mathbf{g}^n} \ .$$

At the end, post-multiplying both sides of this equation with $\mathbf{g}_j$ and then multiplying the resulting expression with $\mathbf{g}_i$ yields

$$(\mathbf{Q})^{.k}_{m}\, g_{kl}\, (\mathbf{Q})^{.l}_{j}\, \mathbf{g}^m = g_{mj} \mathbf{g}^m \quad \Longrightarrow \quad (\mathbf{Q})^{.k}_{i}\, g_{kl}\, (\mathbf{Q})^{.l}_{j} = g_{ij} \ .$$

2. Demonstrate that under the transformation $\bar{\mathbf{g}}_i = \mathbf{Q}\mathbf{g}_i$ or $\mathbf{g}_i = \mathbf{Q}^{\mathrm{T}}\bar{\mathbf{g}}_i$, $i = 1, 2, 3$, the components of the identity tensor $\mathbf{I}$ remain invariant.

   Solution. Given two sets of covariant basis vectors; namely $\{\bar{\mathbf{g}}_i\}$ and $\{\mathbf{g}_i\}$. The corresponding covariant metric coefficients can then be written as $\bar{g}_{ij} = \bar{\mathbf{g}}_i \cdot \bar{\mathbf{g}}_i$ and $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$. Knowing that the covariant metric coefficients represent the covariant components of the identity tensor according to $(5.78)_4$, the goal is now to show that

$$\bar{g}_{ij} = g_{ij} \quad \text{if} \quad \bar{\mathbf{g}}_i = \mathbf{Q}\mathbf{g}_i \ .$$

This amounts to writing

$$\mathbf{I} = \bar{g}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = g_{ij}\bar{\mathbf{g}}^i \otimes \bar{\mathbf{g}}^j \quad \text{if} \quad \bar{\mathbf{g}}_i = \mathbf{Q}\mathbf{g}_i \ .$$

By means of (2.5), (2.51d)$_1$, (2.130)$_1$ and (5.38)$_1$, one can infer that

$$\bar{g}_{ij} = \bar{\mathbf{g}}_i \cdot \bar{\mathbf{g}}_j = \mathbf{Q}\mathbf{g}_i \cdot \mathbf{Q}\mathbf{g}_j = \mathbf{g}_i \cdot \mathbf{Q}^{\mathrm{T}}\mathbf{Q}\mathbf{g}_j = \mathbf{g}_i \cdot \mathbf{I}\mathbf{g}_j = \mathbf{g}_i \cdot \mathbf{g}_j = g_{ij} \ .$$

Exercise 5.4

Let the covariant basis vectors $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{g}_3$ be given by

$$\mathbf{g}_1 = \widehat{\mathbf{e}}_1 + 3\,\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}_2 = -\widehat{\mathbf{e}}_1 + 2\,\widehat{\mathbf{e}}_2 + 0.5\,\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}_3 = \widehat{\mathbf{e}}_2 + 2\,\widehat{\mathbf{e}}_3 \ .$$

In matrix notation, they render

$$[\,\mathbf{g}_1\,] = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \quad , \quad [\,\mathbf{g}_2\,] = \begin{bmatrix} -1 \\ 2 \\ 0.5 \end{bmatrix} \quad , \quad [\,\mathbf{g}_3\,] = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \ .$$

First, consider the vectors

$$\begin{aligned} \mathbf{u} &= \widehat{\mathbf{e}}_1 + 2\,\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3 \ , \\ \mathbf{v} &= 3\,\widehat{\mathbf{e}}_1 - \widehat{\mathbf{e}}_2 + 2\,\widehat{\mathbf{e}}_3 \ , \end{aligned}$$

or, the single-column matrices,

$$[\,\mathbf{u}\,] = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^{\mathrm{T}} \ ,$$

$$[\,\mathbf{v}\,] = \begin{bmatrix} 3 & -1 & 2 \end{bmatrix}^{\mathrm{T}} \ .$$

Then, consider the tensors

$$\begin{aligned} \mathbf{A} = {} & \widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_1 + 2\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_2 - 0.3\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_3 + 0.7\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_1 - 0.8\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_3 \\ & + 1.6\,\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_1 + 1.5\,\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_2 + 1.2\,\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_3 \ , \\ \mathbf{B} = {} & 0.6\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_1 + 1.4\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_2 + 2\,\widehat{\mathbf{e}}_1 \otimes \widehat{\mathbf{e}}_3 + \widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_1 - 2\,\widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_2 - \widehat{\mathbf{e}}_2 \otimes \widehat{\mathbf{e}}_3 \\ & - 1.4\,\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_1 + 0.5\,\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_2 - 0.9\,\widehat{\mathbf{e}}_3 \otimes \widehat{\mathbf{e}}_3 \ , \end{aligned}$$

or, the matrices,

$$[\mathbf{A}] = \begin{bmatrix} 1 & 2 & -0.3 \\ 0.7 & -0.8 & 1 \\ 1.6 & 1.5 & 1.2 \end{bmatrix} ,$$

$$[\mathbf{B}] = \begin{bmatrix} 0.6 & 1.4 & 2 \\ 1 & -2 & -1 \\ -1.4 & 0.5 & -0.9 \end{bmatrix} .$$

Finally, consider the contravariant fourth-order tensor

$$\begin{aligned} \mathbf{A} \odot \mathbf{A} &= \frac{1}{2} \left(\mathbf{A} \boxtimes \mathbf{A} + \mathbf{A} \boxplus \mathbf{A} \right) \\ &= \frac{1}{2} \left(\underline{A}^{ik} \, \underline{A}^{lj} + \underline{A}^{il} \, \underline{A}^{kj} \right) \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k \otimes \mathbf{g}_l \ . \end{aligned}$$

Note that only contravariant components of **A** have been utilized to construct this tensor of rank four.

For in-depth learning vector and tensor algebra in curvilinear coordinates, write a **computer program** to compute the

(a) scalar product $\mathbf{u} \cdot \mathbf{v}$ by use of (5.67a)–(5.67d),
(b) length of **u** by use of $(5.68)_{1-3}$,
(c) cross product $\mathbf{u} \times \mathbf{v}$ by use of $(5.69a)_{1-2}$ and $(5.69b)_{3-4}$,
(d) simple contraction **Au** by use of $(5.84a)_{1-3}$–$(5.84d)_{1-3}$,
(e) composition **AB** by use of $(5.85a)_1$, $(5.85b)_2$, $(5.85c)_3$ and $(5.85d)_4$,
(f) transpose of **A** by use of (5.86a)–(5.86d),
(g) double contraction $\mathbf{A} : \mathbf{B}$ by use of $(5.87)_1$, $(5.87)_5$, $(5.87)_9$ and $(5.87)_{13}$,
(h) trace of **A** by use of $(5.88)_{1-4}$,
(i) determinant of **A** by use of (5.89a)–(5.89d),
(j) cofactor of **A** by use of (5.90a)–(5.90d),
(k) inverse of **A** by use of (5.91a)–(5.91d) and
(l) right mapping $(\mathbf{A} \odot \mathbf{A}) : \mathbf{B}$ by use of the rule $(3.54d)_1$, the Cartesian representation $(3.66a)_4$ and the curvilinear forms $(5.101a)_{1-4}$.

Hint: Start the procedure by constructing the tensors **F** and **C** and subsequently compute the Jacobian J, the contravariant basis $\{\mathbf{g}^i\}$, the covariant metric g_{ij} and the contravariant metric g^{ij}. Then, calculate the components of the given first- and second-order tensors with respect to the different curvilinear bases. Finally, implement all desired relations.

Hint: For implementing each operation, it is strongly recommended to first use the Cartesian components of tensorial variables. This may help better understand the differences regarding computations based on the Cartesian and curvilinear components. Moreover, it helps provide a benchmark for comparing the results.

Solution. Among different programming languages, MATLAB has been chosen in this text due to its popularity. But, it has been tried to develop the codes in a very basic format without using the functions installed with MATLAB. This helps keep their generality as far as possible. The desired code can freely be accessed through https://data.uni-hannover.de/dataset/exercises-tensor-analysis by any serious reader.

References

1. Marsden, J. E., & Hughes, T. J. R. (1983). *Mathematical foundations of elasticity*. Englewood Cliffs, New Jersey: Prentice-Hall.
2. Steinmann, P. (2015). *Geometrical foundations of continuum mechanics: An application to first- and second-order elasticity and elasto-plasticity*. Lecture notes in applied mathematics and mechanics. Berlin, Heidelberg: Springer.
3. Synge, J. L., & Schild, A. (1956). *Tensor calculus*. Toronto: University of Toronto Press.
4. Başar, Y., & Weichert, D. (2000). *Nonlinear continuum mechanics of solids: Fundamental mathematical and physical concepts*. Berlin, Heidelberg: Springer.
5. Hashiguchi, K. (2020). *Nonlinear continuum mechanics for finite elasticity-plasticity: Multiplicative decomposition with subloading surface model*. Amsterdam: Elsevier Science.
6. Brannon, R. M. (2004). *Curvilinear analysis in a Euclidean space*. Albuquerque, NM: University of New Mexico.
7. Simo, J. C., & Pister, K. S. (1984). Remarks on rate constitutive equations for finite deformation problems: Computational implications. *Computer Methods in Applied Mechanics and Engineering, 46,* 201–215.
8. Simo, J. C., Taylor, R. L., & Pister, K. S. (1985). Variational and projection methods for the volume constraint in finite deformation elasto-plasticity. *Computer Methods in Applied Mechanics and Engineering, 51*, 177–208.
9. Ibrahimbegović, A. (1994). Finite elastoplastic deformations of space-curved membranes. *Computer Methods in Applied Mechanics and Engineering, 119*, 371–394.

Chapter 6
Differentiation of Tensor Functions and Representation Theorems

6.1 Tensor Functions and Their Derivatives

Vector and tensor analysis or calculus is prerequisite for many applications in science and engineering. Examples of which include differential geometry, electromagnetic field theory and continuum mechanics. It is an important branch of mathematics that studies differentiation and integration of vector and tensor fields which will extensively be used in this text. This motivates to devote the first part of this chapter as well as the two upcoming chapters of this book to study the fundamental rules of vector and tensor calculus and their applications. Particularly, in this section, the standard rules of differentiation for tensor functions are first introduced. Then, their gradients are established by means of a first-order Taylor series expansion. At the end, some analytical derivatives that frequently appear in this text are approximated by use of finite difference method.

A **tensor function** presents a function whose arguments can be several tensorial variables (with possibly different ranks) while its output will be a scalar, vector or tensor. As an example, one should realize

- ☛☛ $\Phi(\mathbf{A})$ as a scalar-valued function of one tensor variable $\mathbf{A}$, $\Phi(\mathbf{u})$ as a scalar-valued function of one vector variable $\mathbf{u}$ and $\Phi(t)$ as a scalar-valued function of one scalar variable t.
- ☛☛ $\mathbf{v}(\mathbf{A}, \mathbf{u})$ as a vector-valued function of the tensorial variables $\mathbf{A}$ and $\mathbf{u}$, $\mathbf{w}(\mathbf{v}, \mathbf{u})$ as a vector-valued function of the tensorial variables $\mathbf{v}$ and $\mathbf{u}$ and $\mathbf{w}(\mathbf{u}, t)$ as a vector-valued function of the tensorial variables $\mathbf{u}$ and t.
- ☛☛ $\mathbf{B}(\mathbf{A})$ as a tensor-valued function of one tensor variable $\mathbf{A}$, $\mathbf{B}(\mathbf{u})$ as a tensor-valued function of one vector variable $\mathbf{u}$ and $\mathbf{B}(t)$ as a tensor-valued function of one scalar variable t.

All tensor functions are assumed to be **continuous** in this textbook. Considering the vector-valued function $\mathbf{u}(t)$ and the tensor-valued function $\mathbf{A}(t)$, this is indicated by the following relations

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_6

$$\lim_{t \to t_0} [\mathbf{u}(t) - \mathbf{u}(t_0)] = \mathbf{0} \,, \tag{6.1a}$$

$$\lim_{t \to t_0} [\mathbf{A}(t) - \mathbf{A}(t_0)] = \mathbf{O} \,. \tag{6.1b}$$

The tensor functions $\mathbf{u}(t)$ and $\mathbf{A}(t)$ are said to be **differentiable** if the limits

$$\dot{\mathbf{u}} := \lim_{h \to 0} \frac{\mathbf{u}(t+h) - \mathbf{u}(t)}{h} \,, \tag{6.2a}$$

$$\dot{\mathbf{A}} := \lim_{h \to 0} \frac{\mathbf{A}(t+h) - \mathbf{A}(t)}{h} \,, \tag{6.2b}$$

exist and are finite. In these relations, $\dot{\mathbf{u}}$ and $\dot{\mathbf{A}}$ are referred to as the *first derivatives* of $\mathbf{u}(t)$ and $\mathbf{A}(t)$, respectively. They are also sometimes denoted by

$$\mathbf{u}' := \frac{d\mathbf{u}}{dt} \quad , \quad \mathbf{A}' := \frac{d\mathbf{A}}{dt} \,. \tag{6.3}$$

Here, the standard rules of differentiation for scalar functions hold. Their analogues are represented in the following.

Product of a scalar-valued function with a vector-valued function:

$$\overline{\Phi(t)\,\mathbf{u}(t)}^{\bullet} = \dot{\Phi}\mathbf{u} + \Phi\dot{\mathbf{u}} \,, \tag{6.4a}$$

Product of a scalar-valued function with a tensor-valued function:

$$\overline{\Phi(t)\,\mathbf{A}(t)}^{\bullet} = \dot{\Phi}\mathbf{A} + \Phi\dot{\mathbf{A}} \,, \tag{6.4b}$$

Vector addition or subtraction:

$$\overline{\mathbf{u}(t) \pm \mathbf{v}(t)}^{\bullet} = \dot{\mathbf{u}} \pm \dot{\mathbf{v}} \,, \tag{6.4c}$$

Tensor addition or subtraction:

$$\overline{\mathbf{A}(t) \pm \mathbf{B}(t)}^{\bullet} = \dot{\mathbf{A}} \pm \dot{\mathbf{B}} \,, \tag{6.4d}$$

Scalar product between two vector-valued functions:

$$\overline{\mathbf{u}(t) \cdot \mathbf{v}(t)}^{\bullet} = \dot{\mathbf{u}} \cdot \mathbf{v} + \mathbf{u} \cdot \dot{\mathbf{v}} \,, \tag{6.4e}$$

Scalar product between two tensor-valued functions:

$$\overline{\mathbf{A}(t) : \mathbf{B}(t)}^{\bullet} = \dot{\mathbf{A}} : \mathbf{B} + \mathbf{A} : \dot{\mathbf{B}} \,, \tag{6.4f}$$

Single contraction of a tensor-valued function and a vector-valued function:

$$\overline{\mathbf{A}(t)\,\mathbf{u}(t)}^{\bullet} = \dot{\mathbf{A}}\mathbf{u} + \mathbf{A}\dot{\mathbf{u}} \,, \tag{6.4g}$$

Single contraction between two tensor-valued functions:

$$\overline{\mathbf{A}(t)\,\mathbf{B}(t)}^{\bullet} = \dot{\mathbf{A}}\mathbf{B} + \mathbf{A}\dot{\mathbf{B}} \,, \tag{6.4h}$$

Cross product of two vector-valued functions:

$$\overline{\mathbf{u}(t) \times \mathbf{v}(t)}^{\bullet} = \dot{\mathbf{u}} \times \mathbf{v} + \mathbf{u} \times \dot{\mathbf{v}} \,, \tag{6.4i}$$

Tensor product of two vector-valued functions:

$$\dot{\overline{\mathbf{u}(t) \otimes \mathbf{v}(t)}} = \dot{\mathbf{u}} \otimes \mathbf{v} + \mathbf{u} \otimes \dot{\mathbf{v}} \ . \tag{6.4j}$$

Here and elsewhere, a dot over a overline indicates that the quantity under the overline should be differentiated. These differentiation rules can be proved by using (6.2a) and (6.2b) along with basic concepts of differential calculus. For instance, (6.4g) can be verified as follows:

$$\begin{aligned}
\dot{\overline{\mathbf{A}(t)\,\mathbf{u}(t)}} &= \lim_{h\to 0} \frac{\mathbf{A}(t+h)\,\mathbf{u}(t+h) - \mathbf{A}(t)\,\mathbf{u}(t)}{h} \\
&= \lim_{h\to 0} \frac{\mathbf{A}(t+h)\,\mathbf{u}(t+h) + \mathbf{A}(t)\,\mathbf{u}(t+h) - \mathbf{A}(t)\,\mathbf{u}(t+h) - \mathbf{A}(t)\,\mathbf{u}(t)}{h} \\
&= \lim_{h\to 0} \frac{[\mathbf{A}(t+h) - \mathbf{A}(t)]\,\mathbf{u}(t+h)}{h} + \lim_{h\to 0} \frac{\mathbf{A}(t)\,[\mathbf{u}(t+h) - \mathbf{u}(t)]}{h} \\
&= \dot{\mathbf{A}}\mathbf{u} + \mathbf{A}\dot{\mathbf{u}} \ .
\end{aligned}$$

Throughout this textbook, the standard basis $\{\widehat{\mathbf{e}}_i\}$ should be regarded as a set of **fixed** orthonormal vectors whereas an incremental change in an element of either the covariant basis $\{\mathbf{g}_i\}$ or its dual basis $\{\mathbf{g}^i\}$ is not zero in general. Now, consider the Cartesian tensor $\mathbf{A} = A_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$ together with its transpose $\mathbf{A}^{\mathrm{T}} = A_{ij}\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i$. It is then easy to see that

$$\dot{\mathbf{A}} = \dot{A}_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \quad , \quad \dot{\overline{\mathbf{A}^{\mathrm{T}}}} = \dot{A}_{ij}\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_i \ .$$

As a result,

$$\boxed{\dot{\mathbf{A}}^{\mathrm{T}} = \dot{\overline{\mathbf{A}^{\mathrm{T}}}} \ .} \tag{6.5}$$

The flowing identities also hold true

$$\dot{\mathbf{I}} = \mathbf{O} \ , \quad \leftarrow \text{see (2.23) and (5.78)} \tag{6.6a}$$

$$\dot{g}_{ij} \xlongequal[\text{(5.38) and (6.4e)}]{\text{from}} \dot{\mathbf{g}}_i \cdot \mathbf{g}_j + \mathbf{g}_i \cdot \dot{\mathbf{g}}_j \ , \tag{6.6b}$$

$$\dot{g}^{ij} \xlongequal[\text{(5.46) and (6.4e)}]{\text{from}} \dot{\mathbf{g}}^i \cdot \mathbf{g}^j + \mathbf{g}^i \cdot \dot{\mathbf{g}}^j \ . \tag{6.6c}$$

From $(2.78)_2$, $(2.83)_1$, (6.4f) and (6.6a), it follows that

$$\boxed{\dot{\overline{\mathrm{tr}\mathbf{A}}} = \mathbf{I} : \dot{\mathbf{A}} = \mathrm{tr}\dot{\mathbf{A}} \ .} \tag{6.7}$$

Next, the frequently used chain rules of differentiation obey

Chain rule for a vector-valued function of one variable:

$$\frac{d}{dt}\mathbf{u}\left[\Phi\left(t\right)\right] = \frac{d\mathbf{u}}{d\Phi}\frac{d\Phi}{dt} \ , \tag{6.8a}$$

Chain rule for a tensor-valued function of one variable:

$$\frac{d}{dt}\mathbf{A}\left[\Phi\left(t\right)\right] = \frac{d\mathbf{A}}{d\Phi}\frac{d\Phi}{dt} \ , \tag{6.8b}$$

Chain rule for a scalar-valued function of three variables:

$$\frac{d}{dt}\bar{h}\left[\Phi\left(t\right), \mathbf{u}\left(t\right), \mathbf{A}\left(t\right)\right] = \frac{\partial\bar{h}}{\partial\Phi}\frac{d\Phi}{dt} + \frac{\partial\bar{h}}{\partial\mathbf{u}} \cdot \frac{d\mathbf{u}}{dt} + \frac{\partial\bar{h}}{\partial\mathbf{A}} : \frac{d\mathbf{A}}{dt} \ , \tag{6.8c}$$

Chain rule for a vector-valued function of three variables:

$$\frac{d}{dt}\hat{\mathbf{h}}\left[\Phi\left(t\right), \mathbf{u}\left(t\right), \mathbf{A}\left(t\right)\right] = \frac{\partial\hat{\mathbf{h}}}{\partial\Phi}\frac{d\Phi}{dt} + \frac{\partial\hat{\mathbf{h}}}{\partial\mathbf{u}}\frac{d\mathbf{u}}{dt} + \frac{\partial\hat{\mathbf{h}}}{\partial\mathbf{A}} : \frac{d\mathbf{A}}{dt} \ , \tag{6.8d}$$

Chain rule for a tensor-valued function of three variables:

$$\frac{d}{dt}\tilde{\mathbf{H}}\left[\Phi\left(t\right), \mathbf{u}\left(t\right), \mathbf{A}\left(t\right)\right] = \frac{\partial\tilde{\mathbf{H}}}{\partial\Phi}\frac{d\Phi}{dt} + \frac{\partial\tilde{\mathbf{H}}}{\partial\mathbf{u}}\frac{d\mathbf{u}}{dt} + \frac{\partial\tilde{\mathbf{H}}}{\partial\mathbf{A}} : \frac{d\mathbf{A}}{dt} \ . \tag{6.8e}$$

The term $\partial\bar{h}/\partial\Phi$ in the expression (6.8c) denotes the **partial derivative** of $\bar{h}$ with respect to Φ. Therein, $\partial\bar{h}/\partial\Phi$ is a scalar, $\partial\bar{h}/\partial\mathbf{u}$ consistently presents a vector to be (single) contracted with the vector $d\mathbf{u}/dt$ in order to have a scalar, and $\partial\bar{h}/\partial\mathbf{A}$ turns out to be a tensor (since its double contraction with the tensor $d\mathbf{A}/dt$ again delivers a scalar). It is not then difficult to notice that $\partial\hat{\mathbf{h}}/\partial\Phi$ ($\partial\tilde{\mathbf{H}}/\partial\Phi$) is a first-order (second-order) tensor, $\partial\hat{\mathbf{h}}/\partial\mathbf{u}$ ($\partial\tilde{\mathbf{H}}/\partial\mathbf{u}$) presents a second-order (third-order) tensor and $\partial\hat{\mathbf{h}}/\partial\mathbf{A}$ ($\partial\tilde{\mathbf{H}}/\partial\mathbf{A}$) renders a third-order (fourth-order) tensor. Now, one can generally conclude that if a tensor of rank n is a function of a tensor of rank m, then the partial derivative becomes a tensor of rank $m + n$. The partial derivatives in these relations are usually obtained from what are called *directional derivatives*. They are introduced in the following.

6.1.1 Gradient of a Scalar-Valued Function

Approximating differentiable nonlinear functions by linear ones at any point of interest within the domain of definition is of great importance in many scientific and engineering problems. The goal here is thus to approximate a nonlinear scalar-valued function by a linear function.

Let $\bar{h} : \mathrm{R} \times \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3} \times \mathscr{T}_{\mathrm{so}} \longrightarrow \mathrm{R}$ be a nonlinear and sufficiently smooth scalar-valued function of the three tensorial variables Φ, $\mathbf{u}$ and $\mathbf{A}$. The first-order Taylor series of $\bar{h}$ at $(\Phi, \mathbf{u}, \mathbf{A})$ is given by

$$\bar{h}\left(\Phi + d\Phi, \mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}\right) = \bar{h}\left(\Phi, \mathbf{u}, \mathbf{A}\right) + d\bar{h} + o\left(d\Phi, d\mathbf{u}, d\mathbf{A}\right) \ , \tag{6.9}$$

where the **total differential** $d\bar{h}$ is

$$d\bar{h} = \frac{\partial \bar{h}}{\partial \Phi} d\Phi + \frac{\partial \bar{h}}{\partial \mathbf{u}} \cdot d\mathbf{u} + \frac{\partial \bar{h}}{\partial \mathbf{A}} : d\mathbf{A} \ , \tag{6.10}$$

and $o\,(d\Phi, d\mathbf{u}, d\mathbf{A})$ presents the **Landau order symbol**. It contains terms that approach zero faster than $(d\Phi \to 0, d\mathbf{u} \to \mathbf{0}, d\mathbf{A} \to \mathbf{O})$. In the total differential form (6.10), the partial derivative $\partial\bar{h}/\partial\Phi$ is also known as the *gradient* of $\bar{h}$ with respect to Φ which is a scalar. As expected, the derivative of $\bar{h}$ with respect to $\mathbf{u}$ ($\mathbf{A}$) renders a vector (tensor). Note that the reason why the derivatives in (6.10) can be taken stems from the fact that the scalar-valued function $\bar{h}$ was assumed to be sufficiently smooth a priori. Indeed, they can uniquely be determined from (6.9)–(6.10), see (6.15)–(6.16). They can also be obtained from the so-called *directional* (or *Gâteaux*) *derivatives* in an elegant way:

$$\begin{aligned}
D_{\phi}\bar{h}\,(\Phi, \mathbf{u}, \mathbf{A}) &:= \left.\frac{d}{d\varepsilon}\bar{h}\,(\Phi + \varepsilon\phi, \mathbf{u}, \mathbf{A})\right|_{\varepsilon=0} \\
&= \left[\frac{\partial \bar{h}}{\partial\,(\Phi + \varepsilon\phi)} \frac{\partial\,(\Phi + \varepsilon\phi)}{\partial\varepsilon}\right]_{\varepsilon=0} \\
&= \frac{\partial \bar{h}}{\partial \Phi}\phi \ , \qquad (6.11a) \\
D_{\mathbf{v}}\bar{h}\,(\Phi, \mathbf{u}, \mathbf{A}) &:= \left.\frac{d}{d\varepsilon}\bar{h}\,(\Phi, \mathbf{u} + \varepsilon\mathbf{v}, \mathbf{A})\right|_{\varepsilon=0} \\
&= \left[\frac{\partial \bar{h}}{\partial\,(\mathbf{u} + \varepsilon\mathbf{v})} \cdot \frac{\partial\,(\mathbf{u} + \varepsilon\mathbf{v})}{\partial\varepsilon}\right]_{\varepsilon=0} \\
&= \frac{\partial \bar{h}}{\partial \mathbf{u}} \cdot \mathbf{v} \ , \qquad (6.11b) \\
D_{\mathbf{B}}\bar{h}\,(\Phi, \mathbf{u}, \mathbf{A}) &:= \left.\frac{d}{d\varepsilon}\bar{h}\,(\Phi, \mathbf{u}, \mathbf{A} + \varepsilon\mathbf{B})\right|_{\varepsilon=0} \\
&= \left[\frac{\partial \bar{h}}{\partial\,(\mathbf{A} + \varepsilon\mathbf{B})} : \frac{\partial\,(\mathbf{A} + \varepsilon\mathbf{B})}{\partial\varepsilon}\right]_{\varepsilon=0} \\
&= \frac{\partial \bar{h}}{\partial \mathbf{A}} : \mathbf{B} \ . \qquad (6.11c)
\end{aligned}$$

Note that $D_{\mathbf{v}}\bar{h}\,(\Phi, \mathbf{u}, \mathbf{A})$ ($D_{\mathbf{B}}\bar{h}\,(\Phi, \mathbf{u}, \mathbf{A})$) indicates the directional derivative of $\bar{h}$ at $(\Phi, \mathbf{u}, \mathbf{A})$ in the direction of $\mathbf{v}$ ($\mathbf{B}$). It should be noted that the normalized vector $\widehat{\mathbf{v}}$ is often utilized for the definition of directional derivative in (6.11b). As can be seen, the directional derivative of a scalar-valued function is always a scalar. Basically, it is a **linear** operator that satisfies the usual rules of differentiation such as the chain rule, as shown above. This powerful tool keeps only the second term (or linear term) of a Taylor series.

Replacing $(\phi, \mathbf{v}, \mathbf{B})$ by $(d\Phi, d\mathbf{u}, d\mathbf{A})$ in (6.11a)–(6.11c) helps express the total differential (6.10) as

$$d\bar{h} = D_{d\Phi}\bar{h}(\Phi, \mathbf{u}, \mathbf{A}) + D_{d\mathbf{u}}\bar{h}(\Phi, \mathbf{u}, \mathbf{A}) + D_{d\mathbf{A}}\bar{h}(\Phi, \mathbf{u}, \mathbf{A})$$
$$= \frac{d}{d\varepsilon}\bar{h}(\Phi + \varepsilon d\Phi, \mathbf{u} + \varepsilon d\mathbf{u}, \mathbf{A} + \varepsilon d\mathbf{A})\Big|_{\varepsilon=0} . \tag{6.12}$$

Neglecting the reminder $o(d\Phi, d\mathbf{u}, d\mathbf{A})$ helps provide a linear approximation to the nonlinear function $\bar{h}$ around $(\Phi, \mathbf{u}, \mathbf{A})$ according to

$$\bar{h}(\Phi + d\Phi, \mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}) \approx \bar{h}(\Phi, \mathbf{u}, \mathbf{A}) + \underbrace{\frac{\partial\bar{h}}{\partial\Phi}d\Phi + \frac{\partial\bar{h}}{\partial\mathbf{u}} \cdot d\mathbf{u} + \frac{\partial\bar{h}}{\partial\mathbf{A}} : d\mathbf{A}}_{= \frac{d}{d\varepsilon}\bar{h}(\Phi + \varepsilon d\Phi, \mathbf{u} + \varepsilon d\mathbf{u}, \mathbf{A} + \varepsilon d\mathbf{A})\big|_{\varepsilon=0}} . \tag{6.13}$$

The goal is now to express the derivatives $\partial\bar{h}/\partial\mathbf{u}$ and $\partial\bar{h}/\partial\mathbf{A}$ with respect to a basis. Given two Cartesian vectors $\mathbf{u} = u_m\widehat{\mathbf{e}}_m$ and $\mathbf{v} = v_m\widehat{\mathbf{e}}_m$ as well as two Cartesian tensors $\mathbf{A} = A_{mn}\widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n$ and $\mathbf{B} = B_{mn}\widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n$. The directional derivatives (6.11b) and (6.11c) then render

$$\underbrace{\frac{\partial\bar{h}}{\partial\mathbf{u}} \cdot \mathbf{v}}_{= \frac{d}{d\varepsilon}\bar{h}(\Phi, [u_m + \varepsilon v_m]\widehat{\mathbf{e}}_m, \mathbf{A})\big|_{\varepsilon=0}} = \frac{\partial\bar{h}}{\partial u_m}v_m \;, \qquad \underbrace{\frac{\partial\bar{h}}{\partial\mathbf{A}} : \mathbf{B}}_{= \frac{d}{d\varepsilon}\bar{h}(\Phi, \mathbf{u}, [A_{mn} + \varepsilon B_{mn}]\widehat{\mathbf{e}}_m \otimes \widehat{\mathbf{e}}_n)\big|_{\varepsilon=0}} = \frac{\partial\bar{h}}{\partial A_{mn}}B_{mn} \;,$$

which immediately imply that

$$\frac{\partial\bar{h}}{\partial\mathbf{u}} = \frac{\partial\bar{h}}{\partial u_i}\widehat{\mathbf{e}}_i \;, \quad \frac{\partial\bar{h}}{\partial\mathbf{A}} = \frac{\partial\bar{h}}{\partial A_{ij}}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \;.$$

These results by means of the chain rule yield

$$\frac{\partial\bar{h}}{\partial u_i}\widehat{\mathbf{e}}_i = \frac{\partial\bar{h}}{\partial\mathbf{u}} = \frac{\partial\bar{h}}{\partial u_i}\frac{\partial u_i}{\partial\mathbf{u}} \quad \Rightarrow \quad \widehat{\mathbf{e}}_i = \frac{\partial u_i}{\partial\mathbf{u}} \;,$$
$$\frac{\partial\bar{h}}{\partial A_{ij}}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \frac{\partial\bar{h}}{\partial\mathbf{A}} = \frac{\partial\bar{h}}{\partial A_{ij}}\frac{\partial A_{ij}}{\partial\mathbf{A}} \quad \Rightarrow \quad \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \frac{\partial A_{ij}}{\partial\mathbf{A}} \;,$$

which amount to writing $d\mathbf{u} = (du_i)\widehat{\mathbf{e}}_i$ and $d\mathbf{A} = \left(dA_{ij}\right)\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j$. In summary, using the different representations of a first-order tensor in $(1.34)_1$ and (5.64a) and (5.64b) along with the various forms of a second-order tensor according to $(2.19)_2$ and (5.73a)–(5.73d), one will have

$$\frac{\partial\bar{h}}{\partial\mathbf{u}} = \frac{\partial\bar{h}}{\partial u_i}\widehat{\mathbf{e}}_i \quad \text{where} \quad \widehat{\mathbf{e}}_i = \frac{\partial u_i}{\partial\mathbf{u}} \;, \tag{6.14a}$$
$$\frac{\partial\bar{h}}{\partial\mathbf{u}} = \frac{\partial\bar{h}}{\partial\,\underline{u}_i}\mathbf{g}_i \quad \text{where} \quad \mathbf{g}_i = \frac{\partial\,\underline{u}_i}{\partial\mathbf{u}} \;, \tag{6.14b}$$

$$\frac{\partial \bar{h}}{\partial \mathbf{u}} = \frac{\partial \bar{h}}{\partial \underline{u}^i}\mathbf{g}^i \qquad \text{where} \qquad \mathbf{g}^i = \frac{\partial \underline{u}^i}{\partial \mathbf{u}} , \tag{6.14c}$$

$$\frac{\partial \bar{h}}{\partial \mathbf{A}} = \frac{\partial \bar{h}}{\partial A_{ij}}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \qquad \text{where} \qquad \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \frac{\partial A_{ij}}{\partial \mathbf{A}} , \tag{6.14d}$$

$$\frac{\partial \bar{h}}{\partial \mathbf{A}} = \frac{\partial \bar{h}}{\partial \underline{A}_{ij}}\mathbf{g}_i \otimes \mathbf{g}_j \qquad \text{where} \qquad \mathbf{g}_i \otimes \mathbf{g}_j = \frac{\partial \underline{A}_{ij}}{\partial \mathbf{A}} , \tag{6.14e}$$

$$\frac{\partial \bar{h}}{\partial \mathbf{A}} = \frac{\partial \bar{h}}{\partial \underline{A}_i^{\cdot j}}\mathbf{g}_i \otimes \mathbf{g}^j \qquad \text{where} \qquad \mathbf{g}_i \otimes \mathbf{g}^j = \frac{\partial \underline{A}_i^{\cdot j}}{\partial \mathbf{A}} , \tag{6.14f}$$

$$\frac{\partial \bar{h}}{\partial \mathbf{A}} = \frac{\partial \bar{h}}{\partial \underline{A}^i_{\cdot j}}\mathbf{g}^i \otimes \mathbf{g}_j \qquad \text{where} \qquad \mathbf{g}^i \otimes \mathbf{g}_j = \frac{\partial \underline{A}^i_{\cdot j}}{\partial \mathbf{A}} , \tag{6.14g}$$

$$\frac{\partial \bar{h}}{\partial \mathbf{A}} = \frac{\partial \bar{h}}{\partial \underline{A}^{ij}}\mathbf{g}^i \otimes \mathbf{g}^j \qquad \text{where} \qquad \mathbf{g}^i \otimes \mathbf{g}^j = \frac{\partial \underline{A}^{ij}}{\partial \mathbf{A}} . \tag{6.14h}$$

Guided by these relations, one can establish the following rule:

A new rule for index notation

If a subscript (superscript) index appears in the denominator of a partial derivative, it should be regarded as a superscript (subscript) in the summation convention.

Indeed, the partial derivative with respect to a covariant vector results in a contravariant vector and vice versa. This also holds true for tensors.

Moreover, for a given basis, these relations show that the partial derivative relative to a vector or tensor should be taken with respect to only its corresponding components. And the type of a basis vector must consistently be changed.

The procedure to compute the derivatives of a scalar-valued function will be illustrated in the following examples.

The first example here regards a scalar-valued function $\bar{h} : \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3} \times \mathscr{T}_{\mathrm{so}} \longrightarrow \mathrm{R}$ of the form

$$\bar{h}(\mathbf{u}, \mathbf{A}) = \mathbf{u} \cdot \mathbf{u} + \mathbf{A} : \mathbf{A} . \tag{6.15}$$

To begin with, one needs to identify the three terms on the right hand side of (6.9) for the problem at hand:

$$\begin{aligned} \bar{h}(\mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}) &= (\mathbf{u} + d\mathbf{u}) \cdot (\mathbf{u} + d\mathbf{u}) + (\mathbf{A} + d\mathbf{A}) : (\mathbf{A} + d\mathbf{A}) \\ &= \underbrace{\mathbf{u} \cdot \mathbf{u} + \mathbf{A} : \mathbf{A}}_{= \bar{h}(\mathbf{u}, \mathbf{A})} + \underbrace{\mathbf{u} \cdot d\mathbf{u} + d\mathbf{u} \cdot \mathbf{u} + \mathbf{A} : d\mathbf{A} + d\mathbf{A} : \mathbf{A}}_{= 2\mathbf{u} \cdot d\mathbf{u} + 2\mathbf{A} : d\mathbf{A} = d\bar{h}} \\ &\quad + \underbrace{d\mathbf{u} \cdot d\mathbf{u} + d\mathbf{A} : d\mathbf{A}}_{= o(d\mathbf{u}, d\mathbf{A})} . \end{aligned}$$

Guided by (6.10), the desired derivatives will be

$$\underbrace{\frac{\partial}{\partial \mathbf{u}}\left[\mathbf{u}\cdot\mathbf{u}\right] = 2\mathbf{u}}_{\text{or } \partial\left(u_j u_j\right)/\partial u_i = 2u_i} \quad , \quad \underbrace{\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{A}:\mathbf{A}\right] = 2\mathbf{A}}_{\text{or } \partial\left(A_{kl} A_{kl}\right)/\partial A_{ij} = 2A_{ij}} \quad . \tag{6.16}$$

The ambitious reader may want to use the directional derivatives to arrive at these results. Using $(6.12)_2$, the derivative of

$$\begin{aligned}\bar{h}\left(\mathbf{u}+\varepsilon d\mathbf{u}, \mathbf{A}+\varepsilon d\mathbf{A}\right) &= \left(\mathbf{u}+\varepsilon d\mathbf{u}\right)\cdot\left(\mathbf{u}+\varepsilon d\mathbf{u}\right) + \left(\mathbf{A}+\varepsilon d\mathbf{A}\right):\left(\mathbf{A}+\varepsilon d\mathbf{A}\right) \\ &= \varepsilon^0\left(\mathbf{u}\cdot\mathbf{u}+\mathbf{A}:\mathbf{A}\right) + \varepsilon^1\left(\mathbf{u}\cdot d\mathbf{u} + d\mathbf{u}\cdot\mathbf{u} + \mathbf{A}:d\mathbf{A} + d\mathbf{A}:\mathbf{A}\right) \\ &\quad + \varepsilon^2\left(d\mathbf{u}\cdot d\mathbf{u} + d\mathbf{A}:d\mathbf{A}\right) \ ,\end{aligned}$$

with respect to ε at $\varepsilon = 0$ yields the total differential

$$d\bar{h} = 2\mathbf{u}\cdot d\mathbf{u} + 2\mathbf{A}:d\mathbf{A} \ ,$$

from which one can deduce the results in (6.16) .

As another example, consider the scalar-valued function

$$\bar{h}\left(\mathbf{u},\mathbf{A}\right) = |\mathbf{u}| + |\mathbf{A}| + \mathbf{u}\cdot\mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{u}\cdot\mathbf{A}\mathbf{u} \ . \tag{6.17}$$

By following the similar procedures that led to (6.16), one can obtain

$$\underbrace{\frac{\partial\,|\mathbf{u}|}{\partial \mathbf{u}} = \frac{\mathbf{u}}{|\mathbf{u}|}}_{\text{or } \partial\left(u_j u_j\right)^{1/2}/\partial u_i = u_i/\left(u_j u_j\right)^{1/2}} \quad , \quad \underbrace{\frac{\partial\,|\mathbf{A}|}{\partial \mathbf{A}} = \frac{\mathbf{A}}{|\mathbf{A}|}}_{\text{or } \partial\left(A_{kl}A_{kl}\right)^{1/2}/\partial A_{ij} = A_{ij}/\left(A_{kl}A_{kl}\right)^{1/2}} \quad , \tag{6.18a}$$

$$\underbrace{\frac{\partial}{\partial \mathbf{u}}\left[\mathbf{u}\cdot\mathbf{A}\mathbf{u}\right] = \mathbf{A}\mathbf{u} + \mathbf{A}^{\mathrm{T}}\mathbf{u}}_{\text{or } \partial\left(u_k A_{kl} u_l\right)/\partial u_i = A_{il}u_l + u_k A_{ki}} \quad , \quad \underbrace{\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{u}\cdot\mathbf{A}\mathbf{u}\right] = \mathbf{u}\otimes\mathbf{u}}_{\text{or } \partial\left(u_k A_{kl} u_l\right)/\partial A_{ij} = u_i u_j} \ , \quad \swarrow \text{ see Sect. 6.1.4} \tag{6.18b}$$

$$\underbrace{\frac{\partial}{\partial \mathbf{u}}\left[\mathbf{A}\mathbf{u}\cdot\mathbf{A}\mathbf{u}\right] = 2\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{u}}_{\text{or } \partial\left(A_{jk}u_k A_{jl}u_l\right)/\partial u_i = 2A_{ji}A_{jl}u_l} \quad , \quad \underbrace{\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{A}\mathbf{u}\cdot\mathbf{A}\mathbf{u}\right] = 2\mathbf{A}\mathbf{u}\otimes\mathbf{u}}_{\text{or } \partial\left(A_{km}u_m A_{kn}u_n\right)/\partial A_{ij} = 2A_{im}u_m u_j} \quad . \tag{6.18c}$$

The last example here regards the principal scalar invariants (4.17a)–(4.17c). It follows that

$$I_1\left(\mathbf{A}+\varepsilon d\mathbf{A}\right)=\text{tr}\left[\mathbf{A}+\varepsilon d\mathbf{A}\right]=\mathbf{I}:\mathbf{A}+\varepsilon\mathbf{I}:d\mathbf{A}\ , \tag{6.19a}$$

$$\begin{aligned}I_2\left(\mathbf{A}+\varepsilon d\mathbf{A}\right)&=\frac{1}{2}\left[\left(\mathbf{I}:\mathbf{A}+\varepsilon\mathbf{I}:d\mathbf{A}\right)^2-\text{tr}\left(\mathbf{A}+\varepsilon d\mathbf{A}\right)^2\right]\\&=\frac{\varepsilon^0}{2}\left[\left(\mathbf{I}:\mathbf{A}\right)^2-\mathbf{I}:\mathbf{A}^2\right]+\varepsilon^1\left[\left(\mathbf{I}:\mathbf{A}\right)\left(\mathbf{I}:d\mathbf{A}\right)-\mathbf{A}^{\mathrm{T}}:d\mathbf{A}\right]\\&\quad+\frac{\varepsilon^2}{2}\left[\left(\mathbf{I}:d\mathbf{A}\right)\left(\mathbf{I}:d\mathbf{A}\right)-\left(\mathbf{I}:d\mathbf{A}d\mathbf{A}\right)\right]\ ,\end{aligned} \tag{6.19b}$$

$$\begin{aligned}I_3\left(\mathbf{A}+\varepsilon d\mathbf{A}\right)&=\det\left[\mathbf{A}+\varepsilon d\mathbf{A}\right]\\&=\varepsilon^0\det\mathbf{A}+\varepsilon^1\mathbf{A}^{\mathrm{c}}:\left[d\mathbf{A}\right]+\varepsilon^2\mathbf{A}:\left[d\mathbf{A}\right]^{\mathrm{c}}+\varepsilon^3\det\left[d\mathbf{A}\right]\ .\end{aligned} \tag{6.19c}$$

Consequently, the derivatives $dI_i\left(\mathbf{A}+\varepsilon d\mathbf{A}\right)/d\varepsilon,\ i=1,2,3$, at $\varepsilon=0$ help provide the useful relations

$$\underbrace{\frac{\partial I_1}{\partial\mathbf{A}}=\frac{\partial}{\partial\mathbf{A}}\left[\text{tr}\mathbf{A}\right]=\mathbf{I}}_{\text{or }\partial I_1/\partial A_{ij}=\delta_{ij}}\ , \tag{6.20a}$$

$$\underbrace{\frac{\partial I_2}{\partial\mathbf{A}}=\frac{\partial}{\partial\mathbf{A}}\left[\frac{1}{2}I_1^2-\frac{1}{2}\text{tr}\mathbf{A}^2\right]=I_1\mathbf{I}-\mathbf{A}^{\mathrm{T}}}_{\text{or }\partial I_2/\partial A_{ij}=I_1\delta_{ij}-A_{ji}}\ , \tag{6.20b}$$

$$\underbrace{\frac{\partial I_3}{\partial\mathbf{A}}=\frac{\partial}{\partial\mathbf{A}}\left[\det\mathbf{A}\right]=\mathbf{A}^{\mathrm{c}}=\left(\det\mathbf{A}\right)\mathbf{A}^{-\mathrm{T}}=I_3\mathbf{A}^{-\mathrm{T}}}_{\text{or }\partial I_3/\partial A_{ij}=I_3A_{ji}^{-1}}\ . \tag{6.20c}$$

The relations (6.20a)–(6.20c) are extensively used in, for instance, the constitutive response of isotropic hyperelastic solids, see Exercise 6.16. As a consequence of (6.20c), one will have, for any invertible tensor,

$$\boxed{\frac{\partial}{\partial\mathbf{A}}\left[\ln I_3\right]=\frac{\partial}{\partial\mathbf{A}}\left[\ln\left(\det\mathbf{A}\right)\right]=\mathbf{A}^{-\mathrm{T}}\ .} \tag{6.21}$$

Consider now the covariant metric tensor (5.80), i.e. $\mathbf{g}=g_{ij}\mathbf{g}^i\otimes\mathbf{g}^j$, where the determinant of its components renders $\det\left[g_{ij}\right]=J^2$, according to (5.41). Further, consider the contravariant metric tensor $\mathbf{g}^{-1}=g^{ij}\mathbf{g}_i\otimes\mathbf{g}_j$, given in (5.81), for which $\det\left[g^{ij}\right]=J^{-2}$, according to (5.52). Then, by using (6.14e), (6.14h) and (6.20c), taking into account the symmetry of the metric coefficients, one can establish the useful relations

$$\boxed{\frac{\partial J}{\partial \mathbf{g}} = \frac{\partial J}{\partial g_{ij}} \mathbf{g}_i \otimes \mathbf{g}_j \quad \text{with} \quad \frac{\partial J}{\partial g_{ij}} = \frac{J}{2} g^{ij} ,} \tag{6.22}$$

and

$$\boxed{\frac{\partial J}{\partial \mathbf{g}^{-1}} = \frac{\partial J}{\partial g^{ij}} \mathbf{g}^i \otimes \mathbf{g}^j \quad \text{with} \quad \frac{\partial J}{\partial g^{ij}} = -\frac{J}{2} g_{ij} .} \tag{6.23}$$

6.1.2 Gradient of a Vector-Valued Function

Let $\hat{\mathbf{h}} : \mathrm{R} \times \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}} \times \mathscr{T}_{\mathrm{so}} \longrightarrow \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}$ be a nonlinear and sufficiently smooth vector-valued function of the three tensorial variables Φ, $\mathbf{u}$ and $\mathbf{A}$. The first-order Taylor series expansion of $\hat{\mathbf{h}}$ about $(\Phi, \mathbf{u}, \mathbf{A})$ gives

$$\hat{\mathbf{h}} (\Phi + d\Phi, \mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}) = \hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A}) + d\hat{\mathbf{h}} + o (d\Phi, d\mathbf{u}, d\mathbf{A}) \ , \tag{6.24}$$

where the total differential $d\hat{\mathbf{h}}$ presents

$$\boxed{d\hat{\mathbf{h}} = \frac{\partial \hat{\mathbf{h}}}{\partial \Phi} d\Phi + \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{u}} d\mathbf{u} + \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{A}} : d\mathbf{A} ,} \tag{6.25}$$

and each term of the Landau symbol $o (d\Phi, d\mathbf{u}, d\mathbf{A})$ characterizes a vector. Here, the partial derivative of $\hat{\mathbf{h}}$ with respect to Φ, i.e. $\partial\hat{\mathbf{h}}/\partial\Phi$, is a vector. Whereas $\partial\hat{\mathbf{h}}/\partial\mathbf{u}$ $(\partial\hat{\mathbf{h}}/\partial\mathbf{A})$ renders a second-order (third-order) tensor. These gradients can also be obtained from the following directional derivatives:

$$D_{\phi}\hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A}) := \left. \frac{d}{d\varepsilon} \hat{\mathbf{h}} (\Phi + \varepsilon\phi, \mathbf{u}, \mathbf{A}) \right|_{\varepsilon=0} = \frac{\partial \hat{\mathbf{h}}}{\partial \Phi} \phi \ , \tag{6.26a}$$

$$D_{\mathbf{v}}\hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A}) := \left. \frac{d}{d\varepsilon} \hat{\mathbf{h}} (\Phi, \mathbf{u} + \varepsilon\mathbf{v}, \mathbf{A}) \right|_{\varepsilon=0} = \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{u}} \mathbf{v} \ , \tag{6.26b}$$

$$D_{\mathbf{B}}\hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A}) := \left. \frac{d}{d\varepsilon} \hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A} + \varepsilon\mathbf{B}) \right|_{\varepsilon=0} = \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{A}} : \mathbf{B} \ . \tag{6.26c}$$

As expected, the directional derivative of a vector-valued function does not change its order. Moreover, $D_{\mathbf{v}}\hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A})$ $(D_{\mathbf{B}}\hat{\mathbf{h}} (\Phi, \mathbf{u}, \mathbf{A}))$ represents the directional derivative of $\hat{\mathbf{h}}$ at $(\Phi, \mathbf{u}, \mathbf{A})$ in the direction of $\mathbf{v}$ $(\mathbf{B})$. Using (6.26a)–(6.26c), the total differential (6.25) can be rewritten in the compact form

$$d\hat{\mathbf{h}} = \frac{d}{d\varepsilon}\hat{\mathbf{h}}\left(\Phi + \varepsilon d\Phi, \mathbf{u} + \varepsilon d\mathbf{u}, \mathbf{A} + \varepsilon d\mathbf{A}\right)\Big|_{\varepsilon=0} . \tag{6.27}$$

The linearized form of (6.24) then renders

$$\hat{\mathbf{h}}\left(\Phi + d\Phi, \mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}\right) \approx \hat{\mathbf{h}}\left(\Phi, \mathbf{u}, \mathbf{A}\right) + \underbrace{\frac{\partial \hat{\mathbf{h}}}{\partial \Phi} d\Phi + \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{u}} d\mathbf{u} + \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{A}} : d\mathbf{A}}_{= \frac{d}{d\varepsilon}\hat{\mathbf{h}}\left(\Phi + \varepsilon d\Phi, \mathbf{u} + \varepsilon d\mathbf{u}, \mathbf{A} + \varepsilon d\mathbf{A}\right)\big|_{\varepsilon=0}} . \tag{6.28}$$

By following arguments similar to those which led to (6.14a)–(6.14h), one can express the partial derivatives $\partial \hat{\mathbf{h}}/\partial \mathbf{u}$ and $\partial \hat{\mathbf{h}}/\partial \mathbf{A}$ with respect to the standard basis $\{\widehat{\mathbf{e}}_i\}$ as well as the general basis $\{\mathbf{g}_i\}$ and the dual basis $\{\mathbf{g}^i\}$ as

$$\begin{aligned}
\frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{u}} &= \frac{\partial \hat{h}_i}{\partial u_j} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \quad \longleftarrow \text{see (2.19) and (5.73a)-(5.73d)} \\
&= \frac{\partial \underline{\hat{h}}^i}{\partial \underline{u}_j} \mathbf{g}_i \otimes \mathbf{g}_j = \frac{\partial \underline{\hat{h}}^i}{\partial \underline{u}^j} \mathbf{g}_i \otimes \mathbf{g}^j \\
&= \frac{\partial \underline{\hat{h}}_i}{\partial \underline{u}_j} \mathbf{g}^i \otimes \mathbf{g}_j = \frac{\partial \underline{\hat{h}}_i}{\partial \underline{u}^j} \mathbf{g}^i \otimes \mathbf{g}^j ,
\end{aligned} \tag{6.29}$$

$$\begin{aligned}
\frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{A}} &= \frac{\partial \hat{h}_i}{\partial A_{jk}} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \quad \longleftarrow \text{see (3.13) and (5.96a)-(5.96d)} \\
&= \frac{\partial \underline{\hat{h}}^i}{\partial \underline{A}_{jk}} \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k = \frac{\partial \underline{\hat{h}}^i}{\partial \underline{A}^j_{\ .k}} \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}_k \\
&= \frac{\partial \underline{\hat{h}}^i}{\partial \underline{A}^{jk}} \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k = \frac{\partial \underline{\hat{h}}_i}{\partial \underline{A}^{jk}} \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k .
\end{aligned} \tag{6.30}$$

The following example illustrates the procedure to arrive at the partial derivatives of a vector-valued function. ❐

Consider a vector-valued function $\hat{\mathbf{h}} : \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3} \times \mathscr{T}_{\mathrm{so}} \longrightarrow \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$ of the following form

$$\hat{\mathbf{h}}\left(\mathbf{u}, \mathbf{A}\right) = \mathbf{u} + \mathbf{A}\mathbf{u} + \mathbf{A}^2\mathbf{u} . \tag{6.31}$$

To compute the gradients of these three vectorial variables by use of (6.24), one needs to construct $\hat{\mathbf{h}}$ at $(\mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A})$. This is given by

$$\begin{aligned}\hat{\mathbf{h}}(\mathbf{u}+d\mathbf{u},\mathbf{A}+d\mathbf{A}) &= (\mathbf{u}+d\mathbf{u}) + (\mathbf{A}+d\mathbf{A})(\mathbf{u}+d\mathbf{u}) \\ &\quad + \left(\mathbf{A}^2 + \mathbf{A}d\mathbf{A} + d\mathbf{A}\mathbf{A} + d\mathbf{A}d\mathbf{A}\right)(\mathbf{u}+d\mathbf{u}) \\ &= \underbrace{\mathbf{u}+\mathbf{A}\mathbf{u}+\mathbf{A}^2\mathbf{u}}_{=\hat{\mathbf{h}}(\mathbf{u},\mathbf{A})} \\ &\quad + \underbrace{d\mathbf{u}+\mathbf{A}d\mathbf{u}+d\mathbf{A}\mathbf{u}+\mathbf{A}^2d\mathbf{u}+\mathbf{A}d\mathbf{A}\mathbf{u}+d\mathbf{A}\mathbf{A}\mathbf{u}}_{=\mathbf{I}d\mathbf{u}+\mathbf{A}d\mathbf{u}+(\mathbf{I}\otimes\mathbf{u}):d\mathbf{A}+\mathbf{A}^2d\mathbf{u}+(\mathbf{A}\otimes\mathbf{u}):d\mathbf{A}+(\mathbf{I}\otimes\mathbf{A}\mathbf{u}):d\mathbf{A}=d\hat{\mathbf{h}}} \\ &\quad + \underbrace{d\mathbf{A}d\mathbf{u}+\mathbf{A}d\mathbf{A}d\mathbf{u}+d\mathbf{A}\mathbf{A}d\mathbf{u}+d\mathbf{A}d\mathbf{A}\mathbf{u}+d\mathbf{A}d\mathbf{A}d\mathbf{u}}_{=o(d\mathbf{u},d\mathbf{A})}\ ,\end{aligned}$$

from which the desired derivatives, using (6.25), render

$$\underbrace{\frac{\partial\mathbf{u}}{\partial\mathbf{u}}=\mathbf{I}}_{\text{or } \partial u_i/\partial u_j=\delta_{ij}}\quad , \tag{6.32a}$$

$$\underbrace{\frac{\partial}{\partial\mathbf{u}}[\mathbf{A}\mathbf{u}]=\mathbf{A}}_{\text{or } \partial(A_{im}u_m)/\partial u_j=A_{ij}}\quad , \quad \underbrace{\frac{\partial}{\partial\mathbf{A}}[\mathbf{A}\mathbf{u}]=\mathbf{I}\otimes\mathbf{u}}_{\text{or } \partial(A_{im}u_m)/\partial A_{jk}=\delta_{ij}u_k}\quad , \tag{6.32b}$$

$$\underbrace{\frac{\partial}{\partial\mathbf{u}}\left[\mathbf{A}^2\mathbf{u}\right]=\mathbf{A}^2}_{\text{or } \partial(A_{im}A_{mn}u_n)/\partial u_j=A_{im}A_{mj}}\quad , \quad \underbrace{\frac{\partial}{\partial\mathbf{A}}\left[\mathbf{A}^2\mathbf{u}\right]=\mathbf{A}\otimes\mathbf{u}+\mathbf{I}\otimes\mathbf{A}\mathbf{u}}_{\text{or } \partial(A_{im}A_{mn}u_n)/\partial A_{jk}=A_{ij}u_k+\delta_{ij}A_{kn}u_n}\quad . \ \square \tag{6.32c}$$

Using (5.64a)–(5.64b) and (5.78), the widely used derivative of a vector with respect to itself, according to (6.32a), also admits the following representations in indicial notation,

$$\underbrace{\frac{\partial\,\underline{u}^i}{\partial\,\underline{u}_j}=g^{ij}\quad , \quad \frac{\partial\,\underline{u}^i}{\partial\,\underline{u}^j}=\delta^i_j\quad , \quad \frac{\partial\,\underline{u}_i}{\partial\,\underline{u}_j}=\delta^j_i\quad , \quad \frac{\partial\,\underline{u}_i}{\partial\,\underline{u}^j}=g_{ij}}_{\text{recall the components relationships } \underline{u}^i=g^{ij}\,\underline{u}_j\ ,\ \underline{u}^i=\delta^i_j\,\underline{u}^j\ ,\ \underline{u}_i=\delta^j_i\,\underline{u}_j \text{ and } \underline{u}_i=g_{ij}\,\underline{u}^j}\quad . \tag{6.33}$$

The relation $(6.33)_1$ may be viewed as a new definition for the contravariant metric in $(5.46)_1$. In a similar manner, $(6.33)_2$ may be seen as an alternative to $(5.15)_2$.

6.1.3 Gradient of a Tensor-Valued Function

Let $\tilde{\mathbf{H}} : \mathrm{R} \times \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3} \times \mathscr{T}_{\mathrm{so}} \longrightarrow \mathscr{T}_{\mathrm{so}}$ be a nonlinear and sufficiently smooth tensor-valued function of the three tensorial variables Φ, $\mathbf{u}$ and $\mathbf{A}$. The first-order Taylor series expansion of $\tilde{\mathbf{H}}$ around $(\Phi, \mathbf{u}, \mathbf{A})$ renders

$$\tilde{\mathbf{H}}\left(\Phi + d\Phi, \mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}\right) = \tilde{\mathbf{H}}\left(\Phi, \mathbf{u}, \mathbf{A}\right) + d\tilde{\mathbf{H}} + o\left(d\Phi, d\mathbf{u}, d\mathbf{A}\right) \ , \tag{6.34}$$

where the total differential

$$\boxed{d\tilde{\mathbf{H}} = \frac{\partial \tilde{\mathbf{H}}}{\partial \Phi} d\Phi + \frac{\partial \tilde{\mathbf{H}}}{\partial \mathbf{u}} d\mathbf{u} + \frac{\partial \tilde{\mathbf{H}}}{\partial \mathbf{A}} : d\mathbf{A} \ ,} \tag{6.35}$$

as well as the reminder $o\left(d\Phi, d\mathbf{u}, d\mathbf{A}\right)$ represent tensors of rank two. Accordingly, the partial derivative of $\tilde{\mathbf{H}}$ with respect to Φ in (6.35) also characterizes a tensor. Therein, $\partial \tilde{\mathbf{H}} / \partial \mathbf{u}$ and $\partial \tilde{\mathbf{H}} / \partial \mathbf{A}$ consistently present a third- and fourth-order tensor, respectively. One can also determine these gradients in a systematic manner via

$$D_{\phi} \tilde{\mathbf{H}}\left(\Phi, \mathbf{u}, \mathbf{A}\right) := \left. \frac{d}{d\varepsilon} \tilde{\mathbf{H}}\left(\Phi + \varepsilon\phi, \mathbf{u}, \mathbf{A}\right) \right|_{\varepsilon=0} = \frac{\partial \tilde{\mathbf{H}}}{\partial \Phi} \phi \ , \tag{6.36a}$$

$$D_{\mathbf{v}} \tilde{\mathbf{H}}\left(\Phi, \mathbf{u}, \mathbf{A}\right) := \left. \frac{d}{d\varepsilon} \tilde{\mathbf{H}}\left(\Phi, \mathbf{u} + \varepsilon\mathbf{v}, \mathbf{A}\right) \right|_{\varepsilon=0} = \frac{\partial \tilde{\mathbf{H}}}{\partial \mathbf{u}} \mathbf{v} \ , \tag{6.36b}$$

$$D_{\mathbf{B}} \tilde{\mathbf{H}}\left(\Phi, \mathbf{u}, \mathbf{A}\right) := \left. \frac{d}{d\varepsilon} \tilde{\mathbf{H}}\left(\Phi, \mathbf{u}, \mathbf{A} + \varepsilon\mathbf{B}\right) \right|_{\varepsilon=0} = \frac{\partial \tilde{\mathbf{H}}}{\partial \mathbf{A}} : \mathbf{B} \ , \tag{6.36c}$$

or, the useful form,

$$d\tilde{\mathbf{H}} = \left. \frac{d}{d\varepsilon} \tilde{\mathbf{H}}\left(\Phi + \varepsilon d\Phi, \mathbf{u} + \varepsilon d\mathbf{u}, \mathbf{A} + \varepsilon d\mathbf{A}\right) \right|_{\varepsilon=0} \ . \tag{6.37}$$

As a result, the linearized form of (6.34) renders

$$\tilde{\mathbf{H}}\left(\Phi + d\Phi, \mathbf{u} + d\mathbf{u}, \mathbf{A} + d\mathbf{A}\right) \approx \tilde{\mathbf{H}}\left(\Phi, \mathbf{u}, \mathbf{A}\right) + \underbrace{\frac{\partial \hat{\mathbf{H}}}{\partial \Phi} d\Phi + \frac{\partial \hat{\mathbf{H}}}{\partial \mathbf{u}} d\mathbf{u} + \frac{\partial \hat{\mathbf{H}}}{\partial \mathbf{A}} : d\mathbf{A}}_{= \left. \frac{d}{d\varepsilon} \hat{\mathbf{H}}\left(\Phi + \varepsilon d\Phi, \mathbf{u} + \varepsilon d\mathbf{u}, \mathbf{A} + \varepsilon d\mathbf{A}\right) \right|_{\varepsilon=0}} \ . \tag{6.38}$$

Guided by the expressions (6.29) and (6.30), the Cartesian as well as curvilinear forms of $\partial\tilde{\mathbf{H}}/\partial\mathbf{u}$ and $\partial\tilde{\mathbf{H}}/\partial\mathbf{A}$ can be expressed as

$$\begin{aligned}
\frac{\partial\tilde{\mathbf{H}}}{\partial\mathbf{u}} &= \frac{\partial\tilde{H}_{ij}}{\partial u_k}\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k \quad \longleftarrow \text{see (3.13) and (5.96a)-(5.96d)}\\
&= \frac{\partial\,\underline{\tilde{H}}^{ij}}{\partial\,\underline{u}_k}\mathbf{g}_i\otimes\mathbf{g}_j\otimes\mathbf{g}_k = \frac{\partial\,\underline{\tilde{H}}^{i}_{.j}}{\partial\,\underline{u}_k}\mathbf{g}_i\otimes\mathbf{g}^j\otimes\mathbf{g}_k\\
&= \frac{\partial\,\underline{\tilde{H}}^{i}_{.j}}{\partial\,\underline{u}^k}\mathbf{g}_i\otimes\mathbf{g}^j\otimes\mathbf{g}^k = \frac{\partial\,\underline{\tilde{H}}_{ij}}{\partial\,\underline{u}^k}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}^k\ ,
\end{aligned}\tag{6.39}$$

$$\begin{aligned}
\frac{\partial\tilde{\mathbf{H}}}{\partial\mathbf{A}} &= \frac{\partial\tilde{H}_{ij}}{\partial A_{kl}}\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l \quad \longleftarrow \text{see (3.62) and (5.99a)-(5.99h)}\\
&= \frac{\partial\,\underline{\tilde{H}}^{ij}}{\partial\,\underline{A}_{kl}}\mathbf{g}_i\otimes\mathbf{g}_j\otimes\mathbf{g}_k\otimes\mathbf{g}_l = \frac{\partial\,\underline{\tilde{H}}^{ij}}{\partial\,\underline{A}^{k}_{.l}}\mathbf{g}_i\otimes\mathbf{g}_j\otimes\mathbf{g}^k\otimes\mathbf{g}_l = \frac{\partial\,\underline{\tilde{H}}^{ij}}{\partial\,\underline{A}^{kl}}\mathbf{g}_i\otimes\mathbf{g}_j\otimes\mathbf{g}^k\otimes\mathbf{g}^l\\
&= \frac{\partial\,\underline{\tilde{H}}^{i}_{.j}}{\partial\,\underline{A}^{k}_{.l}}\mathbf{g}_i\otimes\mathbf{g}^j\otimes\mathbf{g}^k\otimes\mathbf{g}_l = \frac{\partial\,\underline{\tilde{H}}_{ij}}{\partial\,\underline{A}_{kl}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_k\otimes\mathbf{g}_l = \frac{\partial\,\underline{\tilde{H}}_{ij}}{\partial\,\underline{A}^{k}_{.l}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}^k\otimes\mathbf{g}_l\\
&= \frac{\partial\,\underline{\tilde{H}}_{ij}}{\partial\,\underline{A}^{kl}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}^k\otimes\mathbf{g}^l = \frac{\partial\,\underline{\tilde{H}}^{i}_{.j}}{\partial\,\underline{A}^{kl}}\mathbf{g}_i\otimes\mathbf{g}^j\otimes\mathbf{g}^k\otimes\mathbf{g}^l\ .
\end{aligned}\tag{6.40}$$

The procedure to compute the partial derivatives of a tensor-valued function is illustrated in the following examples. ¶

The first example regards the tensor-valued function

$$\tilde{\mathbf{H}}(\mathbf{A}) = \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3\ . \tag{6.41}$$

The desired derivatives will be calculated here by using the differential form (6.37). Thus,

$$\begin{aligned}
\tilde{\mathbf{H}}(\mathbf{A}+\varepsilon d\mathbf{A}) = \varepsilon^0\left(\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3\right) &\\
+\ \underbrace{\varepsilon^1\left(d\mathbf{A}+d\mathbf{A}\mathbf{A}+\mathbf{A}d\mathbf{A}+d\mathbf{A}\mathbf{A}^2+\mathbf{A}d\mathbf{A}\mathbf{A}+\mathbf{A}^2d\mathbf{A}\right)}_{=\varepsilon^1\left(\mathbf{I}\boxtimes\mathbf{I}+\mathbf{I}\boxtimes\mathbf{A}+\mathbf{A}\boxtimes\mathbf{I}+\mathbf{I}\boxtimes\mathbf{A}^2+\mathbf{A}\boxtimes\mathbf{A}+\mathbf{A}^2\boxtimes\mathbf{I}\right):d\mathbf{A},\ \text{see (3.54b)}} &\\
+\ \varepsilon^2\left(d\mathbf{A}d\mathbf{A}+d\mathbf{A}d\mathbf{A}\mathbf{A}+d\mathbf{A}\mathbf{A}d\mathbf{A}+\mathbf{A}d\mathbf{A}d\mathbf{A}\right) &\\
+\ \varepsilon^3\left(d\mathbf{A}d\mathbf{A}d\mathbf{A}\right)\ , &
\end{aligned}$$

helps obtain

$$\frac{\partial\mathbf{A}}{\partial\mathbf{A}} = \underbrace{\mathbf{I}\boxtimes\mathbf{I}}_{\text{or } \delta_{ik}\delta_{lj} = \partial A_{ij}/\partial A_{kl}} \ \overset{\text{from}}{\underset{(3.57)}{=\!=}}\ \mathbb{I}\ , \tag{6.42a}$$

$$\frac{\partial \mathbf{A}^2}{\partial \mathbf{A}} = \underbrace{\mathbf{I} \boxtimes \mathbf{A} + \mathbf{A} \boxtimes \mathbf{I}}_{\text{or } \delta_{ik} A_{lj} + A_{ik} \delta_{lj} = \partial A^2_{ij} / \partial A_{kl}} , \tag{6.42b}$$

$$\frac{\partial \mathbf{A}^3}{\partial \mathbf{A}} = \underbrace{\mathbf{I} \boxtimes \mathbf{A}^2 + \mathbf{A} \boxtimes \mathbf{A} + \mathbf{A}^2 \boxtimes \mathbf{I}}_{\text{or } \delta_{ik} A^2_{lj} + A_{ik} A_{lj} + A^2_{ik} \delta_{lj} = \partial A^3_{ij} / \partial A_{kl}} . \tag{6.42c}$$

The first result here can be used to derive the last one. This will be carried out in the following for the interested reader:

$$\begin{aligned} d\mathbf{A}^3 &\xlongequal[\text{(6.4h) and (6.8e)}]{\text{from}} \left(\frac{\partial \mathbf{A}}{\partial \mathbf{A}} : d\mathbf{A} \right) \mathbf{A}^2 + \mathbf{A} \left(\frac{\partial \mathbf{A}}{\partial \mathbf{A}} : d\mathbf{A} \right) \mathbf{A} + \mathbf{A}^2 \left(\frac{\partial \mathbf{A}}{\partial \mathbf{A}} : d\mathbf{A} \right) \\ &\xlongequal[\text{(6.42a)}]{\text{from}} \left(\mathbf{I} \boxtimes \mathbf{I} : d\mathbf{A} \right) \mathbf{A}^2 + \mathbf{A} \left(\mathbf{I} \boxtimes \mathbf{I} : d\mathbf{A} \right) \mathbf{A} + \mathbf{A}^2 \left(\mathbf{I} \boxtimes \mathbf{I} : d\mathbf{A} \right) \\ &\xlongequal[\text{(3.40a) and (3.103b)}]{\text{from}} \left[\left(\mathbf{I} \boxtimes \mathbf{I} \right) \mathbf{A}^2 + \mathbf{A} \left(\mathbf{I} \boxtimes \mathbf{I} \right) \mathbf{A} + \mathbf{A}^2 \left(\mathbf{I} \boxtimes \mathbf{I} \right) \right] : d\mathbf{A} \\ &\xlongequal[\text{(2.33) and (3.101b)}]{\text{from}} \left[\mathbf{I} \boxtimes \mathbf{A}^2 + \mathbf{A} \boxtimes \mathbf{A} + \mathbf{A}^2 \boxtimes \mathbf{I} \right] : d\mathbf{A} . \end{aligned} \tag{6.43}$$

The derivation here motivates to establish the following identities. Let $\mathbf{A}$ and $\mathbf{B}$ be two sufficiently smooth tensor-valued functions of one tensor variable $\mathbf{C}$. Then, using the product rule,

$$\frac{\partial}{\partial \mathbf{C}} [\mathbf{AB}] = (\mathbf{I} \boxtimes \mathbf{B}) : \frac{\partial \mathbf{A}}{\partial \mathbf{C}} + (\mathbf{A} \boxtimes \mathbf{I}) : \frac{\partial \mathbf{B}}{\partial \mathbf{C}} . \tag{6.44}$$

This is basically the direct notation of

$$\begin{aligned} \frac{\partial \left[A_{im} B_{mj} \right]}{\partial C_{kl}} &= \delta_{in} B_{mj} \frac{\partial A_{nm}}{\partial C_{kl}} + A_{im} \delta_{nj} \frac{\partial B_{mn}}{\partial C_{kl}} \quad \longleftarrow \text{see Sect. 6.1.4} \\ &= (\mathbf{I} \boxtimes \mathbf{B})_{ijnm} \left(\frac{\partial \mathbf{A}}{\partial \mathbf{C}} \right)_{nmkl} + (\mathbf{A} \boxtimes \mathbf{I})_{ijmn} \left(\frac{\partial \mathbf{B}}{\partial \mathbf{C}} \right)_{mnkl} . \end{aligned} \tag{6.45}$$

See also (6.190a)–(6.190d) with the corresponding derivations.

Further, let $\mathbf{A}\,(\mathbf{B}\,(\mathbf{C}))$ be a sufficiently smooth tensor-valued function of the differentiable tensor function $\mathbf{B}\,(\mathbf{C})$. Then, using the chain rule,

$$\underbrace{\frac{\partial \mathbf{A}}{\partial \mathbf{C}} = \frac{\partial \mathbf{A}}{\partial \mathbf{B}} : \frac{\partial \mathbf{B}}{\partial \mathbf{C}}}_{\text{or } \partial A_{ij} / \partial C_{kl} = \left(\partial A_{ij} / \partial B_{mn} \right) \left(\partial B_{mn} / \partial C_{kl} \right)} . \tag{6.46}$$

Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be three arbitrary tensors. Then, by means of $(2.33)_{1-2}$, $(3.54\text{b})_1$, (3.103b) and $(6.42\text{a})_1$, one can establish

$$\mathbf{B}\frac{\partial \mathbf{A}}{\partial \mathbf{A}}\mathbf{C} : \mathbf{D} = \mathbf{B}\left(\mathbf{I} \boxtimes \mathbf{I}\right)\mathbf{C} : \mathbf{D} = \mathbf{B}\left[\left(\mathbf{I} \boxtimes \mathbf{I}\right) : \mathbf{D}\right]\mathbf{C} = \mathbf{B}\left[\mathbf{IDI}\right]\mathbf{C} = \mathbf{BDC} \;, \tag{6.47}$$

which is eventually the absolute notation of

$$B_{im}\frac{\partial A_{mn}}{\partial A_{kl}}C_{nj}D_{kl} = B_{ik}D_{kl}C_{lj} \;. \tag{6.48}$$

In a similar manner, the following identities are implied

$$\mathbf{B}\frac{\partial \mathbf{A}^2}{\partial \mathbf{A}}\mathbf{C} : \mathbf{D} = \mathbf{BDAC} + \mathbf{BADC} \;, \tag{6.49a}$$

$$\mathbf{B}\frac{\partial \mathbf{A}^3}{\partial \mathbf{A}}\mathbf{C} : \mathbf{D} = \mathbf{BDA}^2\mathbf{C} + \mathbf{BADAC} + \mathbf{BA}^2\mathbf{DC} \;. \tag{6.49b}$$

As another example, consider the tensor-valued function

$$\tilde{\mathbf{H}}\left(\mathbf{A}\right) = \mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{A} + \mathbf{A}\mathbf{A}^{\mathrm{T}} \;. \tag{6.50}$$

It follows that the derivative of

$$\begin{aligned}\tilde{\mathbf{H}}\left(\mathbf{A} + \varepsilon d\mathbf{A}\right) = &\left(\mathbf{A} + \varepsilon d\mathbf{A}\right)^{\mathrm{T}} + \left(\mathbf{A} + \varepsilon d\mathbf{A}\right)^{\mathrm{T}}\left(\mathbf{A} + \varepsilon d\mathbf{A}\right) + \left(\mathbf{A} + \varepsilon d\mathbf{A}\right)\left(\mathbf{A} + \varepsilon d\mathbf{A}\right)^{\mathrm{T}} \\ = &\ \varepsilon^0\left(\mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{A} + \mathbf{A}\mathbf{A}^{\mathrm{T}}\right) \\ &+ \underbrace{\varepsilon^1\left(d\mathbf{A}^{\mathrm{T}} + d\mathbf{A}^{\mathrm{T}}\mathbf{A} + \mathbf{A}^{\mathrm{T}}d\mathbf{A} + d\mathbf{A}\mathbf{A}^{\mathrm{T}} + \mathbf{A}d\mathbf{A}^{\mathrm{T}}\right)}_{= \varepsilon^1\left(\mathbf{I} \boxplus \mathbf{I} + \mathbf{I} \boxplus \mathbf{A} + \mathbf{A}^{\mathrm{T}} \boxtimes \mathbf{I} + \mathbf{I} \boxtimes \mathbf{A}^{\mathrm{T}} + \mathbf{A} \boxplus \mathbf{I}\right) : d\mathbf{A},\ \text{according to (3.54b)-(3.54c)}} \\ &+ \varepsilon^2\left(d\mathbf{A}^{\mathrm{T}}d\mathbf{A} + d\mathbf{A}d\mathbf{A}^{\mathrm{T}}\right) \;,\end{aligned}$$

with respect to ε at $\varepsilon = 0$ helps obtain

$$\frac{\partial \mathbf{A}^{\mathrm{T}}}{\partial \mathbf{A}} = \underbrace{\mathbf{I} \boxplus \mathbf{I}}_{\text{or } \delta_{jk}\delta_{li} = \partial A_{ji}/\partial A_{kl} = \partial A^{\mathrm{T}}_{ij}/\partial A_{kl}} \overset{\text{from}}{\underset{(3.57)}{=\!=}} \bar{\mathbb{I}} \;, \tag{6.51a}$$

$$\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{A}^{\mathrm{T}}\mathbf{A}\right] = \underbrace{\mathbf{I} \boxplus \mathbf{A} + \mathbf{A}^{\mathrm{T}} \boxtimes \mathbf{I}}_{\text{or } \delta_{li}A_{kj} + A_{ki}\delta_{lj} = \partial\left(A_{mi}A_{mj}\right)/\partial A_{kl}} \;, \tag{6.51b}$$

$$\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{A}\mathbf{A}^{\mathrm{T}}\right] = \underbrace{\mathbf{I} \boxtimes \mathbf{A}^{\mathrm{T}} + \mathbf{A} \boxplus \mathbf{I}}_{\text{or } \delta_{ik}A_{jl} + A_{il}\delta_{jk} = \partial\left(A_{im}A_{jm}\right)/\partial A_{kl}} \;. \tag{6.51c}$$

The interested reader may want to use the first result to arrive at the second one:

$$d\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) \xlongequal[\text{(6.4h) and (6.8e)}]{\text{from}} \left(\frac{\partial \mathbf{A}^{\mathrm{T}}}{\partial \mathbf{A}} : d\mathbf{A}\right)\mathbf{A} + \mathbf{A}^{\mathrm{T}}\left(\frac{\partial \mathbf{A}}{\partial \mathbf{A}} : d\mathbf{A}\right)$$
$$\xlongequal[\text{(6.42a) and (6.51a)}]{\text{from}} (\mathbf{I} \boxplus \mathbf{I} : d\mathbf{A})\,\mathbf{A} + \mathbf{A}^{\mathrm{T}}(\mathbf{I} \boxtimes \mathbf{I} : d\mathbf{A})$$
$$\xlongequal[\text{(3.40a) and (3.103b)}]{\text{from}} \left[(\mathbf{I} \boxplus \mathbf{I})\,\mathbf{A} + \mathbf{A}^{\mathrm{T}}(\mathbf{I} \boxtimes \mathbf{I})\right] : d\mathbf{A}$$
$$\xlongequal[\text{(2.33) and (3.101c)}]{\text{from}} \left[\mathbf{I} \boxplus \mathbf{A} + \mathbf{A}^{\mathrm{T}} \boxtimes \mathbf{I}\right] : d\mathbf{A} \,. \tag{6.52}$$

Suppose that $\mathbf{A}$ in (6.41) satisfies $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$. Recall that the symmetric tensor was denoted by $\mathbf{S}$ for the sake of concise notation, see (2.62). Accordingly, the gradients in (6.42a)–(6.42c) translate to

$$\frac{\partial \mathbf{S}}{\partial \mathbf{S}} = \underbrace{\mathbf{I} \odot \mathbf{I}}_{\text{or } \delta_{ik}\delta_{lj} + \delta_{il}\delta_{kj} = 2\partial S_{ij}/\partial S_{kl}} \xlongequal[\text{(3.58a)}]{\text{from}} \frac{1}{2}\left(\mathbb{I} + \bar{\mathbb{I}}\right) , \tag{6.53a}$$

$$\frac{\partial \mathbf{S}^2}{\partial \mathbf{S}} = \underbrace{\mathbf{I} \odot \mathbf{S} + \mathbf{S} \odot \mathbf{I}}_{\text{or } \delta_{ik}S_{lj} + \delta_{il}S_{kj} + S_{ik}\delta_{lj} + S_{il}\delta_{kj} = 2\partial S^2_{ij}/\partial S_{kl}} , \tag{6.53b}$$

$$\frac{\partial \mathbf{S}^3}{\partial \mathbf{S}} = \underbrace{\mathbf{I} \odot \mathbf{S}^2 + \mathbf{S} \odot \mathbf{S} + \mathbf{S}^2 \odot \mathbf{I}}_{\text{or } \delta_{ik}S^2_{lj} + \delta_{il}S^2_{kj} + S_{ik}S_{lj} + S_{il}S_{kj} + S^2_{ik}\delta_{lj} + S^2_{il}\delta_{kj} = 2\partial S^3_{ij}/\partial S_{kl}} . \tag{6.53c}$$

Note that these partial derivatives render **super-symmetric** tensors possessing major and minor symmetries as illustrated in (3.115). Observe that the partial derivative of a generally unsymmetric tensor with respect to itself, according to (6.42a), is constructed by means of the tensor product $\boxtimes$. But, its super-symmetric form, given in (6.53a), is generated by use of $\odot$. This is due to the fact that any symmetric tensor $\mathbf{S}$ satisfies $2d\mathbf{S} = d\mathbf{S} + d\mathbf{S}^{\mathrm{T}}$. And this helps infer that

$$[\mathbf{A} \boxtimes \mathbf{B}] : d\mathbf{S} = [\mathbf{A} \boxtimes \mathbf{B}] : \frac{d\mathbf{S} + d\mathbf{S}^{\mathrm{T}}}{2}$$
$$\xlongequal[\text{(3.107)}]{\text{from}} \frac{1}{2}\left[(\mathbf{A} \boxtimes \mathbf{B}) + (\mathbf{A} \boxtimes \mathbf{B})^{\hat{\mathrm{T}}}\right] : d\mathbf{S}$$
$$\xlongequal[\text{(3.110g)}]{\text{from}} \frac{1}{2}\left[\mathbf{A} \boxtimes \mathbf{B} + \mathbf{A} \boxplus \mathbf{B}\right] : d\mathbf{S}$$
$$\xlongequal[\text{(3.70d)}]{\text{from}} [\mathbf{A} \odot \mathbf{B}] : d\mathbf{S} \,. \tag{6.54}$$

With regard to the various representations of a fourth-order tensor in curvilinear coordinates according to (5.99a)–(5.99h) and by use of (5.78), (6.42a) and (6.53a), one will have the useful relations

$$\frac{\partial \underline{A}^{ij}}{\partial \underline{A}_{kl}} = g^{ik} g^{lj} , \qquad \frac{\partial \underline{S}^{ij}}{\partial \underline{S}_{kl}} = \frac{1}{2} \left(g^{ik} g^{lj} + g^{il} g^{kj} \right) , \tag{6.55a}$$

$$\frac{\partial \underline{A}^{ij}}{\partial \underline{A}^{k}_{.l}} = \delta^{i}_{k} g^{lj} , \qquad \frac{\partial \underline{S}^{ij}}{\partial \underline{S}^{k}_{.l}} = \frac{1}{2} \left(\delta^{i}_{k} g^{lj} + g^{il} \delta^{j}_{k} \right) , \tag{6.55b}$$

$$\frac{\partial \underline{A}^{ij}}{\partial \underline{A}^{kl}} = \delta^{i}_{k} \delta^{j}_{l} , \qquad \frac{\partial \underline{S}^{ij}}{\partial \underline{S}^{kl}} = \frac{1}{2} \left(\delta^{i}_{k} \delta^{j}_{l} + \delta^{i}_{l} \delta^{j}_{k} \right) , \tag{6.55c}$$

$$\frac{\partial \underline{A}^{i}_{.j}}{\partial \underline{A}^{k}_{.l}} = \delta^{i}_{k} \delta^{l}_{j} , \qquad \frac{\partial \underline{S}^{i}_{.j}}{\partial \underline{S}^{k}_{.l}} = \frac{1}{2} \left(\delta^{i}_{k} \delta^{l}_{j} + g^{il} g_{kj} \right) , \tag{6.55d}$$

$$\frac{\partial \underline{A}_{ij}}{\partial \underline{A}_{kl}} = \delta^{k}_{i} \delta^{l}_{j} , \qquad \frac{\partial \underline{S}_{ij}}{\partial \underline{S}_{kl}} = \frac{1}{2} \left(\delta^{k}_{i} \delta^{l}_{j} + \delta^{l}_{i} \delta^{k}_{j} \right) , \tag{6.55e}$$

$$\frac{\partial \underline{A}_{ij}}{\partial \underline{A}^{k}_{.l}} = g_{ik} \delta^{l}_{j} , \qquad \frac{\partial \underline{S}_{ij}}{\partial \underline{S}^{k}_{.l}} = \frac{1}{2} \left(g_{ik} \delta^{l}_{j} + \delta^{l}_{i} g_{kj} \right) , \tag{6.55f}$$

$$\frac{\partial \underline{A}_{ij}}{\partial \underline{A}^{kl}} = g_{ik} g_{lj} , \qquad \frac{\partial \underline{S}_{ij}}{\partial \underline{S}^{kl}} = \frac{1}{2} \left(g_{ik} g_{lj} + g_{il} g_{kj} \right) , \tag{6.55g}$$

$$\frac{\partial \underline{A}^{i}_{.j}}{\partial \underline{A}^{kl}} = \delta^{i}_{k} g_{lj} , \qquad \frac{\partial \underline{S}^{i}_{.j}}{\partial \underline{S}^{kl}} = \frac{1}{2} \left(\delta^{i}_{k} g_{lj} + \delta^{i}_{l} g_{kj} \right) . \tag{6.55h}$$

In what follows, the partial derivative of the inverse of a tensor with respect to itself is characterized.

It is evident that the total differential of the identity tensor vanishes, i.e. $d\mathbf{I} = \mathbf{O}$. An incremental change in $\mathbf{A}\mathbf{A}^{-1}$, using the product rule of differentiation according to (6.4h), thus gives

$$d\mathbf{A}\mathbf{A}^{-1} + \mathbf{A}d\mathbf{A}^{-1} = \mathbf{O} . \tag{6.56}$$

Premultiplying the above result by $\mathbf{A}^{-1}$ leads to

$$d\mathbf{A}^{-1} = -\mathbf{A}^{-1} \left(d\mathbf{A} \right) \mathbf{A}^{-1} = \left(-\mathbf{A}^{-1} \boxtimes \mathbf{A}^{-1} \right) : d\mathbf{A} . \tag{6.57}$$

Thus,

$$\boxed{\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} = \underbrace{-\mathbf{A}^{-1} \boxtimes \mathbf{A}^{-1}}_{\text{or } -A^{-1}_{ik} A^{-1}_{lj} = \partial A^{-1}_{ij} / \partial A_{kl}}} . \tag{6.58}$$

In a similar fashion,

$$\boxed{\frac{\partial \mathbf{A}}{\partial \mathbf{A}^{-1}} = \underbrace{-\mathbf{A} \boxtimes \mathbf{A}}_{\text{or } -A_{ik} A_{lj} = \partial A_{ij} / \partial A^{-1}_{kl}}} . \tag{6.59}$$

The expressions (3.54b), (6.58) and (6.59) help obtain

$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} : \mathbf{B} = -\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}, \qquad \mathbf{B} : \frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} = -\mathbf{A}^{-\mathrm{T}}\mathbf{B}\mathbf{A}^{-\mathrm{T}} , \tag{6.60a}$$

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}^{-1}} : \mathbf{B} = -\mathbf{A}\mathbf{B}\mathbf{A}, \qquad \mathbf{B} : \frac{\partial \mathbf{A}}{\partial \mathbf{A}^{-1}} = -\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{A}^{\mathrm{T}} . \tag{6.60b}$$

It is then a simple exercise to verify that

$$\frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} : \mathbf{A} \otimes \mathbf{A}^{-1} = -\mathbf{A}^{-1} \otimes \mathbf{A}^{-1} , \tag{6.61a}$$

$$\mathbf{A}^{-\mathrm{T}} \otimes \mathbf{A}^{\mathrm{T}} : \frac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} = -\mathbf{A}^{-\mathrm{T}} \otimes \mathbf{A}^{-\mathrm{T}} , \tag{6.61b}$$

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}^{-1}} : \mathbf{A}^{-1} \otimes \mathbf{A} = -\mathbf{A} \otimes \mathbf{A} , \tag{6.61c}$$

$$\mathbf{A}^{\mathrm{T}} \otimes \mathbf{A}^{-\mathrm{T}} : \frac{\partial \mathbf{A}}{\partial \mathbf{A}^{-1}} = -\mathbf{A}^{\mathrm{T}} \otimes \mathbf{A}^{\mathrm{T}} . \tag{6.61d}$$

Moreover,

$$\frac{\partial \mathbf{A}^{-2}}{\partial \mathbf{A}} = \underbrace{-\mathbf{A}^{-1} \boxtimes \mathbf{A}^{-2} - \mathbf{A}^{-2} \boxtimes \mathbf{A}^{-1}}_{\text{or } -A^{-1}_{ik}A^{-2}_{lj} - A^{-2}_{ik}A^{-1}_{lj} = \partial A^{-2}_{ij}/\partial A_{kl}} , \tag{6.62a}$$

$$\frac{\partial \mathbf{A}^{-3}}{\partial \mathbf{A}} = \underbrace{-\mathbf{A}^{-1} \boxtimes \mathbf{A}^{-3} - \mathbf{A}^{-2} \boxtimes \mathbf{A}^{-2} - \mathbf{A}^{-3} \boxtimes \mathbf{A}^{-1}}_{\text{or } -A^{-1}_{ik}A^{-3}_{lj} - A^{-2}_{ik}A^{-2}_{lj} - A^{-3}_{ik}A^{-1}_{lj} = \partial A^{-3}_{ij}/\partial A_{kl}} , \tag{6.62b}$$

$$\frac{\partial \mathbf{A}^{-\mathrm{T}}}{\partial \mathbf{A}} = \underbrace{-\mathbf{A}^{-\mathrm{T}} \boxplus \mathbf{A}^{-\mathrm{T}}}_{\text{or } -A^{-1}_{li}A^{-1}_{jk} = \partial A^{-1}_{ji}/\partial A_{kl} = \partial A^{-\mathrm{T}}_{ij}/\partial A_{kl}} , \tag{6.62c}$$

$$\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1}\right] = \underbrace{-\mathbf{A}^{-\mathrm{T}} \boxplus \left(\mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1}\right) - \left(\mathbf{A}^{-\mathrm{T}}\mathbf{A}^{-1}\right) \boxtimes \mathbf{A}^{-1}}_{\text{or } -A^{-1}_{mk}A^{-1}_{li}A^{-1}_{mj} - A^{-1}_{mi}A^{-1}_{mk}A^{-1}_{lj} = \partial\left(A^{-1}_{mi}A^{-1}_{mj}\right)/\partial A_{kl}} , \tag{6.62d}$$

$$\frac{\partial}{\partial \mathbf{A}}\left[\mathbf{A}^{-1}\mathbf{A}^{-\mathrm{T}}\right] = \underbrace{-\mathbf{A}^{-1} \boxtimes \left(\mathbf{A}^{-1}\mathbf{A}^{-\mathrm{T}}\right) - \left(\mathbf{A}^{-1}\mathbf{A}^{-\mathrm{T}}\right) \boxplus \mathbf{A}^{-\mathrm{T}}}_{\text{or } -A^{-1}_{ik}A^{-1}_{lm}A^{-1}_{jm} - A^{-1}_{im}A^{-1}_{jk}A^{-1}_{lm} = \partial\left(A^{-1}_{im}A^{-1}_{jm}\right)/\partial A_{kl}} . \tag{6.62e}$$

Regarding a symmetric tensor $\mathbf{S}$, the following relations hold true

$$\frac{\partial \mathbf{S}}{\partial \mathbf{S}^{-1}} = \underbrace{-\mathbf{S} \odot \mathbf{S}}_{\text{or } -\left[S_{ik}S_{lj} + S_{il}S_{kj}\right]/2 = \partial S_{ij}/\partial S^{-1}_{kl}} , \tag{6.63a}$$

$$\frac{\partial \mathbf{S}^{-1}}{\partial \mathbf{S}} = \underbrace{-\mathbf{S}^{-1} \odot \mathbf{S}^{-1}}_{\text{or } -\left[S^{-1}_{ik}S^{-1}_{lj} + S^{-1}_{il}S^{-1}_{kj}\right]/2 = \partial S^{-1}_{ij}/\partial S_{kl}} , \tag{6.63b}$$

$$\frac{\partial \mathbf{S}^{-2}}{\partial \mathbf{S}} = \underbrace{-\mathbf{S}^{-1} \odot \mathbf{S}^{-2} - \mathbf{S}^{-2} \odot \mathbf{S}^{-1}}_{\text{or } -\left[S^{-1}_{ik}S^{-2}_{lj} + S^{-1}_{il}S^{-2}_{kj} + S^{-2}_{ik}S^{-1}_{lj} + S^{-2}_{il}S^{-1}_{kj}\right]/2 = \partial S^{-2}_{ij}/\partial S_{kl}} , \tag{6.63c}$$

$$\frac{\partial \mathbf{S}^{-3}}{\partial \mathbf{S}} = \underbrace{-\mathbf{S}^{-1} \odot \mathbf{S}^{-3} - \mathbf{S}^{-2} \odot \mathbf{S}^{-2} - \mathbf{S}^{-3} \odot \mathbf{S}^{-1}}_{\text{or } -\left[S^{-1}_{ik}S^{-3}_{lj} + S^{-1}_{il}S^{-3}_{kj} + S^{-2}_{ik}S^{-2}_{lj} + S^{-2}_{il}S^{-2}_{kj} + S^{-3}_{ik}S^{-1}_{lj} + S^{-3}_{il}S^{-1}_{kj}\right]/2 = \partial S^{-3}_{ij}/\partial S_{kl}} . \tag{6.63d}$$

In light of (3.70d)$_2$ and by using (6.40)$_8$ and (6.63a), the partial derivative of the covariant metric tensor (5.80) with respect to the contravariant metric tensor (5.81) renders

$$\boxed{\frac{\partial \mathbf{g}}{\partial \mathbf{g}^{-1}} = \frac{\partial g_{ij}}{\partial g^{kl}} \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l = -\mathbf{g} \odot \mathbf{g} ,} \tag{6.64}$$

with

$$\boxed{\frac{\partial g_{ij}}{\partial g^{kl}} = -\frac{1}{2}\left(g_{ik}g_{lj} + g_{il}g_{kj}\right) .} \tag{6.65}$$

In a similar fashion,

$$\boxed{\frac{\partial \mathbf{g}^{-1}}{\partial \mathbf{g}} = \underbrace{\frac{\partial g^{ij}}{\partial g_{kl}} \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k \otimes \mathbf{g}_l = -\mathbf{g}^{-1} \odot \mathbf{g}^{-1}}_{\text{with } \partial g^{ij}/\partial g_{kl} = -\left(g^{ik}g^{lj} + g^{il}g^{kj}\right)/2} .} \tag{6.66}$$

6.1.4 *Proof of Tensor Calculus Identities by Index Notation*

Verifying vector and tensor identities involving differentiation in direct notation based on various tensor products is a tedious task and often cumbersome. It is thus preferred to use indicial notation here and elsewhere for convenience. With regard to this, the derivatives of tensor functions can consistently be computed, under the standard rules of differentiation, by means of the principal terms

$$\boxed{\frac{\partial u_i}{\partial u_j} = \delta_{ij} \quad , \quad \frac{\partial A_{ij}}{\partial A_{kl}} = \delta_{ik}\delta_{lj} \quad , \quad \frac{\partial A_{ij}^{-1}}{\partial A_{kl}} = -A_{ik}^{-1} A_{lj}^{-1} \ ,} \tag{6.67}$$

as well as

$$\boxed{\frac{\partial S_{ij}}{\partial S_{kl}} = \frac{1}{2}\left(\delta_{ik}\delta_{lj} + \delta_{il}\delta_{kj}\right) \quad , \quad \frac{\partial S_{ij}^{-1}}{\partial S_{kl}} = -\frac{1}{2}\left(S_{ik}^{-1} S_{lj}^{-1} + S_{il}^{-1} S_{kj}^{-1}\right) \ .} \tag{6.68}$$

To use these relations, for instance, consider the last term in (6.17) whose gradients were presented in (6.18c). Here, they can readily be derived as follows:

$$\begin{aligned}
\frac{\partial\left(A_{jk}u_k A_{jl}u_l\right)}{\partial u_i} &= A_{jk}\frac{\partial u_k}{\partial u_i}A_{jl}u_l + A_{jk}u_k A_{jl}\frac{\partial u_l}{\partial u_i} \\
&= A_{ji}A_{jl}u_l + A_{jk}u_k A_{ji} \\
&= 2\left(\mathbf{A}^{\mathrm{T}}\right)_{ij}(\mathbf{A}\mathbf{u})_j \\
&= 2\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{u}\right)_i \ , \\
\frac{\partial\left(A_{km}u_m A_{kn}u_n\right)}{\partial A_{ij}} &= \frac{\partial A_{km}}{\partial A_{ij}}u_m A_{kn}u_n + A_{km}u_m\frac{\partial A_{kn}}{\partial A_{ij}}u_n \\
&= u_j A_{in}u_n + A_{im}u_m u_j \\
&= 2\,(\mathbf{A}\mathbf{u})_i\,(\mathbf{u})_j \\
&= 2\,(\mathbf{A}\mathbf{u}\otimes\mathbf{u})_{ij} \ .
\end{aligned}$$

Another example is

$$\frac{\partial \mathrm{tr}\ (\mathbf{BAC})}{\partial \mathbf{A}} = \mathbf{B}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} \quad , \quad \frac{\partial \mathrm{tr}\ \left(\mathbf{B}\mathbf{A}^2\mathbf{C}\right)}{\partial \mathbf{A}} = \mathbf{B}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} \ , \tag{6.69}$$

because, for instance,

$$\frac{\partial\left(B_{kl}A_{lm}C_{mk}\right)}{\partial A_{ij}} = B_{kl}\delta_{li}\delta_{jm}C_{mk} = B_{ki}C_{jk} = \left(\mathbf{B}^{\mathrm{T}}\right)_{ik}\left(\mathbf{C}^{\mathrm{T}}\right)_{kj} = \left(\mathbf{B}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\right)_{ij} \ ,$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are arbitrary tensors. It is worthwhile to point out that the curvilinear components of tensorial variables can also be used to verify vector and tensor calculus identities. But, the Cartesian type of components is often preferred for simplicity.

6.1.5 Numerical Differentiation

Analytically computing the partial derivatives of tensor functions is a difficult and tedious task. It often requires a lot of function evaluations which eventually can be error-prone. Examples of which in nonlinear solid mechanics include the consistent tangent modulus in multiplicative plasticity at finite strains and the macroscopic tangent in FE^2 method. On the contrary, numerical differentiation can be regarded as a competitive alternative due to its simplicity and robustness. This motivates to develop some closed-form formulas for approximating the gradients of tensor functions. For a detailed study of numerical differentiation in computational mechanics, see Hughes and Pister [1], Simo and Taylor [2], Miehe [3], Pérez-Foguet et al. [4] and Temizer and Wriggers [5] among many others.

All derivatives here will be approximated by use of **finite differences**. Recall that the **forward**, **backward** and **central** difference approximations of the first derivative of a scalar function of one scalar variable are

$$f'(x) \approx \frac{f(x+\varepsilon) - f(x)}{\varepsilon} := f^{\,\text{for}}(x) \ , \tag{6.70a}$$

$$f'(x) \approx \frac{f(x) - f(x-\varepsilon)}{\varepsilon} := f^{\,\text{bac}}(x) \ , \tag{6.70b}$$

$$f'(x) \approx \frac{f(x+\varepsilon) - f(x-\varepsilon)}{2\varepsilon} := f^{\,\text{cen}}(x) \ , \tag{6.70c}$$

where $\varepsilon \ll 1$ denotes the **perturbation parameter**. One of the important issues in numerical differentiation is the appropriate choice of perturbation parameter since it crucially affects the **true error**

$$\mathsf{err}_{f^\bullet} = f'(x) - f^\bullet(x) \quad \text{where} \quad \bullet = \text{for, bac, cen} \ . \tag{6.71}$$

This total error includes **truncation error** and **round-off error**. The former source of error is a mathematical error that arises from approximating an infinite sum by a finite one. It has basically been introduced as Landau order symbol in, for instance, (6.9). But the latter source of error evolves as a result of subtracting two nearly equal numbers due to the limited ability of digital computers in exactly representing numbers. The truncation error thus decreases when $\varepsilon \to 0$ whereas the round-off error increases when $\varepsilon \to 0$. This reveals the fact that there exists an optimal value minimizing the sum of these sources of error. In forward and backward difference approximations, computations have illustrated that the best choice for ε lies in the interval $\left[10^{-8}\,,\ 10^{-6}\right]$. While this is slightly larger for the central difference.

Guided by (6.70a)–(6.70c), the second-order forward, backward and central difference approximations of the second derivative of $f(x)$ will be

$$f''(x) \approx \frac{f'(x+\varepsilon) - f'(x)}{\varepsilon} := \overline{f}^{\,\text{for}}(x)$$

$$\approx \frac{f(x+2\varepsilon) - 2f(x+\varepsilon) + f(x)}{\varepsilon^2} := \widehat{f}^{\,\text{for}}(x) \;, \tag{6.72a}$$

$$f''(x) \approx \frac{f'(x) - f'(x-\varepsilon)}{\varepsilon} := \overline{f}^{\,\text{bac}}(x)$$

$$\approx \frac{f(x) - 2f(x-\varepsilon) + f(x-2\varepsilon)}{\varepsilon^2} := \widehat{f}^{\,\text{bac}}(x) \;, \tag{6.72b}$$

$$f''(x) \approx \frac{f'(x+\varepsilon) - f'(x-\varepsilon)}{2\varepsilon} := \overline{f}^{\,\text{cen}}(x)$$

$$\approx \frac{f(x+2\varepsilon) - 2f(x) + f(x-2\varepsilon)}{4\varepsilon^2} := \widehat{f}^{\,\text{cen}}(x) \;. \tag{6.72c}$$

And these approximations produce

$$\mathsf{err}_{\overline{f}^{\bullet}} = f''(x) - \overline{f}^{\,\bullet}(x) \;,\; \mathsf{err}_{\widehat{f}^{\bullet}} = f''(x) - \widehat{f}^{\,\bullet}(x) \quad \text{where} \quad \bullet = \text{for, bac, cen} \;. \tag{6.73}$$

Let $\Psi(\mathbf{C})$ be a nonlinear and sufficiently smooth scalar-valued function of one **symmetric** tensor variable $\mathbf{C}$. The goal is now to develop the tensorial analogues of the relations (6.70a)–(6.70c) and (6.72a)–(6.72c). The gradients to be approximated are given by[1]

$$\mathbf{S} = 2\frac{\partial \Psi(\mathbf{C})}{\partial \mathbf{C}} \quad , \quad \mathbb{C} = 2\frac{\partial \mathbf{S}(\mathbf{C})}{\partial \mathbf{C}} = 4\frac{\partial^2 \Psi(\mathbf{C})}{\partial \mathbf{C} \partial \mathbf{C}} \;. \tag{6.74}$$

In alignment with the previous formulas, the variables to be approximated here are distinguished by the superscripts $\bullet$ = for, bac, cen. For instance, the forward difference formula for $\mathbf{S}$ is denoted by $\mathbf{S}^{\,\text{for}}$. The procedure to arrive at $\mathbf{S}^{\,\text{for}}$ is demonstrated in the following (see Exercise 6.17).

Guided by the approximation statement (6.13), one can write

$$\Psi\left(\mathbf{C} + \triangle\mathbf{C}_{(kl)}\right) \approx \Psi(\mathbf{C}) + \frac{1}{2}\mathbf{S} : \triangle\mathbf{C}_{(kl)} \;, \quad kl = 11, 22, 33, 23, 13, 12 \;, \tag{6.75}$$

where the perturbations $\triangle\mathbf{C}_{(kl)}$ are given by

$$\triangle\,\mathbf{C}_{(kl)} = \frac{\varepsilon}{2}\left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l + \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_k\right) \;. \tag{6.76}$$

[1] The tensorial variables Ψ, $\mathbf{S}$ and $\mathbb{C}$ were chosen on purpose. First, consider the scalar-valued function Ψ as the free-energy function in the context of nonlinear solid mechanics. It is thus an amount of energy stored in an elastic material under imposed deformations whose existence is postulated within the context of hyperelasticity. In this regard, the first derivative of Ψ then helps provide the second Piola-Kirchhoff stress $\mathbf{S}$ which renders a symmetric second-order tensor. Finally, the super-symmetric fourth-order tensor $\mathbb{C} = 2\partial\mathbf{S}/\partial\mathbf{C}$ is basically the referential tensor of elasticities. See Exercise 6.16.

The symbol $\triangle$ in these expressions denotes an **actual difference**. It is basically a linear operator similar to the total differential symbol d for which the standard rules of differentiation hold.

Notice that the symmetry of **C** implies the symmetry of **S**. As a result, only six perturbation tests are required to complete the numerical procedure. For this reason, the subscript indices k and l take on values collected in (3.129). The six perturbations $\triangle\mathbf{C}_{(kl)}$ basically provide the six perturbed tensors $\mathbf{C}^{+\varepsilon}_{(kl)}$ as

$$\begin{aligned}\mathbf{C}^{+\varepsilon}_{(kl)} &= \mathbf{C} + \triangle\mathbf{C}_{(kl)} \\ &= \mathbf{C} + \frac{\varepsilon}{2}\left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l + \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_k\right) \ . \end{aligned} \tag{6.77}$$

Now, by means of (2.20), $(2.61)_1$, (2.74a)–(2.74b), (2.79c) and (6.75)–(6.77), the independent Cartesian components of **S** can be obtained via

$$\Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)}\right) - \Psi\left(\mathbf{C}\right) \approx \frac{1}{2}\mathbf{S} : \triangle\mathbf{C}_{(kl)} = \frac{\varepsilon}{4}\left(S_{kl} + S_{lk}\right) = \frac{\varepsilon}{2}S_{kl} \ .$$

Therefore,

$$S_{kl} \approx \frac{2}{\varepsilon}\left[\Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)}\right) - \Psi\left(\mathbf{C}\right)\right] := S^{\text{for}}_{kl} \ . \tag{6.78}$$

To compute the backward and central difference approximations of **S**, one needs to construct six perturbed tensors according to

$$\mathbf{C}^{-\varepsilon}_{(kl)} = \mathbf{C} - \triangle\mathbf{C}_{(kl)} = \mathbf{C} - \frac{\varepsilon}{2}\left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_l + \widehat{\mathbf{e}}_l \otimes \widehat{\mathbf{e}}_k\right) \ . \tag{6.79}$$

By analogy with the procedure that led to (6.78), one will have

$$S_{kl} \approx \frac{2}{\varepsilon}\left[\Psi\left(\mathbf{C}\right) - \Psi\left(\mathbf{C}^{-\varepsilon}_{(kl)}\right)\right] := S^{\text{bac}}_{kl} \ , \tag{6.80a}$$

$$S_{kl} \approx \frac{1}{\varepsilon}\left[\Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)}\right) - \Psi\left(\mathbf{C}^{-\varepsilon}_{(kl)}\right)\right] := S^{\text{cen}}_{kl} \ . \tag{6.80b}$$

It is then a simple exercise to arrive at

$$\begin{aligned}\mathbb{C}_{ijkl} &\approx \frac{2}{\varepsilon}\left[S_{ij}\left(\mathbf{C}^{+\varepsilon}_{(kl)}\right) - S_{ij}\left(\mathbf{C}\right)\right] := \overline{\mathbb{C}}^{\text{for}}_{ijkl} \\ &\approx \frac{4}{\varepsilon^2}\left[\Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)} + \triangle\mathbf{C}_{(ij)}\right) - \Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)}\right)\right. \\ &\quad \left. -\Psi\left(\mathbf{C}^{+\varepsilon}_{(ij)}\right) + \Psi\left(\mathbf{C}\right)\right] := \widehat{\mathbb{C}}^{\text{for}}_{ijkl} \ , \end{aligned} \tag{6.81a}$$

$$\mathbb{C}_{ijkl} \approx \frac{2}{\varepsilon}\left[S_{ij}\left(\mathbf{C}\right) - S_{ij}\left(\mathbf{C}^{-\varepsilon}_{(kl)}\right)\right] := \overline{\mathbb{C}}^{\text{bac}}_{ijkl}$$

$$\approx \frac{4}{\varepsilon^2}\Big[\Psi(\mathbf{C}) - \Psi\left(\mathbf{C}^{-\varepsilon}_{(ij)}\right)$$
$$-\Psi\left(\mathbf{C}^{-\varepsilon}_{(kl)}\right) + \Psi\left(\mathbf{C}^{-\varepsilon}_{(kl)} - \triangle\mathbf{C}_{(ij)}\right)\Big] := \widehat{\mathbb{C}}^{\text{bac}}_{ijkl}\ , \tag{6.81b}$$

$$\mathbb{C}_{ijkl} \approx \frac{1}{\varepsilon}\left[S_{ij}\left(\mathbf{C}^{+\varepsilon}_{(kl)}\right) - S_{ij}\left(\mathbf{C}^{-\varepsilon}_{(kl)}\right)\right] := \overline{\mathbb{C}}^{\text{cen}}_{ijkl}$$
$$\approx \frac{1}{\varepsilon^2}\Big[\Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)} + \triangle\mathbf{C}_{(ij)}\right) - \Psi\left(\mathbf{C}^{+\varepsilon}_{(kl)} - \triangle\mathbf{C}_{(ij)}\right)$$
$$-\Psi\left(\mathbf{C}^{-\varepsilon}_{(kl)} + \triangle\mathbf{C}_{(ij)}\right) + \Psi\left(\mathbf{C}^{-\varepsilon}_{(kl)} - \triangle\mathbf{C}_{(ij)}\right)\Big] := \widehat{\mathbb{C}}^{\text{cen}}_{ijkl}\ . \tag{6.81c}$$

At the end, the quality of these approximations can be examined by

$$\underbrace{\mathsf{err}_{\mathbf{S}^\bullet} = |\mathbf{S}| - |\mathbf{S}^\bullet| \quad , \quad \mathsf{err}_{\overline{\mathbb{C}}^\bullet} = |\mathbb{C}| - \left|\overline{\mathbb{C}}^\bullet\right| \quad , \quad \mathsf{err}_{\widehat{\mathbb{C}}^\bullet} = |\mathbb{C}| - \left|\widehat{\mathbb{C}}^\bullet\right|}_{\text{note that } |\mathbf{S}| = \sqrt{S_{ij}S_{ij}} \text{ and } |\mathbb{C}| = \sqrt{\mathbb{C}_{ijkl}\mathbb{C}_{ijkl}},\ \text{see (2.76) and (3.100)}}\ , \tag{6.82}$$

where $\bullet$ = for, bac, cen.

6.2 Representation Theorems

The representation theorems are widely used in various branches of physics and engineering. Examples of which include solid mechanics and tissue engineering. These are examples of applications of *theory of algebraic invariants* in mechanics of isotropic and anisotropic continuum mediums. The underlying theory aims at finding what is known as *integrity basis* for a given set of tensorial variables and group of orthogonal transformations. An integrity basis is simply a set of polynomials whose every element is **invariant** under the given group of transformations. In principle, one can construct infinitely many algebraic invariants for a finite system of tensors (with possibly different orders). But, by use of some important relations such as the Cayley-Hamilton equation and the so-called *Rivlin's identities*, one can always provide a **finite** number of invariants. Once an integrity basis is generated for a system of tensorial variables, any tensor function of that system, which itself should be invariant under the group of transformations, can be expressed in terms of the elements of that integrity basis. And this basically demonstrates the representation theorem for such an invariant tensor function.

A debatable issue in this context is *redundancy* of the basic invariants. With regard to this, an integrity basis is said to be *irreducible* if none of its subsets is an integrity basis by its own. In other words, a generic element of an irreducible basis cannot be expressed as a linear combination of the remaining ones. It is always desirable to find the so-called *minimal integrity basis* which contains the smallest possible number of members for an invariant representation (note that the elements of such basis are eventually irreducible). An integrity basis is utilized to represent a **polynomial**

function. But for the representation of a general **non-polynomial function**, one needs to use the so-called *functional basis*. A functional basis is called *irreducible* when none of its members is expressible as a single-valued function of the remaining ones. In establishing the functional bases, the main goal is to represent the single-valued functions. And the procedure to develop such bases relies on **geometrical reasoning**. It is known that an integrity basis is always a functional basis but the converse, in general, is not true. Indeed, an irreducible functional basis usually has **fewer** elements than an irreducible integrity basis.

The problem of finding the algebraic invariants has extensively been considered in the past few decades. It is referenced to the pioneering works of Rivlin [6], Pipkin and Rivlin [7], Spencer and Rivlin [8–11], Spencer [12, 13], Smith [14], Wang [15, 16], Betten [17, 18], Boehler [19, 20], Liu [21] and Zheng [22] among the others. The results in this context are usually provided in tables very convenient for use. From these tables, one will simply be able to represent an invariant scalar-, vector- or tensor-valued function of a system of tensorial variables in terms of the elements of the corresponding complete and irreducible functional basis (note that all tables provided in this section deliver functional bases). However, one needs to be familiar with at least some basics of representation theorems. This section is thus devoted to the study of these useful theorems.

6.2.1 Mathematical Preliminaries

To begin with, some definitions and notations are introduced. To this end, consider Cartesian coordinates for convenience. Let

$$\left.\begin{aligned} \alpha &: \mathscr{D} \to \mathscr{R} \\ \mathbf{v} &: \mathscr{D} \to \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}} \\ \mathbf{S} &: \mathscr{D} \to \mathscr{T}_{\mathrm{so}}^{\mathrm{sym}} \\ \mathbf{W} &: \mathscr{D} \to \mathscr{T}_{\mathrm{so}}^{\mathrm{skw}} \end{aligned}\right\} , \tag{6.83}$$

be a scalar-, vector-, symmetric tensor-, and skew-symmetric tensor-valued function, respectively. These tensor functions are defined on

$$\begin{aligned} \mathscr{D} &= (\mathscr{R})^{m_1} \times \left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}\right)^{m_2} \times \left(\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}\right)^{m_3} \times \left(\mathscr{T}_{\mathrm{so}}^{\mathrm{skw}}\right)^{m_4} \\ &= \underbrace{\mathscr{R} \times \cdots \times \mathscr{R}}_{m_1 \text{ times}} \\ &\quad \times \underbrace{\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}} \times \cdots \times \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}}_{m_2 \text{ times}} \times \underbrace{\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}} \times \cdots \times \mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}}_{m_3 \text{ times}} \times \underbrace{\mathscr{T}_{\mathrm{so}}^{\mathrm{skw}} \times \cdots \times \mathscr{T}_{\mathrm{so}}^{\mathrm{skw}}}_{m_4 \text{ times}} , \end{aligned} \tag{6.84}$$

where $\mathscr{R}$ denotes the set of real numbers, $\mathscr{E}_{\mathrm{r}}^{\mathrm{o3}}$ designates the oriented three-dimensional Euclidean vector space, $\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}$ presents the set of all symmetric tensors and $\mathscr{T}_{\text{so}}^{\text{skw}}$ stands for the set of all skew-symmetric tensors. In representation theorems, the tensor functions are classified according to their transformation properties under the action of the full orthogonal group $\mathscr{O}$ or any subgroup $\mathscr{G}$ of $\mathscr{O}$, see (2.132) and (2.133). For the case where $\mathscr{G} = \mathscr{O}$, a function α is said to be an *isotropic scalar-valued function* if, for any $\bar{s} \in (\mathscr{R})^{m_1}$, $\bar{\mathbf{v}} \in \left(\mathscr{E}_{\text{r}}^{\text{o3}}\right)^{m_2}$, $\bar{\mathbf{S}} \in \left(\mathscr{T}_{\text{so}}^{\text{sym}}\right)^{m_3}$ and $\overline{\mathbf{W}} \in \left(\mathscr{T}_{\text{so}}^{\text{skw}}\right)^{m_4}$, it satisfies the following condition

$$\boxed{\alpha\left(\bar{s}, \mathbf{Q}\bar{\mathbf{v}}, \mathbf{Q}\bar{\mathbf{S}}\mathbf{Q}^{\text{T}}, \mathbf{Q}\overline{\mathbf{W}}\mathbf{Q}^{\text{T}}\right) = \alpha\left(\bar{s}, \bar{\mathbf{v}}, \bar{\mathbf{S}}, \overline{\mathbf{W}}\right) \ , \quad \forall \mathbf{Q} \in \mathscr{G} \ .} \tag{6.85}$$

Here, the following abbreviations have been used

$$\left.\begin{array}{l} \bar{s} = \left(s_1, \ldots, s_{m_1}\right) \\ \bar{\mathbf{v}} = \left(\mathbf{v}_1, \ldots, \mathbf{v}_{m_2}\right) \end{array}\right\} \ , \quad \left.\begin{array}{l} \bar{\mathbf{S}} = \left(\mathbf{S}_1, \ldots, \mathbf{S}_{m_3}\right) \\ \overline{\mathbf{W}} = \left(\mathbf{W}_1, \ldots, \mathbf{W}_{m_4}\right) \end{array}\right\} \ , \tag{6.86}$$

where

$$\underbrace{s_i \in \mathscr{R}}_{i=1,2,\ldots,m_1} \ , \quad \underbrace{\mathbf{v}_j \in \mathscr{E}_{\text{r}}^{\text{o3}}}_{j=1,2,\ldots,m_2} \ , \quad \underbrace{\mathbf{S}_k \in \mathscr{T}_{\text{so}}^{\text{sym}}}_{k=1,2,\ldots,m_3} \ , \quad \underbrace{\mathbf{W}_l \in \mathscr{T}_{\text{so}}^{\text{skw}}}_{l=1,2,\ldots,m_4} \ . \tag{6.87}$$

The isotropic scalar-valued functions are also called *isotropic invariants*. As an example, the scalar function $\alpha(\mathbf{A}) = \text{tr}\,\mathbf{A}^k$ is an isotropic invariant:

$$\alpha\left(\mathbf{QAQ}^{\text{T}}\right) = \text{tr}\left(\underbrace{\mathbf{QAQ}^{\text{T}} \cdots \mathbf{QAQ}^{\text{T}}}_{k \text{ time}}\right) = \text{tr}\left(\mathbf{QA}^k\mathbf{Q}^{\text{T}}\right) = \text{tr}\,\mathbf{A}^k = \alpha(\mathbf{A}) \ , \tag{6.88}$$

where (2.33), (2.109g), $(2.130)_1$ and (2.131) have been used. One can now conclude that the principal scalar invariants of a tensor, according to (4.17a)–(4.17c), are isotropic:

$$\boxed{I_1\left(\mathbf{QAQ}^{\text{T}}\right) = I_1(\mathbf{A}) \quad , \quad I_2\left(\mathbf{QAQ}^{\text{T}}\right) = I_2(\mathbf{A}) \quad , \quad I_3\left(\mathbf{QAQ}^{\text{T}}\right) = I_3(\mathbf{A}) \ .} \tag{6.89}$$

The result $(6.89)_3$ can also be verified as follows:

$$\boxed{\alpha\left(\mathbf{QAQ}^{\text{T}}\right) = \det\left(\mathbf{QAQ}^{\text{T}}\right) = (\det\mathbf{Q})^2(\det\mathbf{A}) = \det\mathbf{A} = \alpha(\mathbf{A}) \ ,} \tag{6.90}$$

where (2.99a), $(2.99\text{b})_1$ and (2.134c) have been utilized. Note that a scalar-valued function of the form

$$\alpha(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3) = \text{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3) \ , \tag{6.91}$$

is also an isotropic invariant.

Similarly to (6.85), when $\mathscr{G} = \mathscr{O}$, the vector-valued function $\mathbf{v}$, the symmetric tensor-valued function $\mathbf{S}$ and the skew-symmetric tensor-valued function $\mathbf{W}$ are said to be *isotropic* if

$$\left.\begin{aligned} \mathbf{v}\left(\bar{s}, \mathbf{Q}\bar{\mathbf{v}}, \mathbf{Q}\bar{\mathbf{S}}\mathbf{Q}^{\mathrm{T}}, \mathbf{Q}\bar{\mathbf{W}}\mathbf{Q}^{\mathrm{T}}\right) &= \mathbf{Q}\mathbf{v}\left(\bar{s}, \bar{\mathbf{v}}, \bar{\mathbf{S}}, \bar{\mathbf{W}}\right) \\ \mathbf{S}\left(\bar{s}, \mathbf{Q}\bar{\mathbf{v}}, \mathbf{Q}\bar{\mathbf{S}}\mathbf{Q}^{\mathrm{T}}, \mathbf{Q}\bar{\mathbf{W}}\mathbf{Q}^{\mathrm{T}}\right) &= \mathbf{Q}\mathbf{S}\left(\bar{s}, \bar{\mathbf{v}}, \bar{\mathbf{S}}, \bar{\mathbf{W}}\right)\mathbf{Q}^{\mathrm{T}} \\ \mathbf{W}\left(\bar{s}, \mathbf{Q}\bar{\mathbf{v}}, \mathbf{Q}\bar{\mathbf{S}}\mathbf{Q}^{\mathrm{T}}, \mathbf{Q}\bar{\mathbf{W}}\mathbf{Q}^{\mathrm{T}}\right) &= \mathbf{Q}\mathbf{W}\left(\bar{s}, \bar{\mathbf{v}}, \bar{\mathbf{S}}, \bar{\mathbf{W}}\right)\mathbf{Q}^{\mathrm{T}} \end{aligned}\right\} , \quad \forall \mathbf{Q} \in \mathscr{G} . \tag{6.92}$$

Notice that the scalar arguments of isotropic tensor functions are not affected by any orthogonal transformation and, therefore, they will be dropped in the subsequent developments for convenience.

The tensor functions satisfying (6.85) and (6.92) are referred to as *hemitropic* (or *relative isotropic*) if $\mathscr{G} = \mathscr{O}^+$. Otherwise, they are termed *anisotropic*. As an example, the vector-valued function $\mathbf{v}(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1 \times \mathbf{v}_2$ is **hemitropic** because

$$\begin{aligned} \mathbf{Q}\mathbf{u} \cdot (\mathbf{Q}\mathbf{v}_1 \times \mathbf{Q}\mathbf{v}_2) &\xlongequal[\text{hand from (2.98)}]{\text{on the one}} (\det \mathbf{Q})\, \mathbf{u} \cdot (\mathbf{v}_1 \times \mathbf{v}_2) \\ &= \mathbf{u} \cdot [(\det \mathbf{Q})(\mathbf{v}_1 \times \mathbf{v}_2)] \\ &\xlongequal[\text{hand from (2.51d)}]{\text{on the other}} \mathbf{u} \cdot \left[\mathbf{Q}^{\mathrm{T}} (\mathbf{Q}\mathbf{v}_1 \times \mathbf{Q}\mathbf{v}_2)\right] \\ &= \mathbf{u} \cdot \left[\mathbf{Q}^{\mathrm{T}} \mathbf{v}(\mathbf{Q}\mathbf{v}_1, \mathbf{Q}\mathbf{v}_2)\right] , \end{aligned}$$

helps, using (1.9a) and (1.14), obtain

$$\mathbf{Q}^{\mathrm{T}}\mathbf{v}(\mathbf{Q}\mathbf{v}_1, \mathbf{Q}\mathbf{v}_2) = (\det \mathbf{Q})(\mathbf{v}_1 \times \mathbf{v}_2) = (\det \mathbf{Q})\, \mathbf{v}(\mathbf{v}_1, \mathbf{v}_2) ,$$

or, using (2.5), (2.130) and (2.133),

$$\boxed{\mathbf{v}(\mathbf{Q}\mathbf{v}_1, \mathbf{Q}\mathbf{v}_2) = \mathbf{Q}\mathbf{v}(\mathbf{v}_1, \mathbf{v}_2) .} \tag{6.93}$$

Another example regards the scalar-valued function

$$\alpha(\mathbf{S}_1, \mathbf{S}_2) = \mathrm{tr}(\mathbf{S}_1 \mathbf{L} \mathbf{S}_2) , \tag{6.94}$$

which is **anisotropic** for a generic tensor $\mathbf{L} \in \mathscr{T}_{\mathrm{so}}$.

The set $\mathscr{G}$, possessing the properties of a group, is known as the *symmetry group* of the tensor functions (6.85) and (6.92). One can now say that these functions are *$\mathscr{G}$-invariant*.

Let

$$\Upsilon_{\mathrm{s}} = \left\{ I_1\left(\bar{\mathbf{v}}, \bar{\mathbf{S}}, \bar{\mathbf{W}}\right), \ldots, I_n\left(\bar{\mathbf{v}}, \bar{\mathbf{S}}, \bar{\mathbf{W}}\right) \right\} , \tag{6.95}$$

be an irreducible set of n scalar-valued functions that are all invariant with respect to a given group $\mathscr{G}$ of orthogonal transformations. Note that irreducibility here means that these scalar functions cannot be expressed uniquely in terms of each other. This motivates to call these functions the *basic invariants*. Such a set is called an *integrity basis*. And this helps represent any $\mathscr{G}$-invariant scalar function of the form (6.85) as

$$\boxed{\alpha\left(\overline{\mathbf{v}}, \overline{\mathbf{S}}, \overline{\mathbf{W}}\right) = \bar{\alpha}\left(\Upsilon_{\mathrm{s}}\right) \ .} \tag{6.96}$$

In a similar manner, let $\Upsilon_{\mathbf{v}}$, $\Upsilon_{\mathbf{S}}$ and $\Upsilon_{\mathbf{W}}$ be an irreducible set of $\mathscr{G}$-invariant vector, symmetric tensor and skew-symmetric function, respectively. They are called *generating sets*. And any member of a generating set is referred to as a *generator element*. The $\mathscr{G}$-invariant tensor functions (6.92) then admit the following representations

$$\left.\begin{aligned} \mathbf{v}\left(\overline{\mathbf{v}}, \overline{\mathbf{S}}, \overline{\mathbf{W}}\right) &= \sum_{\check{\mathbf{v}}_p \in \Upsilon_{\mathbf{v}}} \hat{\alpha}_p\left(\Upsilon_{\mathrm{s}}\right) \check{\mathbf{v}}_p \\ \mathbf{S}\left(\overline{\mathbf{v}}, \overline{\mathbf{S}}, \overline{\mathbf{W}}\right) &= \sum_{\check{\mathbf{S}}_q \in \Upsilon_{\mathbf{S}}} \tilde{\alpha}_q\left(\Upsilon_{\mathrm{s}}\right) \check{\mathbf{S}}_q \\ \mathbf{W}\left(\overline{\mathbf{v}}, \overline{\mathbf{S}}, \overline{\mathbf{W}}\right) &= \sum_{\check{\mathbf{W}}_r \in \Upsilon_{\mathbf{W}}} \breve{\alpha}_r\left(\Upsilon_{\mathrm{s}}\right) \check{\mathbf{W}}_r \end{aligned}\right\} , \tag{6.97}$$

where $\hat{\alpha}_p, \tilde{\alpha}_q, \breve{\alpha}_r$ are arbitrary $\mathscr{G}$-invariant scalar functions of the basic invariants and $\check{\mathbf{v}}_p, \check{\mathbf{S}}_q, \check{\mathbf{W}}_r$ present the generator elements. As can be seen, any $\mathscr{G}$-invariant tensor function can be expressed as a linear combination of the generator elements formed from its vectors and (symmetric and antisymmetric) tensor arguments with the coefficients which are arbitrary scalar functions of the basic invariants. And this basically demonstrates the representation theorem for such an $\mathscr{G}$-invariant tensor function. Note that this is enabled by determining the corresponding sets of basic invariants and generators. And this is eventually the main task within the context of representation theorems.

Lemma A *Let $\alpha : \mathscr{O} \to \mathscr{R}$ be a scalar-valued function. Suppose that $\alpha(\mathbf{Q}) = 0$ for any $\mathbf{Q}$ in the full orthogonal group $\mathscr{O}$. Then, the gradient of α at $\mathbf{I}$ is a* ***symmetric*** *tensor whose contraction with any antisymmetric tensor $\mathbf{W} \in \mathscr{T}_{\mathrm{so}}^{skw}$ vanishes:*

$$\left.\frac{\partial \alpha}{\partial \mathbf{Q}}\right|_{\mathbf{I}} : \mathbf{W} = 0 \ . \tag{6.98}$$

This result also holds true for any vector- or tensor-valued function.

Proof For any skew tensor $\mathbf{W} = -\mathbf{W}^{\mathrm{T}}$, one obtains $(\mathbf{I} + \varepsilon\mathbf{W})(\mathbf{I} + \varepsilon\mathbf{W})^{\mathrm{T}} = \mathbf{I} - \varepsilon^2\mathbf{W}^2$ where $\varepsilon \in \mathscr{R}$. When $0 < \varepsilon \ll 1$, one will have $(\mathbf{I} + \varepsilon\mathbf{W})(\mathbf{I} + \varepsilon\mathbf{W})^{\mathrm{T}} \approx \mathbf{I}$ which reveals the fact that the new tensor $\mathbf{I} + \varepsilon\mathbf{W}$ is orthogonal to within a small error of $o(\varepsilon)$, that is,

$$\mathbf{I} + \varepsilon\mathbf{W} = \mathbf{Q}_W + o(\varepsilon) \ . \tag{6.99}$$

Then, α at $\mathbf{I} + \varepsilon\mathbf{W}$ can be expanded via the first-order Taylor series as

$$\alpha\left(\mathbf{I}+\varepsilon\mathbf{W}\right)=\overbrace{\cancel{\alpha\left(\mathbf{I}\right)}}^{=\,0,\ \text{by assumption}}+\left.\frac{\partial\alpha}{\partial\mathbf{Q}}\right|_{\mathbf{I}}:\varepsilon\mathbf{W}+o\left(\varepsilon\right)\ . \tag{6.100}$$

In a similar manner, α at $\mathbf{Q}_W+o\left(\varepsilon\right)$ can be written as

$$\alpha\left(\mathbf{Q}_W+o\left(\varepsilon\right)\right)=\overbrace{\cancel{\alpha\left(\mathbf{Q}_W\right)}}^{=\,0,\ \text{by assumption}}+\left.\frac{\partial\alpha}{\partial\mathbf{Q}}\right|_{\mathbf{Q}_W}:o\left(\varepsilon\right)+o\left(\varepsilon\right)\ . \tag{6.101}$$

With the aid of (2.79h) and (6.99)–(6.101), one finally concludes that $\partial\alpha/\partial\mathbf{Q}$ at $\mathbf{I}$ is symmetric.

A consequence of (6.98) is

$$\boxed{\varepsilon_{ijk}\left(\frac{\partial\alpha}{\partial u_j}u_k+\frac{\partial\alpha}{\partial v_j}v_k\right)=0\ ,} \tag{6.102}$$

where $\alpha\left(\mathbf{u},\mathbf{v}\right)$ is an **isotropic** scalar-valued function. This identity can be verified by defining the scalar function

$$\beta\left(\mathbf{Q}\right)=\alpha\left(\mathbf{Qu},\mathbf{Qv}\right)-\alpha\left(\mathbf{u},\mathbf{v}\right)\ , \tag{6.103}$$

which vanishes for any $\mathbf{Q}\in\mathscr{O}$. Now, suppose that $\bar{\mathbf{u}}=\mathbf{Qu}$ and $\bar{\mathbf{v}}=\mathbf{Qv}$. Then,

$$\frac{\partial\beta}{\partial Q_{ij}}=\underbrace{\frac{\partial\alpha}{\partial\bar{u}_m}\frac{\partial\left[Q_{mn}u_n\right]}{\partial Q_{ij}}}_{=\frac{\partial\alpha}{\partial\bar{u}_m}\delta_{mi}\delta_{nj}u_n=\frac{\partial\alpha}{\partial\bar{u}_i}u_j}+\underbrace{\frac{\partial\alpha}{\partial\bar{v}_m}\frac{\partial\left[Q_{mn}v_n\right]}{\partial Q_{ij}}}_{=\frac{\partial\alpha}{\partial\bar{v}_m}\delta_{mi}\delta_{nj}v_n=\frac{\partial\alpha}{\partial\bar{v}_i}v_j}=\left(\frac{\partial\alpha}{\partial\bar{\mathbf{u}}}\otimes\mathbf{u}+\frac{\partial\alpha}{\partial\bar{\mathbf{v}}}\otimes\mathbf{v}\right)_{ij}\ . \tag{6.104}$$

Notice that

$$\left.\frac{\partial\alpha}{\partial\mathbf{u}}\right|_{\mathbf{I}}=\left.\left(\frac{\partial\alpha}{\partial\bar{\mathbf{u}}}\frac{\partial\bar{\mathbf{u}}}{\partial\mathbf{u}}\right)\right|_{\mathbf{I}}=\left.\left(\frac{\partial\alpha}{\partial\bar{\mathbf{u}}}\mathbf{Q}\right)\right|_{\mathbf{I}}=\left.\frac{\partial\alpha}{\partial\bar{\mathbf{u}}}\right|_{\mathbf{I}}\ . \tag{6.105}$$

Consequently,

$$\left.\frac{\partial\beta}{\partial\mathbf{Q}}\right|_{\mathbf{I}}=\frac{\partial\alpha}{\partial\mathbf{u}}\otimes\mathbf{u}+\frac{\partial\alpha}{\partial\mathbf{v}}\otimes\mathbf{v}\ . \tag{6.106}$$

Guided by (6.98), this second-order tensor is symmetric. Having in mind $(3.16\text{b})_5$, its double contraction with $\mathsf{E}=\varepsilon_{ijk}\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k$ in (3.17) thus gives the zero vector. And this finally provides the desired result (6.102).

Another consequence of (6.98) is

$$\boxed{\varepsilon_{ijk}\left(\frac{\partial\alpha}{\partial u_j}u_k+\frac{\partial\alpha}{\partial A_{jm}}A_{km}+\frac{\partial\alpha}{\partial A_{mj}}A_{mk}\right)=0\ ,} \tag{6.107}$$

where $\alpha(\mathbf{u}, \mathbf{A})$ presents an **isotropic** scalar-valued function. To verify this relation, consider the scalar function

$$\beta(\mathbf{Q}) = \alpha\left(\mathbf{Qu}, \mathbf{QAQ}^{\mathrm{T}}\right) - \alpha(\mathbf{u}, \mathbf{A}) \ , \tag{6.108}$$

which vanishes for any $\mathbf{Q} \in \mathscr{O}$. Let $\overline{\mathbf{u}} = \mathbf{Qu}$ and $\overline{\mathbf{A}} = \mathbf{QAQ}^{\mathrm{T}}$. Then,

$$\begin{aligned}\frac{\partial \beta}{\partial Q_{ij}} &= \underbrace{\frac{\partial \alpha}{\partial \overline{u}_m} \frac{\partial [Q_{mn} u_n]}{\partial Q_{ij}}}_{= \frac{\partial \alpha}{\partial \overline{u}_m} \delta_{mi} \delta_{nj} u_n = \frac{\partial \alpha}{\partial \overline{u}_i} u_j} + \underbrace{\frac{\partial \alpha}{\partial \overline{A}_{mn}} \frac{\partial [Q_{mr} A_{rs} Q_{ns}]}{\partial Q_{ij}}}_{= \frac{\partial \alpha}{\partial \overline{A}_{in}} A_{js} Q_{ns} + \frac{\partial \alpha}{\partial \overline{A}_{mi}} Q_{mr} A_{rj}} \\ &= \left(\frac{\partial \alpha}{\partial \overline{\mathbf{u}}} \otimes \mathbf{u} + \frac{\partial \alpha}{\partial \overline{\mathbf{A}}} \mathbf{QA}^{\mathrm{T}} + \left(\frac{\partial \alpha}{\partial \overline{\mathbf{A}}} \right)^{\mathrm{T}} \mathbf{QA} \right)_{ij} \ . \end{aligned} \tag{6.109}$$

Having in mind (6.105), one should also have

$$\left.\frac{\partial \alpha}{\partial \mathbf{A}}\right|_{\mathbf{I}} = \left.\left(\frac{\partial \alpha}{\partial \overline{\mathbf{A}}} : \frac{\partial \overline{\mathbf{A}}}{\partial \mathbf{A}} \right)\right|_{\mathbf{I}} = \left.\left(\mathbf{Q}^{\mathrm{T}} \frac{\partial \alpha}{\partial \overline{\mathbf{A}}} \mathbf{Q} \right)\right|_{\mathbf{I}} = \left.\frac{\partial \alpha}{\partial \overline{\mathbf{A}}}\right|_{\mathbf{I}} \ ,$$

in order to compute

$$\left.\frac{\partial \beta}{\partial \mathbf{Q}}\right|_{\mathbf{I}} = \frac{\partial \alpha}{\partial \mathbf{u}} \otimes \mathbf{u} + \frac{\partial \alpha}{\partial \mathbf{A}} \mathbf{A}^{\mathrm{T}} + \left(\frac{\partial \alpha}{\partial \mathbf{A}} \right)^{\mathrm{T}} \mathbf{A} \ . \tag{6.110}$$

The symmetry of this result is implied by (6.98). Guided by $(3.16\mathrm{b})_5$, its double contraction with $\mathbf{E} = \varepsilon_{ijk} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k$ in (3.17) should thus be vanished. And the desired result (6.107) follows.

Lemma B *(a) Let* $\mathbf{S} \in \mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}$ *be a symmetric tensor with the following spectral form*

$$\mathbf{S} \xrightarrow[(4.41)]{\text{from}} \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \ , \quad (\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1) \ ,$$

whose three linearly independent eigenvectors form a basis for the three-dimensional vector space. The set $\left\{\mathbf{I}, \mathbf{S}, \mathbf{S}^2\right\}$ *is then linearly independent and*

$$\boxed{\mathrm{Span}\left\{\mathbf{I}, \mathbf{S}, \mathbf{S}^2\right\} = \mathrm{Span}\left\{\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1, \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2, \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3\right\} \ .} \tag{6.111}$$

(b) Let $\mathbf{S} \in \mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}$ *be a symmetric tensor with two distinct eigenvectors whose spectral decomposition renders*

$$\mathbf{S} \xlongequal[\text{(4.45)}]{\text{in light of}} \lambda_1 \, \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \lambda \left(\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 \right) \;, \quad (\lambda_1 \neq \lambda_2 = \lambda_3 = \lambda) \;.$$

The set $\{\mathbf{I}, \mathbf{S}\}$ *is then linearly independent and*

$$\text{Span}\,\{\mathbf{I}, \mathbf{S}\} = \text{Span}\,\{\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1, \mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1\} \;. \tag{6.112}$$

Proof (a) The set of the three tensors $\mathbf{I}$, $\mathbf{S}$ and $\mathbf{S}^2$ is linearly independent if the only solution to

$$a\mathbf{I} + b\mathbf{S} + c\mathbf{S}^2 = 0 \;, \quad (a, b, c \in \mathscr{R}) \;, \tag{6.113}$$

is

$$a = b = c = 0 \;. \tag{6.114}$$

Guided by (4.37a), one can write

$$\mathbf{I} = \sum_{i=1}^{3} \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \quad , \quad \mathbf{S} = \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \quad , \quad \mathbf{S}^2 = \sum_{i=1}^{3} \lambda_i^2 \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \;. \tag{6.115}$$

Substituting (6.115) into (6.113) yields

$$\sum_{i=1}^{3} \left(a + b\lambda_i + c\lambda_i^2 \right) \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i = 0 \;, \tag{6.116}$$

or, using $\widehat{\mathbf{n}}_i \cdot \widehat{\mathbf{n}}_j = \delta_{ij}$,

$$\left.\begin{aligned} a + b\lambda_1 + c\lambda_1^2 = 0 \\ a + b\lambda_2 + c\lambda_2^2 = 0 \\ a + b\lambda_3 + c\lambda_3^2 = 0 \end{aligned}\right\} \;. \tag{6.117}$$

This can be viewed as a homogeneous system of linear algebraic equations for a, b, c, that is,

$$\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \;, \tag{6.118}$$

with

$$\det \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix} = (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2) \;. \leftarrow \begin{array}{c}\text{this represents the} \\ \textbf{Vandermonde}\text{ determinant}\end{array} \tag{6.119}$$

Obviously, the determinant of the matrix of coefficients does not vanish in the present case. And this implies the desired result (6.114).

Next, consider a linear subspace

$$\mathscr{H} = \text{Span}\left\{\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1, \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2, \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3\right\} \ ,$$

of the symmetric tensor space $\mathscr{T}_{\text{so}}^{\text{sym}}$ whose dimension is 3 (note that the set of the three tensors $\widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i, \ i = 1, 2, 3,$ is also linearly independent and necessarily constitutes a basis for $\mathscr{H}$). Recall from $(6.115)_{1-3}$ that the symmetric tensors $\mathbf{I}$, $\mathbf{S}$ and $\mathbf{S}^2$ were expressed as linear combinations of $\widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i, \ i = 1, 2, 3$. And this means that they also belong to $\mathscr{H}$. Indeed, the linearly independent set $\left\{\mathbf{I}, \mathbf{S}, \mathbf{S}^2\right\}$ forms another basis for $\mathscr{H}$ and, therefore, the desired relation (6.111) is followed. This completes the proof of (a) and the reminder of the proof is left as an exercise to be undertaken by the serious reader.

Lemma C *Let* $\mathbf{T}(\mathbf{S})$ *be an isotropic tensor-valued function of a symmetric tensor* $\mathbf{S}$*. Then, the tensors* $\mathbf{S}$ *and* $\mathbf{T}(\mathbf{S})$ *are* ***coaxial*** *(meaning that their eigenvectors are the same, see Exercise 4.4).*

Proof Regarding the algebraic multiplicities of the real eigenvalues of $\mathbf{S}$, three different cases needs to be considered. First, consider the case in which $\mathbf{S}$ possess non-multiple eigenvalues. It thus admits the spectral decomposition

$$\mathbf{S} \xlongequal[(4.41)]{\text{from}} \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \ , \quad (\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1) \ .$$

The goal here is to show that any eigenvector $\widehat{\mathbf{n}}_j$ of $\mathbf{S}$ is simultaneously an eigenvector of $\mathbf{T}(\mathbf{S})$. Let $\mathbf{Q}$ be an orthogonal (as well as symmetric) tensor of the form

$$\mathbf{Q} = \sum_{\substack{i=1 \\ i \neq j}}^{3} \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i - \underbrace{\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j}_{\text{no sum}} = \mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j \ , \tag{6.120}$$

noting that

$$\begin{aligned} \mathbf{Q}\mathbf{Q}^{\mathrm{T}} &= \left(\mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j\right)\left(\mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j\right) \\ &= \mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j + 4\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j \\ &= \mathbf{I} \ , \end{aligned} \tag{6.121}$$

and

$$\mathbf{Q}\widehat{\mathbf{n}}_j = \left(\mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j\right)\widehat{\mathbf{n}}_j = \widehat{\mathbf{n}}_j - 2\left(\widehat{\mathbf{n}}_j \cdot \widehat{\mathbf{n}}_j\right)\widehat{\mathbf{n}}_j = \widehat{\mathbf{n}}_j - 2\widehat{\mathbf{n}}_j = -\widehat{\mathbf{n}}_j \ . \tag{6.122}$$

It should be understood that this choice of $\mathbf{Q}$ leaves all eigenspaces of $\mathbf{S}$ invariant. Consequently, by means of the commutation theorem introduced in Exercise 4.7, the tensors $\mathbf{S}$ and $\mathbf{Q}$ commute:

$$\underbrace{\mathbf{SQ} = \mathbf{QS} \quad \text{or} \quad \mathbf{S} = \mathbf{QSQ}^{\mathrm{T}}}_{\text{one can easily verify that } \mathbf{SQ} = \mathbf{S}\left(\mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j\right) = \mathbf{S} - 2\lambda_j \widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j = \left(\mathbf{I} - 2\widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j\right)\mathbf{S} = \mathbf{QS}} . \tag{6.123}$$

This result, along with the assumption that $\mathbf{T}(\mathbf{S})$ is isotropic, helps obtain

$$\mathbf{T}(\mathbf{S}) = \mathbf{T}\left(\mathbf{QSQ}^{\mathrm{T}}\right) = \mathbf{QT}(\mathbf{S})\,\mathbf{Q}^{\mathrm{T}} \quad \text{or} \quad \mathbf{QT}(\mathbf{S}) = \mathbf{T}(\mathbf{S})\,\mathbf{Q} . \tag{6.124}$$

Postmultiplying both sides of this result by $\widehat{\mathbf{n}}_j$ then gives

$$\mathbf{QT}(\mathbf{S})\,\widehat{\mathbf{n}}_j = \mathbf{T}(\mathbf{S})\,\mathbf{Q}\widehat{\mathbf{n}}_j \quad \text{or} \quad \mathbf{Q}\left(\mathbf{T}(\mathbf{S})\,\widehat{\mathbf{n}}_j\right) = -\left(\mathbf{T}(\mathbf{S})\,\widehat{\mathbf{n}}_j\right) . \tag{6.125}$$

And this means that the orthogonal transformation $\mathbf{Q}$ maps the vector $\mathbf{T}(\mathbf{S})\,\widehat{\mathbf{n}}_j$ into its negative. Guided by $(6.122)_4$, this can only happen when the vectors $\mathbf{T}(\mathbf{S})\,\widehat{\mathbf{n}}_j$ and $\widehat{\mathbf{n}}_j$ are parallel. Thus, $\mathbf{T}(\mathbf{S})\,\widehat{\mathbf{n}}_j$ is a scalar multiple of $\widehat{\mathbf{n}}_j$, that is,

$$\mathbf{T}(\mathbf{S})\,\widehat{\mathbf{n}}_j = \gamma\,\widehat{\mathbf{n}}_j . \tag{6.126}$$

It can easily be shown that premultiplying $\mathbf{QT}(\mathbf{S}) = \mathbf{T}(\mathbf{S})\,\mathbf{Q}$ by $\widehat{\mathbf{n}}_j$ results in $\mathbf{T}^{\mathrm{T}}(\mathbf{S})\,\widehat{\mathbf{n}}_j = \gamma\,\widehat{\mathbf{n}}_j$. And this reveals the fact that $\mathbf{T}(\mathbf{S})$ is a tensor with identical right and left eigenvectors. Thus, an isotropic tensor-valued function of a symmetric tensor represents a **symmetric** tensor, i.e. $\mathbf{T}(\mathbf{S}) = \mathbf{T}^{\mathrm{T}}(\mathbf{S})$, see (6.174a). It then admits the spectral form

$$\mathbf{T}(\mathbf{S}) = \sum_{i=1}^{3} \mu_i(\mathbf{S})\,\widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i . \tag{6.127}$$

The proof for the remaining cases $\lambda_1 \neq \lambda_2 = \lambda_3 = \lambda$ and $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$ is left to be undertaken by the interested reader. In the following, several important representation theorems for isotropic scalar-, vector- and symmetric tensor-valued functions of some tensorial variables are introduced.

6.2.2 Representation Theorem for an Isotropic Scalar-Valued Function of a Vector

A scalar-valued function $\alpha : \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3} \to \mathscr{R}$ is **isotropic** if there exists a function $f : \mathscr{R} \to \mathscr{R}$ such that

$$\boxed{\alpha(\mathbf{v}) = f(\mathbf{v} \cdot \mathbf{v}) \ .} \tag{6.128}$$

Conversely, a function of this form is isotropic.

Proof By means of (2.5), (2.51d), $(2.130)_1$ and (6.85), the converse assertion can simply be checked:

$$\alpha\left(\mathbf{Qv}\right) = f\left(\mathbf{Qv}\cdot\mathbf{Qv}\right) = f\left(\mathbf{v}\cdot\mathbf{Q}^{\mathrm{T}}\mathbf{Qv}\right) = f\left(\mathbf{v}\cdot\mathbf{Iv}\right) = f\left(\mathbf{v}\cdot\mathbf{v}\right) = \alpha\left(\mathbf{v}\right) \ . \tag{6.129}$$

One can now establish (6.128) assuming that $\alpha(\mathbf{Qv}) = \alpha(\mathbf{v})$ holds true. This expression shows that $\alpha(\mathbf{v})$ changes by only changing the magnitude of $\mathbf{v}$. In other words, $\alpha(\mathbf{v})$ is insensitive with respect to the direction of $\mathbf{v}$. In this regard, consider another vector $\mathbf{u}$ with the same length, that is,

$$|\mathbf{v}| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\mathbf{u}\cdot\mathbf{u}} = |\mathbf{u}| \ . \tag{6.130}$$

It then suffices to show that

$$\alpha\left(\mathbf{v}\right) = \alpha\left(\mathbf{u}\right) \ . \tag{6.131}$$

Consider the fact that two vectors of the same magnitude and origin can always coincide by applying an orthogonal transformation. This allows one to write $\mathbf{u} = \mathbf{Qv}$. Thus, by the isotropy condition $\alpha(\mathbf{v}) = \alpha(\mathbf{Qv})$ and the relation $\mathbf{Qv} = \mathbf{u}$, one can finally arrive at the desired result (6.131). At the end, it should be noted that the representation (6.128) makes sense because only the magnitude of a vector remains invariant under an orthogonal transformation.

6.2.3 Representation Theorem for an Isotropic Scalar-Valued Function of a Symmetric Tensor

A scalar-valued function $\alpha : \mathscr{T}_{\mathrm{so}}^{\mathrm{sym}} \to \mathscr{R}$ is **isotropic** if there exists a function $f : \mathscr{R}\times\mathscr{R}\times\mathscr{R} \to \mathscr{R}$ such that

$$\boxed{\alpha\left(\mathbf{S}\right) = f\left(\, I_1\left(\mathbf{S}\right), I_2\left(\mathbf{S}\right), I_3\left(\mathbf{S}\right)\,\right) \ .} \tag{6.132}$$

Conversely, a function of this form is isotropic.

Proof By means of (6.89), the converse assertion can readily be verified as follows:

$$\begin{aligned}\alpha\left(\mathbf{QSQ}^{\mathrm{T}}\right) &= f\left(\, I_1\left(\mathbf{QSQ}^{\mathrm{T}}\right), I_2\left(\mathbf{QSQ}^{\mathrm{T}}\right), I_3\left(\mathbf{QSQ}^{\mathrm{T}}\right)\right) \\ &= f\left(\, I_1\left(\mathbf{S}\right), I_2\left(\mathbf{S}\right), I_3\left(\mathbf{S}\right)\right) \\ &= \alpha\left(\mathbf{S}\right) \ .\end{aligned} \tag{6.133}$$

Assume that $\alpha(\mathbf{S}) = \alpha\left(\mathbf{QSQ}^{\mathrm{T}}\right)$. The goal is now to establish (6.132). Recall from (4.9) and (4.14a)–(4.14c) that the eigenvalues and the principal invariants of a tensor could uniquely determine one another. With regard to this, one here needs to show that $\alpha(\mathbf{S})$ changes when only the principal values of $\mathbf{S}$ change. In other words, $\alpha(\mathbf{S})$ is insensitive with respect to the eigenvectors of $\mathbf{S}$. Without loss of generality, let $\mathbf{S}$ be a symmetric tensor with the three distinct eigenpairs $(\lambda_1, \widehat{\mathbf{n}}_1)$, $(\lambda_2, \widehat{\mathbf{n}}_2)$ and $(\lambda_3, \widehat{\mathbf{n}}_3)$.

Consider another symmetric tensor $\mathbf{T}$ with the same set of eigenvalues. These tensors admit the following spectral representations

$$\mathbf{S} = \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \quad , \quad \mathbf{T} = \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{m}}_i \ . \tag{6.134}$$

Now, it suffices to verify that

$$\alpha\left(\mathbf{S}\right) = \alpha\left(\mathbf{T}\right) \ . \tag{6.135}$$

Let $\mathbf{Q}$ be a tensor of the form

$$\mathbf{Q} = \sum_{i=1}^{3} \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{n}}_i \ , \tag{6.136}$$

which satisfies the orthogonality requirement

$$\begin{aligned} \mathbf{Q}\mathbf{Q}^{\mathrm{T}} &= \left(\sum_{i=1}^{3} \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{n}}_i \right) \left(\sum_{j=1}^{3} \widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{m}}_j \right) \\ &= \sum_{i,j=1}^{3} \delta_{ij} \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{m}}_j \\ &= \sum_{i=1}^{3} \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{m}}_i \\ &= \mathbf{I} \ . \end{aligned} \tag{6.137}$$

Then,

$$\begin{aligned} \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}} &= \left(\sum_{i=1}^{3} \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{n}}_i \right) \left(\sum_{j=1}^{3} \lambda_j \, \widehat{\mathbf{n}}_j \otimes \widehat{\mathbf{n}}_j \right) \left(\sum_{k=1}^{3} \widehat{\mathbf{n}}_k \otimes \widehat{\mathbf{m}}_k \right) \\ &= \sum_{i,j,k=1}^{3} \lambda_j \delta_{ij} \delta_{jk} \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{m}}_k \\ &= \sum_{i=1}^{3} \lambda_i \widehat{\mathbf{m}}_i \otimes \widehat{\mathbf{m}}_i \\ &= \mathbf{T} \ . \end{aligned} \tag{6.138}$$

The isotropy condition $\alpha(\mathbf{S}) = \alpha\left(\mathbf{QSQ}^{\mathrm{T}}\right)$ along with the result $\mathbf{QSQ}^{\mathrm{T}} = \mathbf{T}$ finally implies (6.135). At the end, one should realize that the representation (6.132) makes sense because obviously the principal invariants of a tensor remain invariant under an orthogonal transformation.

Hint: The fact that the three sets $\{I_1(\mathbf{S}), I_2(\mathbf{S}), I_3(\mathbf{S})\}$, $\{\lambda_1(\mathbf{S}), \lambda_2(\mathbf{S}), \lambda_3(\mathbf{S})\}$ and $\left\{\mathrm{tr}(\mathbf{S}), \mathrm{tr}\left(\mathbf{S}^2\right), \mathrm{tr}\left(\mathbf{S}^3\right)\right\}$ can uniquely determine each other helps further represent the theorem (6.132) as

$$\begin{aligned}\alpha(\mathbf{S}) &= \bar{f}\left(\lambda_1(\mathbf{S}), \lambda_2(\mathbf{S}), \lambda_3(\mathbf{S})\right) \\ &= \hat{f}\left(\mathrm{tr}(\mathbf{S}), \mathrm{tr}\left(\mathbf{S}^2\right), \mathrm{tr}\left(\mathbf{S}^3\right)\right) .\end{aligned} \tag{6.139}$$

Hint: Recall from $(4.63\mathrm{a})_2$ and $(4.63\mathrm{c})_2$ that $I_1(\mathbf{W}) = I_3(\mathbf{W}) = 0$. As a result, an isotropic scalar-valued function of a skew tensor should only be represented in terms of its second principal scalar invariant:

$$\boxed{\alpha(\mathbf{W}) = f\left(I_2(\mathbf{W})\right) .} \tag{6.140}$$

The representation theorems, developed so far for isotropic scalar invariants of one tensor variable, can be extended to involve more tensorial arguments. The results have been summarized in Table 6.1. By means of this table, one can readily construct the **functional basis** for any given domain of isotropic invariants, see (6.83)–(6.84) and (6.95). It is important to note that the basic invariants involving more than four tensorial variables are not present in this list since, it can be shown that, they are **redundant**. In the following, some examples are provided for illustration. ❦

The first example regards an isotropic scalar-valued function of a symmetric tensor $\mathbf{S}_1 = \mathbf{S}$ and a skew tensor $\mathbf{W}_1 = \mathbf{W}$ representing

$$\begin{aligned}\alpha(\mathbf{W}, \mathbf{S}) \xlongequal[\text{ninth rows of Table 6.1}]{\text{from the second, third and}} \bar{\alpha}\Big(&\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3, \mathrm{tr}\,\mathbf{W}^2, \mathrm{tr}\left(\mathbf{SW}^2\right), \\ &\mathrm{tr}\left(\mathbf{S}^2\mathbf{W}^2\right), \mathrm{tr}\left(\mathbf{SWS}^2\mathbf{W}^2\right)\Big) . \quad \longleftarrow \text{see Exercise 6.15}\end{aligned} \tag{6.141}$$

Guided by this result, an isotropic scalar-valued function of an arbitrary tensor $\mathbf{A}$ (with $2\mathbf{S} = \mathbf{A} + \mathbf{A}^{\mathrm{T}}$ and $2\mathbf{W} = \mathbf{A} - \mathbf{A}^{\mathrm{T}}$) admits the representation

$$\begin{aligned}\alpha(\mathbf{A}) = \bar{\alpha}\Big(&\mathrm{tr}\,\mathbf{A}, \mathrm{tr}\,\mathbf{A}^2, \mathrm{tr}\,\mathbf{A}^3, \mathrm{tr}\left(\mathbf{AA}^{\mathrm{T}}\right), \mathrm{tr}\left(\mathbf{AA}^{\mathrm{T}}\right)^2, \\ &\mathrm{tr}\left(\mathbf{A}^2\mathbf{A}^{\mathrm{T}}\right), \mathrm{tr}\left[\left(\mathbf{A}^{\mathrm{T}}\right)^2\mathbf{A}^2\mathbf{A}^{\mathrm{T}}\mathbf{A} - \mathbf{A}^2\left(\mathbf{A}^{\mathrm{T}}\right)^2\mathbf{AA}^{\mathrm{T}}\right]\Big) .\end{aligned} \tag{6.142}$$

As another example, consider an isotropic scalar-valued function of the two symmetric tensors $\mathbf{S}_1 = \mathbf{S}$ and $\mathbf{S}_2 = \mathbf{T}$ which is expressible in the form

$$\alpha(\mathbf{S}, \mathbf{T}) \xlongequal[\text{seventh rows of Table 6.1}]{\text{from the second and}} \bar{\alpha}\Big(\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3, \mathrm{tr}\,\mathbf{T}, \mathrm{tr}\,\mathbf{T}^2, \mathrm{tr}\,\mathbf{T}^3, \mathrm{tr}(\mathbf{ST}),$$

$$\operatorname{tr}\left(\mathbf{ST}^2\right), \operatorname{tr}\left(\mathbf{S}^2\mathbf{T}\right), \operatorname{tr}\left(\mathbf{S}^2\mathbf{T}^2\right)\Big) \ . \quad \longleftarrow \text{see Exercise 6.15} \tag{6.143}$$

Hint: For the given domain

$$\mathscr{D} = \left\{(\mathbf{S}, \mathbf{T}) \in \mathscr{T}_{so}^{sym} \times \mathscr{T}_{so}^{sym}\right\} \ ,$$

notice that a scalar invariant of the form tr (**STS**) is not present in Table 6.1. It can explicitly be shown that such an element is redundant. The proof has been given in Exercise 6.13.

As a further example, consider an isotropic scalar-valued function of a vector $\mathbf{v}_1 = \mathbf{v}$ and a symmetric tensor $\mathbf{S}_1 = \mathbf{S}$ which allows the representation

$$\alpha\left(\mathbf{v}, \mathbf{S}\right) \underset{\text{fifth rows of Table 6.1}}{\overset{\text{from the first, second and}}{=\!=\!=\!=}} \bar{\alpha}\left(\mathbf{v}\cdot\mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \mathbf{v}\cdot\mathbf{S}\mathbf{v}, \mathbf{v}\cdot\mathbf{S}^2\mathbf{v}\right) \ . \tag{6.144}$$

The representation (6.143) has been verified in Exercise 6.3. This can be used to prove (6.144) by considering a domain of tensorial variables according to

$$\mathscr{D} = \left\{(\mathbf{S}_1 = \mathbf{v}\otimes\mathbf{v}, \mathbf{S}_2 = \mathbf{S}) \in \mathscr{T}_{so}^{sym} \times \mathscr{T}_{so}^{sym}\right\} \ ,$$

Table 6.1 List of isotropic **scalar** invariants

Variables	Invariant elements
$\mathbf{v}_1$	$\mathbf{v}_1\cdot\mathbf{v}_1$
$\mathbf{S}_1$	$\operatorname{tr}\mathbf{S}_1$, $\operatorname{tr}\mathbf{S}_1^2$, $\operatorname{tr}\mathbf{S}_1^3$
$\mathbf{W}_1$	$\operatorname{tr}\mathbf{W}_1^2$
$\mathbf{v}_1,\ \mathbf{v}_2$	$\mathbf{v}_1\cdot\mathbf{v}_2$
$\mathbf{v}_1,\ \mathbf{S}_1$	$\mathbf{v}_1\cdot\mathbf{S}_1\mathbf{v}_1,\ \mathbf{v}_1\cdot\mathbf{S}_1^2\mathbf{v}_1$
$\mathbf{v}_1,\ \mathbf{W}_1$	$\mathbf{v}_1\cdot\mathbf{W}_1^2\mathbf{v}_1$
$\mathbf{S}_1,\ \mathbf{S}_2$	$\operatorname{tr}(\mathbf{S}_1\mathbf{S}_2),\ \operatorname{tr}(\mathbf{S}_1\mathbf{S}_2^2),\ \operatorname{tr}(\mathbf{S}_1^2\mathbf{S}_2),\ \operatorname{tr}(\mathbf{S}_1^2\mathbf{S}_2^2)$
$\mathbf{W}_1,\ \mathbf{W}_2$	$\operatorname{tr}(\mathbf{W}_1\mathbf{W}_2)$
$\mathbf{S}_1,\ \mathbf{W}_1$	$\operatorname{tr}(\mathbf{S}_1\mathbf{W}_1^2),\ \operatorname{tr}(\mathbf{S}_1^2\mathbf{W}_1^2),\ \operatorname{tr}(\mathbf{S}_1\mathbf{W}_1\mathbf{S}_1^2\mathbf{W}_1^2)$
$\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{S}_1$	$\mathbf{v}_1\cdot\mathbf{S}_1\mathbf{v}_2,\ \mathbf{v}_1\cdot\mathbf{S}_1^2\mathbf{v}_2$
$\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{W}_1$	$\mathbf{v}_1\cdot\mathbf{W}_1\mathbf{v}_2,\ \mathbf{v}_1\cdot\mathbf{W}_1^2\mathbf{v}_2$
$\mathbf{v}_1,\ \mathbf{S}_1,\ \mathbf{S}_2$	$\mathbf{v}_1\cdot\mathbf{S}_1\mathbf{S}_2\mathbf{v}_1$
$\mathbf{v}_1,\ \mathbf{W}_1,\ \mathbf{W}_2$	$\mathbf{v}_1\cdot\mathbf{W}_1\mathbf{W}_2\mathbf{v}_1,\ \mathbf{v}_1\cdot\mathbf{W}_1^2\mathbf{W}_2\mathbf{v}_1,\ \mathbf{v}_1\cdot\mathbf{W}_1\mathbf{W}_2^2\mathbf{v}_1$
$\mathbf{v}_1,\ \mathbf{S}_1,\ \mathbf{W}_1$	$\mathbf{v}_1\cdot\mathbf{W}_1\mathbf{S}_1\mathbf{v}_1,\ \mathbf{v}_1\cdot\mathbf{W}_1\mathbf{S}_1^2\mathbf{v}_1,\ \mathbf{v}_1\cdot\mathbf{W}_1\mathbf{S}_1\mathbf{W}_1^2\mathbf{v}_1$
$\mathbf{S}_1,\ \mathbf{S}_2,\ \mathbf{S}_3$	$\operatorname{tr}(\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3)$
$\mathbf{W}_1,\ \mathbf{W}_2,\ \mathbf{W}_3$	$\operatorname{tr}(\mathbf{W}_1\mathbf{W}_2\mathbf{W}_3)$
$\mathbf{S}_1,\ \mathbf{S}_2,\ \mathbf{W}_1$	$\operatorname{tr}(\mathbf{S}_1\mathbf{S}_2\mathbf{W}_1),\ \operatorname{tr}(\mathbf{S}_1\mathbf{S}_2^2\mathbf{W}_1),\ \operatorname{tr}(\mathbf{S}_1^2\mathbf{S}_2\mathbf{W}_1),\ \operatorname{tr}(\mathbf{S}_1\mathbf{W}_1\mathbf{S}_2\mathbf{W}_1^2)$
$\mathbf{S}_1,\ \mathbf{W}_1,\ \mathbf{W}_2$	$\operatorname{tr}(\mathbf{S}_1\mathbf{W}_1\mathbf{W}_2),\ \operatorname{tr}(\mathbf{S}_1\mathbf{W}_1^2\mathbf{W}_2),\ \operatorname{tr}(\mathbf{S}_1\mathbf{W}_1\mathbf{W}_2^2)$
$\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{S}_1,\ \mathbf{S}_2$	$\mathbf{v}_1\cdot\mathbf{S}_1\mathbf{S}_2\mathbf{v}_2,\ \mathbf{v}_1\cdot\mathbf{S}_2\mathbf{S}_1\mathbf{v}_2$
$\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{W}_1,\ \mathbf{W}_2$	$\mathbf{v}_1\cdot\mathbf{W}_1\mathbf{W}_2\mathbf{v}_2,\ \mathbf{v}_1\cdot\mathbf{W}_2\mathbf{W}_1\mathbf{v}_2$
$\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{S}_1,\ \mathbf{W}_1$	$\mathbf{v}_1\cdot\mathbf{S}_1\mathbf{W}_1\mathbf{v}_2,\ \mathbf{v}_1\cdot\mathbf{W}_1\mathbf{S}_1\mathbf{v}_2$

which helps represent

$$\alpha\left(\mathbf{v}\otimes\mathbf{v},\mathbf{S}\right) \xlongequal[\text{seventh rows of Table 6.1}]{\text{from the second and}} \bar{\alpha}\Big(\operatorname{tr}\left(\mathbf{v}\otimes\mathbf{v}\right)=\mathbf{v}\cdot\mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \operatorname{tr}\left[\left(\mathbf{v}\otimes\mathbf{v}\right)\mathbf{S}\right]=\mathbf{v}\cdot\mathbf{S}\mathbf{v}, \operatorname{tr}\left[\left(\mathbf{v}\otimes\mathbf{v}\right)\mathbf{S}^2\right]=\mathbf{v}\cdot\mathbf{S}^2\mathbf{v}\Big) \ . \tag{6.145}$$

Note that, for instance, the quantity $\operatorname{tr}\left(\mathbf{v}\otimes\mathbf{v}\right)^2=\left(\mathbf{v}\cdot\mathbf{v}\right)^2$ has not been written in the above list of arguments to provide an irreducible representation.

The last example here regards an isotropic scalar-valued function of the two vectors $\mathbf{v}_1=\mathbf{u}$ and $\mathbf{v}_2=\mathbf{v}$ which admits the following representation

$$\alpha\left(\mathbf{u},\mathbf{v}\right) \xlongequal[\text{fourth rows of Table 6.1}]{\text{from the first and}} \bar{\alpha}\left(\mathbf{u}\cdot\mathbf{u},\mathbf{v}\cdot\mathbf{v},\mathbf{u}\cdot\mathbf{v}\right) \ . \quad \longleftarrow \text{see Exercise 6.15} \tag{6.146}$$

This representation makes sense because only the magnitudes of two vectors and the angle between them remain invariant under an orthogonal transformation. The procedure used to establish this representation is explained in Exercise 6.3. Notice that (6.143) can again be utilized to verify (6.146) by considering a domain of tensorial variables according to

$$\mathscr{D}=\left\{\left(\mathbf{S}_1=\mathbf{u}\otimes\mathbf{u},\mathbf{S}_2=\mathbf{v}\otimes\mathbf{v}\right)\in\mathscr{T}_{so}^{sym}\times\mathscr{T}_{so}^{sym}\right\} \ ,$$

and then removing the redundant terms in the resulting set of the basic invariants. ❦

The interested reader may want to use Lemma A to verify the representation (6.146) one more time (see Liu [23]). This is demonstrated in the following.

From (6.102), one will have

$$\frac{\partial\alpha}{\partial u_2}u_3-\frac{\partial\alpha}{\partial u_3}u_2+\frac{\partial\alpha}{\partial v_2}v_3-\frac{\partial\alpha}{\partial v_3}v_2=0 \ , \tag{6.147a}$$

$$\frac{\partial\alpha}{\partial u_3}u_1-\frac{\partial\alpha}{\partial u_1}u_3+\frac{\partial\alpha}{\partial v_3}v_1-\frac{\partial\alpha}{\partial v_1}v_3=0 \ , \tag{6.147b}$$

$$\frac{\partial\alpha}{\partial u_1}u_2-\frac{\partial\alpha}{\partial u_2}u_1+\frac{\partial\alpha}{\partial v_1}v_2-\frac{\partial\alpha}{\partial v_2}v_1=0 \ . \tag{6.147c}$$

From the first-order partial differential equation (6.147c), one then obtains the general solution

$$\alpha=\bar{\alpha}\left(d_1,d_2,d_3\right) \ , \tag{6.148}$$

where $\bar{\alpha}$ presents an arbitrary function of

$$d_1=u_1^2+u_2^2+f_1\left(u_3,v_3\right) \ , \tag{6.149a}$$

$$d_2=v_1^2+v_2^2+f_2\left(u_3,v_3\right) \ , \tag{6.149b}$$

$$d_3=u_1v_1+u_2v_2+f_3\left(u_3,v_3\right) \ . \tag{6.149c}$$

From (6.147b) and (6.148)–(6.149c), it follows that

$$\begin{aligned}&\left(\frac{\partial f_1}{\partial u_3}u_1 - 2u_1u_3 + \frac{\partial f_1}{\partial v_3}v_1\right)\frac{\partial \bar{\alpha}}{\partial d_1} + \left(\frac{\partial f_2}{\partial u_3}u_1 + \frac{\partial f_2}{\partial v_3}v_1 - 2v_1v_3\right)\frac{\partial \bar{\alpha}}{\partial d_2}\\ &+ \left(\frac{\partial f_3}{\partial u_3}u_1 - v_1u_3 + \frac{\partial f_3}{\partial v_3}v_1 - u_1v_3\right)\frac{\partial \bar{\alpha}}{\partial d_3} = 0 \,. \end{aligned} \tag{6.150}$$

The fact that $\partial\bar{\alpha}/\partial d_i, \ i = 1, 2, 3$, are arbitrary now implies that

$$\left(\frac{\partial f_1}{\partial u_3} - 2u_3\right)u_1 + \left(\frac{\partial f_1}{\partial v_3}\right)v_1 = 0 \,, \tag{6.151a}$$

$$\left(\frac{\partial f_2}{\partial u_3}\right)u_1 + \left(\frac{\partial f_2}{\partial v_3} - 2v_3\right)v_1 = 0 \,, \tag{6.151b}$$

$$\left(\frac{\partial f_3}{\partial u_3} - v_3\right)u_1 + \left(\frac{\partial f_3}{\partial v_3} - u_3\right)v_1 = 0 \,. \tag{6.151c}$$

Note that u_1 and v_1 are also arbitrary. And their coefficients in the above relations can vary independently. It is then easy to see that

$$f_1 = u_3^2 \quad , \quad f_2 = v_3^2 \quad , \quad f_3 = u_3v_3 \,. \tag{6.152}$$

Consequently, $d_1 = \mathbf{u}\cdot\mathbf{u}$, $d_2 = \mathbf{v}\cdot\mathbf{v}$ and $d_3 = \mathbf{u}\cdot\mathbf{v}$ help establish the desired representation $\alpha\left(\mathbf{u}, \mathbf{v}\right) = \bar{\alpha}\left(\mathbf{u}\cdot\mathbf{u}, \mathbf{v}\cdot\mathbf{v}, \mathbf{u}\cdot\mathbf{v}\right)$.

6.2.4 Representation Theorem for an Isotropic Vector-Valued Function of a Vector

Let $\mathbf{a} : \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3} \to \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$ be an **isotropic** vector-valued function of a vector. Then, it is necessary and sufficient that it has the following representation

$$\boxed{\mathbf{a}\left(\mathbf{v}\right) = f\left(\mathbf{v}\cdot\mathbf{v}\right)\mathbf{v} \,,} \tag{6.153}$$

where f presents an arbitrary isotropic scalar-valued function.

Proof By means of (2.5), (2.8b), (2.51d), $(2.130)_1$, (6.85) and $(6.92)_1$, the sufficiency can readily be checked:

$$\begin{aligned}\mathbf{a}\left(\mathbf{Qv}\right) &= f\left(\mathbf{Qv}\cdot\mathbf{Qv}\right)\mathbf{Qv} = f\left(\mathbf{v}\cdot\mathbf{Q}^{\mathrm{T}}\mathbf{Qv}\right)\mathbf{Qv}\\ &= f\left(\mathbf{v}\cdot\mathbf{Iv}\right)\mathbf{Qv} = f\left(\mathbf{v}\cdot\mathbf{v}\right)\mathbf{Qv}\\ &= \mathbf{Qa}\left(\mathbf{v}\right) \,.\end{aligned} \tag{6.154}$$

One can now verify (6.153) assuming that the isotropy condition $\mathbf{a}(\mathbf{Qv}) = \mathbf{Qa}(\mathbf{v})$ holds true. Suppose that $\mathbf{v} = \mathbf{0}$. Then, the condition of isotropy implies that $\mathbf{a}(\mathbf{0}) = \mathbf{Qa}(\mathbf{0})$ for any $\mathbf{Q} \in \mathscr{O}$. As a result, $\mathbf{a}(\mathbf{0}) = \mathbf{0}$ and, consequently, (6.153) is fulfilled. Next, suppose that $\mathbf{v} \neq \mathbf{0}$. Then, the vector function $\mathbf{a}$ can additively be decomposed as

$$\mathbf{a}(\mathbf{v}) = \beta_1(\mathbf{v})\,\mathbf{v} + \beta_2(\mathbf{v})\,\mathbf{v}^{\perp} \,, \tag{6.155}$$

where β_1 and β_1 are arbitrary scalar functions and $\mathbf{v}^{\perp}$ denotes some vector orthogonal to $\mathbf{v}$. At this stage, one can choose an orthogonal tensor of the following form

$$\mathbf{Q} = \frac{2}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} \otimes \mathbf{v} - \mathbf{I} \quad \text{satisfying} \quad \mathbf{Qv} = \mathbf{v} \quad \text{and} \quad \mathbf{Qv}^{\perp} = -\mathbf{v}^{\perp} \,. \tag{6.156}$$

Then,

$$\begin{aligned} \mathbf{a}(\mathbf{Qv}) &\overset{\text{from}}{\underset{(6.155)}{=\!=\!=}} \beta_1(\mathbf{Qv})\,\mathbf{Qv} + \beta_2(\mathbf{Qv})\,\mathbf{v}^{\perp} \\ &\overset{\text{from}}{\underset{(6.156)}{=\!=\!=}} \beta_1(\mathbf{v})\,\mathbf{v} + \beta_2(\mathbf{v})\,\mathbf{v}^{\perp} \,, \end{aligned} \tag{6.157a}$$

$$\begin{aligned} \mathbf{Qa}(\mathbf{v}) &\overset{\text{from}}{\underset{(6.155)}{=\!=\!=}} \beta_1(\mathbf{v})\,\mathbf{Qv} + \beta_2(\mathbf{v})\,\mathbf{Qv}^{\perp} \\ &\overset{\text{from}}{\underset{(6.156)}{=\!=\!=}} \beta_1(\mathbf{v})\,\mathbf{v} - \beta_2(\mathbf{v})\,\mathbf{v}^{\perp} \,. \end{aligned} \tag{6.157b}$$

Consequently, the condition of isotropy implies that $\beta_2(\mathbf{v}) = 0$. Moreover, isotropy of $\mathbf{a}(\mathbf{v})$ implies that $\beta_1(\mathbf{v})$ should be an isotropic scalar-valued function. Thus, by (6.128), it can be represented by $\beta_1(\mathbf{v}) = f(\mathbf{v} \cdot \mathbf{v})$.

The goal here is to obtain (6.153) in an alternative way for the interested reader. Note that $\mathbf{v}$ is the generator of the isotropic vector-valued function $\mathbf{a}(\mathbf{v})$ in (6.153). Let $\alpha(\mathbf{u}, \mathbf{v})$ be an isotropic scalar-valued function of the two vectors $\mathbf{u}$ and $\mathbf{v}$. Suppose α is linear in $\mathbf{u}$. Thus, using (6.146), it depends on $\{\mathbf{u} \cdot \mathbf{v}, \mathbf{v} \cdot \mathbf{v}\}$. Notice that the derivative of the elements of this set with respect to $\mathbf{u}$ gives the only generator of $\mathbf{a}(\mathbf{v})$ which is $\mathbf{v}$. This motivates to define a scalar-valued function of the two vectors $\mathbf{u}$ and $\mathbf{v}$ according to

$$\beta(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \hat{\mathbf{a}}(\mathbf{v}) \quad \text{noting that} \quad \hat{\mathbf{a}}(\mathbf{v}) = \frac{\partial \beta(\mathbf{u}, \mathbf{v})}{\partial \mathbf{u}} \,.$$

Here, $\hat{\mathbf{a}}(\mathbf{v})$ presents an isotropic vector-valued function of $\mathbf{v}$. Notice that β represents a linear function of $\mathbf{u}$ which is isotropic owing to

$$\beta(\mathbf{Qu}, \mathbf{Qv}) = \mathbf{Qu} \cdot \hat{\mathbf{a}}(\mathbf{Qv}) = \mathbf{Qu} \cdot \mathbf{Q}\hat{\mathbf{a}}(\mathbf{v}) = \mathbf{u} \cdot \mathbf{Q}^{\mathrm{T}}\mathbf{Q}\hat{\mathbf{a}}(\mathbf{v})$$

Table 6.2 List of isotropic **vector** invariants

Variables	Generator elements
$\mathbf{v}_1$	$\mathbf{v}_1$
$\mathbf{S}_1$	$\mathbf{0}$
$\mathbf{W}_1$	$\mathbf{0}$
$\mathbf{v}_1, \ \mathbf{S}_1$	$\mathbf{S}_1\mathbf{v}_1, \ \mathbf{S}_1^2\mathbf{v}_1$
$\mathbf{v}_1, \ \mathbf{W}_1$	$\mathbf{W}_1\mathbf{v}_1, \ \mathbf{W}_1^2\mathbf{v}_1$
$\mathbf{v}_1, \ \mathbf{S}_1, \ \mathbf{S}_2$	$\mathbf{S}_1\mathbf{S}_2\mathbf{v}_1, \ \mathbf{S}_2\mathbf{S}_1\mathbf{v}_1$
$\mathbf{v}_1, \ \mathbf{W}_1, \ \mathbf{W}_2$	$\mathbf{W}_1\mathbf{W}_2\mathbf{v}_1, \ \mathbf{W}_2\mathbf{W}_1\mathbf{v}_1$
$\mathbf{v}_1, \ \mathbf{S}_1, \ \mathbf{W}_1$	$\mathbf{S}_1\mathbf{W}_1\mathbf{v}_1, \ \mathbf{W}_1\mathbf{S}_1\mathbf{v}_1$

$$= \mathbf{u} \cdot \mathbf{I}\hat{\mathbf{a}}(\mathbf{v}) = \mathbf{u} \cdot \hat{\mathbf{a}}(\mathbf{v}) = \beta(\mathbf{u}, \mathbf{v}) \ .$$

Consequently, this isotropic scalar-valued function of $\{\mathbf{u} \cdot \mathbf{v}, \mathbf{v} \cdot \mathbf{v}\}$ should generally be represented by

$$\beta(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \hat{\mathbf{a}}(\mathbf{v}) = \mathbf{u} \cdot f(\mathbf{v} \cdot \mathbf{v})\,\mathbf{v} \ .$$

Thus, using (1.14), one can conclude that $\hat{\mathbf{a}}(\mathbf{v}) = f(\mathbf{v} \cdot \mathbf{v})\,\mathbf{v}$. See Exercise 6.12 for more consideration on this procedure.

Hint: The above procedure is restricted to only **polynomial representations** (i.e. it delivers integrity bases). Thus, it cannot provide all generators listed in the tables of this text having in mind that irreducible integrity and functional bases generally do not have the same elements. It properly worked here since $\hat{\mathbf{a}}(\mathbf{v}) = f(\mathbf{v} \cdot \mathbf{v})\,\mathbf{v}$ with its generator element $\mathbf{v}$, listed in Table 6.2 as a functional basis, is also an integrity basis.

6.2.5 Representation Theorem for an Isotropic Vector-Valued Function of a Symmetric Tensor

Let $\mathbf{a} : \mathscr{T}_{\text{so}}^{\text{sym}} \to \mathscr{E}_{\text{r}}^{\text{o3}}$ be an **isotropic** vector-valued function of a symmetric tensor. Then, it is necessary and sufficient that it has the following representation

$$\boxed{\mathbf{a}(\mathbf{S}) = \mathbf{0} \ .} \tag{6.158}$$

Proof The sufficiency can easily be checked:

$$\begin{aligned} \mathbf{a}\left(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}\right) &= \mathbf{0} \\ &= \mathbf{Q}\mathbf{0} \\ &= \mathbf{Q}\mathbf{a}(\mathbf{S}) \ . \end{aligned} \tag{6.159}$$

One can now establish (6.158) assuming that the isotropy condition $\mathbf{a}\left(\mathbf{QSQ}^{\mathrm{T}}\right) = \mathbf{Qa}(\mathbf{S})$ holds true. By choosing $\mathbf{Q} = -\mathbf{I}$, taking into account (2.5), $(2.33)_{1-2}$ and (2.55a), one will have

$$\mathbf{a}\left(+\mathbf{ISI}^{\mathrm{T}}\right) = -\mathbf{Ia}(\mathbf{S}) \quad \text{or} \quad \mathbf{a}(\mathbf{S}) = -\mathbf{a}(\mathbf{S}) \ , \tag{6.160}$$

which reveals the fact that any isotropic vector-valued function of a symmetric tensor is nothing but the zero vector.

The established representation theorems for isotropic vector invariants of one tensor variable can be generalized to involve more tensorial arguments. The results have been tabulated in Table 6.2.

As an example, consider an isotropic vector-valued function of the two tensorial variables $\mathbf{v}_1 = \mathbf{v}$ and $\mathbf{S}_1 = \mathbf{S}$ which allows the following representation

$$\mathbf{a}(\mathbf{v}, \mathbf{S}) \xlongequal[\text{fourth rows of Table 6.2}]{\text{from the first, second and}} f_0(\Upsilon_{\mathrm{s}})\,\mathbf{v} + f_1(\Upsilon_{\mathrm{s}})\,\mathbf{Sv} + f_2(\Upsilon_{\mathrm{s}})\,\mathbf{S}^2\mathbf{v} \ , \tag{6.161}$$

where f_0, f_1 and f_2 are arbitrary isotropic scalar invariants of the following set

$$\Upsilon_{\mathrm{s}} \xlongequal[\text{fifth rows of Table 6.1}]{\text{from the first, second and}} \left\{\mathbf{v}\cdot\mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \mathbf{v}\cdot\mathbf{Sv}, \mathbf{v}\cdot\mathbf{S}^2\mathbf{v}\right\} \ . \tag{6.162}$$

The goal here is to obtain (6.161) in an alternative way for the interested reader. As can be seen from (6.161), $\left\{\mathbf{v}, \mathbf{Sv}, \mathbf{S}^2\mathbf{v}\right\}$ is the generating set of the isotropic vector-valued function $\mathbf{a}(\mathbf{v}, \mathbf{S})$. Let $\alpha(\mathbf{u}, \mathbf{v}, \mathbf{S})$ be an isotropic scalar-valued function of the three tensorial variables $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{S}$. Suppose that α is a linear function of $\mathbf{u}$. Thus, by using Table 6.1, it should depend on

$$\Upsilon_{\mathrm{s}}^{\bullet} = \left\{\mathbf{u}\cdot\mathbf{v}, \mathbf{u}\cdot\mathbf{Sv}, \mathbf{u}\cdot\mathbf{S}^2\mathbf{v}, \mathbf{v}\cdot\mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \mathbf{v}\cdot\mathbf{Sv}, \mathbf{v}\cdot\mathbf{S}^2\mathbf{v}\right\} \ .$$

Note that the derivative of the elements of this set with respect to $\mathbf{u}$ provides the generators of $\mathbf{a}(\mathbf{v}, \mathbf{S})$. This motivates to define a scalar-valued function of the three tensorial variables $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{S}$ according to

$$\beta(\mathbf{u}, \mathbf{v}, \mathbf{S}) = \mathbf{u}\cdot\hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) \quad \text{noting that} \quad \hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) = \frac{\partial\beta(\mathbf{u}, \mathbf{v}, \mathbf{S})}{\partial\mathbf{u}} \ .$$

Here, $\hat{\mathbf{a}}(\mathbf{v}, \mathbf{S})$ is an isotropic vector-valued function of $\mathbf{v}$ and $\mathbf{S}$. Notice that β is a linear function of $\mathbf{u}$ which satisfies the isotropy condition

$$\begin{aligned}
\beta\left(\mathbf{Qu}, \mathbf{Qv}, \mathbf{QSQ}^{\mathrm{T}}\right) &= \mathbf{Qu}\cdot\hat{\mathbf{a}}\left(\mathbf{Qv}, \mathbf{QSQ}^{\mathrm{T}}\right) = \mathbf{Qu}\cdot\mathbf{Q}\hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) \\
&= \mathbf{u}\cdot\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) = \mathbf{u}\cdot\mathbf{I}\hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) \\
&= \mathbf{u}\cdot\hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) = \beta(\mathbf{u}, \mathbf{v}, \mathbf{S}) \ .
\end{aligned}$$

Consequently, this isotropic function of $\Upsilon_s^{\bullet}$ should generally be of the following form

$$\begin{aligned}\beta(\mathbf{u}, \mathbf{v}, \mathbf{S}) = \mathbf{u} \cdot \hat{\mathbf{a}}(\mathbf{v}, \mathbf{S}) = \mathbf{u} \cdot f_0 \left(\mathbf{v} \cdot \mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \mathbf{v} \cdot \mathbf{S}\mathbf{v}, \mathbf{v} \cdot \mathbf{S}^2\mathbf{v}\right) \mathbf{v} \\ + \mathbf{u} \cdot f_1 \left(\mathbf{v} \cdot \mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \mathbf{v} \cdot \mathbf{S}\mathbf{v}, \mathbf{v} \cdot \mathbf{S}^2\mathbf{v}\right) \mathbf{S}\mathbf{v} \\ + \mathbf{u} \cdot f_2 \left(\mathbf{v} \cdot \mathbf{v}, \operatorname{tr}\mathbf{S}, \operatorname{tr}\mathbf{S}^2, \operatorname{tr}\mathbf{S}^3, \mathbf{v} \cdot \mathbf{S}\mathbf{v}, \mathbf{v} \cdot \mathbf{S}^2\mathbf{v}\right) \mathbf{S}^2\mathbf{v} \ .\end{aligned}$$

Thus, by using (1.14), one can arrive at the representation (6.161).

6.2.6 Representation Theorem for an Isotropic Symmetric Tensor-Valued Function of a Vector

Let $\mathbf{T} : \mathscr{E}_r^{o3} \to \mathscr{T}_{so}^{sym}$ be an **isotropic** tensor-valued function of a vector. Then, it is necessary and sufficient that it has the following representation

$$\boxed{\mathbf{T}(\mathbf{v}) = f_1(\mathbf{v} \cdot \mathbf{v})\, \mathbf{I} + f_2(\mathbf{v} \cdot \mathbf{v})\, \mathbf{v} \otimes \mathbf{v} \ ,} \tag{6.163}$$

where f_1 and f_2 are arbitrary isotropic scalar-valued functions.

Proof The sufficiency can readily be checked:

$$\begin{aligned}\mathbf{T}(\mathbf{Q}\mathbf{v}) &= f_1(\mathbf{Q}\mathbf{v} \cdot \mathbf{Q}\mathbf{v})\, \mathbf{I} + f_2(\mathbf{Q}\mathbf{v} \cdot \mathbf{Q}\mathbf{v})\, \mathbf{Q}\mathbf{v} \otimes \mathbf{Q}\mathbf{v} \\ &= f_1\left(\mathbf{v} \cdot \mathbf{Q}^T\mathbf{Q}\mathbf{v}\right) \mathbf{Q}\mathbf{I}\mathbf{Q}^T + f_2\left(\mathbf{v} \cdot \mathbf{Q}^T\mathbf{Q}\mathbf{v}\right) \mathbf{Q}\mathbf{v} \otimes \mathbf{v}\mathbf{Q}^T \\ &= \mathbf{Q}\left[f_1(\mathbf{v} \cdot \mathbf{I}\mathbf{v})\, \mathbf{I} + f_2(\mathbf{v} \cdot \mathbf{I}\mathbf{v})\, \mathbf{v} \otimes \mathbf{v}\right] \mathbf{Q}^T \\ &= \mathbf{Q}\left[f_1(\mathbf{v} \cdot \mathbf{v})\, \mathbf{I} + f_2(\mathbf{v} \cdot \mathbf{v})\, \mathbf{v} \otimes \mathbf{v}\right] \mathbf{Q}^T \\ &= \mathbf{Q}\mathbf{T}(\mathbf{v})\, \mathbf{Q}^T \ .\end{aligned} \tag{6.164}$$

One can now establish (6.163) assuming that $\mathbf{T}(\mathbf{Q}\mathbf{v}) = \mathbf{Q}\mathbf{T}(\mathbf{v})\,\mathbf{Q}^T$ holds true. Suppose one is given a scalar-valued function $\alpha(\mathbf{u}, \mathbf{v})$ of the form

$$\alpha(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{T}(\mathbf{v})\, \mathbf{v} \ , \tag{6.165}$$

which is linear in $\mathbf{u}$ and isotropic:

$$\begin{aligned}\alpha(\mathbf{Q}\mathbf{u}, \mathbf{Q}\mathbf{v}) = \mathbf{Q}\mathbf{u} \cdot \mathbf{T}(\mathbf{Q}\mathbf{v})\, \mathbf{Q}\mathbf{v} &= \mathbf{Q}\mathbf{u} \cdot \mathbf{Q}\mathbf{T}(\mathbf{v})\, \mathbf{Q}^T\mathbf{Q}\mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{Q}^T\mathbf{Q}\mathbf{T}(\mathbf{v})\, \mathbf{I}\mathbf{v} = \mathbf{u} \cdot \mathbf{I}\mathbf{T}(\mathbf{v})\, \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{T}(\mathbf{v})\, \mathbf{v} = \alpha(\mathbf{u}, \mathbf{v}) \ .\end{aligned} \tag{6.166}$$

Consequently, using the first and fourth rows of Table 6.1, this isotropic function can generally be represented by

$$\alpha(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \lambda(\mathbf{v} \cdot \mathbf{v})\, \mathbf{v} \, . \tag{6.167}$$

From (1.14), (6.165) and (6.167), one can conclude that

$$\mathbf{T}(\mathbf{v})\, \mathbf{v} = \lambda(\mathbf{v} \cdot \mathbf{v})\, \mathbf{v} \, , \tag{6.168}$$

which reveals the fact that $(\lambda, \mathbf{v})$ or $(\lambda, \widehat{\mathbf{v}} = \mathbf{v}/|\mathbf{v}|)$ is an eigenpair of the symmetric tensor $\mathbf{T}$ (note that this eigenvalue problem can also be obtained by defining an isotropic vector-valued function of the form $\mathbf{a}(\mathbf{v}) = \mathbf{T}(\mathbf{v})\, \mathbf{v}$ and then considering $\mathbf{a}(\mathbf{v}) = \lambda(\mathbf{v} \cdot \mathbf{v})\, \mathbf{v}$ according to (6.153)).

Suppose $\mathbf{T}$ is a tensor with the three distinct eigenvalues λ, μ and ν. Denoting by $\{\widehat{\mathbf{v}}, \widehat{\mathbf{u}}, \widehat{\mathbf{w}}\}$ the corresponding orthonormal set of eigenvectors, the spectral decomposition of $\mathbf{T}$ is given by

$$\mathbf{T}(\mathbf{v}) \xlongequal[\text{(4.41)}]{\text{from}} \lambda(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \mu(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}} + \nu(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{w}} \otimes \widehat{\mathbf{w}} \, . \tag{6.169}$$

Let $\mathbf{Q}$ be a tensor of the form

$$\mathbf{Q} = \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{u}} \otimes \widehat{\mathbf{w}} + \widehat{\mathbf{w}} \otimes \widehat{\mathbf{u}} \, , \tag{6.170}$$

which satisfies the orthogonality condition

$$\begin{aligned} \mathbf{Q}\mathbf{Q}^{\mathrm{T}} &= (\widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{u}} \otimes \widehat{\mathbf{w}} + \widehat{\mathbf{w}} \otimes \widehat{\mathbf{u}})\,(\widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{w}} \otimes \widehat{\mathbf{u}} + \widehat{\mathbf{u}} \otimes \widehat{\mathbf{w}}) \\ &= \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}} + \widehat{\mathbf{w}} \otimes \widehat{\mathbf{w}} = \mathbf{I} \, , \end{aligned} \tag{6.171}$$

and

$$\mathbf{Q}\widehat{\mathbf{v}} = \widehat{\mathbf{v}} \quad \text{or} \quad \mathbf{Q}\mathbf{v} = \mathbf{v} \, , \tag{6.172a}$$

$$\mathbf{Q}\mathbf{T}(\mathbf{v})\, \mathbf{Q}^{\mathrm{T}} = \lambda(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \nu(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}} + \mu(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{w}} \otimes \widehat{\mathbf{w}} \, . \tag{6.172b}$$

Making use of the relations (6.169), $(6.172\text{a})_1$ and (6.172b), the isotropy requirement $\mathbf{T}(\mathbf{Q}\mathbf{v}) = \mathbf{Q}\mathbf{T}(\mathbf{v})\, \mathbf{Q}^{\mathrm{T}}$ then implies that $\mu(\mathbf{v} \cdot \mathbf{v}) = \nu(\mathbf{v} \cdot \mathbf{v})$. As a result, (6.169) can be rewritten as

$$\begin{aligned} \mathbf{T}(\mathbf{v}) &= \lambda(\mathbf{v} \cdot \mathbf{v})\, \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \mu(\mathbf{v} \cdot \mathbf{v})\, [\mathbf{I} - \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}}] \\ &= \underbrace{\mu(\mathbf{v} \cdot \mathbf{v})}_{:= f_1(\mathbf{v} \cdot \mathbf{v})} \mathbf{I} + \underbrace{\frac{\lambda(\mathbf{v} \cdot \mathbf{v}) - \mu(\mathbf{v} \cdot \mathbf{v})}{\mathbf{v} \cdot \mathbf{v}}}_{:= f_2(\mathbf{v} \cdot \mathbf{v})} \mathbf{v} \otimes \mathbf{v} \, . \end{aligned} \tag{6.173}$$

And this completes the proof (note that the cases in which $\mathbf{T}$ possess repeated eigenvalues are trivial).

6.2.7 Representation Theorem for an Isotropic Symmetric Tensor-Valued Function of a Symmetric Tensor

Let $\mathbf{T} : \mathscr{T}_{\text{so}}^{\text{sym}} \to \mathscr{T}_{\text{so}}^{\text{sym}}$ be an **isotropic** symmetric tensor-valued function of a symmetric tensor. Then, it is necessary and sufficient that it has the following representation

$$\mathbf{T}(\mathbf{S}) = f_0 \mathbf{I} + f_1 \mathbf{S} + f_2 \mathbf{S}^2 \qquad \text{if} \quad \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \tag{6.174a}$$

$$\mathbf{T}(\mathbf{S}) = \bar{f}_0 \mathbf{I} + \bar{f}_1 \mathbf{S} \qquad \text{if} \quad \lambda_1 \neq \lambda_2 = \lambda_3 = \lambda \ , \tag{6.174b}$$

$$\mathbf{T}(\mathbf{S}) = \hat{f}_0 \mathbf{I} \qquad \text{if} \quad \lambda_1 = \lambda_2 = \lambda_3 = \lambda \ , \tag{6.174c}$$

where the coefficients are arbitrary functions of the principal scalar invariants $I_1(\mathbf{S})$, $I_2(\mathbf{S})$ and $I_3(\mathbf{S})$ given in (4.17a)–(4.17c).

Proof Since the spectrum of a symmetric tensor may contain a real multiple eigenvalue, three different cases need to be considered. First, let $\mathbf{S}$ be a tensor with three distinct eigenvalues; namely, $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1$. By means of (2.33), $(2.130)_{1-2}$, (6.85), (6.89) and $(6.92)_2$, the sufficiency can readily be verified:

$$\begin{aligned} \mathbf{T}\left(\mathbf{QSQ}^{\mathrm{T}}\right) &= f_0 \mathbf{I} + f_1 \mathbf{QSQ}^{\mathrm{T}} + f_2 \mathbf{QSQ}^{\mathrm{T}}\mathbf{QSQ}^{\mathrm{T}} \\ &= f_0 \mathbf{QIQ}^{\mathrm{T}} + f_1 \mathbf{QSQ}^{\mathrm{T}} + f_2 \mathbf{QSISQ}^{\mathrm{T}} \\ &= \mathbf{Q}\left[f_0 \mathbf{I} + f_1 \mathbf{S} + f_2 \mathbf{S}^2\right]\mathbf{Q}^{\mathrm{T}} \\ &= \mathbf{Q}\left[\mathbf{T}(\mathbf{S})\right]\mathbf{Q}^{\mathrm{T}} \ . \end{aligned} \tag{6.175}$$

The goal is now to verify the representation (6.174a) assuming that the isotropy condition $\mathbf{T}\left(\mathbf{QSQ}^{\mathrm{T}}\right) = \mathbf{QT}(\mathbf{S})\,\mathbf{Q}^{\mathrm{T}}$ holds true. Guided by Lemma C, $\mathbf{S}$ and any isotropic symmetric tensor-valued function of it, say $\mathbf{T}(\mathbf{S})$, have eigenvectors in common. Thus, by the spectral theorem, they represent

$$\mathbf{S} = \sum_{i=1}^{3} \lambda_i(\mathbf{S})\, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \quad , \quad \mathbf{T}(\mathbf{S}) = \sum_{i=1}^{3} \mu_i(\mathbf{S})\, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \ .$$

By means of (6.111), one can infer that there exist three scalars $f_0(\mathbf{S})$, $f_1(\mathbf{S})$ and $f_2(\mathbf{S})$ such that

$$\mathbf{T}(\mathbf{S}) = f_0(\mathbf{S})\,\mathbf{I} + f_1(\mathbf{S})\,\mathbf{S} + f_2(\mathbf{S})\,\mathbf{S}^2 \ . \tag{6.176}$$

Now, one should show that these scalars are isotropic invariants. The isotropy condition $\mathbf{T}(\mathbf{S}) = \mathbf{Q}^{\mathrm{T}}\mathbf{T}\left(\mathbf{QSQ}^{\mathrm{T}}\right)\mathbf{Q}$ implies that

$$\begin{aligned} &\left[f_0(\mathbf{S}) - f_0\left(\mathbf{QSQ}^{\mathrm{T}}\right)\right]\mathbf{I} + \left[f_1(\mathbf{S}) - f_1\left(\mathbf{QSQ}^{\mathrm{T}}\right)\right]\mathbf{S} \\ &\quad + \left[f_2(\mathbf{S}) - f_2\left(\mathbf{QSQ}^{\mathrm{T}}\right)\right]\mathbf{S}^2 = \mathbf{O} \ . \end{aligned} \tag{6.177}$$

Recall from (a) of Lemma B that the set $\{\mathbf{I}, \mathbf{S}, \mathbf{S}^2\}$ was linearly independent. Thus, the relation (6.177) implies that $f_k(\mathbf{S}) = f_k(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}})$, $k = 0, 1, 2$. Guided by (6.132), they can be represented in terms of the elements of the set $\{I_1(\mathbf{S}), I_2(\mathbf{S}), I_3(\mathbf{S})\}$, see (6.139).

Next, consider the case in which $\mathbf{S}$ has exactly two distinct eigenvalues; namely, λ_1 and $\lambda_2 = \lambda_3 = \lambda$. Similarly to (6.175), the sufficiency can easily be verified. Let $\mathbf{T}(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}) = \mathbf{Q}\mathbf{T}(\mathbf{S})\mathbf{Q}^{\mathrm{T}}$. The goal is then to verify (6.174b). In this case, the spectral formula for $\mathbf{S}$ represents $\mathbf{S} = \lambda_1 \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \lambda(\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1)$. And the characteristic spaces of $\mathbf{S}$ render Span $\{\widehat{\mathbf{n}}_1\}$ and Span $\{\widehat{\mathbf{m}} \mid \widehat{\mathbf{m}} \cdot \widehat{\mathbf{n}}_1 = 0\}$. By Lemma C, these subspaces should be contained in the characteristic spaces of $\mathbf{T}(\mathbf{S})$. Consequently,

$$\mathbf{T}(\mathbf{S}) = \mu_1(\mathbf{S})[\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1] + \mu(\mathbf{S})[\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1] \ .$$

Now, guided by (6.112), one can arrive at the desired result $\mathbf{T}(\mathbf{S}) = \bar{f}_0\mathbf{I} + \bar{f}_1\mathbf{S}$. It only remains to verify that $\bar{f}_0$ and $\bar{f}_1$ are isotropic invariants. The isotropy condition $\mathbf{T}(\mathbf{S}) = \mathbf{Q}^{\mathrm{T}}\mathbf{T}(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}})\mathbf{Q}$ implies that

$$\left[\bar{f}_0(\mathbf{S}) - \bar{f}_0(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}})\right]\mathbf{I} + \left[\bar{f}_1(\mathbf{S}) - \bar{f}_1(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}})\right]\mathbf{S} = \mathbf{O} \ . \tag{6.178}$$

Having in mind that the set $\{\mathbf{I}, \mathbf{S}\}$ is linearly independent, one can then conclude that $\bar{f}_k(\mathbf{S}) = \bar{f}_k(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}})$, $k = 0, 1$. Note that the principal scalar invariants can deliver different values regardless of having a multiple eigenvalue or not. Thus, the coefficients should still be functions of $\{I_1(\mathbf{S}), I_2(\mathbf{S}), I_3(\mathbf{S})\}$.

Finally, consider the case in which the symmetric tensor $\mathbf{S}$ has only one distinct eigenvalue; namely, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. In this case, using (4.46), $\mathbf{S} = \lambda\mathbf{I}$. And this means that the whole vector space, $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$, is the characteristic space of $\mathbf{S}$. By Lemma C, $\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$ should also be the characteristic space of $\mathbf{T}(\mathbf{S})$. The representation (6.174c) then follows. And this completes the proof.

The goal here is to obtain (6.174a) in an alternative way for the interested reader. Note that $\{\mathbf{I}, \mathbf{S}, \mathbf{S}^2\}$ is the generating set of the isotropic symmetric tensor-valued function $\mathbf{T}(\mathbf{S})$ in (6.174a). Let $\alpha(\mathbf{U}, \mathbf{S})$ be an isotropic scalar-valued function of the two symmetric tensors $\mathbf{U}$ and $\mathbf{S}$. Suppose that α is a linear function of $\mathbf{U}$. Thus, guided by Table 6.1, it only depends on $\{\mathrm{tr}\,\mathbf{U}, \mathrm{tr}(\mathbf{U}\mathbf{S}), \mathrm{tr}(\mathbf{U}\mathbf{S}^2), \mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\}$. Notice that the derivative of the elements of this set with respect to $\mathbf{U}$ gives the tensor generators of $\mathbf{T}(\mathbf{S})$. This motivates to define a scalar-valued function of the two symmetric tensors $\mathbf{U}$ and $\mathbf{S}$ via

$$\beta(\mathbf{U}, \mathbf{S}) = \mathbf{U} : \hat{\mathbf{T}}(\mathbf{S}) \quad \text{noting that} \quad \hat{\mathbf{T}}(\mathbf{S}) = \frac{\partial \beta(\mathbf{U}, \mathbf{S})}{\partial \mathbf{U}} \ .$$

Here, $\hat{\mathbf{T}}(\mathbf{S})$ denotes an isotropic symmetric tensor-valued function of $\mathbf{S}$. Accordingly, the above linear function of $\mathbf{U}$ satisfies the isotropy requirement

$$\begin{aligned}\beta\left(\mathbf{QUQ}^{\mathrm{T}}, \mathbf{QSQ}^{\mathrm{T}}\right) &= \mathbf{QUQ}^{\mathrm{T}} : \hat{\mathbf{T}}\left(\mathbf{QSQ}^{\mathrm{T}}\right) = \mathbf{QUQ}^{\mathrm{T}} : \mathbf{Q}\hat{\mathbf{T}}\left(\mathbf{S}\right)\mathbf{Q}^{\mathrm{T}} \\ &= \mathbf{U} : \mathbf{Q}^{\mathrm{T}}\mathbf{Q}\hat{\mathbf{T}}\left(\mathbf{S}\right)\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{U} : \mathbf{I}\hat{\mathbf{T}}\left(\mathbf{S}\right)\mathbf{I} \\ &= \mathbf{U} : \hat{\mathbf{T}}\left(\mathbf{S}\right) = \beta\left(\mathbf{U}, \mathbf{S}\right) \ .\end{aligned}$$

Consequently, this isotropic function of $\left\{\mathrm{tr}\,\mathbf{U}, \mathrm{tr}\left(\mathbf{US}\right), \mathrm{tr}\left(\mathbf{US}^2\right), \mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\right\}$ must generally admit the following form

$$\begin{aligned}\beta\left(\mathbf{U}, \mathbf{S}\right) = \mathbf{U} : \hat{\mathbf{T}}\left(\mathbf{S}\right) = \mathbf{U} : h_0\left(\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\right)\mathbf{I} \\ + \mathbf{U} : h_1\left(\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\right)\mathbf{S} \\ + \mathbf{U} : h_2\left(\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\right)\mathbf{S}^2 \ .\end{aligned}$$

Thus, by (2.81) and the fact that $h_k\left(\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\right) = f_k\left(I_1\left(\mathbf{S}\right), I_2\left(\mathbf{S}\right), I_3\left(\mathbf{S}\right)\right)$, $k = 0, 1, 2$, the representation (6.174a) follows.

In a similar fashion, the interested reader can arrive at the representation (6.174a) one more time. To show this, let $\hat{\mathbf{T}}\left(\mathbf{S}\right)$ be an isotropic symmetric tensor-valued function of $\mathbf{S}$ and construct a vector-valued function of $\mathbf{v}$ and $\mathbf{S}$ via

$$\mathbf{a}\left(\mathbf{v}, \mathbf{S}\right) = \hat{\mathbf{T}}\left(\mathbf{S}\right)\mathbf{v} \quad \text{noting that} \quad \hat{\mathbf{T}}\left(\mathbf{S}\right) = \frac{\partial \mathbf{a}\left(\mathbf{v}, \mathbf{S}\right)}{\partial \mathbf{v}} \ .$$

It can readily be shown that this **linear** function of $\mathbf{v}$ is isotropic. Guided by (6.161)–(6.162), it can thus be represented by

$$\mathbf{a}\left(\mathbf{v}, \mathbf{S}\right) = \left[h_0\left(\Upsilon_{\mathrm{s}}\right)\mathbf{I} + h_1\left(\Upsilon_{\mathrm{s}}\right)\mathbf{S} + h_2\left(\Upsilon_{\mathrm{s}}\right)\mathbf{S}^2\right]\mathbf{v} \quad \text{where} \quad \Upsilon_{\mathrm{s}} = \left\{\mathrm{tr}\,\mathbf{S}, \mathrm{tr}\,\mathbf{S}^2, \mathrm{tr}\,\mathbf{S}^3\right\} \ .$$

By (2.6), the representation (6.174a) then follows. As discussed, this procedure cannot establish all functional bases. It worked here because the representation (6.174a) is not only an integrity basis but also a functional basis.

The relation (6.174a) may be used to represent the most general form of an isotropic fourth-order tensor with minor and major symmetries. This is demonstrated in the following.

Let $\mathbf{T}$ be an **isotropic** symmetric tensor function of a symmetric tensor $\mathbf{S}$. Further, suppose $\mathbf{T}$ is linear in $\mathbf{S}$. Then, it should be expressible in the form

$$\mathbf{T}\left(\mathbf{S}\right) = \lambda\left(\mathrm{tr}\mathbf{S}\right)\mathbf{I} + 2\mu\mathbf{S} \ . \tag{6.179}$$

Recall from (3.33) that a fourth-order tensor was a linear mapping from $\mathscr{T}_{\mathrm{so}}$ into itself. The linear transformation (6.179) can then be rewritten as

$$\mathbf{T}\left(\mathbf{S}\right) = \mathbb{C} : \mathbf{S} \ . \tag{6.180}$$

This implies the minor symmetries $\mathbb{C}_{ijkl} = \mathbb{C}_{jikl} = \mathbb{C}_{ijlk}$ owing to $\mathbf{T} = \mathbf{T}^{\mathrm{T}}$ and $\mathbf{S} = \mathbf{S}^{\mathrm{T}}$. Thus, $\mathbb{C}$ is a linear transformation from $\mathscr{T}_{\mathrm{so}}^{\mathrm{sym}}$ into itself. Now, the isotropy condition $\mathbf{T}\left(\mathbf{QSQ}^{\mathrm{T}}\right) = \mathbf{QT}(\mathbf{S})\,\mathbf{Q}^{\mathrm{T}}$ helps write

$$T_{ij}\left(\mathbf{QSQ}^{\mathrm{T}}\right) = \mathbb{C}_{ijkl}\left(\mathbf{QSQ}^{\mathrm{T}}\right)_{kl} = \mathbb{C}_{ijkl} Q_{km} S_{mn} Q_{ln}$$
$$= Q_{ik} T_{kl}(\mathbf{S})\, Q_{jl} = Q_{ik} \mathbb{C}_{klmn} S_{mn} Q_{jl} \ ,$$

which implies

$$\underbrace{Q_{km} Q_{ln} \mathbb{C}_{ijkl} = Q_{ik} Q_{jl} \mathbb{C}_{klmn}}_{\text{or } Q_{km} Q_{rm} Q_{ln} Q_{sn} \mathbb{C}_{ijkl} = Q_{ik} Q_{jl} Q_{rm} Q_{sn} \mathbb{C}_{klmn} \ \text{ or } \ \delta_{kr} \delta_{ls} \mathbb{C}_{ijkl} = Q_{ik} Q_{jl} Q_{rm} Q_{sn} \mathbb{C}_{klmn}} \ ,$$

or

$$\boxed{\mathbb{C}_{ijkl} = Q_{im} Q_{jn} Q_{ko} Q_{lp} \mathbb{C}_{mnop} \ .} \tag{6.181}$$

Guided by (3.72), one can thus infer that $\mathbb{C}$ is an **isotropic** fourth-order tensor. Finally, by comparing (6.179) and (6.180), one can arrive at the following representation formula

$$\boxed{\underbrace{\mathbb{C} = \lambda \mathbf{I} \otimes \mathbf{I} + 2\mu \mathbf{I} \odot \mathbf{I}}_{\text{or } \mathbb{C}_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu \left(\delta_{ik} \delta_{lj} + \delta_{il} \delta_{kj}\right)} \ ,} \quad \longleftarrow \text{see (3.160)} \tag{6.182}$$

for an isotropic fourth-order tensor $\mathbb{C}$ which possesses the minor symmetries $\mathbb{C}_{ijkl} = \mathbb{C}_{jikl} = \mathbb{C}_{ijlk}$. Notice that in this representation, $\mathbb{C}$ also possesses the major symmetries $\mathbb{C}_{ijkl} = \mathbb{C}_{klij}$.

The representation theorems established so far for isotropic symmetric tensor functions of one tensor variable can be extended to involve more tensorial arguments. The results have been demonstrated in Table 6.3.

As an example, consider an isotropic symmetric tensor-valued function of the two tensorial variables $\mathbf{v}_1 = \mathbf{v}$ and $\mathbf{S}_1 = \mathbf{S}$ which admits the following representation

$$\mathbf{T}(\mathbf{v}, \mathbf{S}) \xlongequal[\text{and sixth rows of Table 6.3}]{\text{from the first, second, third}} f_0(\Upsilon_{\mathrm{s}})\,\mathbf{I} + f_1(\Upsilon_{\mathrm{s}})\,\mathbf{S} + f_2(\Upsilon_{\mathrm{s}})\,\mathbf{S}^2 + f_3(\Upsilon_{\mathrm{s}})\,\mathbf{v} \otimes \mathbf{v}$$
$$+ f_4(\Upsilon_{\mathrm{s}})\,(\mathbf{v} \otimes \mathbf{Sv} + \mathbf{Sv} \otimes \mathbf{v}) + f_5(\Upsilon_{\mathrm{s}})\left(\mathbf{v} \otimes \mathbf{S}^2\mathbf{v} + \mathbf{S}^2\mathbf{v} \otimes \mathbf{v}\right) \ , \tag{6.183}$$

where $f_k,\ k = 0, 1, 2, 3, 4, 5,$ are arbitrary isotropic scalar invariants of the set (6.162).

Hint: The generators $\mathbf{v} \otimes \mathbf{S}^2\mathbf{v} + \mathbf{S}^2\mathbf{v} \otimes \mathbf{v}$ and $\mathbf{Sv} \otimes \mathbf{Sv}$ are equivalent in the sense that they can uniquely determine each other, see (6.236).

Table 6.3 List of isotropic **symmetric tensor** invariants

Variables	Generator elements
0	$\mathbf{I}$
$\mathbf{v}_1$	$\mathbf{v}_1 \otimes \mathbf{v}_1$
$\mathbf{S}_1$	$\mathbf{S}_1, \ \mathbf{S}_1^2$
$\mathbf{W}_1$	$\mathbf{W}_1^2$
$\mathbf{v}_1, \ \mathbf{v}_2$	$\mathbf{v}_1 \otimes \mathbf{v}_2 + \mathbf{v}_2 \otimes \mathbf{v}_1$
$\mathbf{v}_1, \ \mathbf{S}_1$	$\mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{v}_1 + \mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{v}_1, \ \mathbf{v}_1 \otimes \mathbf{S}_1^2\mathbf{v}_1 + \mathbf{S}_1^2\mathbf{v}_1 \otimes \mathbf{v}_1$
$\mathbf{v}_1, \ \mathbf{W}_1$	$\mathbf{v}_1 \otimes \mathbf{W}_1\mathbf{v}_1 + \mathbf{W}_1\mathbf{v}_1 \otimes \mathbf{v}_1, \ \mathbf{W}_1\mathbf{v}_1 \otimes \mathbf{W}_1\mathbf{v}_1, \ \mathbf{W}_1\mathbf{v}_1 \otimes \mathbf{W}_1^2\mathbf{v}_1 + \mathbf{W}_1^2\mathbf{v}_1 \otimes \mathbf{W}_1\mathbf{v}_1$
$\mathbf{S}_1, \ \mathbf{S}_2$	$\mathbf{S}_1\mathbf{S}_2 + \mathbf{S}_2\mathbf{S}_1, \ \mathbf{S}_1^2\mathbf{S}_2 + \mathbf{S}_2\mathbf{S}_1^2, \ \mathbf{S}_2^2\mathbf{S}_1 + \mathbf{S}_1\mathbf{S}_2^2$
$\mathbf{W}_1, \ \mathbf{W}_2$	$\mathbf{W}_1\mathbf{W}_2 + \mathbf{W}_2\mathbf{W}_1, \ \mathbf{W}_1\mathbf{W}_2^2 - \mathbf{W}_2\mathbf{W}_1^2, \ \mathbf{W}_1^2\mathbf{W}_2 - \mathbf{W}_2\mathbf{W}_1^2$
$\mathbf{S}_1, \ \mathbf{W}_1$	$\mathbf{S}_1\mathbf{W}_1 - \mathbf{W}_1\mathbf{S}_1, \ \mathbf{W}_1\mathbf{S}_1\mathbf{W}_1, \ \mathbf{S}_1^2\mathbf{W}_1 - \mathbf{W}_1\mathbf{S}_1^2, \ \mathbf{W}_1\mathbf{S}_1\mathbf{W}_1^2 - \mathbf{W}_1^2\mathbf{S}_1\mathbf{W}_1$
$\mathbf{v}_1, \ \mathbf{v}_2, \ \mathbf{S}_1$	$\mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{v}_2 + \mathbf{S}_1\mathbf{v}_2 \otimes \mathbf{v}_1, \ \mathbf{v}_2 \otimes \mathbf{S}_1\mathbf{v}_1 + \mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{v}_2$
$\mathbf{v}_1, \ \mathbf{v}_2, \ \mathbf{W}_1$	$\mathbf{v}_1 \otimes \mathbf{W}_1\mathbf{v}_2 + \mathbf{W}_1\mathbf{v}_2 \otimes \mathbf{v}_1, \ \mathbf{v}_2 \otimes \mathbf{W}_1\mathbf{v}_1 + \mathbf{W}_1\mathbf{v}_1 \otimes \mathbf{v}_2$

The goal is now to derive a representation formula for an isotropic third-order tensor function of a vector variable, $\mathsf{A}(\mathbf{v})$, which is symmetric it its first two indices, i.e. $\mathsf{A}_{ijk} = \mathsf{A}_{jik}$.

To begin with, let $\mathbf{T}$ be an **isotropic** symmetric tensor function of the two vectors $\mathbf{u}$ and $\mathbf{v}$. Then, by reading off from Tables 6.1 and 6.3,

$$\begin{aligned}\mathbf{T}(\mathbf{u}, \mathbf{v}) = {} & \bar{h}_0(\mathbf{u}\cdot\mathbf{u}, \mathbf{v}\cdot\mathbf{v}, \mathbf{u}\cdot\mathbf{v})\,\mathbf{I} + \bar{h}_1(\mathbf{u}\cdot\mathbf{u}, \mathbf{v}\cdot\mathbf{v}, \mathbf{u}\cdot\mathbf{v})\,\mathbf{u}\otimes\mathbf{u} \\ & + \bar{h}_2(\mathbf{u}\cdot\mathbf{u}, \mathbf{v}\cdot\mathbf{v}, \mathbf{u}\cdot\mathbf{v})\,\mathbf{v}\otimes\mathbf{v} \\ & + \bar{h}_3(\mathbf{u}\cdot\mathbf{u}, \mathbf{v}\cdot\mathbf{v}, \mathbf{u}\cdot\mathbf{v})\,[\mathbf{u}\otimes\mathbf{v} + \mathbf{v}\otimes\mathbf{u}] \ ,\end{aligned} \tag{6.184}$$

where $\bar{h}_k, \ k = 0, 1, 2, 3,$ are arbitrary isotropic scalar-valued functions. Further, suppose that $\mathbf{T}$ is linear in $\mathbf{u}$. Then,

$$\begin{aligned}\mathbf{T}(\mathbf{u}, \mathbf{v}) = {} & (\mathbf{u}\cdot\mathbf{v})\, f(\mathbf{v}\cdot\mathbf{v})\,\mathbf{I} + (\mathbf{u}\cdot\mathbf{v})\, g(\mathbf{v}\cdot\mathbf{v})\,\mathbf{v}\otimes\mathbf{v} \\ & + h(\mathbf{v}\cdot\mathbf{v})\,[\mathbf{u}\otimes\mathbf{v} + \mathbf{v}\otimes\mathbf{u}] \ ,\end{aligned} \tag{6.185}$$

where f, g and h are arbitrary isotropic scalar-valued functions. Now, recall from (3.1) that a third-order tensor was introduced as a linear transformation from $\mathscr{E}_r^{o3}$ into $\mathscr{T}_{so}$. This allows one to rewrite the linear mapping (6.185) as

$$\mathbf{T}(\mathbf{u}, \mathbf{v}) = \mathsf{A}(\mathbf{v})\,\mathbf{u} \ , \tag{6.186}$$

noting that the symmetry of $\mathbf{T}$ naturally implies that $\mathsf{A}_{ijk} = \mathsf{A}_{jik}$. The isotropy condition $\mathbf{T}(\mathbf{Q}\mathbf{u}, \mathbf{Q}\mathbf{v}) = \mathbf{Q}\mathbf{T}(\mathbf{u}, \mathbf{v})\,\mathbf{Q}^{\mathrm{T}}$ then results in

$$T_{ij}(\mathbf{Qu}, \mathbf{Qv}) = \mathsf{A}_{ijk}(\mathbf{Qv})(\mathbf{Qu})_k = \mathsf{A}_{ijk}(\mathbf{Qv}) Q_{km} u_m$$
$$= Q_{ik} T_{kl}(\mathbf{u}, \mathbf{v}) Q_{jl} = Q_{ik} \mathsf{A}_{klm}(\mathbf{v}) u_m Q_{jl} \ ,$$

which implies that

$$\underbrace{Q_{km} \mathsf{A}_{ijk}(\mathbf{Qv}) = Q_{ik} Q_{jl} \mathsf{A}_{klm}(\mathbf{v})}_{\text{or } Q_{km} Q_{nm} \mathsf{A}_{ijk}(\mathbf{Qv}) = Q_{ik} Q_{jl} Q_{nm} \mathsf{A}_{klm}(\mathbf{v}) \text{ or } \delta_{kn} \mathsf{A}_{ijk}(\mathbf{Qv}) = Q_{ik} Q_{jl} Q_{nm} \mathsf{A}_{klm}(\mathbf{v})} \ ,$$

or

$$\boxed{\mathsf{A}_{ijk}(\mathbf{Qv}) = Q_{il} Q_{jm} Q_{kn} \mathsf{A}_{lmn}(\mathbf{v}) \ .} \tag{6.187}$$

Consequently, using (3.24), one can deduce that A is an **isotropic** third-order tensor. Comparing (6.185) and (6.186) helps finally express an isotropic third-order tensor function of a vector, $\mathsf{A}(\mathbf{v})$, possessing the symmetry $\mathsf{A}_{ijk} = \mathsf{A}_{jik}$ in the following form

$$\mathsf{A}_{ijk}(\mathbf{v}) = f(\mathbf{v} \cdot \mathbf{v}) \left[(\mathbf{I} \otimes \mathbf{v})_{ijk} \right] + g(\mathbf{v} \cdot \mathbf{v}) \left[(\mathbf{v} \otimes \mathbf{v} \otimes \mathbf{v})_{ijk} \right]$$
$$+ h(\mathbf{v} \cdot \mathbf{v}) \left[(\mathbf{v} \otimes \mathbf{I})_{ijk} + (\mathbf{v} \otimes \mathbf{I})_{jik} \right] \ . \tag{6.188}$$

At the end, for completeness, the isotropic generating sets of a skew-symmetric tensor function are tabulated in Table 6.4. As can be seen, the zero tensor is the only generator of an antisymmetric tensor function of a vector (or symmetric tensor).

6.3 Exercises

Exercise 6.1

Let $\mathbf{Q} = \mathbf{Q}(t)$ be a given orthogonal tensor-valued function of a scalar variable such as time t. Verify that $\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}$ is a skew-symmetric tensor satisfying the property

$$\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\left(\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}\right)^{\mathrm{T}} \ . \tag{6.189}$$

Solution. Recall that any orthogonal tensor satisfies $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$. The first derivative of this relation then gives

$$\xrightarrow[\text{(2.130)}]{\text{from}} \quad \dot{\overline{\mathbf{Q}\mathbf{Q}^{\mathrm{T}}}} = \dot{\mathbf{I}}$$
$$\xrightarrow[\text{(6.4h) and (6.6a)}]{\text{from}} \quad \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}\dot{\overline{\mathbf{Q}^{\mathrm{T}}}} = \mathbf{O}$$
$$\xrightarrow[\text{(2.11c) and (2.11d)}]{\text{from}} \quad \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\dot{\overline{\mathbf{Q}^{\mathrm{T}}}}$$

$$\xrightarrow[(6.5)]{\text{from}} \quad \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\dot{\mathbf{Q}}^{\mathrm{T}}$$

$$\xrightarrow[(2.55b) \text{ and } (2.55d)]{\text{from}} \quad \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\left(\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}\right)^{\mathrm{T}} .$$

Exercise 6.2

Let $\mathbf{A}(t)$ be an invertible tensor. Then, show that

$$\overline{\det \dot{\mathbf{A}}(t)} = \det \mathbf{A}(t) \left[\mathbf{A}^{-\mathrm{T}}(t) : \dot{\mathbf{A}}(t)\right] = \det \mathbf{A}(t) \operatorname{tr}\left[\mathbf{A}^{-1}(t)\, \dot{\mathbf{A}}(t)\right] .$$

Finally, verify this relation for the following matrix of the given tensor

$$[\mathbf{A}(t)] = \begin{bmatrix} 1/2 & 0 & 0 \\ t^{-\alpha} & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} ,$$

where α denotes a scalar.

Solution. The first derivative of the determinant of an invertible tensor $\mathbf{A}(t)$ is

Table 6.4 List of isotropic **skew-symmetric tensor** invariants

Variables	Generator elements
$\mathbf{v}_1$	$\mathbf{O}$
$\mathbf{S}_1$	$\mathbf{O}$
$\mathbf{W}_1$	$\mathbf{W}_1$
$\mathbf{v}_1$, $\mathbf{v}_2$	$\mathbf{v}_1 \otimes \mathbf{v}_2 - \mathbf{v}_2 \otimes \mathbf{v}_1$
$\mathbf{v}_1$, $\mathbf{S}_1$	$\mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{v}_1 - \mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{v}_1$, $\mathbf{v}_1 \otimes \mathbf{S}_1^2\mathbf{v}_1 - \mathbf{S}_1^2\mathbf{v}_1 \otimes \mathbf{v}_1$, $\mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{S}_1^2\mathbf{v}_1 - \mathbf{S}_1^2\mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{v}_1$
$\mathbf{v}_1$, $\mathbf{W}_1$	$\mathbf{v}_1 \otimes \mathbf{W}_1\mathbf{v}_1 - \mathbf{W}_1\mathbf{v}_1 \otimes \mathbf{v}_1$, $\mathbf{v}_1 \otimes \mathbf{W}_1^2\mathbf{v}_1 - \mathbf{W}_1^2\mathbf{v}_1 \otimes \mathbf{v}_1$
$\mathbf{S}_1$, $\mathbf{S}_2$	$\mathbf{S}_1\mathbf{S}_2 - \mathbf{S}_2\mathbf{S}_1$, $\mathbf{S}_1\mathbf{S}_2^2 - \mathbf{S}_2^2\mathbf{S}_1$, $\mathbf{S}_1^2\mathbf{S}_2 - \mathbf{S}_2\mathbf{S}_1^2$, $\mathbf{S}_1\mathbf{S}_2\mathbf{S}_1^2 - \mathbf{S}_1^2\mathbf{S}_2\mathbf{S}_1$, $\mathbf{S}_2\mathbf{S}_1\mathbf{S}_2^2 - \mathbf{S}_2^2\mathbf{S}_1\mathbf{S}_2$
$\mathbf{W}_1$, $\mathbf{W}_2$	$\mathbf{W}_1\mathbf{W}_2 - \mathbf{W}_2\mathbf{W}_1$
$\mathbf{S}_1$, $\mathbf{W}_1$	$\mathbf{S}_1\mathbf{W}_1 + \mathbf{W}_1\mathbf{S}_1$, $\mathbf{S}_1\mathbf{W}_1^2 - \mathbf{W}_1^2\mathbf{S}_1$
$\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{S}_1$	$\mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{v}_2 - \mathbf{S}_1\mathbf{v}_2 \otimes \mathbf{v}_1$, $\mathbf{v}_2 \otimes \mathbf{S}_1\mathbf{v}_1 - \mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{v}_2$
$\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{W}_1$	$\mathbf{v}_1 \otimes \mathbf{W}_1\mathbf{v}_2 - \mathbf{W}_1\mathbf{v}_2 \otimes \mathbf{v}_1$, $\mathbf{v}_2 \otimes \mathbf{W}_1\mathbf{v}_1 - \mathbf{W}_1\mathbf{v}_1 \otimes \mathbf{v}_2$
$\mathbf{v}_1$, $\mathbf{S}_1$, $\mathbf{S}_2$	$\mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{S}_2\mathbf{v}_1 - \mathbf{S}_2\mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{v}_1$, $\mathbf{S}_1\mathbf{S}_2\mathbf{v}_1 \otimes \mathbf{v}_1 - \mathbf{v}_1 \otimes \mathbf{S}_1\mathbf{S}_2\mathbf{v}_1$, $\mathbf{S}_2\mathbf{S}_1\mathbf{v}_1 \otimes \mathbf{v}_1 - \mathbf{v}_1 \otimes \mathbf{S}_2\mathbf{S}_1\mathbf{v}_1$
$\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$	$\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3 - \mathbf{S}_3\mathbf{S}_2\mathbf{S}_1$, $\mathbf{S}_2\mathbf{S}_3\mathbf{S}_1 - \mathbf{S}_1\mathbf{S}_3\mathbf{S}_2$, $\mathbf{S}_3\mathbf{S}_1\mathbf{S}_2 - \mathbf{S}_2\mathbf{S}_1\mathbf{S}_3$

$$\overline{\det \mathbf{A}(t)}^{\cdot} \xlongequal[\text{(6.8c)}]{\text{from}} \frac{d\,[\det \mathbf{A}]}{d\mathbf{A}} : \dot{\mathbf{A}}$$
$$\xlongequal[\text{(6.20c)}]{\text{from}} (\det \mathbf{A})\, \mathbf{A}^{-\mathrm{T}} : \dot{\mathbf{A}}$$
$$\xlongequal[\text{(2.55b) and (2.84)}]{\text{from}} (\det \mathbf{A})\, \mathbf{I} : \mathbf{A}^{-1}\dot{\mathbf{A}}$$
$$\xlongequal[\text{(2.83)}]{\text{from}} (\det \mathbf{A})\, \mathrm{tr}\left[\mathbf{A}^{-1}(t)\, \dot{\mathbf{A}}(t)\right] .$$

The goal is now to check $\overline{\det \mathbf{A}}^{\cdot} = \det \mathbf{A} \mathrm{tr}\left(\mathbf{A}^{-1}\dot{\mathbf{A}}\right)$ for the given matrix. By use of (1.47), (1.52), (1.79), (2.114) and (2.121), the determinant and inverse of this matrix render

$$\det[\mathbf{A}] = 1 \quad , \quad \left[\mathbf{A}^{-1}\right] = \begin{bmatrix} 2 & 0 & 0 \\ -2t^{-\alpha} & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} .$$

And its first derivative, using (6.2b), is

$$\left[\dot{\mathbf{A}}\right] = \begin{bmatrix} 0 & 0 & 0 \\ -\alpha t^{-\alpha-1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} .$$

Obviously, $\overline{\det[\mathbf{A}]}^{\cdot} = \dot{1} = 0$ and $\mathrm{tr}\left([\mathbf{A}]^{-1}\left[\dot{\mathbf{A}}\right]\right) = 0$.

Exercise 6.3

Let Φ be a sufficiently smooth scalar-valued function and $\mathbf{u}$, $\mathbf{v}$ be sufficiently smooth vector-valued functions of one vector variable $\mathbf{w}$. Further, let Ψ be a sufficiently smooth scalar-valued function and $\mathbf{A}$, $\mathbf{B}$ be sufficiently smooth tensor-valued functions of one tensor variable $\mathbf{C}$. Then, verify the important identities

$$\frac{\partial}{\partial \mathbf{w}}[\Phi \mathbf{u} \cdot \mathbf{v}] = (\mathbf{u} \cdot \mathbf{v}) \frac{\partial \Phi}{\partial \mathbf{w}} + \Phi \left[\left(\frac{\partial \mathbf{u}}{\partial \mathbf{w}} \right)^{\mathrm{T}} \mathbf{v} + \left(\frac{\partial \mathbf{v}}{\partial \mathbf{w}} \right)^{\mathrm{T}} \mathbf{u} \right] , \tag{6.190a}$$

$$\frac{\partial}{\partial \mathbf{w}}[\Phi \mathbf{u}] = \mathbf{u} \otimes \frac{\partial \Phi}{\partial \mathbf{w}} + \Phi \frac{\partial \mathbf{u}}{\partial \mathbf{w}} , \tag{6.190b}$$

$$\frac{\partial}{\partial \mathbf{C}}[\Psi \mathbf{A} : \mathbf{B}] = (\mathbf{A} : \mathbf{B}) \frac{\partial \Psi}{\partial \mathbf{C}} + \Psi \left[\mathbf{B} : \frac{\partial \mathbf{A}}{\partial \mathbf{C}} + \mathbf{A} : \frac{\partial \mathbf{B}}{\partial \mathbf{C}} \right] , \tag{6.190c}$$

$$\frac{\partial}{\partial \mathbf{C}}[\Psi \mathbf{A}] = \mathbf{A} \otimes \frac{\partial \Psi}{\partial \mathbf{C}} + \Psi \frac{\partial \mathbf{A}}{\partial \mathbf{C}} . \tag{6.190d}$$

Solution. Of special interest here is to work with the components of tensors. The procedure to be followed relies on indicial notation as a powerful tool for proving identities in tensor analysis.

By use of the product rule of differentiation along with $(1.38)_7$, $(2.24)_5$, (2.49), $(6.14a)_1$ and $(6.29)_1$, one can write

$$\frac{\partial}{\partial w_i}\left[\Phi u_j v_j\right] = \underbrace{u_j v_j \frac{\partial \Phi}{\partial w_i}}_{\text{or } (\mathbf{u}\cdot\mathbf{v})\,(\partial\Phi/\partial\mathbf{w})_i} + \underbrace{\Phi \frac{\partial u_j}{\partial w_i} v_j}_{\text{or } \Phi\,(\partial\mathbf{u}/\partial\mathbf{w})_{ji}\,(\mathbf{v})_j = \Phi\,(\partial\mathbf{u}/\partial\mathbf{w})^{\mathrm{T}}_{ij}\,(\mathbf{v})_j} + \underbrace{\Phi u_j \frac{\partial v_j}{\partial w_i}}_{\text{or } \Phi\,(\mathbf{u})_j\,(\partial\mathbf{v}/\partial\mathbf{w})_{ji} = \Phi\,(\partial\mathbf{v}/\partial\mathbf{w})^{\mathrm{T}}_{ij}\,(\mathbf{u})_j} ,$$

$$\frac{\partial}{\partial w_j}\left[\Phi u_i\right] = \underbrace{u_i \frac{\partial \Phi}{\partial w_j}}_{\text{or } (\mathbf{u})_i\,(\partial\Phi/\partial\mathbf{w})_j = (\mathbf{u}\otimes\partial\Phi/\partial\mathbf{w})_{ij}} + \underbrace{\Phi \frac{\partial u_i}{\partial w_j}}_{\text{or } \Phi\,(\partial\mathbf{u}/\partial\mathbf{w})_{ij}} .$$

In a similar manner, the product rule in conjunction with $(2.75)_4$, $(3.66b)_4$, (3.70a), $(6.14d)_1$ and $(6.40)_1$ help establish

$$\frac{\partial}{\partial C_{ij}}\left[\Psi A_{kl} B_{kl}\right] = \underbrace{A_{kl} B_{kl} \frac{\partial \Psi}{\partial C_{ij}}}_{\text{or } (\mathbf{A}:\mathbf{B})\,(\partial\Psi/\partial\mathbf{C})_{ij}} + \Psi \underbrace{B_{kl} \frac{\partial A_{kl}}{\partial C_{ij}}}_{\text{or } (\mathbf{B})_{kl}\,(\partial\mathbf{A}/\partial\mathbf{C})_{klij} = (\mathbf{B}:\partial\mathbf{A}/\partial\mathbf{C})_{ij}} + \Psi \underbrace{A_{kl} \frac{\partial B_{kl}}{\partial C_{ij}}}_{\text{or } (\mathbf{A})_{kl}\,(\partial\mathbf{B}/\partial\mathbf{C})_{klij} = (\mathbf{A}:\partial\mathbf{B}/\partial\mathbf{C})_{ij}} ,$$

$$\frac{\partial}{\partial C_{kl}}\left[\Psi A_{ij}\right] = \underbrace{A_{ij} \frac{\partial \Psi}{\partial C_{kl}}}_{\text{or } (\mathbf{A})_{ij}\,(\partial\Psi/\partial\mathbf{C})_{kl} = (\mathbf{A}\otimes\partial\Psi/\partial\mathbf{C})_{ijkl}} + \underbrace{\Psi \frac{\partial A_{ij}}{\partial C_{kl}}}_{\text{or } \Psi\,(\partial\mathbf{A}/\partial\mathbf{C})_{ijkl}} .$$

Exercise 6.4

Verify the following identities

$$\frac{\partial}{\partial \mathbf{u}}\left[\frac{\mathbf{u}}{|\mathbf{u}|}\right] = \frac{1}{|\mathbf{u}|}\left[\mathbf{I} - \frac{\mathbf{u}}{|\mathbf{u}|} \otimes \frac{\mathbf{u}}{|\mathbf{u}|}\right] , \tag{6.191a}$$

$$\frac{\partial}{\partial \mathbf{A}}\left[\frac{\mathbf{A}}{|\mathbf{A}|}\right] = \frac{1}{|\mathbf{A}|}\left[\mathbf{I} \boxtimes \mathbf{I} - \frac{\mathbf{A}}{|\mathbf{A}|} \otimes \frac{\mathbf{A}}{|\mathbf{A}|}\right] , \tag{6.191b}$$

$$\frac{\partial}{\partial \mathbf{u}}\left[\frac{\mathbf{A}\mathbf{u}}{|\mathbf{A}\mathbf{u}|}\right] = \frac{1}{|\mathbf{A}\mathbf{u}|}\left[\mathbf{A} - \frac{\mathbf{A}\mathbf{u}}{|\mathbf{A}\mathbf{u}|} \otimes \frac{\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{u}}{|\mathbf{A}\mathbf{u}|}\right] , \tag{6.191c}$$

$$\frac{\partial}{\partial \mathbf{A}}\left[\frac{\mathbf{Au}}{|\mathbf{Au}|}\right] = \frac{1}{|\mathbf{Au}|}\left[\mathbf{I}\otimes\mathbf{u} - \frac{\mathbf{Au}}{|\mathbf{Au}|}\otimes\frac{\mathbf{Au}}{|\mathbf{Au}|}\otimes\mathbf{u}\right] . \tag{6.191d}$$

Solution. Relying on the previously established identities, the desired relations will be proved in direct notation here.
The expression (6.191a):

$$\begin{aligned}\frac{\partial}{\partial \mathbf{u}}\left[|\mathbf{u}|^{-1}\,\mathbf{u}\right] &\xlongequal[\text{(6.190b)}]{\text{from}} \mathbf{u}\otimes\frac{\partial\,|\mathbf{u}|^{-1}}{\partial \mathbf{u}} + |\mathbf{u}|^{-1}\,\frac{\partial \mathbf{u}}{\partial \mathbf{u}} \\ &\xlongequal[\text{(6.18a) and (6.32a)}]{\text{from}} \mathbf{u}\otimes\left[-\,|\mathbf{u}|^{-2}\,\frac{\mathbf{u}}{|\mathbf{u}|}\right] + |\mathbf{u}|^{-1}\,\mathbf{I} \\ &\xlongequal[\text{(2.16a) and (2.16b)}]{\text{from}} \frac{1}{|\mathbf{u}|}\left[\mathbf{I} - \frac{\mathbf{u}}{|\mathbf{u}|}\otimes\frac{\mathbf{u}}{|\mathbf{u}|}\right] .\end{aligned}$$

The expression (6.191b):

$$\begin{aligned}\frac{\partial}{\partial \mathbf{A}}\left[|\mathbf{A}|^{-1}\,\mathbf{A}\right] &\xlongequal[\text{(6.190d)}]{\text{from}} \mathbf{A}\otimes\frac{\partial\,|\mathbf{A}|^{-1}}{\partial \mathbf{A}} + |\mathbf{A}|^{-1}\,\frac{\partial \mathbf{A}}{\partial \mathbf{A}} \\ &\xlongequal[\text{(6.18a) and (6.42a)}]{\text{from}} \mathbf{A}\otimes\left[-\,|\mathbf{A}|^{-2}\,\frac{\mathbf{A}}{|\mathbf{A}|}\right] + |\mathbf{A}|^{-1}\,\mathbf{I}\boxtimes\mathbf{I} \\ &\xlongequal[\text{(3.53a) and (3.53b)}]{\text{from}} \frac{1}{|\mathbf{A}|}\left[\mathbf{I}\boxtimes\mathbf{I} - \frac{\mathbf{A}}{|\mathbf{A}|}\otimes\frac{\mathbf{A}}{|\mathbf{A}|}\right] .\end{aligned}$$

The expression (6.191c):

$$\begin{aligned}\frac{\partial}{\partial \mathbf{u}}\left[\frac{\mathbf{Au}}{|\mathbf{Au}|}\right] &\xlongequal[\text{(1.11) and (6.190b)}]{\text{from}} \mathbf{Au}\otimes\frac{\partial\,(\mathbf{Au}\cdot\mathbf{Au})^{-1/2}}{\partial \mathbf{u}} + |\mathbf{Au}|^{-1}\,\frac{\partial \mathbf{Au}}{\partial \mathbf{u}} \\ &\xlongequal[\text{(1.11) and (6.32b)}]{\text{from}} \mathbf{Au}\otimes\left[-\frac{1}{2\,|\mathbf{Au}|\,|\mathbf{Au}|\,|\mathbf{Au}|}\,\frac{\partial\,(\mathbf{Au}\cdot\mathbf{Au})}{\partial \mathbf{u}}\right] + \frac{\mathbf{A}}{|\mathbf{Au}|} \\ &\xlongequal[\text{(2.16a)-(2.16b) and (6.18c)}]{\text{from}} \frac{1}{|\mathbf{Au}|}\left[\mathbf{A} - \frac{\mathbf{Au}}{|\mathbf{Au}|}\otimes\frac{\mathbf{A}^{\mathrm{T}}\mathbf{Au}}{|\mathbf{Au}|}\right] .\end{aligned}$$

The expression (6.191d):

$$\begin{aligned}\frac{\partial}{\partial \mathbf{A}}\left[\frac{\mathbf{Au}}{|\mathbf{Au}|}\right] &\xlongequal[\text{in light of (6.190d)}]{\text{from (1.11) and}} \mathbf{Au}\otimes\frac{\partial\,(\mathbf{Au}\cdot\mathbf{Au})^{-1/2}}{\partial \mathbf{A}} + |\mathbf{Au}|^{-1}\,\frac{\partial \mathbf{Au}}{\partial \mathbf{A}} \\ &\xlongequal[\text{(1.11) and (6.32b)}]{\text{from}} \mathbf{Au}\otimes\left[-\frac{1}{2\,|\mathbf{Au}|\,|\mathbf{Au}|\,|\mathbf{Au}|}\,\frac{\partial\,(\mathbf{Au}\cdot\mathbf{Au})}{\partial \mathbf{A}}\right] + \frac{\mathbf{I}\otimes\mathbf{u}}{|\mathbf{Au}|} \\ &\xlongequal[\text{(3.7a)-(3.7c) and (6.18c)}]{\text{from}} \frac{1}{|\mathbf{Au}|}\left[\mathbf{I} - \frac{\mathbf{Au}}{|\mathbf{Au}|}\otimes\frac{\mathbf{Au}}{|\mathbf{Au}|}\right]\otimes\mathbf{u} .\end{aligned}$$

Exercise 6.5

First, show that the derivatives of the powers of a tensor, $\mathbf{A}^n \ (n = 1, 2, \ldots)$, with respect to itself can be computed via the following relations

$$\frac{\partial \mathbf{A}^n}{\partial \mathbf{A}} = \sum_{k=0}^{n-1} \mathbf{A}^k \boxtimes \mathbf{A}^{n-k-1} = \sum_{m=1}^{n} \mathbf{A}^{m-1} \boxtimes \mathbf{A}^{n-m} \ . \quad \longleftarrow \text{see (3.49b) and (3.70b)} \tag{6.192}$$

Then, verify that

$$\frac{\partial \mathbf{A}^{-n}}{\partial \mathbf{A}} = -\sum_{k=0}^{n-1} \mathbf{A}^{k-n} \boxtimes \mathbf{A}^{-k-1} = -\sum_{m=1}^{n} \mathbf{A}^{m-n-1} \boxtimes \mathbf{A}^{-m} \ . \tag{6.193}$$

Finally, use these results to establish the important relations

$$\frac{\partial}{\partial \mathbf{A}} \left[\mathrm{tr}\mathbf{A}^n \right] = n \left(\mathbf{A}^{n-1} \right)^{\mathrm{T}} \ , \tag{6.194a}$$

$$\frac{\partial}{\partial \mathbf{A}} \left[\mathrm{tr}\mathbf{A}^{-n} \right] = -n \left(\mathbf{A}^{-n-1} \right)^{\mathrm{T}} \ . \tag{6.194b}$$

Solution. Here, different procedures will be used to derive the desired relations. First, by means of (2.36), $(2.37)_1$, $(3.54\text{b})_1$ and (6.36c), one will have

$$\begin{aligned}
\frac{\partial \mathbf{A}^n}{\partial \mathbf{A}} : \mathbf{B} &= \frac{d}{d\varepsilon} \left[(\mathbf{A} + \varepsilon \mathbf{B})^n \right] \Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon} \Bigg[\underbrace{(\mathbf{A} + \varepsilon \mathbf{B}) (\mathbf{A} + \varepsilon \mathbf{B}) \cdots (\mathbf{A} + \varepsilon \mathbf{B})}_{n \text{ times}} \Bigg] \Bigg|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon} \left[\varepsilon^0 \mathbf{A}^n + \varepsilon^1 \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{B} \mathbf{A}^{n-k-1} + \varepsilon^2 \cdots \right] \Bigg|_{\varepsilon=0} \\
&= \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{B} \mathbf{A}^{n-k-1} = \sum_{k=0}^{n-1} \mathbf{A}^k \boxtimes \mathbf{A}^{n-k-1} : \mathbf{B} \ .
\end{aligned}$$

Then, guided by the expressions (6.56) and (6.57), the total differential of the identity $\mathbf{A}^{-n}\mathbf{A}^n = \mathbf{I}$ gives

$$d\mathbf{A}^{-n} = -\mathbf{A}^{-n} \left(d\mathbf{A}^n \right) \mathbf{A}^{-n} = \underbrace{-\left(\mathbf{A}^{-n} \boxtimes \mathbf{A}^{-n} \right)}_{= \partial \mathbf{A}^{-n} / \partial \mathbf{A}^n} : d\mathbf{A}^n \ .$$

Substituting $(6.192)_1$ into the above result, taking into account $(2.37)_1$ and $(3.94\text{b})_4$, will lead to the second required relation:

$$
\begin{aligned}
d\mathbf{A}^{-n} &= \left[\frac{\partial \mathbf{A}^{-n}}{\partial \mathbf{A}}\right] : d\mathbf{A} \\
&= \left[\frac{\partial \mathbf{A}^{-n}}{\partial \mathbf{A}^{n}} : \frac{\partial \mathbf{A}^{n}}{\partial \mathbf{A}}\right] : d\mathbf{A} \\
&= \left[-\left(\mathbf{A}^{-n} \boxtimes \mathbf{A}^{-n}\right) : \left(\sum_{k=0}^{n-1} \mathbf{A}^{k} \boxtimes \mathbf{A}^{n-k-1}\right)\right] : d\mathbf{A} \\
&= \left[-\sum_{k=0}^{n-1} \mathbf{A}^{k-n} \boxtimes \mathbf{A}^{-k-1}\right] : d\mathbf{A} \, .
\end{aligned}
$$

Note that either $(6.192)_1$ or $(6.193)_1$ could be verified by mathematical induction. This remains to be done by the ambitious reader.

The aim is now to calculate the derivatives of the first invariant of the monomials of a tensor, $\mathrm{tr}\mathbf{A}^n$ $(n = 1, 2, \ldots)$, with respect to itself. By means of (2.33), $(2.37)_1$, $(2.55\text{d})_1$, (2.78), (3.36), $(3.54\text{b})_2$ and $(6.192)_1$, one will have

$$
\begin{aligned}
d\left[\mathbf{I} : \mathbf{A}^{n}\right] &= \frac{\partial \left[\mathbf{I} : \mathbf{A}^{n}\right]}{\partial \mathbf{A}} : d\mathbf{A} \\
&= \underbrace{\mathbf{A}^{n} : \frac{\partial \mathbf{I}}{\partial \mathbf{A}} : d\mathbf{A}}_{= \mathbf{A}^{n} : \mathbb{O} : d\mathbf{A} = \mathbf{A}^{n} : \mathbf{O} = 0} + \mathbf{I} : \frac{\partial \mathbf{A}^{n}}{\partial \mathbf{A}} : d\mathbf{A} \\
&= \mathbf{I} : \sum_{k=0}^{n-1} \mathbf{A}^{k} \boxtimes \mathbf{A}^{n-k-1} : d\mathbf{A} \\
&= \sum_{k=0}^{n-1} \left(\mathbf{A}^{k}\right)^{\mathrm{T}} \mathbf{I} \left(\mathbf{A}^{n-k-1}\right)^{\mathrm{T}} : d\mathbf{A} \\
&= \sum_{k=0}^{n-1} \left(\mathbf{A}^{n-k-1} \mathbf{A}^{k}\right)^{\mathrm{T}} : d\mathbf{A} \\
&= \underbrace{n \left(\mathbf{A}^{n-1}\right)^{\mathrm{T}}}_{= \partial \left[\mathbf{I} : \mathbf{A}^{n}\right] / \partial \mathbf{A}} : d\mathbf{A} \, .
\end{aligned}
$$

In a similar fashion,

$$
\begin{aligned}
d\left[\mathbf{I} : \mathbf{A}^{-n}\right] &= \frac{\partial \left[\mathbf{I} : \mathbf{A}^{-n}\right]}{\partial \mathbf{A}} : d\mathbf{A} \\
&= \mathbf{I} : \frac{\partial \mathbf{A}^{-n}}{\partial \mathbf{A}} : d\mathbf{A}
\end{aligned}
$$

$$= -\sum_{k=0}^{n-1} \left(\mathbf{A}^{k-n}\right)^{\mathrm{T}} \mathbf{I} \left(\mathbf{A}^{-k-1}\right)^{\mathrm{T}} : d\mathbf{A}$$
$$= \underbrace{-n \left(\mathbf{A}^{-n-1}\right)^{\mathrm{T}}}_{= \partial \left[\mathbf{I} : \mathbf{A}^{-n}\right] / \partial \mathbf{A}} : d\mathbf{A} \ .$$

Exercise 6.6

From (2.190)$_1$ and (6.20c)$_2$, it simply follows that

$$\frac{\partial I_3}{\partial \mathbf{A}} = \left(\mathbf{A}^2\right)^{\mathrm{T}} - I_1 \mathbf{A}^{\mathrm{T}} + I_2 \mathbf{I} \ . \tag{6.195}$$

Obtain this important result by differentiating the Cayley-Hamilton equation (4.21) with respect to $\mathbf{A}$. This may be viewed as an alternative derivation for the gradient of the determinant of a tensor, see (6.19c).

Moreover, by using (1.80) obtain the similar result to present another alternative derivation. See Liu [23] and Holzapfel [44] for the other derivations.

Solution. To begin with, one needs to take the trace of the Cayley-Hamilton equation. Considering the linearity of the trace operator along with (2.90)–(2.91) reveals

$$3I_3 = \mathrm{tr}\mathbf{A}^3 - I_1 \mathrm{tr}\mathbf{A}^2 + I_2 \mathrm{tr}\mathbf{A} \ .$$

Now, using (2.11g), (6.20a)$_{1-2}$, (6.20b)$_{1-2}$ and (6.194a),

$$3\frac{\partial I_3}{\partial \mathbf{A}} = \underbrace{\frac{\partial \mathrm{tr}\mathbf{A}^3}{\partial \mathbf{A}}}_{= 3\left(\mathbf{A}^2\right)^{\mathrm{T}}} - \underbrace{\frac{\partial I_1}{\partial \mathbf{A}}}_{= \mathbf{I}} \mathrm{tr}\mathbf{A}^2 - I_1 \underbrace{\frac{\partial \mathrm{tr}\mathbf{A}^2}{\partial \mathbf{A}}}_{= 2\mathbf{A}^{\mathrm{T}}} + \underbrace{\frac{\partial I_2}{\partial \mathbf{A}}}_{= I_1 \mathbf{I} - \mathbf{A}^{\mathrm{T}}} \mathrm{tr}\mathbf{A} + I_2 \underbrace{\frac{\partial \mathrm{tr}\mathbf{A}}{\partial \mathbf{A}}}_{= \mathbf{I}}$$
$$= 3\left(\mathbf{A}^2\right)^{\mathrm{T}} - 3I_1 \mathbf{A}^{\mathrm{T}} + \left(I_2 + I_1^2 - \mathrm{tr}\mathbf{A}^2\right) \mathbf{I}$$
$$= 3\left(\mathbf{A}^2\right)^{\mathrm{T}} - 3I_1 \mathbf{A}^{\mathrm{T}} + 3I_2 \mathbf{I} \ .$$

The interested reader may want to arrive at this result by considering the gradient of

$$\det \mathbf{A} = \frac{1}{6} \varepsilon_{uvk} \varepsilon_{lmn} A_{ul} A_{vm} A_{kn} \ ,$$

as follows:

$$\frac{\partial}{\partial A_{ij}} \left[\det \mathbf{A}\right] = \frac{1}{6} \varepsilon_{uvk} \varepsilon_{lmn} \delta_{ui} \delta_{jl} A_{vm} A_{kn} + \frac{1}{6} \varepsilon_{uvk} \varepsilon_{lmn} A_{ul} \delta_{vi} \delta_{jm} A_{kn}$$

$$+\frac{1}{6}\varepsilon_{uvk}\varepsilon_{lmn}A_{ul}A_{vm}\delta_{ki}\delta_{jn}$$

$$= \underbrace{\frac{1}{6}\varepsilon_{ivk}\varepsilon_{jmn}A_{vm}A_{kn}}_{=\frac{1}{6}\varepsilon_{iuk}\varepsilon_{jmn}A_{um}A_{kn}} + \underbrace{\frac{1}{6}\varepsilon_{uik}\varepsilon_{ljn}A_{ul}A_{kn}}_{=\frac{1}{6}\varepsilon_{iuk}\varepsilon_{jln}A_{ul}A_{kn}=\frac{1}{6}\varepsilon_{iuk}\varepsilon_{jmn}A_{um}A_{kn}}$$

$$+ \underbrace{\frac{1}{6}\varepsilon_{uvi}\varepsilon_{lmj}A_{ul}A_{vm}}_{=\frac{1}{6}\varepsilon_{iuv}\varepsilon_{jlm}A_{ul}A_{vm}=\frac{1}{6}\varepsilon_{iuk}\varepsilon_{jmn}A_{um}A_{kn}}$$

$$= \underbrace{\frac{1}{2}\varepsilon_{iuk}\varepsilon_{jmn}A_{um}A_{kn}}_{=\frac{1}{2}\varepsilon_{iul}\varepsilon_{jmn}A_{um}A_{ln}=\frac{1}{2}\varepsilon_{ikl}\varepsilon_{jmn}A_{km}A_{ln}}$$

$$= (\mathbf{A}^{\mathrm{c}})_{ij} = A^2_{ji} - I_1 A_{ji} + I_2\delta_{ij} \ ,$$

where the expressions (1.36), (1.53), $(2.117)_1$, $(2.190)_2$, $(6.14\text{d})_1$ and $(6.42\text{a})_2$ along with the product rule of differentiation have been used.

Exercise 6.7

The goal of this exercise is to show the prominent role of **exponential map** in the solution of linear ordinary tensorial differential equations. To this end, consider the exponential tensor function

$$\mathbf{exp}\,(t\mathbf{A}) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\mathbf{A}^n \ , \qquad \longleftarrow \text{see (2.39)} \tag{6.196}$$

where t is a scalar such as time. If t is constant, then the derivative of $\mathbf{exp}\,(t\mathbf{A})$ takes the following form

$$\frac{d}{d\mathbf{A}}\left[\mathbf{exp}\,(t\mathbf{A})\right] = \sum_{n=1}^{\infty}\frac{t^n}{n!}\frac{d\mathbf{A}^n}{d\mathbf{A}} = \sum_{n=1}^{\infty}\frac{t^n}{n!}\sum_{m=1}^{n}\mathbf{A}^{m-1}\boxtimes\mathbf{A}^{n-m} \ . \tag{6.197}$$

And if $\mathbf{A}$ remains unchanged, then the gradient of $\mathbf{exp}\,(t\mathbf{A})$ becomes

$$\frac{d}{dt}\left[\mathbf{exp}\,(t\mathbf{A})\right] = \sum_{n=1}^{\infty}\frac{n}{n!}t^{n-1}\mathbf{A}^n$$

$$= \sum_{n=1}^{\infty}\frac{t^{n-1}}{(n-1)!}\mathbf{A}^n$$

$$
\begin{aligned}
&= \mathbf{A} \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} \mathbf{A}^{n-1} \\
&= \mathbf{A} \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbf{A}^k \\
&= \underbrace{\mathbf{A} \mathbf{exp}\,(t\mathbf{A}) = \mathbf{exp}\,(t\mathbf{A})\,\mathbf{A}}_{\text{note that } \mathbf{A} \text{ and } \mathbf{exp}\,(t\mathbf{A}) \text{ obviously commute}} \quad .
\end{aligned} \tag{6.198}
$$

In multivariable calculus, the expression (6.198) reminds that the matrix exponential can successfully be used for solving systems of linear ordinary differential equations. Now, consider the following initial-value problem

$$
\dot{\mathbf{Y}}\,(t) = \mathbf{A}\mathbf{Y}\,(t) \quad \text{subject to} \quad \mathbf{Y}\,(0) = \mathbf{Y}_0 \ , \tag{6.199}
$$

whose solution is

$$
\mathbf{Y}\,(t) = \mathbf{exp}\,(t\mathbf{A})\,\mathbf{Y}_0 \ . \tag{6.200}
$$

Show that the solution of

$$
\dot{\mathbf{Y}}\,(t) + \mathbf{A}\mathbf{Y}\,(t) + \mathbf{Y}\,(t)\,\mathbf{B} = \mathbf{C}\,(t) \quad \text{with the initial condition} \quad \mathbf{Y}\,(0) = \mathbf{Y}_0 \ , \tag{6.201}
$$

can be written as

$$
\begin{aligned}
\mathbf{Y}\,(t) = &\int_0^t \mathbf{exp}\left(-\hat{t}\mathbf{A}\right) \mathbf{C}\left(t - \hat{t}\right) \mathbf{exp}\left(-\hat{t}\,\mathbf{B}\right) d\hat{t} \\
&+ \mathbf{exp}\,(-t\mathbf{A})\,\mathbf{Y}_0\,\mathbf{exp}\,(-t\mathbf{B}) \ .
\end{aligned} \tag{6.202}
$$

Solution. By use of the rule $(3.54\text{b})_1$ and the identity $\mathbf{A}\mathbf{exp}\,(t\mathbf{A}) = \mathbf{exp}\,(t\mathbf{A})\,\mathbf{A}$, the key point here is to rewrite (6.201) as

$$
\underbrace{\left[\mathbf{exp}\,(t\mathbf{A}) \boxtimes \mathbf{exp}\,(t\mathbf{B})\right] : \left[\dot{\mathbf{Y}} + \mathbf{A}\mathbf{Y} + \mathbf{Y}\mathbf{B}\right]}_{= \mathbf{exp}\,(t\mathbf{A})\,\dot{\mathbf{Y}}\mathbf{exp}\,(t\mathbf{B}) + \mathbf{A}\mathbf{exp}\,(t\mathbf{A})\,\mathbf{Y}\mathbf{exp}\,(t\mathbf{B}) + \mathbf{exp}\,(t\mathbf{A})\,\mathbf{Y}\mathbf{exp}\,(t\mathbf{B})\,\mathbf{B}} = \underbrace{\left[\mathbf{exp}\,(t\mathbf{A}) \boxtimes \mathbf{exp}\,(t\mathbf{B})\right] : [\mathbf{C}]}_{= \mathbf{exp}\,(t\mathbf{A})\,\mathbf{C}\mathbf{exp}\,(t\mathbf{B})} \ ,
$$

or, in the more convenient form,

$$
\frac{d}{dt}\left[\mathbf{exp}\,(t\mathbf{A})\,\mathbf{Y}\mathbf{exp}\,(t\mathbf{B})\right] = \mathbf{exp}\,(t\mathbf{A})\,\mathbf{C}\mathbf{exp}\,(t\mathbf{B}) \ .
$$

It follows that

$$
\mathbf{exp}\,(t\mathbf{A})\,\mathbf{Y}\,(t)\,\mathbf{exp}\,(t\mathbf{B}) = \int_0^t \mathbf{exp}\left(\hat{t}\mathbf{A}\right) \mathbf{C}\left(\hat{t}\right) \mathbf{exp}\left(\hat{t}\,\mathbf{B}\right) d\hat{t} + \mathbf{Y}_0 \ ,
$$

with the aid of (2.33), (2.105) and (2.210), delivers the desired result. Note that (6.200) can be resulted from (6.202) by setting $\mathbf{C} = \mathbf{O}$, $\mathbf{B} = \mathbf{O}$ and $\mathbf{A} = -\hat{\mathbf{A}}$.

Exercise 6.8

Given the scalar-valued function

$$f(t) = \det\left[\mathbf{exp}(t\mathbf{A})\right] . \tag{6.203}$$

By computing its derivative, deduce that the solution of the resulting differential equation is the identity $\det\left[\mathbf{exp}(\mathbf{A})\right] = \exp(\mathrm{tr}\mathbf{A})$. This may be viewed as an alternative derivation for this important expression verified previously in Exercise 4.6.

Solution. The solution of the linear ordinary differential equation

$$\begin{aligned}
\frac{d\det\left[\mathbf{exp}(t\mathbf{A})\right]}{dt} &\xlongequal[\text{the chain rule}]{\text{using}} \frac{d\det\left[\mathbf{exp}(t\mathbf{A})\right]}{d\mathbf{exp}(t\mathbf{A})} : \frac{d\mathbf{exp}(t\mathbf{A})}{dt} \\
&\xlongequal[\text{(6.20c) and (6.198)}]{\text{from}} \det\left[\mathbf{exp}(t\mathbf{A})\right]\left[\mathbf{exp}(t\mathbf{A})\right]^{-\mathrm{T}} : \mathbf{A}\mathbf{exp}(t\mathbf{A}) \\
&\xlongequal[\text{(2.110) and in view of (2.210)}]{\text{from}} \det\left[\mathbf{exp}(t\mathbf{A})\right]\left[\mathbf{exp}(-t\mathbf{A})\right]^{\mathrm{T}} : \mathbf{A}\mathbf{exp}(t\mathbf{A}) \\
&\xlongequal[\text{(2.55b) and (2.84)}]{\text{from}} \det\left[\mathbf{exp}(t\mathbf{A})\right]\mathbf{I} : \mathbf{A}\mathbf{exp}(t\mathbf{A})\,\mathbf{exp}(-t\mathbf{A}) \\
&\xlongequal[\text{(2.33), (2.82) and in view of (2.207c)}]{\text{from}} (\mathrm{tr}\mathbf{A})\det\left[\mathbf{exp}(t\mathbf{A})\right] ,
\end{aligned}$$

is

$$\underbrace{\ln\left(\det\left[\mathbf{exp}(t\mathbf{A})\right]\right) = \mathrm{tr}\mathbf{A}t}_{\text{note that } \ln\left(\det\left[\mathbf{exp}(0\mathbf{A})\right]\right) = \ln(\det[\mathbf{I}]) = \ln(1) = 0} .$$

Exponentiating this result then yields

$$\det\left[\mathbf{exp}(t\mathbf{A})\right] = \exp(\mathrm{tr}\mathbf{A}t) .$$

Finally, substituting $t = 1$ into this expression delivers the required result.

Exercise 6.9

Suppose that $\mathbf{S}$ is a symmetric second-order tensor with the spectral representations (4.41), $(4.45)_3$ and (4.46). Then, verify the useful relations

$$\frac{\partial \lambda_a}{\partial \mathbf{S}} = \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \ , \ a = 1, 2, 3, \qquad \text{if} \quad \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \tag{6.204a}$$

$$\frac{\partial \lambda_1}{\partial \mathbf{S}} = \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 \ , \ 2\frac{\partial \lambda}{\partial \mathbf{S}} = \mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 \qquad \text{if} \quad \lambda_1 \neq \lambda_2 = \lambda_3 = \lambda \ , \tag{6.204b}$$

$$3\frac{\partial \lambda}{\partial \mathbf{S}} = \mathbf{I} \qquad \text{if} \quad \lambda_1 = \lambda_2 = \lambda_3 = \lambda \ . \tag{6.204c}$$

Solution. To this end, let $\mathbf{S}$ be a symmetric tensor that is known in its contravariant components, i.e. $\mathbf{S} = \underline{S}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j$. And let $\widehat{\mathbf{n}}_a$ be a unit vector that is known in its covariant components, i.e. $\widehat{\mathbf{n}}_a = \underline{\widehat{n}}_{i\ a} \mathbf{g}^i$.

For the first case, suppose that $\mathbf{S}$ is a tensor with exactly three distinct eigenpairs $(\lambda_a, \widehat{\mathbf{n}}_a)\ ,\ a = 1, 2, 3$. To begin with, consider the following eigenvalue problem for such a tensor (Ibrahimbegović [24])

$$\left[\mathbf{S} - \lambda_a \mathbf{g}^{-1}\right] \widehat{\mathbf{n}}_a = \mathbf{0} \quad \text{or} \quad \left[\underline{S}^{ij} - \lambda_a g^{ij}\right] \underline{\widehat{n}}_{j\ a} = 0 \ , \tag{6.205}$$

which helps represent

$$\mathbf{S} = \sum_{a=1}^{3} \lambda_a \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \otimes \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \quad \text{or} \quad \underline{S}^{ij} = \sum_{a=1}^{3} \lambda_a g^{ik} \, \underline{\widehat{n}}_{k\ a} g^{jl} \, \underline{\widehat{n}}_{l\ a} \ , \tag{6.206}$$

noting that

$$\mathbf{n}_a \cdot \mathbf{g}^{-1} \mathbf{n}_b = \delta_{ab} \quad \text{or} \quad \underline{\widehat{n}}_{i\ a} g^{ij} \, \underline{\widehat{n}}_{j\ b} = \delta_{ab} \ . \tag{6.207}$$

The derivation of (6.204a) basically relies on computing the total differential of $(6.205)_1$, that is,

$$(d\mathbf{S})\, \widehat{\mathbf{n}}_a + \mathbf{S}\, (d\, \widehat{\mathbf{n}}_a) = d\lambda_a \mathbf{g}^{-1} \widehat{\mathbf{n}}_a + \lambda_a \left(d\mathbf{g}^{-1}\right) \widehat{\mathbf{n}}_a + \lambda_a \mathbf{g}^{-1}\, (d\, \widehat{\mathbf{n}}_a) \ .$$

Postmultiplying this expression by $\widehat{\mathbf{n}}_a$ and subsequently using $(6.57)_1$ then gives

$$(d\mathbf{S})\, \widehat{\mathbf{n}}_a \cdot \widehat{\mathbf{n}}_a = d\lambda_a \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \cdot \widehat{\mathbf{n}}_a - \lambda_a \mathbf{g}^{-1}\, (d\mathbf{g})\, \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \cdot \widehat{\mathbf{n}}_a \quad (a = 1, 2, 3; \ \text{no sum}) \ ,$$

since

$$\begin{aligned} \mathbf{S}\, (d\, \widehat{\mathbf{n}}_a) \cdot \widehat{\mathbf{n}}_a - \lambda_a \mathbf{g}^{-1}\, (d\, \widehat{\mathbf{n}}_a) \cdot \widehat{\mathbf{n}}_a &= (d\, \widehat{\mathbf{n}}_a) \cdot \mathbf{S} \widehat{\mathbf{n}}_a - \lambda_a\, (d\, \widehat{\mathbf{n}}_a) \cdot \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \\ &= (d\, \widehat{\mathbf{n}}_a) \cdot \left[\mathbf{S} \widehat{\mathbf{n}}_a - \lambda_a \mathbf{g}^{-1} \widehat{\mathbf{n}}_a\right] \\ &= 0 \ . \end{aligned}$$

Using $(2.79c)_1$ and $(6.207)_1$, one can further have

$$\underbrace{\widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a : (d\mathbf{S}) + \lambda_a \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \cdot (d\mathbf{g})\, \mathbf{g}^{-1} \widehat{\mathbf{n}}_a = d\lambda_a}_{\text{or } d\lambda_a = \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a : (d\mathbf{S}) + \lambda_a \mathbf{g}^{-1} \widehat{\mathbf{n}}_a \otimes \mathbf{g}^{-1} \widehat{\mathbf{n}}_a : (d\mathbf{g})} \ .$$

Consequently, by $\lambda_a = \lambda_a(\mathbf{S}, \mathbf{g})$ with $d\lambda_a = (\partial\lambda_a/\partial\mathbf{S}) : d\mathbf{S} + (\partial\lambda_a/\partial\mathbf{g}) : d\mathbf{g}$, one can conclude that

$$\frac{\partial\lambda_a}{\partial\mathbf{S}} = \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a, \quad \frac{\partial\lambda_a}{\partial\mathbf{g}} = \lambda_a \mathbf{g}^{-1}\widehat{\mathbf{n}}_a \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_a \ . \tag{6.208}$$

Next, consider the case of double coalescence, i.e. $\lambda_1 \neq \lambda_2 = \lambda_3 = \lambda$. In this case, by contracting

$$\mathbf{S} = \lambda_1 \mathbf{g}^{-1}\widehat{\mathbf{n}}_1 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_1 + \lambda\left(\mathbf{g}^{-1} - \mathbf{g}^{-1}\widehat{\mathbf{n}}_1 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_1\right) \ ,$$

with the symmetric tensors $\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1$ and $\mathbf{g}$, one can arrive at

$$\lambda_1 = \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 : \mathbf{S}, \quad 2\lambda + \lambda_1 = \mathbf{g} : \mathbf{S} \ .$$

By rewriting (6.211b) in the covariant form $(\lambda_1 - \lambda)\,\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 = \mathbf{gSg} - \lambda\mathbf{g}$, it should not be difficult to see that

$$\begin{aligned} d\lambda_1 &= \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 : d\mathbf{S} + (\lambda_1 - \lambda)^{-1}(\mathbf{SgS} - \lambda\mathbf{S}) : d\mathbf{g} \ , \\ 2d\lambda &= (\mathbf{g} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1) : d\mathbf{S} - (\lambda_1 - \lambda)^{-1}(\mathbf{SgS} - \lambda_1\mathbf{S}) : d\mathbf{g} \ . \end{aligned}$$

Thus,

$$\frac{\partial\lambda_1}{\partial\mathbf{S}} = \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1, \qquad \frac{\partial\lambda_1}{\partial\mathbf{g}} = (\lambda_1 - \lambda)^{-1}(\mathbf{SgS} - \lambda\mathbf{S}) \ , \tag{6.209a}$$

$$2\frac{\partial\lambda}{\partial\mathbf{S}} = \mathbf{g} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1, \qquad 2\frac{\partial\lambda}{\partial\mathbf{g}} = -(\lambda_1 - \lambda)^{-1}(\mathbf{SgS} - \lambda_1\mathbf{S}) \ . \tag{6.209b}$$

The interested reader may want to use (6.208), that is,

$$\begin{aligned} \frac{\partial\lambda_1}{\partial\mathbf{S}} &= \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1, \quad \frac{\partial\lambda_1}{\partial\mathbf{g}} = \lambda_1 \mathbf{g}^{-1}\widehat{\mathbf{n}}_1 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_1 \ , \\ \frac{\partial\lambda_2}{\partial\mathbf{S}} + \frac{\partial\lambda_3}{\partial\mathbf{S}} &= \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \ , \\ \frac{\partial\lambda_2}{\partial\mathbf{g}} + \frac{\partial\lambda_3}{\partial\mathbf{g}} &= \lambda_2 \mathbf{g}^{-1}\widehat{\mathbf{n}}_2 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_2 + \lambda_3 \mathbf{g}^{-1}\widehat{\mathbf{n}}_3 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_3 \ . \end{aligned}$$

to alternatively obtain the results (6.209a)–(6.209b). This can be achieved by taking the limit as $\lambda_2 - \lambda$ and $\lambda_3 - \lambda$ go to zero considering the fact that

$$\begin{aligned} \mathbf{g} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 &= \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \ , \\ \mathbf{SgS} - \lambda\mathbf{S} &= \lambda_1(\lambda_1 - \lambda)\,\mathbf{g}^{-1}\widehat{\mathbf{n}}_1 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_1 \ , \end{aligned}$$

$$\mathbf{SgS} - \lambda_1 \mathbf{S} = \lambda\,(\lambda - \lambda_1)\left[\mathbf{g}^{-1}\widehat{\mathbf{n}}_2 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_2 + \mathbf{g}^{-1}\widehat{\mathbf{n}}_3 \otimes \mathbf{g}^{-1}\widehat{\mathbf{n}}_3\right] .$$

Finally, consider the case of triple coalescence, i.e. $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. In this case, $\mathbf{S} = \lambda \mathbf{g}^{-1}$ implies $3\lambda = \mathbf{S} : \mathbf{g}$ or $3\lambda = \underline{S}^{ij} g_{ij}$. Consequently, by using (6.55c)$_2$ and (6.55f)$_2$,

$$3\frac{\partial \lambda}{\partial \mathbf{S}} = \mathbf{g} \quad , \quad 3\frac{\partial \lambda}{\partial \mathbf{g}} = \mathbf{S} \,. \tag{6.210}$$

These results can also be obtained from (6.208). In this case, one can write

$$\begin{aligned}
\frac{\partial \lambda_1}{\partial \mathbf{S}} + \frac{\partial \lambda_2}{\partial \mathbf{S}} + \frac{\partial \lambda_3}{\partial \mathbf{S}} &= \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \ , \\
\frac{\partial \lambda_1}{\partial \mathbf{g}} + \frac{\partial \lambda_2}{\partial \mathbf{g}} + \frac{\partial \lambda_3}{\partial \mathbf{g}} &= \mathbf{g}^{-1}\left[\lambda_1 \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \lambda_2 \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + \lambda_3 \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3\right]\mathbf{g}^{-1} \ .
\end{aligned}$$

Consequently, in the limit when $\lambda_1 \to \lambda$, $\lambda_2 \to \lambda$ and $\lambda_3 \to \lambda$, the desired results in (6.210) immediately follow.

Exercise 6.10

Let $\mathbf{S}$ be a symmetric Cartesian tensor with the spectral formulas (4.41) and (4.45)$_3$. Then, show that the symmetric tensor $\widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i$ (no summation) can be represented in closed form according to (see Simo and Taylor [25] and Saleeb et al. [26])

$$\begin{aligned}
\widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i &= \frac{\mathbf{S}^2 - [I_1(\mathbf{S}) - \lambda_i]\,\mathbf{S} + I_3(\mathbf{S})\,\lambda_i^{-1}\mathbf{I}}{D_i} \\
&= \frac{(\mathbf{S} - \lambda_j \mathbf{I})\,(\mathbf{S} - \lambda_k \mathbf{I})}{(\lambda_i - \lambda_j)\,(\lambda_i - \lambda_k)} \quad \leftarrow \text{see (4.40)} \qquad \text{if} \quad \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ ,
\end{aligned} \tag{6.211a}$$

$$\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 = \frac{\mathbf{S} - \lambda \mathbf{I}}{\lambda_1 - \lambda} \qquad \text{if} \quad \lambda_1 \neq \lambda_2 = \lambda_3 = \lambda \ , \tag{6.211b}$$

where the scalar $D_i = 2\lambda_i^2 - I_1(\mathbf{S})\,\lambda_i + I_3(\mathbf{S})\,\lambda_i^{-1}$ should be nonzero and (i, j, k) is $(1, 2, 3)$, $(2, 3, 1)$ or $(3, 1, 2)$. Recall from (4.9) that the eigenvalues could also be obtained in closed form in terms of the principal scalar invariants. These relations are important from the computational point of view because they may help avoid the explicit computation of the eigenvalues and eigenvectors. For further discussions, see Appendix A of de Souza Neto et al. [27].

Solution. First, let $\mathbf{S}$ be a symmetric tensor with exactly three distinct eigenvalues, i.e. $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1$. Then, by means of the product rule of differentiation, the sensitivity of

$$\lambda_i^3 - I_1 \lambda_i^2 + I_2 \lambda_i - I_3 = 0 \ , \qquad \leftarrow \text{see (4.6)}$$

with respect to $\mathbf{S}$ gives

$$\frac{\partial \lambda_i}{\partial \mathbf{S}} \left(3\lambda_i^2 - 2I_1\lambda_i + I_2 \right) = \lambda_i^2 \frac{\partial I_1}{\partial \mathbf{S}} - \lambda_i \frac{\partial I_2}{\partial \mathbf{S}} + \frac{\partial I_3}{\partial \mathbf{S}} \ ,$$

or, using (6.20a)–(6.20c),

$$\frac{\partial \lambda_i}{\partial \mathbf{S}} \left(3\lambda_i^2 - 2I_1\lambda_i + I_2 \right) = \lambda_i^2 \mathbf{I} - \lambda_i \left[I_1 \mathbf{I} - \mathbf{S} \right] + I_3 \mathbf{S}^{-1} \ .$$

Introducing $I_2 = -\lambda_i^2 + I_1\lambda_i + I_3\lambda_i^{-1}$ and $I_3\mathbf{S}^{-1} = \mathbf{S}^2 - I_1\mathbf{S} + I_2\mathbf{I}$ (from the Cayley-Hamilton theorem) into the above relation yields

$$\frac{\partial \lambda_i}{\partial \mathbf{S}} = \frac{\mathbf{S}^2 - \left[I_1 - \lambda_i \right] \mathbf{S} + \left[\lambda_i^2 - I_1\lambda_i + I_2 = I_3\lambda_i^{-1} \right] \mathbf{I}}{2\lambda_i^2 - I_1\lambda_i + I_3\lambda_i^{-1}} \ .$$

Consequently, by (6.204a), the desired relation $(6.211\text{a})_1$ follows. Note that this useful result can also be written in an alternative form. This is demonstrated in the following.

By using $\mathbf{S} = \sum_{i=1}^{3} \lambda_i \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i$, $\mathbf{I} = \sum_{i=1}^{3} \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i$ and $\widehat{\mathbf{n}}_i \cdot \widehat{\mathbf{n}}_j = \delta_{ij}$, one will have

$$\begin{aligned} \mathbf{S} - \lambda_1 \mathbf{I} &= (\lambda_2 - \lambda_1) \, \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 + (\lambda_3 - \lambda_1) \, \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \ , \\ \mathbf{S} - \lambda_2 \mathbf{I} &= (\lambda_1 - \lambda_2) \, \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + (\lambda_3 - \lambda_2) \, \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 \ , \end{aligned}$$

and, consequently,

$$\underbrace{(\mathbf{S} - \lambda_1 \mathbf{I}) \, (\mathbf{S} - \lambda_2 \mathbf{I}) = (\lambda_3 - \lambda_1) \, (\lambda_3 - \lambda_2) \, \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3}_{\text{or } \widehat{\mathbf{n}}_3 \otimes \widehat{\mathbf{n}}_3 = \left[(\mathbf{S} - \lambda_1 \mathbf{I}) \, (\mathbf{S} - \lambda_2 \mathbf{I}) \right] / \left[(\lambda_3 - \lambda_1) \, (\lambda_3 - \lambda_2) \right]} \ .$$

In a similar manner,

$$\begin{aligned} \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 &= \frac{(\mathbf{S} - \lambda_2 \mathbf{I}) \, (\mathbf{S} - \lambda_3 \mathbf{I})}{(\lambda_1 - \lambda_2) \, (\lambda_1 - \lambda_3)} \ , \\ \widehat{\mathbf{n}}_2 \otimes \widehat{\mathbf{n}}_2 &= \frac{(\mathbf{S} - \lambda_1 \mathbf{I}) \, (\mathbf{S} - \lambda_3 \mathbf{I})}{(\lambda_2 - \lambda_1) \, (\lambda_2 - \lambda_3)} \ . \end{aligned}$$

Next, consider the case of double coalescence, i.e. $\lambda_1 \neq \lambda_2 = \lambda_3 = \lambda$. In this case, from $\mathbf{S} = \lambda_1 \, \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \lambda \, (\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1)$, one can simply write

$$\underbrace{\mathbf{S} - \lambda \mathbf{I} = (\lambda_1 - \lambda)\, \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1}_{\text{or } \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 = (\mathbf{S} - \lambda \mathbf{I}) \,/\, (\lambda_1 - \lambda)} \ .$$

Exercise 6.11

Let **S** be a symmetric second-order tensor. First, show that a symmetric tensor function of the form $\mathbf{T}(\mathbf{S}) = (\mathbf{I} + \mathbf{S})^{-1}$ is isotropic. Then, using (6.174a)–(6.174c), derive the following representation formulas

$$(\mathbf{I} + \mathbf{S})^{-1} = f_0 \mathbf{I} + f_1 \mathbf{S} + f_2 \mathbf{S}^2 \qquad \text{if} \quad \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \tag{6.212a}$$

$$(\mathbf{I} + \mathbf{S})^{-1} = \bar{f}_0 \mathbf{I} + \bar{f}_1 \mathbf{S} \qquad \text{if} \quad \lambda_1 \neq \lambda_2 = \lambda_3 = \lambda \ , \tag{6.212b}$$

$$(\mathbf{I} + \mathbf{S})^{-1} = \hat{f}_0 \mathbf{I} \qquad \text{if} \quad \lambda_1 = \lambda_2 = \lambda_3 = \lambda \ , \tag{6.212c}$$

where

$$f_0 = \frac{1 + \lambda_1 + \lambda_2 + \lambda_3 + \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3}{(1 + \lambda_1)(1 + \lambda_2)(1 + \lambda_3)} \ , \tag{6.213a}$$

$$f_1 = -\frac{1 + \lambda_1 + \lambda_2 + \lambda_3}{(1 + \lambda_1)(1 + \lambda_2)(1 + \lambda_3)} \ , \tag{6.213b}$$

$$f_2 = \frac{1}{(1 + \lambda_1)(1 + \lambda_2)(1 + \lambda_3)} \ , \tag{6.213c}$$

$$\bar{f}_0 = \frac{1 + \lambda_1 + \lambda}{(1 + \lambda_1)(1 + \lambda)} \ , \tag{6.213d}$$

$$\bar{f}_1 = -\frac{1}{(1 + \lambda_1)(1 + \lambda)} \ , \tag{6.213e}$$

$$\hat{f}_0 = \frac{1}{1 + \lambda} \ . \tag{6.213f}$$

Solution. For the first part, by using $\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1}$, one can verify that

$$\begin{aligned} \mathbf{T}\left(\mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}\right) &= \left(\mathbf{Q}\mathbf{I}\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}\right)^{-1} \\ &= \left(\mathbf{Q}\,(\mathbf{I} + \mathbf{S})\,\mathbf{Q}^{\mathrm{T}}\right)^{-1} \\ &= \mathbf{Q}^{-\mathrm{T}}\,(\mathbf{I} + \mathbf{S})^{-1}\,\mathbf{Q}^{-1} \\ &= \mathbf{Q}\mathbf{T}\,(\mathbf{S})\,\mathbf{Q}^{\mathrm{T}} \ . \end{aligned} \tag{6.214}$$

In the following, three different cases are considered since the principal values may be repeated.

First, suppose that **S** is a tensor with non-multiple eigenvalues. Then, with the aid of (6.115) and (6.174a), the isotropic symmetric tensor function $\mathbf{T}(\mathbf{S})$ is expressible in the form

$$\mathbf{T}(\mathbf{S}) = \sum_{i=1}^{3} \left(f_0 + f_1\lambda_i + f_2\lambda_i^2 \right) \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \; . \tag{6.215}$$

On the other hand, by using (4.37b) and having in mind that any symmetric tensor has identical right and left eigenvectors, one can write

$$(\mathbf{I} + \mathbf{S})^{-1} = \sum_{i=1}^{3} (1 + \lambda_i)^{-1} \, \widehat{\mathbf{n}}_i \otimes \widehat{\mathbf{n}}_i \; . \tag{6.216}$$

It in not then difficult to see that

$$\underbrace{\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix}}_{:= [A]} \underbrace{\begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix}}_{:= [b]} = \underbrace{\begin{bmatrix} (1+\lambda_1)^{-1} \\ (1+\lambda_2)^{-1} \\ (1+\lambda_3)^{-1} \end{bmatrix}}_{:= [c]} , \quad \leftarrow \text{notice that } [A] \text{ renders a } \textbf{Vandermonde} \text{ matrix, see (6.119)}$$

or $[b] = [A]^{-1}\,[c]$ with

$$[A]^{-1} = \begin{bmatrix} \frac{\lambda_2\lambda_3}{(\lambda_3-\lambda_1)(\lambda_2-\lambda_1)} & \frac{-\lambda_1\lambda_3}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)} & \frac{\lambda_1\lambda_2}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \\ \frac{-(\lambda_2+\lambda_3)}{(\lambda_3-\lambda_1)(\lambda_2-\lambda_1)} & \frac{(\lambda_1+\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)} & \frac{-(\lambda_1+\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \\ \frac{1}{(\lambda_3-\lambda_1)(\lambda_2-\lambda_1)} & \frac{-1}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)} & \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \end{bmatrix} ,$$

helps provide (6.213a)–(6.213c).

Next, let $\mathbf{S}$ be a tensor with a simple and multiple eigenvalue. In this case, one will have

$$\begin{aligned} \mathbf{T}(\mathbf{S}) &= \bar{f}_0\mathbf{I} + \bar{f}_1\mathbf{S} \\ &= \left(\bar{f}_0 + \bar{f}_1\lambda_1\right) \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + \left(\bar{f}_0 + \bar{f}_1\lambda\right)\left(\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1\right) , \end{aligned} \tag{6.217a}$$

$$(\mathbf{I} + \mathbf{S})^{-1} = (1+\lambda_1)^{-1} \, \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1 + (1+\lambda)^{-1}\left(\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1\right) \; . \tag{6.217b}$$

Note that $\widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1$ and $\mathbf{I} - \widehat{\mathbf{n}}_1 \otimes \widehat{\mathbf{n}}_1$ in the above relations are linearly independent. Consequently,

$$\begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda \end{bmatrix} \begin{bmatrix} \bar{f}_0 \\ \bar{f}_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{1+\lambda_1} \\ \frac{1}{1+\lambda} \end{bmatrix} \text{ or } \begin{bmatrix} \bar{f}_0 \\ \bar{f}_1 \end{bmatrix} = \begin{bmatrix} \lambda & -\lambda_1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{(\lambda-\lambda_1)(1+\lambda_1)} \\ \frac{1}{(\lambda-\lambda_1)(1+\lambda)} \end{bmatrix} ,$$

simply gives the desired coefficients (6.213d)–(6.213e).

Finally, consider the case in which all eigenvalues of $\mathbf{S}$ are identical. In this case,

$$\mathbf{T}(\mathbf{S}) = \hat{f}_0 \mathbf{I} \quad , \quad (\mathbf{I} + \mathbf{S})^{-1} = (1 + \lambda)^{-1} \mathbf{I} \,. \tag{6.218}$$

And this immediately implies (6.213f).

At the end, the interested reader may want to compute $(\mathbf{I} + \mathbf{S})^{-1}$ via the representation (6.212a) and the identity $(\mathbf{I} + \mathbf{S})^{-1} (\mathbf{I} + \mathbf{S}) = \mathbf{I}$. By using the Cayley-Hamilton equation, the result is (Hoger and Carlson [28])

$$(\mathbf{I} + \mathbf{S})^{-1} = \frac{\mathbf{S}^2 - [1 + I_1(\mathbf{S})] \mathbf{S} + [1 + I_1(\mathbf{S}) + I_2(\mathbf{S})] \mathbf{I}}{1 + I_1(\mathbf{S}) + I_2(\mathbf{S}) + I_3(\mathbf{S})} \,. \tag{6.219}$$

Exercise 6.12

Let **S**, **T** and **U** be three symmetric second-order tensors. Suppose one is given the isotropic **integrity basis**

$$\begin{aligned} \Upsilon_{\mathrm{s}}^{\mathrm{ib}} = \big\{ & \mathrm{tr}\mathbf{S}, \mathrm{tr}\mathbf{S}^2, \mathrm{tr}\mathbf{S}^3, \mathrm{tr}\mathbf{T}, \mathrm{tr}\mathbf{T}^2, \mathrm{tr}\mathbf{T}^3, \mathrm{tr}\mathbf{U}, \mathrm{tr}\mathbf{U}^2, \mathrm{tr}\mathbf{U}^3, \\ & \mathrm{tr}(\mathbf{ST}), \mathrm{tr}(\mathbf{ST}^2), \mathrm{tr}(\mathbf{S}^2\mathbf{T}), \mathrm{tr}(\mathbf{S}^2\mathbf{T}^2), \\ & \mathrm{tr}(\mathbf{SU}), \mathrm{tr}(\mathbf{SU}^2), \mathrm{tr}(\mathbf{S}^2\mathbf{U}), \mathrm{tr}(\mathbf{S}^2\mathbf{U}^2), \\ & \mathrm{tr}(\mathbf{TU}), \mathrm{tr}(\mathbf{TU}^2), \mathrm{tr}(\mathbf{T}^2\mathbf{U}), \mathrm{tr}(\mathbf{T}^2\mathbf{U}^2), \\ & \mathrm{tr}(\mathbf{STU}), \mathrm{tr}(\mathbf{S}^2\mathbf{TU}), \mathrm{tr}(\mathbf{T}^2\mathbf{US}), \mathrm{tr}(\mathbf{U}^2\mathbf{ST}), \\ & \mathrm{tr}(\mathbf{S}^2\mathbf{T}^2\mathbf{U}), \mathrm{tr}(\mathbf{T}^2\mathbf{U}^2\mathbf{S}), \mathrm{tr}(\mathbf{U}^2\mathbf{S}^2\mathbf{T}) \big\} \,, \end{aligned} \tag{6.220}$$

for the agencies **S**, **T** and **U** (Zheng [29]). Then, derive a representation formula for an isotropic **polynomial** symmetric tensor function of **S** and **T** according to

$$\begin{aligned} \mathbf{S}^{\mathrm{ib}}(\mathbf{S}, \mathbf{T}) = & f_0(\hat{\Upsilon}_{\mathrm{s}}) \mathbf{I} + f_1(\hat{\Upsilon}_{\mathrm{s}}) \mathbf{S} + f_2(\hat{\Upsilon}_{\mathrm{s}}) \mathbf{S}^2 + f_3(\hat{\Upsilon}_{\mathrm{s}}) \mathbf{T} \\ & + f_4(\hat{\Upsilon}_{\mathrm{s}}) \mathbf{T}^2 + f_5(\hat{\Upsilon}_{\mathrm{s}}) [\mathbf{ST} + \mathbf{TS}] + f_6(\hat{\Upsilon}_{\mathrm{s}}) [\mathbf{S}^2\mathbf{T} + \mathbf{TS}^2] \\ & + f_7(\hat{\Upsilon}_{\mathrm{s}}) [\mathbf{T}^2\mathbf{S} + \mathbf{ST}^2] + f_8(\hat{\Upsilon}_{\mathrm{s}}) [\mathbf{S}^2\mathbf{T}^2 + \mathbf{T}^2\mathbf{S}^2] \,, \end{aligned} \tag{6.221}$$

where $f_k,\ k = 0, \ldots, 8$, are isotropic polynomial invariants of

$$\begin{aligned} \hat{\Upsilon}_{\mathrm{s}} = \big\{ & \mathrm{tr}\mathbf{S}, \mathrm{tr}\mathbf{S}^2, \mathrm{tr}\mathbf{S}^3, \mathrm{tr}\mathbf{T}, \mathrm{tr}\mathbf{T}^2, \mathrm{tr}\mathbf{T}^3, \\ & \mathrm{tr}(\mathbf{ST}), \mathrm{tr}(\mathbf{ST}^2), \mathrm{tr}(\mathbf{S}^2\mathbf{T}), \mathrm{tr}(\mathbf{S}^2\mathbf{T}^2) \big\} \,. \end{aligned}$$ ← see (6.143)

Comparing (6.220) with one obtained from Table 6.1 shows that an irreducible functional basis contains fewer elements than an integrity basis. In the above expressions, the elements colored in blue show the extra members.

Solution. The procedure used here follows from the pioneering work of Pipkin and Rivlin [7], see also Chap. 8 of Boehler [20]. And it should be used to only represent the polynomial tensor functions. To begin with, one needs to determine an appropriate integrity basis.

Let $\beta\,(\mathbf{S}, \mathbf{T}, \mathbf{U})$ be a scalar-valued function of the three symmetric tensors **S**, **T** and **U**; defined by,

$$\beta\,(\mathbf{S}, \mathbf{T}, \mathbf{U}) = \mathbf{U} : \mathbf{S}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T}) \quad \text{noting that} \quad \mathbf{S}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T}) = \frac{\partial \beta\,(\mathbf{S}, \mathbf{T}, \mathbf{U})}{\partial \mathbf{U}} \ . \tag{6.222}$$

Here, $\mathbf{S}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T})$ presents an **isotropic** symmetric tensor-valued function of **S** and **T**. Consequently, this linear function of **U** fulfills the isotropy condition:

$$\begin{aligned}
\beta\left(\mathbf{QSQ}^{\mathrm{T}}, \mathbf{QTQ}^{\mathrm{T}}, \mathbf{QUQ}^{\mathrm{T}}\right) &= \mathbf{QUQ}^{\mathrm{T}} : \mathbf{S}^{\mathrm{ib}}\left(\mathbf{QSQ}^{\mathrm{T}}, \mathbf{QTQ}^{\mathrm{T}}\right) \\
&= \mathbf{QUQ}^{\mathrm{T}} : \mathbf{QS}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T})\,\mathbf{Q}^{\mathrm{T}} \\
&= \mathbf{U} : \mathbf{Q}^{\mathrm{T}}\mathbf{QS}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T})\,\mathbf{Q}^{\mathrm{T}}\mathbf{Q} \\
&= \mathbf{U} : \mathbf{IS}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T})\,\mathbf{I} \\
&= \mathbf{U} : \mathbf{S}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T}) \\
&= \beta\,(\mathbf{S}, \mathbf{T}, \mathbf{U}) \ .
\end{aligned} \tag{6.223}$$

Thus, by removing the terms involving $\mathbf{U}^2$ in (6.220), it should depend on

$$\begin{aligned}
\bar{\Upsilon}_{\mathrm{s}}^{\mathrm{ib}} = \big\{&\mathbf{U} : \mathbf{I}, \mathbf{U} : \mathbf{S}, \mathbf{U} : \mathbf{S}^2, \mathbf{U} : \mathbf{T}, \mathbf{U} : \mathbf{T}^2, \\
&\underbrace{\mathbf{U} : \mathbf{ST}, \mathbf{U} : \mathbf{S}^2\mathbf{T}, \mathbf{U} : \mathbf{T}^2\mathbf{S}, \mathbf{U} : \mathbf{S}^2\mathbf{T}^2, \hat{\Upsilon}_{\mathrm{s}}\big\}}_{\text{note that, e.g., } \operatorname{tr}(\mathbf{STU}) = \mathbf{U} : \mathbf{ST} = \mathbf{U} : (\mathbf{ST} + \mathbf{TS})/2} \ .
\end{aligned} \tag{6.224}$$

This helps express (6.222) as

$$\begin{aligned}
\beta\,(\mathbf{S}, \mathbf{T}, \mathbf{U}) = \mathbf{U} : \mathbf{S}^{\mathrm{ib}}\,(\mathbf{S}, \mathbf{T}) = \mathbf{U} : \Big\{ & f_0\left(\hat{\Upsilon}_{\mathrm{s}}\right)\mathbf{I} + f_1\left(\hat{\Upsilon}_{\mathrm{s}}\right)\mathbf{S} + f_2\left(\hat{\Upsilon}_{\mathrm{s}}\right)\mathbf{S}^2 \\
& + f_3\left(\hat{\Upsilon}_{\mathrm{s}}\right)\mathbf{T} + f_4\left(\hat{\Upsilon}_{\mathrm{s}}\right)\mathbf{T}^2 + f_5\left(\hat{\Upsilon}_{\mathrm{s}}\right)[\mathbf{ST} + \mathbf{TS}] \\
& + f_6\left(\hat{\Upsilon}_{\mathrm{s}}\right)\left[\mathbf{S}^2\mathbf{T} + \mathbf{TS}^2\right] + f_7\left(\hat{\Upsilon}_{\mathrm{s}}\right)\left[\mathbf{T}^2\mathbf{S} + \mathbf{ST}^2\right] \\
& + f_8\left(\hat{\Upsilon}_{\mathrm{s}}\right)\left[\mathbf{S}^2\mathbf{T}^2 + \mathbf{T}^2\mathbf{S}^2\right]\Big\} \ ,
\end{aligned}$$

which implies the desired representation (6.221).

Exercise 6.13

By means of the Cayley-Hamilton equation (4.21), obtain the following two variables tensor identity

$$\mathbf{ABA} + \mathbf{A}^2\mathbf{B} + \mathbf{BA}^2 = \mathbf{G}(\mathbf{A}, \mathbf{B}) \ , \tag{6.225}$$

where

$$\begin{aligned}\mathbf{G} &= (\mathrm{tr}\,\mathbf{A})(\mathbf{AB} + \mathbf{BA}) + (\mathrm{tr}\,\mathbf{B})\,\mathbf{A}^2 \\ &\quad - [\mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B} - \mathrm{tr}(\mathbf{AB})]\,\mathbf{A} - \frac{(\mathrm{tr}\,\mathbf{A})^2 - \mathrm{tr}\,\mathbf{A}^2}{2}\mathbf{B} \\ &\quad + \left[\mathrm{tr}\left(\mathbf{A}^2\mathbf{B}\right) - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}(\mathbf{AB}) + \frac{(\mathrm{tr}\,\mathbf{A})^2 - \mathrm{tr}\,\mathbf{A}^2}{2}\mathrm{tr}\,\mathbf{B}\right]\mathbf{I} \ ,\end{aligned} \tag{6.226}$$

as well as the three fields tensor identity

$$\begin{aligned}&\mathbf{ABC} + \mathbf{ACB} + \mathbf{BCA} \\ &\quad + \mathbf{BAC} + \mathbf{CAB} + \mathbf{CBA} = \mathbf{H}(\mathbf{A}, \mathbf{B}, \mathbf{C}) \ ,\end{aligned} \tag{6.227}$$

where

$$\begin{aligned}\mathbf{H} &= \frac{1}{2}[\mathbf{G}(\mathbf{A}+\mathbf{C}, \mathbf{B}) - \mathbf{G}(\mathbf{A}-\mathbf{C}, \mathbf{B})] \\ &= (\mathrm{tr}\,\mathbf{C})(\mathbf{AB} + \mathbf{BA}) + (\mathrm{tr}\,\mathbf{A})(\mathbf{BC} + \mathbf{CB}) \\ &\quad + (\mathrm{tr}\,\mathbf{B})(\mathbf{AC} + \mathbf{CA}) - [\mathrm{tr}\,\mathbf{B}\,\mathrm{tr}\,\mathbf{C} - \mathrm{tr}(\mathbf{BC})]\,\mathbf{A} \\ &\quad - [\mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{C} - \mathrm{tr}(\mathbf{AC})]\,\mathbf{B} - [\mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B} - \mathrm{tr}(\mathbf{AB})]\,\mathbf{C} \\ &\quad + [\mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B}\,\mathrm{tr}\,\mathbf{C} - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}(\mathbf{BC}) - \mathrm{tr}\,\mathbf{B}\,\mathrm{tr}(\mathbf{AC}) \\ &\quad - \mathrm{tr}\,\mathbf{C}\,\mathrm{tr}(\mathbf{AB}) + \mathrm{tr}(\mathbf{ABC}) + \mathrm{tr}(\mathbf{BAC})]\,\mathbf{I} \ .\end{aligned} \tag{6.228}$$

Here, **A**, **B** and **C** denote arbitrary second-order tensors. The above expressions are known as the *Rivlin's identities* (Rivlin [6]) which are beneficial for representation of isotropic tensor-valued functions. See Scheidler [30] for an application to kinematics of continuum bodies.

Solution. The Rivlin's identities can be provided in some alternative ways. Here, they are obtained by using the partial derivatives of the Cayley-Hamilton relation $\mathbf{A}^3 - I_1\mathbf{A}^2 + I_2\mathbf{A} - I_3\mathbf{I} = \mathbf{O}$ with respect to **A**. From (6.20a)–(6.20c) and (6.195), it follows that

$$\mathbb{O} = \frac{\partial \mathbf{A}^3}{\partial \mathbf{A}} - \underbrace{\mathbf{A}^2 \otimes \frac{\partial I_1}{\partial \mathbf{A}}}_{= \mathbf{A}^2 \otimes \mathbf{I}} - I_1 \frac{\partial \mathbf{A}^2}{\partial \mathbf{A}}$$

$$+ \underbrace{\mathbf{A} \otimes \frac{\partial I_2}{\partial \mathbf{A}}}_{= \mathbf{A} \otimes \left(I_1 \mathbf{I} - \mathbf{A}^{\mathrm{T}}\right)} + I_2 \frac{\partial \mathbf{A}}{\partial \mathbf{A}} - \underbrace{\mathbf{I} \otimes \frac{\partial I_3}{\partial \mathbf{A}}}_{= \mathbf{I} \otimes \left(I_3 \mathbf{A}^{-\mathrm{T}}\right) = \mathbf{I} \otimes \left(\mathbf{A}^{\mathrm{T}} \mathbf{A}^{\mathrm{T}} - I_1 \mathbf{A}^{\mathrm{T}} + I_2 \mathbf{I}\right)} .$$

The double contraction of this expression with an arbitrary tensor **B**, taking into account (3.36) and (6.47)–(6.49b), will lead to

$$\mathbf{O} = \underbrace{\frac{\partial \mathbf{A}^3}{\partial \mathbf{A}} : \mathbf{B}}_{= \mathbf{B}\mathbf{A}^2 + \mathbf{A}\mathbf{B}\mathbf{A} + \mathbf{A}^2\mathbf{B}} - \underbrace{\left(\mathbf{A}^2 \otimes \mathbf{I}\right) : \mathbf{B}}_{= \mathbf{A}^2 (\mathbf{I} : \mathbf{B}) = (\mathrm{tr}\,\mathbf{B})\, \mathbf{A}^2} - I_1 \underbrace{\frac{\partial \mathbf{A}^2}{\partial \mathbf{A}} : \mathbf{B}}_{= \mathbf{B}\mathbf{A} + \mathbf{A}\mathbf{B}}$$
$$+ \underbrace{\left[\mathbf{A} \otimes \left(I_1 \mathbf{I} - \mathbf{A}^{\mathrm{T}}\right)\right] : \mathbf{B}}_{= \mathbf{A}\left[I_1 (\mathrm{tr}\,\mathbf{B}) - \mathrm{tr}\,(\mathbf{A}\mathbf{B})\right]} + I_2 \underbrace{\frac{\partial \mathbf{A}}{\partial \mathbf{A}} : \mathbf{B}}_{= \mathbf{B}} - \underbrace{\left[\mathbf{I} \otimes \left(\mathbf{A}^{\mathrm{T}} \mathbf{A}^{\mathrm{T}} - I_1 \mathbf{A}^{\mathrm{T}} + I_2 \mathbf{I}\right)\right] : \mathbf{B}}_{= \mathbf{I}\left[\mathrm{tr}\left(\mathbf{A}^2\mathbf{B}\right) - I_1 \mathrm{tr}\,(\mathbf{A}\mathbf{B}) + I_2 \mathrm{tr}\,\mathbf{B}\right]} .$$

This result along with (4.17a) and $(4.17\text{b})_2$ delivers the desired relations (6.225) and (6.226). Note that the Cayley-Hamilton equation (4.21) can be resulted from the tensor identity (6.225) by setting $\mathbf{B} = \mathbf{A}$.

By repeating the above procedure, i.e. calculating the sensitivity of (6.225) with respect to **A** and subsequently contracting the resulting expression with a generic tensor **C**, one can arrive at

$$\mathbf{O} = \underbrace{\frac{\partial\, [\mathbf{A}\mathbf{B}\mathbf{A}]}{\partial \mathbf{A}} : \mathbf{C}}_{= (\mathbf{C}\mathbf{B}\mathbf{A} + \mathbf{A}\mathbf{B}\mathbf{C})} + \underbrace{\frac{\partial \left[\mathbf{A}^2\mathbf{B}\right]}{\partial \mathbf{A}} : \mathbf{C}}_{= (\mathbf{C}\mathbf{A} + \mathbf{A}\mathbf{C})\,\mathbf{B}} + \underbrace{\frac{\partial \left[\mathbf{B}\mathbf{A}^2\right]}{\partial \mathbf{A}} : \mathbf{C}}_{= \mathbf{B}\,(\mathbf{C}\mathbf{A} + \mathbf{A}\mathbf{C})}$$
$$- \underbrace{\frac{\partial\, [(\mathrm{tr}\,\mathbf{A})\, (\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A})]}{\partial \mathbf{A}} : \mathbf{C}}_{= (\mathrm{tr}\,\mathbf{C})\,(\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A}) + (\mathrm{tr}\,\mathbf{A})\,(\mathbf{C}\mathbf{B} + \mathbf{B}\mathbf{C})} - \underbrace{\frac{\partial \left[(\mathrm{tr}\,\mathbf{B})\, \mathbf{A}^2\right]}{\partial \mathbf{A}} : \mathbf{C}}_{= (\mathrm{tr}\,\mathbf{B})\,(\mathbf{C}\mathbf{A} + \mathbf{A}\mathbf{C})}$$
$$- \underbrace{\frac{\partial\, \{[\mathrm{tr}\,(\mathbf{A}\mathbf{B}) - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B}]\, \mathbf{A}\}}{\partial \mathbf{A}} : \mathbf{C}}_{\substack{= \left\{\mathbf{A} \otimes \frac{\partial [\mathrm{tr}(\mathbf{A}\mathbf{B}) - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B}]}{\partial \mathbf{A}}\right\} : \mathbf{C} + [\mathrm{tr}\,(\mathbf{A}\mathbf{B}) - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B}]\, \frac{\partial \mathbf{A}}{\partial \mathbf{A}} : \mathbf{C} \\ = \mathbf{A}\,\{\mathrm{tr}\,(\mathbf{B}\mathbf{C}) - \mathrm{tr}\,\mathbf{B}\,\mathrm{tr}\,\mathbf{C}\} + [\mathrm{tr}\,(\mathbf{A}\mathbf{B}) - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{B}]\,\mathbf{C}}}$$
$$+ \underbrace{\frac{\partial \left\{\left[(\mathrm{tr}\,\mathbf{A})^2 - \mathrm{tr}\,\mathbf{A}^2\right] \mathbf{B}\right\}}{2\partial \mathbf{A}} : \mathbf{C}}_{= \left\{\mathbf{B} \otimes \frac{\partial \left[(\mathrm{tr}\,\mathbf{A})^2 - \mathrm{tr}\,\mathbf{A}^2\right]}{2\partial \mathbf{A}}\right\} : \mathbf{C} = \mathbf{B}\left\{\left[(\mathrm{tr}\,\mathbf{A})\,\mathbf{I} - \mathbf{A}^{\mathrm{T}}\right] : \mathbf{C}\right\} = [\mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,\mathbf{C} - \mathrm{tr}\,(\mathbf{A}\mathbf{C})]\,\mathbf{B}}$$
$$- \left\{ \underbrace{\frac{\partial \left[\mathrm{tr}\left(\mathbf{A}^2\mathbf{B}\right)\right]}{\partial \mathbf{A}} : \mathbf{C}}_{= \left[\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} + \mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\right] : \mathbf{C} = \mathrm{tr}\,(\mathbf{A}\mathbf{B}\mathbf{C}) + \mathrm{tr}\,(\mathbf{B}\mathbf{A}\mathbf{C})} - \underbrace{\frac{\partial\, [\mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,(\mathbf{A}\mathbf{B})]}{\partial \mathbf{A}} : \mathbf{C}}_{= \mathrm{tr}\,\mathbf{C}\,\mathrm{tr}\,(\mathbf{A}\mathbf{B}) + \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,(\mathbf{B}\mathbf{C})} \right.$$

$$+ \underbrace{\frac{\partial \left\{ \left[(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2 \right] \operatorname{tr} \mathbf{B} \right\}}{2 \partial \mathbf{A}} : \mathbf{C}}_{= (\operatorname{tr} \mathbf{B}) \frac{\partial \left[(\operatorname{tr} \mathbf{A})^2 - \operatorname{tr} \mathbf{A}^2 \right]}{2 \partial \mathbf{A}} : \mathbf{C} = [\operatorname{tr} \mathbf{A} \operatorname{tr} \mathbf{C} - \operatorname{tr} (\mathbf{AC})] \operatorname{tr} \mathbf{B}} \Bigg\} \mathbf{I} ,$$

where the equations (2.55b), $(2.55d)_1$, $(2.83)_1$, $(2.84)_2$, (3.33), $(6.20a)_2$, $(6.20b)_2$, $(6.47)_4$, (6.49a), $(6.69)_{1-2}$ and (6.190d) have been used. From this result, one can simply arrive at the three variables tensor identity (6.227). Note that (6.225) can be recovered from (6.227) by setting $\mathbf{C} = \mathbf{A}$.

For an n-dimensional space, the expressions (6.225) and (6.227) respectively translate to

$$\underbrace{\mathbf{O} = \sum_{k=1}^{n} \sum_{i=1}^{k} (-1)^{k-i} \, I_{k-i} \, (\mathbf{A}) \, \mathbf{A}^{n-k} \left\{ \left[\operatorname{tr} \left(\mathbf{A}^{i-1} \mathbf{B} \right) \right] \mathbf{I} - \mathbf{B} \mathbf{A}^{i-1} \right\} ,}_{\text{note that } I_0 (\mathbf{A}) = 1 \text{ and } \mathbf{A}^0 = \mathbf{I}} \tag{6.229}$$

and

$$\begin{aligned} \mathbf{O} = \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} \sum_{j=1}^{k-i} & (-1)^{k-i-j} \, I_{k-i-j} \, (\mathbf{A}) \, \mathbf{A}^{n-k} \left\{ \mathbf{C} \mathbf{A}^{i-1} \mathbf{B} \mathbf{A}^{j-1} + \mathbf{B} \mathbf{A}^{j-1} \mathbf{C} \mathbf{A}^{i-1} \right. \\ & - \left[\operatorname{tr} \left(\mathbf{A}^{j-1} \mathbf{B} \right) \right] \mathbf{C} \mathbf{A}^{i-1} - \left[\operatorname{tr} \left(\mathbf{A}^{j-1} \mathbf{C} \right) \right] \mathbf{B} \mathbf{A}^{i-1} \\ & \left. + \left[\operatorname{tr} \left(\mathbf{A}^{i-1} \mathbf{B} \right) \operatorname{tr} \left(\mathbf{A}^{j-1} \mathbf{C} \right) - \operatorname{tr} \left(\mathbf{A}^{i-1} \mathbf{B} \mathbf{A}^{j-1} \mathbf{C} \right) \right] \mathbf{I} \right\} , \end{aligned} \tag{6.230}$$

which are referred to as the *generalized Rivlin's identities* (see Dui and Chen [31] and also Chap. 6 of Itskov [32]).

Hint: The interested reader can derive the Rivlin's identities in an alternative way by applying the so-called *polarization operator* to the Cayley-Hamilton theorem. That is, by replacing the tensor $\mathbf{A}$ in the Cayley-Hamilton relation by $\mathbf{A} + \varepsilon \mathbf{B}$, where $\varepsilon \in \mathscr{R}$, $\mathbf{B} \in \mathscr{T}_{\text{so}}$, and subsequently computing the partial derivative of the resulting expression with respect to ε at $\varepsilon = 0$, one can identically derive (6.225) (recall from (6.11a)–(6.11c) that this method was introduced as the directional derivative). See Lew [33] for more details. Similar result can be achieved by replacing $\mathbf{A}$ in the Cayley-Hamilton equation by $\mathbf{A} + \mathbf{B}$ and $\mathbf{A} - \mathbf{B}$ and then subtracting the resulting expressions.

Exercise 6.14

This exercise aims at finding some redundant elements in the representation of isotropic tensor functions by means of the Cayley-Hamilton theorem and/or Rivlin's identities.

The problem is well understood by introducing the **product** of tensor variables with some respective definitions. To begin with, consider N tensors $\mathbf{A}_1, \ldots, \mathbf{A}_N$, none of which is the identity tensor. And let $\boldsymbol{\Pi}$ be a product of these tensors defined by

$$\boldsymbol{\Pi} = \mathbf{M}_1^{\alpha_1} \mathbf{M}_2^{\alpha_2} \cdots \mathbf{M}_n^{\alpha_n} , \tag{6.231}$$

where α's are positive integers and $\mathbf{M}_1, \mathbf{M}_2, \ldots$ are a selection from the tensors $\mathbf{A}_1, \ldots, \mathbf{A}_N$ including repetition but no two adjacent tensors are allowed to be the same in the sequence $\mathbf{M}_1, \mathbf{M}_2, \ldots$. The following definitions are introduced:

- ☛☞ Any of $\mathbf{M}_1, \mathbf{M}_2, \ldots$ is a *product factor* of $\boldsymbol{\Pi}$ and the corresponding exponent $\alpha_i,\ i = 1, 2, \ldots, n$ presents the *power* of that factor in $\boldsymbol{\Pi}$.
- ☛☞ The *total degree* (or simply *degree*) of $\boldsymbol{\Pi}$ is the sum of all α's.
- ☛☞ The *partial degree* of $\boldsymbol{\Pi}$ in a variable is the sum of the exponents of that variable (note that the partial degree of the identity tensor is zero).
- ☛☞ The *extension* of $\boldsymbol{\Pi}$ is the number of factors occurring in the product (note that the extension of the identity tensor is zero).

As an example, the partial degree of $\mathbf{A}_2\mathbf{A}_1\mathbf{A}_2^2\mathbf{A}_3^4$ is 1, 3 and 4 in $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$, respectively. As another example, the extension of $\mathbf{A}_1\mathbf{A}_2^2\mathbf{A}_3^3$ is 3 while the extension of $\mathbf{A}_2\mathbf{A}_1\mathbf{A}_2^2\mathbf{A}_3^2$ is 4.

A **tensor polynomial** in $\mathbf{A}_1, \ldots, \mathbf{A}_N$ is defined to be a linear combination of products of $\boldsymbol{\Pi}$-type with coefficients which are expressible as polynomials in traces of $\mathbf{A}_1, \ldots, \mathbf{A}_N$. The extension of a polynomial is introduced as that of its term of largest extension. And the highest degree in products of such a polynomial is called its degree. Note that the partial degree of a polynomial is defined in a similar manner. For instance, consider the tensor polynomial

$$\alpha \left(\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1\right) + \beta \left(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_1 + \mathbf{A}_2\mathbf{A}_1\mathbf{A}_2\right) + \gamma \left(\mathbf{A}_1^2\mathbf{A}_2^2 + \mathbf{A}_2^2\mathbf{A}_1^2\right) ,$$

whose extension (degree) is 3 (4) and its partial degree in $\mathbf{A}_1$ ($\mathbf{A}_2$) renders 2 (2).

If a tensor polynomial is replaced by an exactly equal polynomial of lower extension, then such a polynomial is said to be *contracted*. For instance, the *contraction* of $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_1$ to a polynomial of extension 2 has been illustrated in the tensor identity (6.225).

A tensor polynomial is *reducible* when it can be replaced by an identically equal polynomial of lower degree. A well-known example is $\mathbf{A}^3$ which, by using the Cayley-Hamilton equation, is expressible as a polynomial of degree 2. As a result, any tensor product containing a variable to the nth power is reducible when $n \geq 3$. Notice that the products in Tables 6.3 and 6.4 do not contain any factor to the power of 3 (or greater than 3). As another example, consider the Rivlin's identity (6.225) which helps one *reduce* $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_1 + \mathbf{A}_1^2\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_1^2$ to a polynomial of degree 2.

The following notation

$$\boldsymbol{\Pi} \equiv \mathbf{O}\ , \tag{6.232}$$

is used to indicate that the product $\boldsymbol{\Pi}$ is reducible. In a similar manner, the reducibility of its trace is denoted by

$$\mathrm{tr}\,\boldsymbol{\Pi} \equiv 0\ . \tag{6.233}$$

Let $\boldsymbol{\Pi}$ and $\tilde{\boldsymbol{\Pi}}$ be two tensor products formed from any number of tensors with the same partial degree in each factor. Then, $\boldsymbol{\Pi}$ ($\mathrm{tr}\,\boldsymbol{\Pi}$) is said to be *equivalent* to $\tilde{\boldsymbol{\Pi}}$ ($\mathrm{tr}\,\tilde{\boldsymbol{\Pi}}$) if and only if $\boldsymbol{\Pi} + \tilde{\boldsymbol{\Pi}}$ ($\mathrm{tr}\,\boldsymbol{\Pi} + \mathrm{tr}\,\tilde{\boldsymbol{\Pi}}$) is reducible. This is denoted by

$$\boldsymbol{\Pi} \equiv \tilde{\boldsymbol{\Pi}} \quad , \quad \mathrm{tr}\,\boldsymbol{\Pi} \equiv \mathrm{tr}\,\tilde{\boldsymbol{\Pi}}\ . \tag{6.234}$$

One is thus allowed to use $\tilde{\boldsymbol{\Pi}}$ instead of $\boldsymbol{\Pi}$ in the representation theorems and vice versa.

As an example, consider the two symmetric tensors $\mathbf{S}$ and $\mathbf{v} \otimes \mathbf{v}$. Let

$$\boldsymbol{\Pi}_{\mathrm{s}} = \mathbf{S}\,(\mathbf{v} \otimes \mathbf{v})\,\mathbf{S} = \mathbf{S}\mathbf{v} \otimes \mathbf{S}\mathbf{v} \quad , \quad \tilde{\boldsymbol{\Pi}}_{\mathrm{s}} = \mathbf{v} \otimes \mathbf{S}^2\mathbf{v} + \mathbf{S}^2\mathbf{v} \otimes \mathbf{v}\ . \tag{6.235}$$

Note that the partial degree of $\boldsymbol{\Pi}_{\mathrm{s}}$ (or $\tilde{\boldsymbol{\Pi}}_{\mathrm{s}}$) is 2 and 1 in $\mathbf{S}$ and $\mathbf{v} \otimes \mathbf{v}$, respectively. And the extension of $\boldsymbol{\Pi}_{\mathrm{s}}$ is 3 whereas the extension of $\tilde{\boldsymbol{\Pi}}_{\mathrm{s}}$ renders 2. By using (6.225), one then has

$$\boldsymbol{\Pi}_{\mathrm{s}} + \tilde{\boldsymbol{\Pi}}_{\mathrm{s}} = \mathbf{S}\mathbf{v} \otimes \mathbf{S}\mathbf{v} + \left(\mathbf{v} \otimes \mathbf{S}^2\mathbf{v} + \mathbf{S}^2\mathbf{v} \otimes \mathbf{v}\right) = \mathbf{G}\,(\mathbf{S}, \mathbf{v} \otimes \mathbf{v})\ . \tag{6.236}$$

According to the above considerations, $\boldsymbol{\Pi}_{\mathrm{s}} + \tilde{\boldsymbol{\Pi}}_{\mathrm{s}}$ is reducible. Thus, $\boldsymbol{\Pi}_{\mathrm{s}}$ is equivalent to $\tilde{\boldsymbol{\Pi}}_{\mathrm{s}}$. And this reveals the fact that $\tilde{\boldsymbol{\Pi}}_{\mathrm{s}}$ in the sixth row of Table 6.3 can be replaced by $\boldsymbol{\Pi}_{\mathrm{s}}$.

In what follows, the tensors $\mathbf{A}_1, \ldots, \mathbf{A}_7$ are respectively denoted by $\mathbf{A}, \mathbf{B}, \ldots, \mathbf{G}$ for convenience.

The following lemma is frequently used in the subsequent developments.

Lemma D *Suppose* $\boldsymbol{\Pi}$ *is reducible. Then, the quantity tr* $\left(\tilde{\mathbf{A}}\boldsymbol{\Pi}\right)$ *(or tr* $\left(\boldsymbol{\Pi}\tilde{\mathbf{A}}\right)$*) is reducible for a (nonzero and non-identity) tensor* $\tilde{\mathbf{A}}$.

As a quick example, the invariant $\mathrm{tr}\left(\mathbf{A}^3\mathbf{B}\right)$ should not be considered for the representation of an isotropic function. Another example regards

$$\mathrm{tr}\,[\mathbf{D}(\mathbf{ABC} + \mathbf{ACB} + \mathbf{BCA} + \mathbf{BAC} + \mathbf{CAB} + \mathbf{CBA})] \equiv 0\ , \tag{6.237}$$

which relies on the fact that the left hand side of (6.227) is reducible.

1. Show that $\mathbf{ABA}^2 + \mathbf{A}^2\mathbf{BA}$ is reducible.
Let $\mathbf{S}$, $\mathbf{T}$ and $\mathbf{U}$ be three symmetric tensors. Then, by use of the above result, show that $\mathrm{tr}\left(\mathbf{STS}^2\mathbf{U}\right)$ and $\mathrm{tr}\left(\mathbf{S}^2\mathbf{TSU}\right)$ are reducible. Moreover, deduce that $\mathbf{Sv} \otimes \mathbf{S}^2\mathbf{v} + \mathbf{S}^2\mathbf{v} \otimes \mathbf{Sv}$ should not be considered as a generator for the isotropic symmetric tensor-valued functions of $\mathbf{v}$ and $\mathbf{S}$.

Solution. Consider the three fields tensor identity (6.227). Let $\mathbf{C} = \mathbf{A}^2$. Then,

$$\mathbf{ABA}^2 + \mathbf{A}^2\mathbf{BA} = -2\left(\mathbf{A}^3\mathbf{B} + \mathbf{BA}^3\right) + \mathbf{H}\left(\mathbf{A}, \mathbf{B}, \mathbf{A}^2\right) \ . \tag{6.238}$$

Substituting for $\mathbf{A}^3$ from the Cayley-Hamilton theorem on the right hand side helps write

$$\begin{aligned}\mathbf{ABA}^2 + \mathbf{A}^2\mathbf{BA} = (\mathrm{tr}\,\mathbf{A})\,\mathbf{ABA} &+ [\mathrm{tr}\,(\mathbf{AB})]\,\mathbf{A}^2 \\ &+ \left[\mathrm{tr}\left(\mathbf{A}^2\mathbf{B}\right) - \mathrm{tr}\,\mathbf{A}\,\mathrm{tr}\,(\mathbf{AB})\right]\mathbf{A} \\ &- (\det\mathbf{A})\,\mathbf{B} + (\det\mathbf{A})\,(\mathrm{tr}\,\mathbf{B})\,\mathbf{I} \ .\end{aligned} \tag{6.239}$$

Thus, the degree 4 polynomial $\mathbf{ABA}^2 + \mathbf{A}^2\mathbf{BA}$ has been reduced to a polynomial of degree 3. With the aid of (6.225), it can also be contracted to a polynomial of extension 2.
Guided by Lemma D, one can deduce that

$$\mathrm{tr}\left[\left(\mathbf{ABA}^2 + \mathbf{A}^2\mathbf{BA}\right)\mathbf{C}\right] \equiv 0 \ . \tag{6.240}$$

Now, suppose that $\mathbf{A} = \mathbf{S}$, $\mathbf{B} = \mathbf{T}$ and $\mathbf{C} = \mathbf{U}$ are symmetric tensors. Then,

$$\underbrace{\mathrm{tr}\left(\mathbf{STS}^2\mathbf{U}\right) = \mathrm{tr}\left(\mathbf{S}^2\mathbf{TSU}\right)}_{\text{since } \delta_{ij}\,(\mathbf{S})_{il}\,(\mathbf{T})_{lk}\left(\mathbf{S}^2\right)_{km}(\mathbf{U})_{mj} = (\mathbf{S})_{il}\,(\mathbf{T})_{lk}\left(\mathbf{S}^2\right)_{km}(\mathbf{U})_{mi} = \delta_{nm}\left(\mathbf{S}^2\right)_{nk}(\mathbf{T})_{kl}\,(\mathbf{S})_{li}\,(\mathbf{U})_{im}} \ . \tag{6.241}$$

And this helps deduce that

$$\mathrm{tr}\left(\mathbf{STS}^2\mathbf{U}\right) = \mathrm{tr}\left(\mathbf{S}^2\mathbf{TSU}\right) \equiv 0 \ . \tag{6.242}$$

Replacing $\mathbf{T}$ by $\mathbf{v} \otimes \mathbf{v}$ in the result $\mathbf{STS}^2 + \mathbf{S}^2\mathbf{TS} \equiv 0$ immediately implies the reducibility of $\mathbf{Sv} \otimes \mathbf{S}^2\mathbf{v} + \mathbf{S}^2\mathbf{v} \otimes \mathbf{Sv}$.

2. Prove the reducibility of the product $\mathbf{A}^2\mathbf{BC}^2$.
By means of this result, show that the trace of the product $\mathbf{S}^2\mathbf{T}^2\mathbf{U}^2$ is reducible when all of these tensors are symmetric.

Solution. To express $\mathbf{A}\left(\mathbf{ABC}^2\right)$ as a sum of products of lower degrees, one may rearrange the tensor identity (6.225) according to

$$\mathbf{ABC}^2 = \mathbf{G}\,(\mathbf{C}, \mathbf{AB}) - \mathbf{CABC} - \mathbf{C}^2\mathbf{AB} \ , \tag{6.243}$$

and find

$$
\begin{aligned}
\mathbf{A}^2\mathbf{B}\mathbf{C}^2 &= \mathbf{A}\mathbf{G}\left(\mathbf{C}, \mathbf{A}\mathbf{B}\right) \\
&\quad - \underbrace{\left[\mathbf{A}\mathbf{C}\mathbf{A}\right]\mathbf{B}\mathbf{C}}_{=\left[\mathbf{G}\left(\mathbf{A},\mathbf{C}\right) - \mathbf{A}^2\mathbf{C} - \mathbf{C}\mathbf{A}^2\right]\mathbf{B}\mathbf{C}} - \underbrace{\left[\mathbf{A}\mathbf{C}^2\mathbf{A}\right]\mathbf{B}}_{=\left[\mathbf{G}\left(\mathbf{A},\mathbf{C}^2\right) - \mathbf{A}^2\mathbf{C}^2 - \mathbf{C}^2\mathbf{A}^2\right]\mathbf{B}} \\
&= \mathbf{A}\mathbf{G}\left(\mathbf{C}, \mathbf{A}\mathbf{B}\right) - \mathbf{G}\left(\mathbf{A}, \mathbf{C}\right)\mathbf{B}\mathbf{C} - \mathbf{G}\left(\mathbf{A}, \mathbf{C}^2\right)\mathbf{B} \\
&\quad + \underbrace{\mathbf{A}^2\left[\mathbf{C}\mathbf{B}\mathbf{C}\right]}_{=\mathbf{A}^2\left[\mathbf{G}\left(\mathbf{C},\mathbf{B}\right) - \cancel{\mathbf{C}^2\mathbf{B}} - \mathbf{B}\mathbf{C}^2\right]} + \underbrace{\mathbf{C}\left[\mathbf{A}^2\mathbf{B}\right]\mathbf{C}}_{=\left[\mathbf{G}\left(\mathbf{C},\mathbf{A}^2\mathbf{B}\right) - \cancel{\mathbf{C}^2\mathbf{A}^2\mathbf{B}} - \mathbf{A}^2\mathbf{B}\mathbf{C}^2\right]} \\
&\quad + \cancel{\mathbf{A}^2\mathbf{C}^2\mathbf{B}} + \cancel{\mathbf{C}^2\mathbf{A}^2\mathbf{B}} \\
&= -2\mathbf{A}^2\mathbf{B}\mathbf{C}^2 + \mathbf{A}\mathbf{G}\left(\mathbf{C}, \mathbf{A}\mathbf{B}\right) - \mathbf{G}\left(\mathbf{A}, \mathbf{C}\right)\mathbf{B}\mathbf{C} \\
&\quad - \mathbf{G}\left(\mathbf{A}, \mathbf{C}^2\right)\mathbf{B} + \mathbf{A}^2\mathbf{G}\left(\mathbf{C}, \mathbf{B}\right) + \mathbf{G}\left(\mathbf{C}, \mathbf{A}^2\mathbf{B}\right) .
\end{aligned}
\tag{6.244}
$$

Considering (6.226), the product $3\mathbf{A}^2\mathbf{B}\mathbf{C}^2$ can thus be expressed as a polynomial of degree 4. And this immediately implies that

$$\mathbf{A}^2\mathbf{B}\mathbf{A}^2 \equiv 0 . \tag{6.245}$$

As a result, one can conclude that $\mathbf{S}^2\mathbf{v} \otimes \mathbf{S}^2\mathbf{v}$ should not be considered in the generating set of an isotropic symmetric tensor invariant of $\mathbf{v}$ and $\mathbf{S} = \mathbf{S}^{\mathrm{T}}$.
From the result $\mathbf{A}^2\mathbf{B}\mathbf{C}^2 \equiv 0$ and Lemma D, one can write

$$\mathrm{tr}\left(\mathbf{A}^2\mathbf{B}\mathbf{C}^2\mathbf{D}\right) \equiv 0 . \tag{6.246}$$

Let $\mathbf{A} = \mathbf{S}$, $\mathbf{B} = \mathbf{T}^2$, $\mathbf{C} = \mathbf{U}^2$ and $\mathbf{D} = \mathbf{S}$ be symmetric tensors. From (6.237), it then follows that

$$
\begin{aligned}
&\underbrace{\mathrm{tr}\left(\mathbf{S}^2\mathbf{T}^2\mathbf{U}^2\right)}_{=\mathbf{I}:\mathbf{S}^2\mathbf{T}^2\mathbf{U}^2 = \mathbf{S}^2:\mathbf{T}^2\mathbf{U}^2} + \underbrace{\mathrm{tr}\left(\mathbf{S}^2\mathbf{U}^2\mathbf{T}^2\right)}_{=\mathbf{I}:\mathbf{S}^2\mathbf{U}^2\mathbf{T}^2 = \mathbf{S}^2:\mathbf{U}^2\mathbf{T}^2 = \mathbf{S}^2:\mathbf{T}^2\mathbf{U}^2} \\
&+ \underbrace{\mathrm{tr}\left(\mathbf{S}\mathbf{T}^2\mathbf{U}^2\mathbf{S}\right)}_{=\mathbf{I}:\mathbf{S}\mathbf{T}^2\mathbf{U}^2\mathbf{S} = \mathbf{S}:\mathbf{T}^2\mathbf{U}^2\mathbf{S} = \mathbf{S}^2:\mathbf{T}^2\mathbf{U}^2} + \underbrace{\mathrm{tr}\left(\mathbf{S}\mathbf{T}^2\mathbf{S}\mathbf{U}^2\right)}_{\text{this term is reducible}} \\
&+ \underbrace{\mathrm{tr}\left(\mathbf{S}\mathbf{U}^2\mathbf{S}\mathbf{T}^2\right)}_{\text{this quantity is also reducible}} + \underbrace{\mathrm{tr}\left(\mathbf{S}\mathbf{U}^2\mathbf{T}^2\mathbf{S}\right)}_{=\mathbf{I}:\mathbf{S}\mathbf{U}^2\mathbf{T}^2\mathbf{S} = \mathbf{S}^2:\mathbf{U}^2\mathbf{T}^2 = \mathbf{S}^2:\mathbf{T}^2\mathbf{U}^2} \equiv 0 .
\end{aligned}
$$

And this implies the desired result

$$\mathrm{tr}\left(\mathbf{S}^2\mathbf{T}^2\mathbf{U}^2\right) \equiv 0 . \tag{6.247}$$

3. Show that tr (**ABCDEFG**) is reducible, i.e. no product has a total degree greater than **six**.

Solution. From (6.246), one will have

$$\mathrm{tr}\left(\mathbf{A}^2\mathbf{B}\mathbf{C}^2\mathbf{G}\right) \equiv 0 \quad \text{or} \quad \mathrm{tr}\left(\mathbf{A}^2\mathbf{C}\mathbf{E}^2\mathbf{G}\right) \equiv 0 \,. \tag{6.248}$$

Replacing $\mathbf{E}$ in $(6.248)_2$ by $\mathbf{E}+\mathbf{F}$ leads to

$$\mathrm{tr}\left(\mathbf{A}^2\mathbf{CEFG} + \mathbf{A}^2\mathbf{CFEG}\right) \equiv 0 \,. \tag{6.249}$$

One then has

$$\begin{aligned}
&\underbrace{\mathrm{tr}\left(\mathbf{A}^2\mathbf{CDEFG} + \cancel{\mathbf{A}^2\mathbf{CDFEG}}\right)}_{\text{if } \mathbf{C} \text{ is replaced by } \mathbf{CD} \text{ in (6.249)}} \\
&+ \underbrace{\mathrm{tr}\left(\mathbf{A}^2\mathbf{CDEFG} + \cancel{\mathbf{A}^2\mathbf{CFDEG}}\right)}_{\text{by setting } \mathbf{E} = \mathbf{DE} \text{ in (6.249)}} \\
&- \underbrace{\mathrm{tr}\left(\cancel{\mathbf{A}^2\mathbf{CDFEG}} + \cancel{\mathbf{A}^2\mathbf{CFDEG}}\right)}_{\text{by choosing } \mathbf{E} = -\mathbf{D} \text{ and, subsequently, setting } \mathbf{G} = \mathbf{EG} \text{ in (6.249)}} \equiv 0 \,,
\end{aligned}$$

or

$$\mathrm{tr}\left(\mathbf{A}^2\mathbf{CDEFG}\right) \equiv 0 \,. \tag{6.250}$$

It is then easy to see that

$$\begin{aligned}
\mathrm{tr}\,(\mathbf{ABCDEFG}) &\equiv -\mathrm{tr}\,(\mathbf{BACDEFG}) \\
&\equiv +\mathrm{tr}\,(\mathbf{BCADEFG}) \\
&\equiv -\mathrm{tr}\,(\mathbf{BCDAEFG}) \,.
\end{aligned} \tag{6.251}$$

This result states that the trace of the degree 7 product $\boldsymbol{\Pi} = \mathbf{ABCDEFG}$ is equivalent to (i) the negative of the trace of an odd permutation of the adjacent factors in $\boldsymbol{\Pi}$ and (ii) the trace of any product obtained by an even permutation of the adjacent factors.

In (6.237), replace $\mathbf{D}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ by $\mathbf{G}$, $\mathbf{AB}$, $\mathbf{CD}$, $\mathbf{EF}$, respectively. Using $(2.87)_1$ and $(6.251)_{1-3}$, one then obtains

$$\begin{aligned}
&\mathrm{tr}\,(\mathbf{ABCDEFG}) + \underbrace{\mathrm{tr}\,(\mathbf{ABEFCDG})}_{\equiv\, \mathrm{tr}\,(\mathbf{ABCDEFG})} + \underbrace{\mathrm{tr}\,(\mathbf{CDEFABG})}_{\equiv\, \mathrm{tr}\,(\mathbf{CDABEFG})\, \equiv\, \mathrm{tr}\,(\mathbf{ABCDEFG})} \\
&+ \underbrace{\mathrm{tr}\,(\mathbf{CDABEFG})}_{\equiv\, \mathrm{tr}\,(\mathbf{ABCDEFG})} + \underbrace{\mathrm{tr}\,(\mathbf{EFABCDG})}_{\equiv\, \mathrm{tr}\,(\mathbf{ABEFCDG})\, \equiv\, \mathrm{tr}\,(\mathbf{ABCDEFG})} \\
&+ \underbrace{\mathrm{tr}\,(\mathbf{EFCDABG})}_{\equiv\, \mathrm{tr}\,(\mathbf{EFABCDG})\, \equiv\, \mathrm{tr}\,(\mathbf{ABEFCDG})\, \equiv\, \mathrm{tr}\,(\mathbf{ABCDEFG})} \equiv 0 \,,
\end{aligned}$$

or

$$\mathrm{tr}\,(\mathbf{ABCDEFG}) \equiv 0 \,. \tag{6.252}$$

This exercise presented an important application of the Rivlin's identities in the representation theorems. It was explicitly shown that how these tensor identities help recognize some redundant elements. If one is willing to generate the invariant and generator elements intrinsically, infinitely many of them can be constructed for a given set of variables. Indeed, these powerful tools impose some restrictions and, therefore, one is left with a finite set of (scalar or tensor) invariants. However, in general, there is no guarantee to find an irreducible set of invariants for a given set of variables, if these identities are only the main concerns. Hence, having in mind the outcomes of these identities, a rigorous procedure should be developed to establish the complete and irreducible sets of invariants as tabulated in the well-known tables. This is briefly discussed in the following exercise.

Exercise 6.15

Verify (6.141), (6.143) and (6.146).

Solution. The procedure used here to establish the desired representations follows from the pioneering work of Smith [14] which aims at finding a **functional basis** for a given system of tensorial variables. The interested reader should also consult Smith [34] and Wang [35]. In this context, the goal is to determine a set of basic isotropic invariants according to (6.95) such that, once the orientation and sense of the coordinate axes are specified, the following equations

$$I_k\left(\overline{\mathbf{v}}, \overline{\mathbf{S}}, \overline{\mathbf{W}}\right) = \phi_k \ , \quad k = 1, 2, \ldots, n \ , \tag{6.253}$$

have a **unique** solution for the components of $\mathbf{v}_1, \ldots, \mathbf{W}_{m_4}$ where ϕ_k denotes the prescribed value of I_k. Thus, by having the values of the components of $\mathbf{v}_1, \ldots, \mathbf{W}_{m_4}$ in a suitably chosen reference frame, one will be able to represent any isotropic invariant $\alpha\left(\overline{\mathbf{v}}, \overline{\mathbf{S}}, \overline{\mathbf{W}}\right)$ as a **single-valued function** of the basic scalar invariants $I_1, \ldots I_n$. To this end, consider the Cartesian coordinates for convenience.

To verify (6.141), consider the domain

$$\mathscr{D} = \left\{(\mathbf{S}_1 = \mathbf{S}, \mathbf{W}_1 = \mathbf{W}) \in \mathscr{T}_{\text{so}}^{\text{sym}} \times \mathscr{T}_{\text{so}}^{\text{skw}}\right\} \ .$$

Let $\boldsymbol{\omega}$ be the axial vector of $\mathbf{W}$. In matrix notation, they render

$$[\mathbf{W}] = \begin{bmatrix} 0 & \hat{W}_{12} & \hat{W}_{13} \\ -\hat{W}_{12} & 0 & \hat{W}_{23} \\ -\hat{W}_{13} & -\hat{W}_{23} & 0 \end{bmatrix} \ , \quad [\boldsymbol{\omega}] = \begin{bmatrix} -\hat{W}_{23} \\ \hat{W}_{13} \\ -\hat{W}_{12} \end{bmatrix} \ .$$

To begin with, one can orient the 1 axis in such a way that it lies along the direction of $\boldsymbol{\omega}$. Consequently,

$$[\mathbf{W}] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & W_{23} \\ 0 & -W_{23} & 0 \end{bmatrix} \; , \; [\boldsymbol{\omega}] = \begin{bmatrix} -W_{23} \\ 0 \\ 0 \end{bmatrix} \; , \; \left[\mathbf{W}^2\right] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -W_{23}^2 & 0 \\ 0 & 0 & -W_{23}^2 \end{bmatrix} \; .$$

This helps determine, up to a sign, the value of W_{23} from the equation $\operatorname{tr}\mathbf{W}^2 = -2W_{23}^2$. One can choose the sense of the 2 axis in such a way that W_{23} is always positive and, therefore, its sign is fixed. Now, the goal is to determine the six independent components of the symmetric tensor $\mathbf{S}$. By having the values of W_{23} and

$$\operatorname{tr}\left(\mathbf{S}\mathbf{W}^2\right) \; , \; \operatorname{tr}\mathbf{S} \; , \; \operatorname{tr}\mathbf{S}^2 \; , \; \operatorname{tr}\left(\mathbf{S}^2\mathbf{W}^2\right) \; , \; \operatorname{tr}\left(\mathbf{S}\mathbf{W}\mathbf{S}^2\mathbf{W}^2\right) \; , \; \operatorname{tr}\mathbf{S}^3 \; , \tag{6.254}$$

one can obtain the values of

$$\begin{aligned} & S_{22} + S_{33} \; , \; S_{11} \; , \; S_{12}^2 + S_{13}^2 \; , \; S_{22}^2 + S_{33}^2 + 2S_{23}^2 \; , \\ & \left(S_{12}^2 - S_{13}^2\right) S_{23} - \left(S_{22} - S_{33}\right) S_{12} S_{13} \; , \\ & \left(S_{22} - S_{33}\right)\left(S_{12}^2 - S_{13}^2\right) + 4 S_{12} S_{13} S_{23} \; . \end{aligned} \tag{6.255}$$

To proceed, consider the lower right hand 2 by 2 matrix of $\mathbf{S}$, that is,

$$\begin{bmatrix} S_{22} & S_{23} \\ S_{23} & S_{33} \end{bmatrix} \; , \tag{6.256}$$

whose characteristic equation is given by

$$\lambda^2 - \left(S_{22} + S_{33}\right)\lambda + \frac{1}{2}\left[\left(S_{22} + S_{33}\right)^2 - \left(S_{22}^2 + S_{33}^2 + 2S_{23}^2\right)\right] = 0 \; . \tag{6.257}$$

With the known values of $S_{22} + S_{33}$ and $S_{22}^2 + S_{33}^2 + 2S_{23}^2$, one can solve (6.257) to provide the principal values of (6.256). Suppose that $\mathbf{S}$ is a tensor with non-multiple eigenvalues and $S_{12}^2 + S_{13}^2 \neq 0$. The reference frame can now be rotated about the 1 axis until the 2 and 3 axes coincide with the principal directions of the matrix (6.256). In this frame, $S_{23} = 0$ and the values of S_{22} and S_{33} are known (having in mind that $S_{22} \neq S_{33}$ by assumption). Consequently, using the known values of the quantities in (6.255), the values of S_{12}^2, S_{13}^2 and $S_{12}S_{13}$ are known. Suppose that $S_{12} \neq 0$. Then, one can choose the sense of the 1 axis such that S_{12} is always positive. Thus, S_{12}^2 determines S_{12} and subsequently S_{13} can be obtained from $S_{12}S_{13}$. When $S_{12} = 0$, the sense of the 1 axis can be chosen so that $S_{13} > 0$. And it is thus determined from S_{13}^2.

Consider the case in which $S_{12}^2 + S_{13}^2 = 0$ (or $S_{12} = S_{13} = 0$) and the 2 by 2 matrix (6.256) has two distinct eigenvalues. In this case, one can still rotate the reference frame about the 1 axis until the 2 and 3 axes lie along the principal directions of such a matrix. In this frame, the off-diagonal element S_{23} vanishes and the values of the diagonal elements S_{22} and S_{33} are determined from the known values of $S_{22} + S_{33}$ and $\left(S_{22} + S_{33}\right)^2 - 2S_{22}S_{33}$.

Consider now the case in which the 2 by 2 matrix (6.256) has a repeated eigenvalue. In this case, $S_{23} = 0$ and the value of $S_{22} = S_{33}$ is immediately determined from $S_{22} + S_{33}$ having in mind that S_{11} is known. Further, suppose that $S_{12}^2 + S_{13}^2 \neq 0$. Then, the reference frame can be rotated about the 1 axis until $S_{13} = 0$. Consequently, S_{12} is determined up to a sign. The sense of the 1 axis can accordingly be chosen such that $S_{12} > 0$. Finally, suppose that $S_{12}^2 + S_{13}^2 = 0$. Then, $[\mathbf{S}]$ will be a diagonal matrix with known entries.

It is important to point out that the procedure outlined above should be used when there exists a single skew tensor and all of the vectors are zero in the given domain of tensorial variables. One can thus involve more symmetric tensors and continue to determine their components by considering more appropriate basic scalar invariants.

Next, attention is focused on finding the basic isotropic invariants of the two symmetric tensors $\mathbf{S}$ and $\mathbf{T}$. Note that the procedure employed in the following to verify (6.143) is only applicable when all of the vectors and skew tensors are zero in the given list of tensor arguments. Let at least one of these tensors, say $\mathbf{S}$, possess three simple eigenvalues. Given the values of $\operatorname{tr}\mathbf{S}$, $\operatorname{tr}\mathbf{S}^2$ and $\operatorname{tr}\mathbf{S}^3$, one can solve the characteristic equation

$$\lambda^3 - (\operatorname{tr}\mathbf{S})\,\lambda^2 + \frac{(\operatorname{tr}\mathbf{S})^2 - \operatorname{tr}\mathbf{S}^2}{2}\lambda - \frac{2\operatorname{tr}\mathbf{S}^3 - 3\operatorname{tr}\mathbf{S}\operatorname{tr}\mathbf{S}^2 + (\operatorname{tr}\mathbf{S})^3}{6} = 0\,, \tag{6.258}$$

to obtain the three distinct principal values $\lambda_1 := S_{11}$, $\lambda_2 := S_{22}$ and $\lambda_3 := S_{33}$ where $S_{11} > S_{22} > S_{33}$. The reference frame can then be oriented such that the axes lie along the principal directions of $\mathbf{S}$. In this new frame, the matrix form of $\mathbf{S}$ renders

$$[\mathbf{S}] = \begin{bmatrix} S_{11} & 0 & 0 \\ 0 & S_{22} & 0 \\ 0 & 0 & S_{33} \end{bmatrix} . \tag{6.259}$$

Now, the goal is to determine the six independent components of $\mathbf{T}$ with respect to the orthonormal basis constructed from the eigenvectors of $\mathbf{S}$. Given the values of $\operatorname{tr}\mathbf{T}$, $\operatorname{tr}(\mathbf{ST})$ and $\operatorname{tr}\left(\mathbf{S}^2\mathbf{T}\right)$, one can obtain the values of the diagonal elements T_{11}, T_{22} and T_{33}. Subsequently, with the aid of these values and $\operatorname{tr}\mathbf{T}^2$, $\operatorname{tr}\left(\mathbf{ST}^2\right)$ and $\operatorname{tr}\left(\mathbf{S}^2\mathbf{T}^2\right)$, one can calculate T_{12}^2, T_{13}^2 and T_{23}^2. These values along with $\operatorname{tr}\mathbf{T}^3$ help compute $T_{12}T_{13}T_{23}$. To uniquely determine the values of off-diagonal elements, the following cases need to be considered:

☞ Suppose $\mathbf{T}$ is a symmetric tensor whose off-diagonal elements are all nonzero. In this case, the sense of the 2 and 3 axes can be chosen such that T_{12} and T_{13} are positive. Then, the known values of T_{12}^2 and T_{13}^2 respectively determine T_{12} and T_{13}. Consequently, T_{23} is obtained from $T_{12}T_{13}T_{23}$.

☞ Suppose $\mathbf{T}$ is a symmetric tensor with only one zero off-diagonal element, say $T_{23} = 0$. In the present case, one can also choose the sense of the 2 and 3 axes so that $T_{12} > 0$ and $T_{13} > 0$. The positive components T_{12} and T_{13} are then obtained from the known values of T_{12}^2 and T_{13}^2, respectively.

☞☛ Suppose **T** is a symmetric tensor with only one nonzero off-diagonal element, say $T_{12} \neq 0$. The positiveness of T_{12} is then guaranteed by appropriately choosing the sense of the 2 axis. As a result, the value of T_{12} is determined from the known value of T_{12}^2.

The case in which both **S** and **T** have repeated eigenvalues is left as an exercise. One should finally note that the problem can consistently be continued by involving more symmetric tensors.

Finally, attention is focused on characterizing the basic isotropic invariants of the two vectors **u** and **v**. It is important to note that the procedure established here to verify (6.146) is only applicable when all of the vectors in the given domain of tensorial variables are **coplanar** but there exist two non-collinear vectors among them. Suppose **u** and **v** are these two vectors satisfying

$$\det \begin{bmatrix} \mathbf{u}\cdot\mathbf{u} & \mathbf{u}\cdot\mathbf{v} \\ \mathbf{u}\cdot\mathbf{v} & \mathbf{v}\cdot\mathbf{v} \end{bmatrix} \neq 0 \; . \quad \leftarrow \text{this guaranties that } \mathbf{u} \text{ and } \mathbf{v} \text{ are } \textbf{non-collinear} \tag{6.260}$$

Then, one can choose the orientation and sense of the 1 and 2 axes so that

$$[\mathbf{u}] = \begin{bmatrix} u_1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad [\mathbf{v}] = \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} \; , \tag{6.261}$$

where $u_1 > 0$ and $v_2 > 0$. The known value of $\mathbf{u} \cdot \mathbf{u}$ helps determine the value of u_1 and, consequently, $\mathbf{u} \cdot \mathbf{v}$ yields v_1. It is then easy to evaluate v_2 from $\mathbf{v} \cdot \mathbf{v}$. Thus, any isotropic scalar-valued function of **u** and **v** should be constructed from the three quantities $\mathbf{u} \cdot \mathbf{u}$, $\mathbf{v} \cdot \mathbf{v}$ and $\mathbf{u} \cdot \mathbf{v}$.

The problem can be continued by involving more tensorial variables. For instance, consider an extra skew tensor **W** having in mind that all components in (6.261) are known. The goal is thus to determine the three unknown components W_{12}, W_{13} and W_{23}. To begin with, consider the known value of $\mathbf{u} \cdot \mathbf{W}\mathbf{v}$ which helps obtain the value of W_{12}. Then, the known values of u_1, v_1, v_2, W_{12} and

$$\mathbf{u} \cdot \mathbf{W}^2\mathbf{u} \, , \; \mathbf{u} \cdot \mathbf{W}^2\mathbf{v} \, , \; \mathbf{v} \cdot \mathbf{W}^2\mathbf{v} \; ,$$

help compute the quantities

$$W_{13}^2 \, , \; W_{13}W_{23} \, , \; W_{23}^2 \; .$$

Suppose that $W_{13} \neq 0$. One can now choose the sense of the 3 axis so that $W_{13} > 0$ and thus determined. Consequently, $W_{13}W_{23}$ gives W_{23}. The case in which $W_{13} = 0$ but $W_{23} \neq 0$ can be treated in a similar manner. One is now in a position to generally express any isotropic scalar function of $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{W}$ as a single-valued function of the 8 isotropic invariants:

$$\alpha\left(\mathbf{u},\mathbf{v},\mathbf{W}\right)=\bar{\alpha}\left(\mathbf{u}\cdot\mathbf{u},\mathbf{v}\cdot\mathbf{v},\mathbf{u}\cdot\mathbf{v},\operatorname{tr}\mathbf{W}^2,\right.$$
$$\left.\mathbf{u}\cdot\mathbf{W}\mathbf{v},\mathbf{u}\cdot\mathbf{W}^2\mathbf{u},\mathbf{u}\cdot\mathbf{W}^2\mathbf{v},\mathbf{v}\cdot\mathbf{W}^2\mathbf{v}\right) . \tag{6.262}$$

Let

$$\mathscr{D}=\left\{(\mathbf{S}_1=\mathbf{u}\otimes\mathbf{u},\mathbf{S}_2=\mathbf{v}\otimes\mathbf{v},\mathbf{W}_1=\mathbf{W})\in\mathscr{T}_{\text{so}}^{\text{sym}}\times\mathscr{T}_{\text{so}}^{\text{sym}}\times\mathscr{T}_{\text{so}}^{\text{skw}}\right\} . \tag{6.263}$$

Then, by reading off from Table 6.1 and subsequently removing the resulting redundant terms, one can arrive at the representation (6.262) one more time. And this basically shows the consistency between the elements of the well-established tables widely used within the context of representation theorems.

Exercise 6.16

The goal of this exercise is to introduce some important strain and stress measures and characterize their functional relationships within the context of *nonlinear solid mechanics*. For a detailed account on these concepts, the reader is referred to the classical treatises of Truesdell and Toupin [36], Eringen [37], Truesdell and Noll [38], Eringen [39], Truesdell [40], Marsden and Hughes [41], Ogden [42], Chadwick [43] and more recent expositions of Holzapfel [44], Ibrahimbegović [45] and Steinmann [46]. For a detailed account on **finite element methods** solving nonlinear problems of continua, the interested reader is referred to Oden [47], Zienkiewicz and Taylor [48], Bathe [49], Crisfield [50], Bonet and Wood [51], Belytschko et al. [52] and Wriggers [53] among many others.

Solid mechanics constitutes a major part of mechanical, civil, aerospace, biomedical and nuclear engineering that includes theory of elasticity - with linear and nonlinear classifications - as well as inelasticity which mainly includes plasticity, viscoelasticity and damage. It is a branch of continuum mechanics that aims at describing the macroscopic behavior of solids under the action of mechanical, thermal, chemical, electrical, magnetic and tribological loads.

Three main features, i.e. experiment, theory and computation, are usually required to solve a real solid mechanics problem. Therein, experimental data demonstrates the real behavior of material and is served as a basis to show the correctness of the outputs of other parts, the theoretical side presents the basic mathematical structure that aims at modeling the physics of the problem and the computational part tries to provide a proper solution to differential equations that naturally arise in the theoretical aspect of the problem. The set of equations developed within the theoretical side are represented in the language of tensor analysis. And this greatly shows the dominant role of tensor calculus as the mathematical underpinning of solid mechanics.

In nonlinear continuum mechanics of solids, one deals with (at least) two configurations. The so-called *reference configuration* which is usually chosen to be the *initial* (or *undeformed*) *configuration* without any loading scenario and *spatial* (or *current* or *deformed*) *configuration* where the material has undergone some deformation process in response to the applied loads.

Suppose that all tensor quantities here are expressed with respect to the Cartesian basis vectors. Let $\mathbf{C}$ and $\mathbf{b}$ be two symmetric positive definite tensors; defined by,

$$\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} \quad , \quad \mathbf{b} = \mathbf{F}\mathbf{F}^{\mathrm{T}} \ , \tag{6.264}$$

where $\mathbf{F}$ presents an invertible tensor, i.e. $\det \mathbf{F} \neq 0$, which is generally unsymmetric. The so-called *polar decomposition* for $\mathbf{F}$ reads

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{v}\mathbf{R} \ , \tag{6.265}$$

where $\mathbf{R}$ is a (rigid) rotation tensor and $\mathbf{U}$, $\mathbf{v}$ also represent two symmetric positive definite tensors. Within the context of nonlinear continuum mechanics, the tensors $\mathbf{C}$, $\mathbf{b}$, $\mathbf{F}$, $\mathbf{U}$ and $\mathbf{v}$ are called the *right Cauchy-Green strain tensor*, *left Cauchy-Green strain* (or *Finger*) *tensor*, *deformation gradient*, *right stretch tensor* and *left stretch tensor*, respectively. These strain measures are introduced in order to characterize the finite deformation behavior of solids. See Fig. 6.1 for a geometrical interpretation.

For the sake of consistency, it is widely agreed to use uppercase letters for tensor quantities acting in the reference configuration. And the lowercase letters are utilized when the tensor quantities are computed in the current configuration. This convention also applies to the corresponding indices. For instance, $\mathbf{C}$ is a *referential* tensor with C_{AB} and $\mathbf{b}$ presents a *spatial* strain measure with b_{ij}. In this regard, $\mathbf{F}$ with F_{iA} is called a *two-point tensor* because it interacts between the basic configurations.

The relations (6.264) and (6.265) imply that

$$\mathbf{C} = \mathbf{U}^2 \quad , \quad \mathbf{b} = \mathbf{v}^2 \ . \tag{6.266}$$

And

$$\mathbf{R} = R_{iA}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{E}}_A = \sum_{a=1}^{3} \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_a \ , \tag{6.267a}$$

$$\mathbf{U} = U_{AB}\widehat{\mathbf{E}}_A \otimes \widehat{\mathbf{E}}_B = \sum_{a=1}^{3} \lambda_a \, \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \ , \tag{6.267b}$$

$$\mathbf{v} = v_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \sum_{a=1}^{3} \lambda_a \, \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \ , \tag{6.267c}$$

help provide

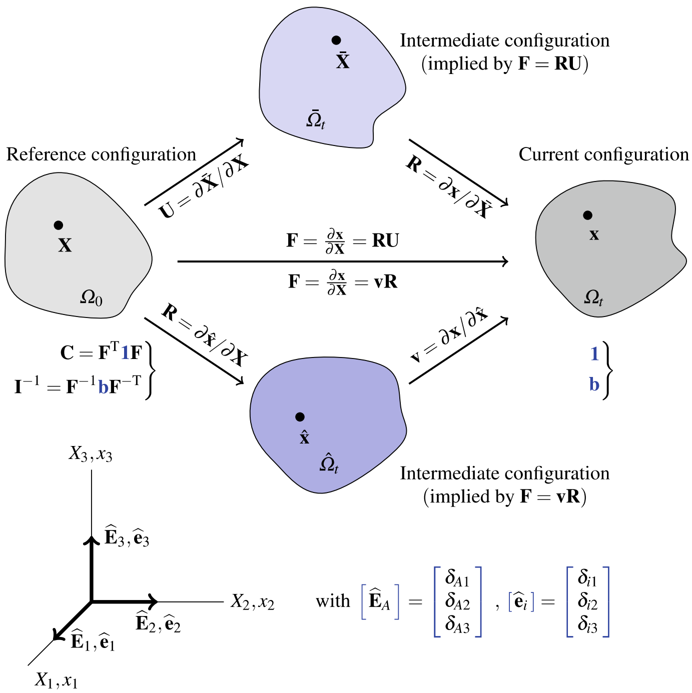


Fig. 6.1 Reference, two intermediate and current configurations

$$\mathbf{F} = F_{iA}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{E}}_A = \sum_{a=1}^{3} \lambda_a\, \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_a \ , \tag{6.268a}$$

$$\mathbf{C} = C_{AB}\widehat{\mathbf{E}}_A \otimes \widehat{\mathbf{E}}_B = \sum_{a=1}^{3} \lambda_a^2\, \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \ , \tag{6.268b}$$

$$\mathbf{b} = b_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j = \sum_{a=1}^{3} \lambda_a^2\, \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \ , \tag{6.268c}$$

where $\lambda_a > 0,\ a = 1, 2, 3$, represent the *principal stretches*, the set $\left\{\widehat{\mathbf{N}}_a\right\}$ indicates the *principal referential* (or *material* or *Lagrangian*) *directions*, the set $\{\widehat{\mathbf{n}}_a\}$ presents the *principal spatial* (or *Eulerian*) *directions* and the set $\left\{\widehat{\mathbf{E}}_A\right\}$ ($\{\widehat{\mathbf{e}}_i\}$) stands for the material (spatial) standard basis (note that these Cartesian bases are the same from the computational point of view, see (1.41)). For consistency, the identity tensors in the material and spatial descriptions are denoted by

$$\mathbf{I} = \delta_{AB}\widehat{\mathbf{E}}_A \otimes \widehat{\mathbf{E}}_B \quad , \quad \mathbf{1} = \delta_{ij}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ . \tag{6.269}$$

Within the context of *hyperelasticity*, which is widely used in nonlinear solid mechanics, the existence of a scalar-valued function of the deformation gradient is postulated:

$$\Psi = \Psi_{\mathbf{F}}(\mathbf{F}) \ . \tag{6.270}$$

This is called the *stored-energy function* (or *strain-energy function*). It is a special function which obeys the

- ☞☞ *normalization condition* $\Psi_{\mathbf{F}}(\mathbf{I}) = 0$ which means that no energy is saved within the material when there is no loading or when the material is unloaded to its initial placement in the reference configuration,
- ☞☞ physical observation that the energy stored in the material increases with deformation, i.e. $\Psi_{\mathbf{F}}(\mathbf{F}) \geq 0$,
- ☞☞ *growth conditions* $\Psi_{\mathbf{F}}(\mathbf{F}) \to +\infty$ when either $\det \mathbf{F} \to 0^+$ or $\det \mathbf{F} \to +\infty$ which physically means the material cannot sustain the loading which causes its compression to a point (with vanishing volume) or its expansion to the infinite range, and
- ☞☞ condition of *polyconvexity* to ensure the global existence of solutions, see Ball [54] and Šilhavý [55].

Moreover, the strain-energy function needs to be *objective* which leads to its some alternative forms. The *principle of material objectivity* states that the response of material should be independent of the observer. As a result, Ψ must remain invariant under any (proper) orthogonal transformation of the current configuration, i.e.

$$\Psi(\mathbf{F}) = \Psi(\mathbf{QF}) \ . \tag{6.271}$$

By setting $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$ and considering the polar decomposition $\mathbf{F} = \mathbf{RU}$, the energy function (6.270) can now be represented in the canonical form

$$\Psi = \Psi_{\mathbf{F}}(\mathbf{F}) = \Psi_{\mathbf{U}}(\mathbf{U}) \ . \tag{6.272}$$

Consequently, using $\mathbf{U} = \sqrt{\mathbf{C}}$,

$$\Psi = \Psi_{\mathbf{F}}(\mathbf{F}) = \Psi_{\mathbf{U}}(\mathbf{U}) = \Psi_{\mathbf{C}}(\mathbf{C}) \ . \tag{6.273}$$

By isotropy assumption, i.e. when there is no preferred direction within the material, the strain-energy function is further restricted to

$$\Psi(\mathbf{F}) = \Psi\left(\mathbf{FQ}^{\mathrm{T}}\right) \ . \tag{6.274}$$

This guarantees that the energy function remains unchanged under any (proper) orthogonal transformation of the reference configuration. Thus, the material response will be similar in all directions. Note that the objectivity requirement is an axiom which should always be satisfied but the isotropy condition is only valid for some specific materials called *isotropic materials*. As a consequence of isotropy,

$$\Psi\left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right) = \Psi\left(\mathbf{Q}\mathbf{F}^{\mathrm{T}}\mathbf{F}\mathbf{Q}^{\mathrm{T}}\right) \xlongequal[\text{and using } \mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathbf{U}^2]{\text{by setting } \mathbf{Q} = \mathbf{R}} \Psi\left(\mathbf{R}\mathbf{U}\mathbf{U}\mathbf{R}^{\mathrm{T}}\right) = \Psi\left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right) .$$

Thus, in the context of *isotropic hyperelasticity*, the strain-energy function admits the equivalent forms

$$\begin{aligned} \Psi &\xlongequal[\text{hyperelasticity}]{\text{by}} \Psi_{\mathbf{F}}(\mathbf{F}) \\ &\xlongequal[\text{objectivity}]{\text{by}} \Psi_{\mathbf{U}}(\mathbf{U}) \\ &\xlongequal[\mathbf{U} = \sqrt{\mathbf{C}}]{\text{by}} \Psi_{\mathbf{C}}(\mathbf{C}) \\ &\xlongequal[\text{isotropy}]{\text{by}} \Psi_{\mathbf{b}}(\mathbf{b}) \\ &\xlongequal[\mathbf{b} = \mathbf{v}^2]{\text{by}} \Psi_{\mathbf{v}}(\mathbf{v}) . \end{aligned} \tag{6.275}$$

Moreover, by (6.132) and $(6.139)_1$, it also admits the following representations

$$\Psi = \Psi_I (I_1, I_2, I_3) = \Psi_\lambda (\lambda_1, \lambda_2, \lambda_3) , \tag{6.276}$$

where $I_i,\ i = 1, 2, 3$, are the principal scalar invariants of either $\mathbf{C}$ or $\mathbf{b}$ owing to

$$I_1(\mathbf{C}) = I_1(\mathbf{b}) \ , \ I_2(\mathbf{C}) = I_2(\mathbf{b}) \ , \ I_3(\mathbf{C}) = I_3(\mathbf{b}) \ . \quad \leftarrow \text{see (4.17a)–(4.17c)} \tag{6.277}$$

To this end, consider an isotropic hyperelastic material. The sensitivity of the energy function with respect to the strain tensors will provide the stress measures. This is demonstrated in the following. ✆

First, consider the **two-point** generally unsymmetric tensor

$$\mathbf{P} = \frac{\partial \Psi_{\mathbf{F}}}{\partial \mathbf{F}} , \tag{6.278}$$

which is known as the *first Piola-Kirchhoff stress tensor*.

The so-called *Biot stress tensor*, which renders a symmetric **referential** tensor, is then given by

$$\mathbf{T}_{\mathrm{B}} = \frac{\partial \Psi_{\mathbf{U}}}{\partial \mathbf{U}} . \tag{6.279}$$

Note that, in general, the Biot stress tensor is not symmetric. Here, $\mathbf{T}_{\mathrm{B}} = \mathbf{T}_{\mathrm{B}}^{\mathrm{T}}$ is implied by the isotropy assumption $\Psi_{\mathbf{C}}(\mathbf{C}) = \Psi_I(I_1, I_2, I_3)$, see (6.291b). Thus, the above relation should generally be written as $\mathrm{sym}\mathbf{T}_{\mathrm{B}} = \partial \Psi_{\mathbf{U}} / \partial \mathbf{U}$.

Next, the very popular *second Piola-Kirchhoff stress tensor*, which also represents a symmetric **material** tensor, is defined by

$$\mathbf{S} = 2\frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} \ . \tag{6.280}$$

Another important stress measure which can be introduced at this point is the symmetric **spatial** *Cauchy* (or *true*) *stress tensor*. It is introduced through the following expression

$$\boldsymbol{\sigma} = 2J^{-1}\frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}}\mathbf{b} = 2J^{-1}\mathbf{b}\frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \ , \tag{6.281}$$

where

$$J = \det \mathbf{F} \ , \tag{6.282}$$

presents the *volumetric Jacobian*. The above specific form for $\boldsymbol{\sigma}$ is only valid for isotropic hyperelastic materials because, in general, $\Psi_{\mathbf{C}}(\mathbf{C}) \neq \Psi_{\mathbf{b}}(\mathbf{b})$.

The last stress measure considered here regards the *Kirchhoff stress tensor*. It resides in the **current** configuration and represents a symmetric tensor. It is often convenient to work with such a stress variable; defined by,

$$\boldsymbol{\tau} = J\boldsymbol{\sigma} \ . \tag{6.283}$$

The expressions (6.278)–(6.281) which interrelate the stress and strain measures are known as the *constitutive equations*. They basically help determine the state of stress at any point of interest in the continuum medium.

The alternate stress tensors $\mathbf{P}$, sym$\mathbf{T}_{\mathrm{B}}$, $\mathbf{S}$ and $\boldsymbol{\tau}$ are said to be *work conjugate* to the strain rate measures $\dot{\mathbf{F}}$, $\dot{\mathrm{U}}$, $\dot{\mathbf{C}}$ and $\mathbf{d}$, respectively, where $\mathbf{d} = \mathbf{F}^{-\mathrm{T}}\dot{\mathbf{C}}\mathbf{F}^{-1}/2$ is called the *rate of deformation tensor*. In this regard, for instance, the quantities $\mathbf{P}$ and $\dot{\mathbf{F}}$ constitute a *work conjugate pair*. For more considerations on work conjugate pairs, see Atluri [56]. ✆

Consider the quantity $\mathbf{P}$ which basically represents a **nonlinear** tensor function of $\mathbf{F}$. Thus, any change in the strain measure $\mathbf{F}$ naturally leads to a change in its conjugate stress tensor $\mathbf{P}$. Notice that the derivative of $\mathbf{P}$ with respect to $\mathbf{F}$ provides a fourth-order tensor. Such a tensor represents what is known as *elasticity tensor* (or *tangent modulus*). Each stress measure has its own elasticity tensor. The tangent moduli play a crucial role in iterative solution techniques such as **Newton-Raphson method** which is widely used in computational mechanics. This motivates to introduce the most important elasticity tensors in the following.

First, consider the mixed fourth-order tensor

$$\mathbb{A} = \frac{\partial \mathbf{P}}{\partial \mathbf{F}} = \frac{\partial^2 \Psi_{\mathbf{F}}}{\partial \mathbf{F} \partial \mathbf{F}} \ , \tag{6.284}$$

which is known as the *first elasticity tensor*. It has only the major symmetries $\mathbb{A}_{iAjB} = \mathbb{A}_{jBiA}$ for hyperelastic materials. In this case, one deals with only 45 independent components.

Then, consider the tangent modulus

$$\mathbb{C} = 2\frac{\partial \mathbf{S}}{\partial \mathbf{C}} = 4\frac{\partial^2 \Psi_{\mathbf{C}}}{\partial \mathbf{C} \partial \mathbf{C}} , \tag{6.285}$$

which is known as the *referential* (or *material* or *Lagrangian* or *second*) *elasticity tensor*. It possess the minor symmetries $\mathbb{C}_{ABCD} = \mathbb{A}_{BACD} = \mathbb{A}_{ABDC}$. For hyperelasticity, it additionally has the major symmetries $\mathbb{C}_{ABCD} = \mathbb{A}_{CDAB}$. One is thus left with 21 independent components. Guided by (3.115), $\mathbb{C}$ eventually renders a super-symmetric tensor for hyperelastic materials. It is worth mentioning that $\mathbb{C}$ is a positive semi-definite tensor, i.e. $\mathbf{S} : \mathbb{C} : \mathbf{S} \geq 0$ for any symmetric tensor $\mathbf{S}$ (Gurtin et al. [57]).

Finally, the spatial counterpart of the referential tangent modulus is defined through the so-called *Piola transformation*

$$\underbrace{\mathbb{c}_{ijkl} = J^{-1} F_{iA} F_{jB} F_{kC} F_{lD} \mathbb{C}_{ABCD}}_{\text{note that } F_{iA}\widehat{\mathbf{e}}_i = \mathbf{F}\widehat{\mathbf{E}}_A \text{ and } \lambda_a \widehat{\mathbf{n}}_a = \mathbf{F}\widehat{\mathbf{N}}_a} . \tag{6.286}$$

This is called the *spatial* (or *Eulerian*) *elasticity tensor* which has also the minor symmetries. And it is super-symmetric when the material is hyperelastic. Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be two symmetric material tensors. Further, let $\mathbb{C}$ be a fourth-order tensor of the form $\mathbf{S}_1 \otimes \mathbf{S}_2$ or $\mathbf{S}_1 \odot \mathbf{S}_2$. Accordingly, by the Piola transformation,

$$\begin{cases} \text{if } \mathbb{C} = \mathbf{S}_1 \otimes \mathbf{S}_2 \text{ then } J\mathbb{c} = \left(\mathbf{F}\mathbf{S}_1\mathbf{F}^{\mathrm{T}}\right) \otimes \left(\mathbf{F}\mathbf{S}_2\mathbf{F}^{\mathrm{T}}\right) \\ \text{if } \mathbb{C} = \mathbf{S}_1 \odot \mathbf{S}_2 \text{ then } J\mathbb{c} = \left(\mathbf{F}\mathbf{S}_1\mathbf{F}^{\mathrm{T}}\right) \odot \left(\mathbf{F}\mathbf{S}_2\mathbf{F}^{\mathrm{T}}\right) \end{cases} . \quad \longleftarrow \text{see (6.302b) and (6.302c)} \tag{6.287}$$

For an isotropic hyperelastic material, the spatial elasticity tensor admits the explicit representation (Miehe [58])

$$\mathbb{c} = J^{-1}\mathbf{b}\tilde{\mathbb{c}}\mathbf{b} \quad \text{where} \quad \tilde{\mathbb{c}} = 4\frac{\partial^2 \Psi_{\mathbf{b}}}{\partial \mathbf{b} \partial \mathbf{b}} . \tag{6.288}$$

Let $\mathbf{s}_1$ and $\mathbf{s}_2$ be two symmetric spatial tensors. Further, let $\tilde{\mathbb{c}}$ be a fourth-order tensor of the form $\mathbf{s}_1 \otimes \mathbf{s}_2$ or $\mathbf{s}_1 \odot \mathbf{s}_2$. Accordingly, using (3.101a) and $(3.101d)_2$,

$$\begin{cases} \text{if } \tilde{\mathbb{c}} = \mathbf{s}_1 \otimes \mathbf{s}_2 \text{ then } J\mathbb{c} = (\mathbf{b}\mathbf{s}_1) \otimes (\mathbf{s}_2\mathbf{b}) \\ \text{if } \tilde{\mathbb{c}} = \mathbf{s}_1 \odot \mathbf{s}_2 \text{ then } J\mathbb{c} = (\mathbf{b}\mathbf{s}_1) \odot (\mathbf{s}_2\mathbf{b}) \end{cases} . \quad \longleftarrow \text{see (6.304)} \tag{6.289}$$

1. Show that the introduced stress tensors are related by[2]

$$\underbrace{\mathbf{P} = \mathbf{R}\mathbf{T}_{\mathrm{B}} = \mathbf{F}\mathbf{S} = \boldsymbol{\tau}\mathbf{F}^{-\mathrm{T}}}_{\text{note that these relations hold for all materials}} \quad . \tag{6.290}$$

Moreover, by using (6.276), show that they can be represented by

$$\begin{aligned} \mathbf{P} &= \gamma_1 \mathbf{F} + \gamma_2 \mathbf{F}\mathbf{C} + \gamma_3 \mathbf{F}^{-\mathrm{T}} \\ &= \sum_{a=1}^{3} P_a \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_a \qquad \text{if } \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \end{aligned} \tag{6.291a}$$

$$\begin{aligned} \mathbf{T}_{\mathrm{B}} &= \gamma_1 \mathbf{U} + \gamma_2 \mathbf{U}^3 + \gamma_3 \mathbf{U}^{-1} \\ &= \sum_{a=1}^{3} T_{\mathrm{B}\,a} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \qquad \text{if } \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \end{aligned} \tag{6.291b}$$

$$\begin{aligned} \mathbf{S} &= \gamma_1 \mathbf{I} + \gamma_2 \mathbf{C} + \gamma_3 \mathbf{C}^{-1} \\ &= \sum_{a=1}^{3} S_a \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \qquad \text{if } \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \end{aligned} \tag{6.291c}$$

$$\begin{aligned} \boldsymbol{\tau} &= \gamma_1 \mathbf{b} + \gamma_2 \mathbf{b}^2 + \gamma_3 \mathbf{1} \\ &= \sum_{a=1}^{3} \tau_a \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \qquad \text{if } \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \ , \end{aligned} \tag{6.291d}$$

where

$$P_a = T_{\mathrm{B}\,a} = \frac{\partial \Psi_\lambda}{\partial \lambda_a} \quad , \quad S_a = \frac{1}{\lambda_a} \frac{\partial \Psi_\lambda}{\partial \lambda_a} \quad , \quad \tau_a = \lambda_a \frac{\partial \Psi_\lambda}{\partial \lambda_a} \ , \tag{6.292}$$

and

$$\gamma_1 = 2 \frac{\partial \Psi_I}{\partial I_1} + 2 I_1 \frac{\partial \Psi_I}{\partial I_2} \quad , \quad \gamma_2 = -2 \frac{\partial \Psi_I}{\partial I_2} \quad , \quad \gamma_3 = 2 I_3 \frac{\partial \Psi_I}{\partial I_3} \ . \tag{6.293}$$

Solution. The problem can simply be solved by computing the time rate of change of the strain-energy function. And the procedure used to derive the desired relations greatly relies on the product and chain rules of differentiation. First, consider $\mathbf{C}$ and its (canonical) argument $\mathbf{F}$, i.e. $\mathbf{C}(\mathbf{F}) = \mathbf{F}^{\mathrm{T}}\mathbf{F}$. On the one hand, the rate of change of $(6.273)_3$ delivers

[2] In the literature, the stress tensors in (6.290) are defined in advance and then the constitutive relations (6.278)–(6.281) are implied by the *second law of thermodynamics*. In this regard, loosely speaking, this exercise can be viewed as an inverse problem.

$$\dot{\Psi} = \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \underbrace{\dot{C}_{AB}}_{= \overline{F_{iA} F_{iB}}^{\cdot}} = \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \left(\dot{F}_{iA} F_{iB} + F_{iA} \dot{F}_{iB} \right)$$
$$= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \dot{F}_{iA} F_{iB} + \underbrace{\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BA}}}_{= \partial \Psi_{\mathbf{C}} / \partial C_{AB}} F_{iB} \dot{F}_{iA} = 2 \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \dot{F}_{iA} F_{iB} \ . \tag{6.294}$$

On the other hand, the rate of change of (6.273)$_1$ is $\dot{\Psi} = (\partial \Psi_{\mathbf{F}} / \partial F_{iA}) \, \dot{F}_{iA}$. Guided by (2.81) and knowing that the tensors $\mathbf{F}, \dot{\mathbf{F}}$ can be chosen arbitrarily, one then has

$$\frac{2 \partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \cancel{\dot{F}_{iA}} F_{iB} = \frac{\partial \Psi_{\mathbf{F}}}{\partial F_{iA}} \cancel{\dot{F}_{iA}} \Rightarrow \begin{cases} F_{iB} \frac{2 \partial \Psi_{\mathbf{C}}}{\partial C_{BA}} = \frac{\partial \Psi_{\mathbf{F}}}{\partial F_{iA}} & \text{in index notation} \\ \mathbf{F} \frac{2 \partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} = \frac{\partial \Psi_{\mathbf{F}}}{\partial \mathbf{F}} & \text{in direct notation} \end{cases} . \tag{6.295}$$

Having in mind (6.278) and (6.280), one can finally conclude that $\mathbf{P} = \mathbf{FS}$. The ambitious reader may want to arrive at this result in an alternative way:

$$\begin{aligned}
(\mathbf{P})_{iA} &= \frac{\partial \Psi_{\mathbf{F}}}{\partial F_{iA}} \\
&= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BC}} \frac{\partial C_{BC}}{\partial F_{iA}} \\
&= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BC}} \frac{\partial \left(F_{jB} F_{jC} \right)}{\partial F_{iA}} \\
&= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BC}} \underbrace{\frac{\partial F_{jB}}{\partial F_{iA}}}_{= \delta_{ji} \delta_{AB}, \text{ by (6.67)}} F_{jC} + \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BC}} F_{jB} \underbrace{\frac{\partial F_{jC}}{\partial F_{iA}}}_{= \delta_{ji} \delta_{AC}} \\
&= \underbrace{\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AC}}}_{= \partial \Psi_{\mathbf{C}} / \partial C_{CA}} F_{iC} + \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BA}} F_{iB} \\
&= F_{iB} \frac{2 \partial \Psi_{\mathbf{C}}}{\partial C_{BA}} \\
&= (\mathbf{FS})_{iA} \ .
\end{aligned} \tag{6.296}$$

In a similar manner,

$$\begin{aligned}
\left(\frac{\partial \Psi_{\mathbf{U}}}{\partial \mathbf{U}} \right)_{AB} &= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MN}} \frac{\partial C_{MN}}{\partial U_{AB}} \\
&= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MN}} \frac{\partial \left(U_{MC} U_{CN} \right)}{\partial U_{AB}}
\end{aligned}$$

$$= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MN}} \underbrace{\frac{\partial U_{MC}}{\partial U_{AB}}}_{= \frac{1}{2}\left(\delta_{MA}\delta_{BC} + \delta_{MB}\delta_{AC}\right),\ \text{according to (6.68)}} U_{CN} + \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MN}} U_{MC} \underbrace{\frac{\partial U_{CN}}{\partial U_{AB}}}_{= \frac{1}{2}\left(\delta_{CA}\delta_{BN} + \delta_{CB}\delta_{AN}\right)}$$

$$= \underbrace{\frac{1}{2}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AN}} U_{BN}}_{= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AN}} U_{NB}} + \underbrace{\frac{1}{2}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{BN}} U_{AN}}_{= U_{AN}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{NB}}} + \underbrace{\frac{1}{2}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MB}} U_{MA}}_{= U_{AM}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MB}}} + \underbrace{\frac{1}{2}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{MA}} U_{MB}}_{= \frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AM}} U_{MB}}$$

$$= \underbrace{\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{AN}} U_{NB} + U_{AN}\frac{\partial \Psi_{\mathbf{C}}}{\partial C_{NB}}}_{\text{note that } \mathbf{U} = \sqrt{\mathbf{C}} \text{ and } \partial \Psi_{\mathbf{C}}/\partial \mathbf{C} \text{ commute as a result of isotropy, see (6.291c)}}$$

$$= \left(\mathbf{U}\frac{2\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}}\right)_{AB} .$$

Consequently,

$$\mathbf{T}_{\mathrm{B}} = \mathbf{U}\mathbf{S} = \mathbf{R}^{\mathrm{T}}\mathbf{R}\mathbf{U}\mathbf{S} = \mathbf{R}^{\mathrm{T}}\mathbf{F}\mathbf{S} = \mathbf{R}^{\mathrm{T}}\mathbf{P} . \tag{6.297}$$

Similarly to (6.296), one will have

$$\frac{\partial \Psi_{\mathbf{F}}}{\partial F_{iA}} = \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{jk}} \frac{\partial b_{jk}}{\partial F_{iA}}$$

$$= \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{jk}} \frac{\partial \left(F_{jB} F_{kB}\right)}{\partial F_{iA}}$$

$$= \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{jk}} \underbrace{\frac{\partial F_{jB}}{\partial F_{iA}}}_{= \delta_{ji}\delta_{AB}} F_{kB} + \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{jk}} F_{jB} \underbrace{\frac{\partial F_{kB}}{\partial F_{iA}}}_{= \delta_{ki}\delta_{AB}}$$

$$= \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{ik}} F_{kA} + \underbrace{\frac{\partial \Psi_{\mathbf{b}}}{\partial b_{ji}}}_{= \partial \Psi_{\mathbf{b}}/\partial b_{ij}} F_{jA}$$

$$= 2\frac{\partial \Psi_{\mathbf{b}}}{\partial b_{ik}} F_{kA} .$$

Postmultiplying the above relation by F_{jA} gives the desired functional relationship

$$\underbrace{\frac{\partial \Psi_{\mathbf{F}}}{\partial F_{iA}} F_{jA} = 2\frac{\partial \Psi_{\mathbf{b}}}{\partial b_{ik}} \left[F_{kA} F_{jA} = b_{kj}\right] \quad \text{or} \quad \frac{\partial \Psi_{\mathbf{F}}}{\partial \mathbf{F}} \mathbf{F}^{\mathrm{T}} = 2\frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \mathbf{b}}_{\text{or, using (6.278), (6.281), (6.283) and (6.296), } \boldsymbol{\tau} = \mathbf{P}\mathbf{F}^{\mathrm{T}} = \mathbf{F}\mathbf{S}\mathbf{F}^{\mathrm{T}}} . \tag{6.298}$$

Consistent with this, one can explicitly express $\boldsymbol{\tau}$ in terms of $\mathbf{v}$ as

$$\boldsymbol{\tau} = \frac{\partial \Psi_{\mathbf{v}}}{\partial \mathbf{v}} \mathbf{v} = \mathbf{v} \frac{\partial \Psi_{\mathbf{v}}}{\partial \mathbf{v}} \ , \tag{6.299}$$

because, using (6.68) along with the product and chain rules,

$$\begin{aligned}
\left(\frac{\partial \Psi_{\mathbf{v}}}{\partial \mathbf{v}} \right)_{ij} &= \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{mn}} \frac{\partial b_{mn}}{\partial v_{ij}} \\
&= \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{mn}} \frac{(v_{mk} v_{kn})}{\partial v_{ij}} \\
&= \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{mn}} \underbrace{\frac{\partial v_{mk}}{\partial v_{ij}}}_{= \frac{1}{2}(\delta_{mi}\delta_{jk} + \delta_{mj}\delta_{ik})} v_{kn} + \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{mn}} v_{mk} \underbrace{\frac{\partial v_{kn}}{\partial v_{ij}}}_{= \frac{1}{2}(\delta_{ki}\delta_{jn} + \delta_{kj}\delta_{in})} \\
&= \underbrace{\frac{1}{2} \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{in}} v_{jn}}_{= \frac{\partial \Psi}{\partial b_{in}} v_{nj}} + \underbrace{\frac{1}{2} \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{jn}} v_{in}}_{= v_{in} \frac{\partial \Psi}{\partial b_{nj}}} + \underbrace{\frac{1}{2} \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{mj}} v_{mi}}_{= v_{im} \frac{\partial \Psi}{\partial b_{mj}}} + \underbrace{\frac{1}{2} \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{mi}} v_{mj}}_{= \frac{\partial \Psi}{\partial b_{im}} v_{mj}} \\
&= \underbrace{\frac{\partial \Psi_{\mathbf{b}}}{\partial b_{in}} v_{nj} + v_{in} \frac{\partial \Psi_{\mathbf{b}}}{\partial b_{nj}}}_{\text{note that } \mathbf{v} = \sqrt{\mathbf{b}} \text{ commutes with } \partial \Psi_{\mathbf{b}} / \partial \mathbf{b} \text{ by isotropy, see (6.291d)}} \\
&= 2 \left(\frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \mathbf{v} \right)_{ij} \ .
\end{aligned}$$

Note that the important relationship $\boldsymbol{\tau} = \mathbf{F}\mathbf{S}\mathbf{F}^{\mathrm{T}}$ can also be resulted from the rate form $(\partial \Psi_{\mathbf{C}} / \mathbf{C}) : \dot{\mathbf{C}} = (\partial \Psi_{\mathbf{b}} / \mathbf{b}) : \dot{\mathbf{b}}$ as follows:

$$\begin{aligned}
& \frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} : \dot{\mathbf{F}}^{\mathrm{T}} \mathbf{F} + \frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} : \mathbf{F}^{\mathrm{T}} \dot{\mathbf{F}} = \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} : \dot{\mathbf{F}} \mathbf{F}^{\mathrm{T}} + \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} : \mathbf{F} \dot{\mathbf{F}}^{\mathrm{T}} \\
\Longrightarrow \ & \frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} \mathbf{F}^{\mathrm{T}} : \dot{\mathbf{F}}^{\mathrm{T}} + \mathbf{F} \frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} : \dot{\mathbf{F}} = \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \mathbf{F} : \dot{\mathbf{F}} + \mathbf{F}^{\mathrm{T}} \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} : \dot{\mathbf{F}}^{\mathrm{T}} \\
\Longrightarrow \ & 2\mathbf{F} \frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} : \dot{\mathbf{F}} = 2 \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \mathbf{F} : \dot{\mathbf{F}} \quad \longleftarrow \text{note that } \mathbf{F} \text{ and } \dot{\mathbf{F}} \text{ can be chosen arbitrarily} \\
\Longrightarrow \ & 2\mathbf{F} \frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} \mathbf{F}^{\mathrm{T}} = 2 \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \mathbf{F} \mathbf{F}^{\mathrm{T}} = 2 \frac{\partial \Psi_{\mathbf{b}}}{\partial \mathbf{b}} \mathbf{b} \ ,
\end{aligned}$$

where $(2.79a)_1$, $(2.79b)_{1-2}$, (2.81) and $(6.264)_{1-2}$ along with the product rule of differentiation have been used.

At this stage, consider the isotropy assumption (6.276) to verify the constitutive relations $(6.291c)_{1-2}$. Then, by $(6.20a)_2$, $(6.20b)_2$, $(6.20c)_4$, (6.204a) and $(6.268b)_2$ along with the chain rule, one can write

$$\mathbf{S} = 2\frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} = 2\sum_{a=1}^{3} \frac{\partial \Psi_I}{\partial I_a}\frac{\partial I_a}{\partial \mathbf{C}}$$

$$= 2\frac{\partial \Psi_I}{\partial I_1}\underbrace{\frac{\partial I_1}{\partial \mathbf{C}}}_{=\mathbf{I}} + 2\frac{\partial \Psi_I}{\partial I_2}\underbrace{\frac{\partial I_2}{\partial \mathbf{C}}}_{=I_1\mathbf{I}-\mathbf{C}} + 2\frac{\partial \Psi_I}{\partial I_3}\underbrace{\frac{\partial I_3}{\partial \mathbf{C}}}_{=I_3\mathbf{C}^{-1}} ,$$

$$\mathbf{S} = 2\frac{\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} = 2\sum_{a=1}^{3} \frac{\partial \Psi_\lambda}{\partial \lambda_a^2}\frac{\partial \lambda_a^2}{\partial \mathbf{C}}$$

$$= 2\sum_{a=1}^{3} \frac{\partial \Psi_\lambda}{\partial \lambda_a^2}\widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a = \sum_{a=1}^{3} \frac{\partial \Psi_\lambda}{\lambda_a \partial \lambda_a}\widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a .$$

The remaining desired relations can then be shown in a straightforward manner. This is left to be undertaken by the reader. Moreover, the ambitious reader may want to solve the problem for the cases in which the principal stretches are repeated.

2. Verify that the first elasticity tensor is related to the material tensor of elasticity through the following expression

$$\mathbb{A}_{iBkD} = \delta_{ik} S_{BD} + \mathbb{C}_{ABCD} F_{iA} F_{kC} , \tag{6.300}$$

which immediately implies that

$$\mathbb{A}_{iBkD} = \delta_{ik} F_{Bj}^{-1} \tau_{jl} F_{Dl}^{-1} + J \mathbb{c}_{ijkl} F_{Bj}^{-1} F_{Dl}^{-1} . \tag{6.301}$$

Moreover, for an isotropic hyperelastic solid characterized by the stored-energy function $\Psi = \Psi_I(I_1, I_2, I_3)$, show that the most general form for the introduced elasticity tensors are

$$\mathbb{A} = \frac{\partial \mathbf{P}}{\partial \mathbf{F}} = \frac{\partial^2 \Psi_I(I_1, I_2, I_3)}{\partial \mathbf{F} \partial \mathbf{F}} = \sum_{i=1}^{8} \delta_i \overline{\mathbb{A}}_i + \sum_{i=1}^{3} \gamma_i \widehat{\mathbb{A}}_i$$

$$= \delta_1 \mathbf{F} \otimes \mathbf{F} + \delta_2 (\mathbf{FC} \otimes \mathbf{F} + \mathbf{F} \otimes \mathbf{FC}) + \delta_3 \left(\mathbf{F}^{-\mathrm{T}} \otimes \mathbf{F} + \mathbf{F} \otimes \mathbf{F}^{-\mathrm{T}}\right)$$

$$+ \delta_4 \mathbf{FC} \otimes \mathbf{FC} + \delta_5 \left(\mathbf{F}^{-\mathrm{T}} \otimes \mathbf{FC} + \mathbf{FC} \otimes \mathbf{F}^{-\mathrm{T}}\right) + \delta_6 \mathbf{F}^{-\mathrm{T}} \otimes \mathbf{F}^{-\mathrm{T}}$$

$$+ \delta_7 \left(\mathbf{F}^{-\mathrm{T}} \boxplus \mathbf{F}^{-\mathrm{T}} + \mathbf{1} \boxtimes \mathbf{C}^{-1}\right) + \delta_8 (\mathbf{F} \boxplus \mathbf{F} + \mathbf{b} \boxtimes \mathbf{I})$$

$$+ \mathbf{1} \boxtimes \left(\gamma_1 \mathbf{I} + \gamma_2 \mathbf{C} + \gamma_3 \mathbf{C}^{-1}\right) , \quad \nwarrow \text{ see (3.70b)-(3.70c)} \tag{6.302a}$$

$$\mathbb{C} = 2\frac{\partial \mathbf{S}}{\partial \mathbf{C}} = 4\frac{\partial^2 \Psi_I(I_1, I_2, I_3)}{\partial \mathbf{C} \partial \mathbf{C}} = \sum_{i=1}^{8} \delta_i \overline{\mathbb{C}}_i$$

$$= \delta_1 \mathbf{I} \otimes \mathbf{I} + \delta_2 (\mathbf{I} \otimes \mathbf{C} + \mathbf{C} \otimes \mathbf{I}) + \delta_3 \left(\mathbf{I} \otimes \mathbf{C}^{-1} + \mathbf{C}^{-1} \otimes \mathbf{I}\right)$$

$$
\begin{aligned}
&+ \delta_4 \mathbf{C} \otimes \mathbf{C} + \delta_5 \left(\mathbf{C} \otimes \mathbf{C}^{-1} + \mathbf{C}^{-1} \otimes \mathbf{C} \right) + \delta_6 \mathbf{C}^{-1} \otimes \mathbf{C}^{-1} \\
&+ \delta_7 \mathbf{C}^{-1} \odot \mathbf{C}^{-1} + \delta_8 \mathbf{I} \odot \mathbf{I} \ , \quad \leftarrow \text{see (3.70d)}
\end{aligned}
\tag{6.302b}
$$

$$
\begin{aligned}
J\mathbb{c} = 4\mathbf{b} \frac{\partial^2 \Psi_I}{\partial \mathbf{b} \partial \mathbf{b}} \mathbf{b} &= \sum_{i=1}^{8} \delta_i \overline{\mathbb{c}}_i \\
&= \delta_1 \mathbf{b} \otimes \mathbf{b} + \delta_2 \left(\mathbf{b} \otimes \mathbf{b}^2 + \mathbf{b}^2 \otimes \mathbf{b} \right) + \delta_3 \left(\mathbf{b} \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{b} \right) \\
&\quad + \delta_4 \mathbf{b}^2 \otimes \mathbf{b}^2 + \delta_5 \left(\mathbf{b}^2 \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{b}^2 \right) + \delta_6 \mathbf{1} \otimes \mathbf{1} \\
&\quad + \delta_7 \mathbf{1} \odot \mathbf{1} + \delta_8 \mathbf{b} \odot \mathbf{b} \ ,
\end{aligned}
\tag{6.302c}
$$

where

$$
\delta_1 = 4 \left(\frac{\partial^2 \Psi_I}{\partial I_1 \partial I_1} + 2 I_1 \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_2} + \frac{\partial \Psi_I}{\partial I_2} + I_1^2 \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_2} \right) \ , \tag{6.303a}
$$

$$
\delta_2 = -4 \left(\frac{\partial^2 \Psi_I}{\partial I_1 \partial I_2} + I_1 \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_2} \right) \ , \tag{6.303b}
$$

$$
\delta_3 = 4 \left(I_3 \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_3} + I_1 I_3 \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_3} \right) \quad , \quad \delta_4 = 4 \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_2} \ , \tag{6.303c}
$$

$$
\delta_5 = -4 I_3 \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_3} \quad , \quad \delta_6 = 4 \left(I_3 \frac{\partial \Psi_I}{\partial I_3} + I_3^2 \frac{\partial^2 \Psi_I}{\partial I_3 \partial I_3} \right) \ , \tag{6.303d}
$$

$$
\delta_7 = -4 I_3 \frac{\partial \Psi_I}{\partial I_3} \quad , \quad \delta_8 = -4 \frac{\partial \Psi_I}{\partial I_2} \ , \tag{6.303e}
$$

and $\gamma_i, \ i = 1, 2, 3$, are already defined in (6.293).

Solution. Recall that the first and second Piola-Kirchhoff stress tensors were related by $\mathbf{P}(\mathbf{F}) = \mathbf{F}\mathbf{S}(\mathbf{C}(\mathbf{F}))$ for which $(\mathbf{C})_{CE} = (\mathbf{F})_{mC} (\mathbf{F})_{mE}$. The derivative of this relation with respect to the deformation gradient helps provide the desired result (6.300). Thus, making use of the product and chain rules of differentiation,

$$
\begin{aligned}
\left(\frac{\partial^2 \Psi_{\mathbf{F}}}{\partial \mathbf{F} \partial \mathbf{F}} \right)_{iBkD} &= \frac{\partial}{\partial F_{kD}} \left(F_{iA} \frac{2 \partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \right) \\
&= \underbrace{\frac{\partial F_{iA}}{\partial F_{kD}}}_{= \delta_{ik} \delta_{DA}} \frac{2 \partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \\
&\quad + F_{iA} \frac{\partial}{\partial C_{CE}} \left(\frac{2 \partial \Psi_{\mathbf{C}}}{\partial C_{AB}} \right) \qquad \underbrace{\frac{\partial \left(F_{mC} F_{mE} \right)}{\partial F_{kD}}}_{= \delta_{mk} \delta_{DC} F_{mE} + F_{mC} \delta_{mk} \delta_{DE} = \delta_{DC} F_{kE} + \delta_{DE} F_{kC}}
\end{aligned}
$$

$$= \delta_{ik} \frac{2\partial \Psi_{\mathbf{C}}}{\partial C_{DB}} + F_{iA} \underbrace{\left(\frac{2\partial^2 \Psi_{\mathbf{C}}}{\partial C_{DE} \partial C_{AB}} F_{kE} + \frac{2\partial^2 \Psi_{\mathbf{C}}}{\partial C_{CD} \partial C_{AB}} F_{kC} \right)}_{= \frac{4\partial^2 \Psi_{\mathbf{C}}}{\partial C_{AB} \partial C_{CD}} F_{kC}}$$

$$= \delta_{ik} \left(\frac{2\partial \Psi_{\mathbf{C}}}{\partial \mathbf{C}} \right)_{BD} + F_{iA} \left(\frac{4\partial^2 \Psi_{\mathbf{C}}}{\partial \mathbf{C} \partial \mathbf{C}} \right)_{ABCD} F_{kC} \ .$$

Introducing $S_{BD} = F_{Bj}^{-1} \tau_{jl} F_{Dl}^{-1}$ and $\mathbb{C}_{ABCD} = F_{Am}^{-1} F_{Bj}^{-1} F_{Cn}^{-1} F_{Dl}^{-1} J \mathbb{c}_{mjnl}$ into (6.300) then readily leads to (6.301).

At this stage, consider the isotropic energy function $\Psi = \Psi_I(I_1, I_2, I_3)$ with the material stress tensor $\mathbf{S} = \gamma_1 \mathbf{I} + \gamma_2 \mathbf{C} + \gamma_3 \mathbf{C}^{-1}$ to represent the second elasticity tensor (6.302b). The desired result follows in detail from the following step by step computation

$$(\mathbb{C})_{ABCD} = 2 \frac{\partial}{\partial C_{CD}} \left\{ 2 \left[\left(\frac{\partial \Psi_I}{\partial I_1} + I_1 \frac{\partial \Psi_I}{\partial I_2} \right) \delta_{AB} - \frac{\partial \Psi_I}{\partial I_2} C_{AB} + I_3 \frac{\partial \Psi_I}{\partial I_3} C_{AB}^{-1} \right] \right\}$$

$$= 4 \underbrace{\delta_{AB} \frac{\partial}{\partial C_{CD}} \left(\frac{\partial \Psi_I}{\partial I_1} \right)}_{= \delta_{AB} \left[\frac{\partial^2 \Psi_I}{\partial I_1 \partial I_1} \frac{\partial I_1}{\partial C_{CD}} + \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_2} \frac{\partial I_2}{\partial C_{CD}} + \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_3} \frac{\partial I_3}{\partial C_{CD}} \right]}$$

$$= \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_1} \delta_{AB} \delta_{CD} + \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_2} \left(I_1 \delta_{AB} \delta_{CD} - \delta_{AB} C_{CD} \right) + \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_3} I_3 \delta_{AB} C_{CD}^{-1}$$

$$+ 4 \frac{\partial \Psi_I}{\partial I_2} \underbrace{\delta_{AB} \frac{\partial I_1}{\partial C_{CD}}}_{= \delta_{AB} \delta_{CD}}$$

$$+ 4 I_1 \underbrace{\delta_{AB} \frac{\partial}{\partial C_{CD}} \left(\frac{\partial \Psi_I}{\partial I_2} \right)}_{= \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_2} \delta_{AB} \delta_{CD} + \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_2} \left(I_1 \delta_{AB} \delta_{CD} - \delta_{AB} C_{CD} \right) + \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_3} I_3 \delta_{AB} C_{CD}^{-1}}$$

$$- 4 \underbrace{C_{AB} \frac{\partial}{\partial C_{CD}} \left(\frac{\partial \Psi_I}{\partial I_2} \right)}_{= \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_2} C_{AB} \delta_{CD} + \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_2} \left(I_1 C_{AB} \delta_{CD} - C_{AB} C_{CD} \right) + \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_3} I_3 C_{AB} C_{CD}^{-1}}$$

$$- 4 \frac{\partial \Psi_I}{\partial I_2} \underbrace{\frac{\partial C_{AB}}{\partial C_{CD}}}_{= \frac{1}{2} \left(\delta_{AC} \delta_{DB} + \delta_{AD} \delta_{CB} \right)} + 4 \frac{\partial \Psi_I}{\partial I_3} \underbrace{C_{AB}^{-1} \frac{\partial I_3}{\partial C_{CD}}}_{= I_3 C_{AB}^{-1} C_{CD}^{-1}}$$

$$+ 4 I_3 \underbrace{C_{AB}^{-1} \frac{\partial}{\partial C_{CD}} \left(\frac{\partial \Psi_I}{\partial I_3} \right)}_{= \frac{\partial^2 \Psi_I}{\partial I_1 \partial I_3} C_{AB}^{-1} \delta_{CD} + \frac{\partial^2 \Psi_I}{\partial I_2 \partial I_3} \left(I_1 C_{AB}^{-1} \delta_{CD} - C_{AB}^{-1} C_{CD} \right) + \frac{\partial^2 \Psi_I}{\partial I_3 \partial I_3} I_3 C_{AB}^{-1} C_{CD}^{-1}}$$

$$+ 4I_3 \frac{\partial \Psi_I}{\partial I_3} \underbrace{\frac{\partial C^{-1}_{AB}}{\partial C_{CD}}}_{= -\frac{1}{2}\left(C^{-1}_{AC}C^{-1}_{DB} + C^{-1}_{AD}C^{-1}_{CB}\right)} .$$

In a similar manner, $\tilde{\mathbb{c}}$ in (6.288) takes the form

$$\begin{aligned}\tilde{\mathbb{c}} = {} & \delta_1 \mathbf{1} \otimes \mathbf{1} + \delta_2 (\mathbf{1} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{1}) + \delta_3 \left(\mathbf{1} \otimes \mathbf{b}^{-1} + \mathbf{b}^{-1} \otimes \mathbf{1}\right) \\ & + \delta_4 \mathbf{b} \otimes \mathbf{b} + \delta_5 \left(\mathbf{b} \otimes \mathbf{b}^{-1} + \mathbf{b}^{-1} \otimes \mathbf{b}\right) + \delta_6 \mathbf{b}^{-1} \otimes \mathbf{b}^{-1} \\ & + \delta_7 \mathbf{b}^{-1} \odot \mathbf{b}^{-1} + \delta_8 \mathbf{1} \odot \mathbf{1} \, . \end{aligned} \tag{6.304}$$

Pre- and postmultiplying the above relation by $\mathbf{b}$, taking into account (6.289), leads to the desired result $(6.302\text{c})_3$. At the end, it should not be difficult to establish $(6.302\text{a})_4$ by using $(6.291\text{c})_1$, (6.300) and $(6.302\text{b})_4$.

3. Let $\Psi_\lambda(\lambda_1, \lambda_2, \lambda_3)$ be a stored-energy function describing isotropic response of a hyperelastic solid. Then, derive the following spectral formulas for the introduced elasticity tensors

$$\begin{aligned}\mathbb{A} = {} & \sum_{a,\, b=1}^{3} \frac{\partial P_a}{\partial \lambda_b} \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{n}}_b \otimes \widehat{\mathbf{N}}_b \\ & + \sum_{\substack{a,\, b=1 \\ b \neq a}}^{3} \frac{\lambda_a P_b - \lambda_b P_a}{\lambda_b^2 - \lambda_a^2} \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{n}}_b \otimes \widehat{\mathbf{N}}_a \\ & + \sum_{\substack{a,\, b=1 \\ b \neq a}}^{3} \left(\frac{P_a}{\lambda_a} + \lambda_a^2 \frac{\lambda_b^{-1} P_b - \lambda_a^{-1} P_a}{\lambda_b^2 - \lambda_a^2} \right) \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{N}}_b \, , \end{aligned} \tag{6.305a}$$

$$\begin{aligned}\mathbb{C} = {} & \sum_{a,\, b=1}^{3} \frac{1}{\lambda_b} \frac{\partial S_a}{\partial \lambda_b} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{N}}_b \\ & + \sum_{\substack{a,\, b=1 \\ b \neq a}}^{3} \frac{S_b - S_a}{\lambda_b^2 - \lambda_a^2} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \otimes \left[\widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b + \otimes \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{N}}_a \right] \, , \end{aligned} \tag{6.305b}$$

$$\begin{aligned}J\mathbb{c} = {} & \sum_{a,b=1}^{3} \frac{\partial \tau_a}{\partial \varepsilon_b} \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_b \otimes \widehat{\mathbf{n}}_b - \sum_{a=1}^{3} 2\tau_a \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_a \\ & + \sum_{\substack{a,\, b=1 \\ b \neq a}}^{3} \frac{\lambda_a^2 \tau_b - \lambda_b^2 \tau_a}{\lambda_b^2 - \lambda_a^2} \widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_b \otimes \left[\widehat{\mathbf{n}}_a \otimes \widehat{\mathbf{n}}_b + \widehat{\mathbf{n}}_b \otimes \widehat{\mathbf{n}}_a \right] \, , \end{aligned} \tag{6.305c}$$

where $\varepsilon_a = \log \lambda_a$ and $\tau_a = \lambda_a^2 S_a = \lambda_a P_a = \lambda_a \partial \Psi_\lambda / \partial \lambda_a$ according to (6.292).

Solution. Attention here is focused on representing the material elasticity tensor (6.305b) by computing the time rate of change of the strain measure $\mathbf{C}$ and its conjugate stress tensor $\mathbf{S}$. The main assumption is that $\mathbf{U}$ (or $\mathbf{C}$) has exactly three distinct principal values.

To begin with, consider the orthonormal set of principal referential directions, i.e. $\{\widehat{\mathbf{N}}_1, \widehat{\mathbf{N}}_2, \widehat{\mathbf{N}}_3\}$, which can be expressed with respect to the material standard basis $\{\widehat{\mathbf{E}}_1, \widehat{\mathbf{E}}_2, \widehat{\mathbf{E}}_3\}$ by the linear transformation $\widehat{\mathbf{N}}_a = \mathbf{Q}\widehat{\mathbf{E}}_a,\ a = 1, 2, 3$. One then has $\mathbf{Q} = \sum_{b=1}^{3} \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{E}}_b$ which satisfies $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$. The time rate of change of $\widehat{\mathbf{N}}_a$ is then given by

$$\underbrace{\dot{\widehat{\mathbf{N}}}_a = \dot{\mathbf{Q}}\widehat{\mathbf{E}}_a = \left(\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}\right)\left(\mathbf{Q}\widehat{\mathbf{E}}_a\right) = \boldsymbol{\Omega}\widehat{\mathbf{N}}_a}_{\text{note that } \overline{\mathbf{Q}\mathbf{Q}^{\mathrm{T}}}^{\cdot} = \dot{\mathbf{I}} \text{ implies } \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\dot{\mathbf{Q}}^{\mathrm{T}} \text{ or } \boldsymbol{\Omega} = -\boldsymbol{\Omega}^{\mathrm{T}}} \quad , \tag{6.306}$$

with

$$\underbrace{\boldsymbol{\Omega} = \sum_{c=1}^{3} \dot{\widehat{\mathbf{N}}}_c \otimes \widehat{\mathbf{N}}_c = \sum_{a,b=1}^{3} \Omega_{ab}\widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b}_{\text{note that } \Omega_{ab} = -\Omega_{ba} \text{ and } \Omega_{aa} = \Omega_{bb} = 0} \ . \tag{6.307}$$

Recall from (6.291c)$_2$ that $\mathbf{S} = \sum_{c=1}^{3} S_c\widehat{\mathbf{N}}_c \otimes \widehat{\mathbf{N}}_c$. It then follows that

$$\begin{aligned}
\boldsymbol{\Omega}\mathbf{S} &= \sum_{\substack{a,b=1\\ b\neq a}}^{3} \Omega_{ab}\widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \left(\sum_{c=1}^{3} S_c\widehat{\mathbf{N}}_c \otimes \widehat{\mathbf{N}}_c\right) \\
&= \sum_{\substack{a,b=1\\ b\neq a}}^{3} \sum_{c=1}^{3} \Omega_{ab} S_c \delta_{bc} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_c = \sum_{\substack{a,b=1\\ b\neq a}}^{3} \Omega_{ab} S_b \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \ .
\end{aligned} \tag{6.308}$$

In a similar manner,

$$\mathbf{S}\boldsymbol{\Omega}^{\mathrm{T}} = (\boldsymbol{\Omega}\mathbf{S})^{\mathrm{T}} = \sum_{\substack{a,b=1\\ b\neq a}}^{3} \Omega_{ab} S_b \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{N}}_a = \sum_{\substack{a,b=1\\ b\neq a}}^{3} \Omega_{ba} S_a \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \ . \tag{6.309}$$

The relations (6.291c)$_2$, (6.308)$_3$ and (6.309)$_3$ can now be used to evaluate the rate of change in the second Piola-Kirchhoff stress tensor as follows:

$$\dot{\mathbf{S}} = \underbrace{\sum_{a=1}^{3} \dot{S}_a \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a}_{= \sum_{a,b=1}^{3} \frac{\partial S_a}{\partial \lambda_b} \dot{\lambda}_b \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a} + \underbrace{\sum_{a=1}^{3} S_a \left(\dot{\widehat{\mathbf{N}}}_a = \boldsymbol{\Omega}\widehat{\mathbf{N}}_a\right) \otimes \widehat{\mathbf{N}}_a}_{= \boldsymbol{\Omega}\mathbf{S}} + \underbrace{\sum_{a=1}^{3} S_a \widehat{\mathbf{N}}_a \otimes \boldsymbol{\Omega}\widehat{\mathbf{N}}_a}_{= \mathbf{S}\boldsymbol{\Omega}^{\mathrm{T}}}$$

$$= \sum_{a,\,b=1}^{3} \frac{\partial S_a}{\partial \lambda_b} \dot{\lambda}_b \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a + \sum_{\substack{a,\,b=1 \\ b \neq a}}^{3} \underbrace{\left(\Omega_{ab} S_b + \Omega_{ba} S_a \right)}_{= \Omega_{ab} \left(S_b - S_a \right)} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \; . \tag{6.310}$$

In a similar manner, the rate of $\mathbf{C} = \sum_{a=1}^{3} \lambda_a^2 \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a$ takes the form

$$\begin{aligned} \dot{\mathbf{C}} = \sum_{a,\,b=1}^{3} & \underbrace{2\lambda_a \delta_{ab} \dot{\lambda}_b}_{\substack{= \widehat{\mathbf{N}}_a \cdot \dot{\mathbf{C}} \widehat{\mathbf{N}}_a = \dot{C}_{aa} \text{ which represent} \\ \text{the normal components (diagonal elements)}}} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \\ + \sum_{\substack{a,\,b=1 \\ b \neq a}}^{3} & \underbrace{\Omega_{ab} \left(\lambda_b^2 - \lambda_a^2 \right)}_{\substack{= \widehat{\mathbf{N}}_a \cdot \dot{\mathbf{C}} \widehat{\mathbf{N}}_b = \dot{C}_{ab} \text{ which provide} \\ \text{the shear components (off-diagonal elements)}}} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \; . \end{aligned} \tag{6.311}$$

To proceed, consider a general super-symmetric fourth-order tensor defined by

$$\mathbb{C} = \sum_{\bar{a}=1}^{3} \sum_{\bar{b}=1}^{3} \sum_{c=1}^{3} \sum_{d=1}^{3} \mathbb{C}_{\bar{a}\bar{b}cd} \widehat{\mathbf{N}}_{\bar{a}} \otimes \widehat{\mathbf{N}}_{\bar{b}} \otimes \widehat{\mathbf{N}}_c \otimes \widehat{\mathbf{N}}_d \; . \tag{6.312}$$

Then,

$$\begin{aligned} \mathbb{C} : \frac{1}{2} \dot{\mathbf{C}} = & \sum_{\bar{a},\,\bar{b}=1}^{3} \sum_{c=1}^{3} \mathbb{C}_{\bar{a}\bar{b}cc} \lambda_c \dot{\lambda}_c \widehat{\mathbf{N}}_{\bar{a}} \otimes \widehat{\mathbf{N}}_{\bar{b}} \\ & + \frac{1}{2} \sum_{\bar{a},\,\bar{b}=1}^{3} \sum_{\substack{c,\,d=1 \\ d \neq c}}^{3} \mathbb{C}_{\bar{a}\bar{b}cd} \Omega_{cd} \left(\lambda_d^2 - \lambda_c^2 \right) \widehat{\mathbf{N}}_{\bar{a}} \otimes \widehat{\mathbf{N}}_{\bar{b}} \; , \end{aligned} \tag{6.313}$$

and, consequently,

$$\begin{aligned} \underbrace{\widehat{\mathbf{N}}_a \cdot \left[\mathbb{C} : \frac{1}{2} \dot{\mathbf{C}} \right] \widehat{\mathbf{N}}_b}_{a,\; b \,=\, 1,2,3} = & \sum_{c=1}^{3} \mathbb{C}_{abcc} \lambda_c \dot{\lambda}_c \\ & + \sum_{\substack{c,\,d=1 \\ c \neq d}}^{3} \frac{1}{2} \mathbb{C}_{abcd} \Omega_{cd} \left(\lambda_d^2 - \lambda_c^2 \right) \; . \end{aligned} \tag{6.314}$$

From (6.310), one can also have

$$\underbrace{\widehat{\mathbf{N}}_a . \left[\dot{\mathbf{S}} \right] \widehat{\mathbf{N}}_b}_{a,\, b \,=\, 1,2,3} = \sum_{c=1}^{3} \frac{\partial S_a}{\partial \lambda_c} \dot{\lambda}_c \delta_{ab} + \underbrace{\Omega_{ab} \left(S_b - S_a \right)}_{\text{no summation}} \ . \tag{6.315}$$

In what follows, the goal is to determine the spectral components of $\mathbb{C}$ by enforcing $\widehat{\mathbf{N}}_a \cdot \left[\mathbb{C} : \left(\dot{\mathbf{C}}/2 \right) - \dot{\mathbf{S}} \right] \widehat{\mathbf{N}}_b = 0$. To do so, one needs to consider the following cases:
(i) If $a = b$:

$$\sum_{c=1}^{3} \left[\mathbb{C}_{aacc} \lambda_c - \frac{\partial S_a}{\partial \lambda_c} \right] \dot{\lambda}_c + \sum_{\substack{c,\, d=1 \\ c \neq d}}^{3} \frac{1}{2} \mathbb{C}_{aacd} \Omega_{cd} \left(\lambda_d^2 - \lambda_c^2 \right) = 0 \ .$$

Since $\dot{\lambda}_c$ and Ω_{cd} are arbitrary in the above equation, one can obtain

$$\underbrace{\mathbb{C}_{aacc} = \frac{1}{\lambda_c} \frac{\partial S_a}{\partial \lambda_c}}_{= \frac{1}{\lambda_c} \frac{\partial}{\partial \lambda_c} \left(\frac{1}{\lambda_a} \frac{\partial \Psi_\lambda}{\partial \lambda_a} \right) = - \frac{1}{\lambda_a^2 \lambda_c} \delta_{ac} + \frac{1}{\lambda_a \lambda_c} \frac{\partial^2 \Psi_\lambda}{\partial \lambda_a \partial \lambda_c} = \mathbb{C}_{ccaa}} \quad , \quad \underbrace{\mathbb{C}_{aacd} = 0}_{\text{when } c \neq d} \ . \tag{6.316}$$

(ii) If $a \neq b$:

$$\sum_{c=1}^{3} \mathbb{C}_{abcc} \lambda_c \dot{\lambda}_c + \sum_{\substack{c,\, d=1 \\ c \neq d}}^{3} \frac{1}{2} \mathbb{C}_{abcd} \Omega_{cd} \left(\lambda_d^2 - \lambda_c^2 \right) = \Omega_{ab} \left(S_b - S_a \right) \ .$$

Considering the arbitrariness of $\dot{\lambda}_c$ then implies that

$$\mathbb{C}_{abcc} = 0 \ . \tag{6.317}$$

Subsequently, the fact that Ω_{cd} is arbitrary helps conclude that

$$\text{when } c = a \ , \ d = b : \quad \mathbb{C}_{abab} = \frac{S_b - S_a}{\lambda_b^2 - \lambda_a^2} \ , \tag{6.318}$$

where the minor symmetries of $\mathbb{C}$ have been taken into account. Notice that the components not written in the above relation are identically zero. Consistent with (3.128), the specified components in matrix notation thus render

$$\left[\mathbb{C}_{abcd} \right] = \begin{bmatrix} \mathbb{C}_{1111} & \mathbb{C}_{1122} & \mathbb{C}_{1133} & 0 & 0 & 0 \\ & \mathbb{C}_{2222} & \mathbb{C}_{2233} & 0 & 0 & 0 \\ & & \mathbb{C}_{3333} & 0 & 0 & 0 \\ & & & \mathbb{C}_{2323} & 0 & 0 \\ & sym. & & & \mathbb{C}_{1313} & 0 \\ & & & & & \mathbb{C}_{1212} \end{bmatrix} \ . \tag{6.319}$$

The super-symmetric fourth-order tensor (6.312) finally takes the form

$$\mathbb{C} = \underbrace{\sum_{a,\,b=1}^{3} \frac{1}{\lambda_b} \frac{\partial S_a}{\partial \lambda_b} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{N}}_b}_{:= \, \mathbb{C}_{\text{mat}}, \ \text{note that this } \textbf{material} \text{ portion evolves due to variations of the eigenvalues of the stress tensor}}$$
$$+ \underbrace{\sum_{\substack{a,\,b=1 \\ b \neq a}}^{3} \frac{S_b - S_a}{\lambda_b^2 - \lambda_a^2} \widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b \otimes \left[\widehat{\mathbf{N}}_a \otimes \widehat{\mathbf{N}}_b + \widehat{\mathbf{N}}_b \otimes \widehat{\mathbf{N}}_a \right]}_{:= \, \mathbb{C}_{\text{geo}}, \ \text{note that the evolution of this } \textbf{geometric} \text{ contribution is implied by rotations of the eigenvectors of the stress tensor}} \ . \tag{6.320}$$

This completes the proof. It is then a simple exercise to represent (6.305a) and (6.305c). At the end, the reader is referred to Betsch and Stein [59] for derivation of the spectral form of elasticity tensor in *multiplicative plasticity* at finite deformations. See also Betsch and Steinmann [60] for derivation of the spectral form of tangent operator in generalized eigenvalue problems.

All strain and stress measures as well as tangent moduli introduced in this exercise were expressed in Cartesian coordinates. Here, for completeness, the curvilinear representation of such tensorial variables is briefly discussed. In this context, the ambitious reader is referred to a highly mathematical work by Marsden and Hughes [41] for a detailed exposition. See also Simo and Marsden [61] and Sansour [62] among many others.
To begin with, consider the material (spatial) point $\mathbf{X} = \hat{\mathbf{X}} \left(\Theta^1, \Theta^2, \Theta^3 \right)$ ($\mathbf{x} = \hat{\mathbf{x}} \left(\Theta^1, \Theta^2, \Theta^3 \right)$). And let $\{\mathbf{G}_A\}$ ($\{\mathbf{g}_i\}$) be a set of covariant basis vectors in the reference (current) configuration. In continuum mechanics, the two-point strain measure $\mathbf{F}$ naturally represents a contra-covariant tensor. It is given by

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \frac{\partial \mathbf{x}}{\partial \Theta^i} \frac{\partial \Theta^i}{\partial \mathbf{X}} = \underline{F}^i_{\ .A} \mathbf{g}_i \otimes \mathbf{G}^A \quad \text{where} \quad \underline{F}^i_{\ .A} = \delta^i_A \ .$$

Consistent with this, the rigid rotation tensor $\mathbf{R}$ and the stretch tensors $\mathbf{U}$ and $\mathbf{v}$ can be introduced as contra-covariant tensors:

$$\left. \begin{aligned} \mathbf{R} &= \underline{R}^i_{\ .A} \mathbf{g}_i \otimes \mathbf{G}^A \\ \mathbf{U} &= \underline{U}^A_{\ .B} \mathbf{G}_A \otimes \mathbf{G}^B \\ \mathbf{v} &= \underline{v}^i_{\ .j} \mathbf{g}_i \otimes \mathbf{g}^j \end{aligned} \right\} \ .$$

Now, the right Cauchy-Green strain measure $\mathbf{C}$ renders a covariant tensor

$$\mathbf{C} = \underline{F}^i_{\ .A} g_{ij} \, \underline{F}^j_{\ .B} \mathbf{G}^A \otimes \mathbf{G}^B \quad \text{where} \quad g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j \ ,$$

while the Finger tensor $\mathbf{b}$ is a contravariant tensor

$$\mathbf{b} = \underline{F}^{i}_{\,.\,A} G^{AB} \, \underline{F}^{j}_{\,.\,B} \mathbf{g}_i \otimes \mathbf{g}_j \quad \text{where} \quad G^{AB} = \mathbf{G}^A \cdot \mathbf{G}^B \ .$$

noting that, for consistency, they are sometimes written by $\mathbf{C} = \mathbf{F}^{\mathrm{T}} \mathbf{g} \mathbf{F}$ and $\mathbf{b} = \mathbf{F} \mathbf{G}^{-1} \mathbf{F}^{\mathrm{T}}$.
In this context, the free-energy function consistently depends on the metric tensors in addition to the strain measures. Consider an isotropic material and let

$$\begin{aligned} \Psi &= \Psi_1 \left[\mathbf{g}, \mathbf{F}, \mathbf{G} \right] \\ &= \Psi_2 \left[\mathbf{C} \left(\mathbf{g}, \mathbf{F} \right), \mathbf{G} \right] \\ &= \Psi_3 \left[\mathbf{g}, \mathbf{b} \left(\mathbf{F}, \mathbf{G} \right) \right] \ . \end{aligned}$$

Then, the popular stress tensors

$$\begin{aligned} \mathbf{P} &= \frac{\partial \Psi_1}{\partial \mathbf{F}} = \frac{\partial \Psi_1}{\partial \, \underline{F}^{i}_{\,.\,A}} \mathbf{g}^i \otimes \mathbf{G}_A \ , \\ \mathbf{S} &= 2 \frac{\partial \Psi_2}{\partial \mathbf{C}} = 2 \frac{\partial \Psi_2}{\partial \, \underline{C}_{AB}} \mathbf{G}_A \otimes \mathbf{G}_B \ , \\ \boldsymbol{\tau} &= 2 \mathbf{g}^{-1} \frac{\partial \Psi_3}{\partial \mathbf{b}} \mathbf{b} = 2 g^{ik} \frac{\partial \Psi_3}{\partial \, \underline{b}^{kl}} \, \underline{b}^{lj} \mathbf{g}_i \otimes \mathbf{g}_j \ , \end{aligned}$$

are related by $\mathbf{g}^{-1} \mathbf{P} = \mathbf{F} \mathbf{S} = \boldsymbol{\tau} \mathbf{F}^{-\mathrm{T}}$. Note that the Kirchhoff stress tensor can also be written as

$$\boldsymbol{\tau} = 2 \frac{\partial \Psi_3}{\partial \mathbf{g}} = 2 \frac{\partial \Psi_3}{\partial g_{ij}} \mathbf{g}_i \otimes \mathbf{g}_j \ ,$$

which is known as the *Doyle-Ericksen formula*.
Next, the elasticity tensors are given by

$$\mathbb{A} = \frac{\partial^2 \Psi_1}{\partial \mathbf{F} \partial \mathbf{F}} \quad , \quad \mathbb{C} = 4 \frac{\partial^2 \Psi_2}{\partial \mathbf{C} \partial \mathbf{C}} \quad , \quad J \mathbb{c} = 4 \mathbf{b} \frac{\partial^2 \Psi_3}{\partial \mathbf{b} \partial \mathbf{b}} \mathbf{b} = 4 \frac{\partial^2 \Psi_3}{\partial \mathbf{g} \partial \mathbf{g}} \ .$$

At the end, it should not be difficult to establish the relationships

$$\begin{aligned} \underline{\mathbb{A}}_{i\ .\ k\ .}^{\ .\ B\ .\ D} &= g_{ik} \, \underline{S}^{BD} + \underline{\mathbb{C}}^{ABCD} \, \underline{F}^{j}_{\,.\,A} \, \underline{F}^{l}_{\,.\,C} \, g_{ij} g_{kl} \\ &= g_{ik} \left(\underline{F}^{-1} \right)^{B}_{\,.\,j} \underline{\tau}^{jl} \left(\underline{F}^{-1} \right)^{D}_{\,.\,l} + J \, \underline{\mathbb{c}}^{mjnl} \left(\underline{F}^{-1} \right)^{B}_{\,.\,j} \left(\underline{F}^{-1} \right)^{D}_{\,.\,l} g_{im} g_{kn} \ . \end{aligned}$$

Exercise 6.17

The goal of this exercise is to implement numerical differentiation of tensor-valued tensor functions of one symmetric tensor variable by using various difference schemes introduced in Sect. 6.1.5. Suppose one is given

$$\Psi\left(\mathbf{C}\right) = \frac{\lambda}{2}\left(\ln J\right)^2 - \mu \ln J + \frac{\mu}{2}\left(\text{tr}\mathbf{C} - 3\right) \ , \tag{6.321}$$

where the matrix form of this tensor, $[\mathbf{C}]$, and the square root of its determinant, $J = \sqrt{\det\left[\mathbf{C}\right]}$, render

$$[\mathbf{C}] = \begin{bmatrix} 1.2 & 0.54 & 0.78 \\ 0.54 & 2.1 & 1.9 \\ 0.78 & 1.9 & 3.9 \end{bmatrix} = \left[\mathbf{C}^{\mathrm{T}}\right] \quad , \quad J = 2.1637 \ ,$$

at a given material point of a continuum body. In (6.321), the constants $\lambda = 100$ Mpa and $\mu = 1$ Mpa present the Lamé parameters, see (3.158) and (3.159). In the context of nonlinear solid mechanics, the isotropic scalar-valued function (6.321) renders a simple **hyperelastic model** known as the *compressible neo-Hookean material*, see Simo and Pister [63].

First, show that the gradients in (6.74) now take the form

$$\mathbf{S} = 2\frac{\partial\Psi}{\partial\mathbf{C}} = \lambda\left(\ln J\right)\mathbf{C}^{-1} + \mu\left(\mathbf{I} - \mathbf{C}^{-1}\right) \ , \tag{6.322a}$$

$$\mathbb{C} = 2\frac{\partial\mathbf{S}}{\partial\mathbf{C}} = \lambda\left[\mathbf{C}^{-1}\otimes\mathbf{C}^{-1} - \left(2\ln J\right)\mathbf{C}^{-1}\odot\mathbf{C}^{-1}\right] + 2\mu\mathbf{C}^{-1}\odot\mathbf{C}^{-1} \ . \tag{6.322b}$$

Then, write a **computer program** to approximate these analytical derivatives by use of (6.78) and (6.80a)–(6.81c). At the end of the code, evaluate the total error of each approximation by means of (6.82).

Solution. With the aid of $(6.20\text{a})_2$, $(6.20\text{c})_4$, (6.63b) and (6.190d), the desired relations (6.322a)–(6.322b) are verified as follows:

$$\mathbf{S} = \cancel{2}\lambda\left(\ln J\right)\underbrace{\frac{\partial\left(\ln J\right)}{\partial J^2}}_{=1/\left(\cancel{2}\cancel{J^2}\right)}\underbrace{\frac{\partial J^2}{\partial\mathbf{C}}}_{=\cancel{J^2}\mathbf{C}^{-1}} - \cancel{2}\mu\underbrace{\frac{\partial\left(\ln J\right)}{\partial J^2}}_{=1/\left(\cancel{2}\cancel{J^2}\right)}\underbrace{\frac{\partial J^2}{\partial\mathbf{C}}}_{=\cancel{J^2}\mathbf{C}^{-1}} + \mu\underbrace{\frac{\partial\text{tr}\mathbf{C}}{\partial\mathbf{C}}}_{=\mathbf{I}} \ ,$$

$$\mathbb{C} = \lambda\mathbf{C}^{-1}\otimes\underbrace{\frac{\partial\left(2\ln J\right)}{\partial\mathbf{C}}}_{=\mathbf{C}^{-1}} + 2\lambda\left(\ln J\right)\underbrace{\frac{\partial\mathbf{C}^{-1}}{\partial\mathbf{C}}}_{=-\mathbf{C}^{-1}\odot\mathbf{C}^{-1}} - 2\mu\underbrace{\frac{\partial\mathbf{C}^{-1}}{\partial\mathbf{C}}}_{=-\mathbf{C}^{-1}\odot\mathbf{C}^{-1}} \ .$$

The code is available at https://data.uni-hannover.de/dataset/exercises-tensor-analysis for free.

References

1. Hughes, T. J. R., & Pister, K. S. (1978). Consistent linearization in mechanics of solids and structures. *Computers & Structures, 8*, 391–397.
2. Simo, J. C., & Taylor, R. L. (1985). Consistent tangent operators for rate-independent elastoplasticity. *Computer Methods in Applied Mechanics and Engineering, 48*, 101–118.
3. Miehe, C. (1996). Numerical computation of algorithmic (consistent) tangent moduli in large-strain computational inelasticity. *Computer Methods in Applied Mechanics and Engineering, 134*, 223–240.
4. Pérez-Foguet, A., Rodríguez-Ferran, A., & Huerta, A. (2000). Numerical differentiation for non-trivial consistent tangent matrices: An application to the MRS-Lade model. *International Journal for Numerical Methods in Engineering, 48*, 159–184.
5. Temizer, İ, & Wriggers, P. (2008). On the computation of the macroscopic tangent for multiscale volumetric homogenization problems. *Computer Methods in Applied Mechanics and Engineering, 198*, 495–510.
6. Rivlin, R. S. (1955). Further remarks on the stress-deformation relations for isotropic materials. *Journal of Rational Mechanics and Analysis, 4*, 681–702.
7. Pipkin, A. C., & Rivlin, R. S. (1959). The formulation of constitutive equations in continuum physics. I. *Archive for Rational Mechanics and Analysis, 4*, 129–144.
8. Spencer, A. J. M., & Rivlin, R. S. (1958). The theory of matrix polynomials and its application to the mechanics of isotropic continua. *Archive for Rational Mechanics and Analysis, 2*, 309–336.
9. Spencer, A. J. M., & Rivlin, R. S. (1958). Finite integrity bases for five or fewer symmetric 3×3 matrices. *Archive for Rational Mechanics and Analysis, 2*, 435–446.
10. Spencer, A. J. M., & Rivlin, R. S. (1959). Further results in the theory of matrix polynomials. *Archive for Rational Mechanics and Analysis, 4*, 214–230.
11. Spencer, A. J. M., & Rivlin, R. S. (1962). Isotropic integrity bases for vectors and second-order tensors: Part I. *Archive for Rational Mechanics and Analysis, 9*, 45–63.
12. Spencer, A. J. M. (1961). The invariants of six symmetric 3×3 matrices. *Archive for Rational Mechanics and Analysis, 7*, 64–77.
13. Spencer, A.J.M. (1971). *Theory of invariants*. In: A. C. Eringen (Ed.) (Vol. 1, pp. 239–353). Continuum Physics. Academic Press
14. Smith, G. F. (1971). On isotropic functions of symmetric tensors, skew-symmetric tensors and vectors. *International Journal of Engineering Science, 9*, 899–916.
15. Wang, C.-C. (1969). On representations for isotropic functions: Part I. Isotropic functions of symmetric tensors and vectors. *Archive for Rational Mechanics and Analysis, 33*, 249–267.
16. Wang, C.-C. (1969). On representations for isotropic functions: Part II. Isotropic functions of skew-symmetric tensors, symmetric tensors, and vectors. *Archive for Rational Mechanics and Analysis, 33*, 268–287.
17. Betten, J. (1982). Integrity basis for a second-order and a fourth-order tensor. *International Journal of Mathematics and Mathematical Sciences, 5*, 87–96.
18. Betten, J. (1987). Irreducible invariants of fourth-order tensors. *Mathematical Modelling, 8*, 29–33.
19. Boehler, J. P. (1977). On Irreducible Representations for Isotropic Scalar Functions. *Zamm, 57*, 323–327.
20. Boehler, J. P. (1987). *Applications of tensor functions in solid mechanics*. Vienna: Springer.
21. Liu, I-Shih. (1982). On representations of anisotropic invariants. *International Journal of Engineering Science, 20*, 1099–1109.
22. Zheng, Q. S. (1993). On transversely isotropic, orthotropic and relative isotropic functions of symmetric tensors, skew-symmetric tensors and vectors. Part I: Two dimensional orthotropic and relative isotropic functions and three dimensional relative isotropic functions. *International Journal of Engineering Science, 31*, 1399–1409.
23. Liu, I. -S. (2002). *Continuum mechanics*. Berlin, Heidelberg: Springer.

24. Ibrahimbegović, A. (1994). Equivalent spatial and material descriptions of finite deformation elastoplasticity in principal axes. *International Journal of Solids and Structures, 31*, 3027–3040.
25. Simo, J. C., & Taylor, R. L. (1991). Quasi-incompressible finite elasticity in principal stretches: Continuum basis and numerical algorithms. *Computer Methods in Applied Mechanics and Engineering, 85*, 273–310.
26. Saleeb, A. F., Chang, T. Y. P., & Arnold, S. M. (1992). On the development of explicit robust schemes for implementation of a class of hyperelastic models in large-strain analysis of rubbers. *International Journal for Numerical Methods in Engineering, 33*, 1237–1249.
27. de Souza Neto, E. A., Perić, D., & Owen, D. R. J. (2008). *Computational methods for plasticity: Theory and applications*. Wiley.
28. Hoger, A., & Carlson, D. E. (1984). Determination of the stretch and rotation in the polar decomposition of the deformation gradient. *Quarterly of Applied Mathematics, 42*, 113–117.
29. Zheng, Q. S. (1994). Theory of representations for tensor functions-a unified invariant approach to constitutive equations. *Applied Mechanics Reviews, 47*, 545–587.
30. Scheidler, M. J. (1994). The tensor equation AX+XA=Φ(A, H), with applications to kinematics of continua. *Journal of Elasticity, 36*, 117–153.
31. Dui, G., & Chen, Y.-C. (2004). A note on Rivlin's identities and their extension. *Journal of Elasticity, 76*, 107–112.
32. Itskov, M. (2015). *Tensor algebra and tensor analysis for engineers: With applications to continuum mechanics* (4th ed.). Cham: Springer.
33. Lew, J. S. (1966). The generalized Cayley-Hamilton theorem in n dimensions. *Zeitschrift für angewandte Mathematik und Physik (ZAMP), 17*, 650–653.
34. Smith, G. F. (1970). On a fundamental error in two papers of C.-C. Wang, "On representations for isotropic functions, parts I and II." *Archive for Rational Mechanics and Analysis, 36*, 161–165.
35. Wang, C.-C. (1970). A new representation theorem for isotropic functions: An answer to Professor G. F. Smith's criticism of my papers on representations for isotropic functions. *Archive for Rational Mechanics and Analysis, 36*, 166–197.
36. Truesdell, C., & Toupin, R. (1960). The classical field theories. In: S. S. Flügge (Ed.), *Encyclopedia of physics* (Vol. III/1, pp. 226–793). Berlin, Heidelberg: Springer.
37. Eringen, A. C. (1962). *Nonlinear theory of continuous media*. New York, London: McGraw-Hill.
38. Truesdell, C., & Noll, W. (1965). The non-linear field theories of mechanics. In: S. Flügge (Ed.), *Encyclopedia of physics* (Vol. III/3, pp. 1–603). Berlin, Heidelberg: Springer.
39. Eringen, A.C. (1975). *Continuum physics. Volume 2-Continuum mechanics of single-substance bodies*. Academic Press.
40. Truesdell, C. (1977). *A first course in rational continuum mechanics* (Vol. 1). Academic Press.
41. Marsden, J. E., & Hughes, T. J. R. (1983). *Mathematical foundations of elasticity*. Englewood Cliffs, New Jersey: Prentice-Hall.
42. Ogden, R. W. (1984). *Non-linear elastic deformations*. Chichester: Ellis Horwood.
43. Chadwick, P. (1999). *Continuum mechanics: Concise theory and problems, second corrected and enlarged edition*. Mineola, New York: Dover Publications.
44. Holzapfel, G. A. (2000). *Nonlinear solid mechanics: A continuum approach for engineering*. Wiley.
45. Ibrahimbegović, A. (2009). *Nonlinear solid mechanics: Theoretical formulations and finite element solution methods*. Dordrecht: Springer.
46. Steinmann, P. (2015). *Geometrical foundations of continuum mechanics: An application to first- and second-order elasticity and Elasto-plasticity*. Lecture notes in applied mathematics and mechanicsBerlin, Heidelberg: Springer.
47. Oden, J. T. (1972). *Finite elements of nonlinear continua*. New York: McGraw-Hill.
48. Zienkiewicz, O. C., & Taylor, R.L. (1991). *The finite element method, fourth edition, volume 2: Solid and fluid mechanics, dynamics and non-linearity*. McGraw-Hill, London.
49. Bathe, K. J. (1996). *Finite element procedures*. Englewood Cliffs, New Jersey: Prentice-Hall.

50. Crisfield, M. A. (1997). *Non-linear finite element analysis of solids and structures, volume 2: Advanced topics*. Wiley.
51. Bonet, J., & Wood, R. D. (1997). *Nonlinear continuum mechanics for finite element analysis*. New York: Cambridge University Press.
52. Belytschko, T., Liu, W. K., & Moran, B. (2000). *Nonlinear finite elements for continua and structures*. Wiley.
53. Wriggers, P. (2008). *Nonlinear finite element methods*. Berlin, Heidelberg: Springer.
54. Ball, J. M. (1976). Convexity conditions and existence theorems in nonlinear elasticity. *Archive for Rational Mechanics and Analysis, 63*, 337–403.
55. Šilhavý, M. (1997). *The mechanics and thermodynamics of continuous media*. Berlin, Heidelberg: Springer.
56. Atluri, S. N. (1984). Alternate stress and conjugate strain measures, and mixed variational formulations involving rigid rotations, for computational analyses of finitely deformed solids, with application to plates and shells-I. *Computers & Structures, 18*, 93–116.
57. Gurtin, M. E., Fried, E., & Anand, L. (2010). *The mechanics and thermodynamics of continua*. Cambridge: Cambridge University Press.
58. Miehe, C. (1994). Aspects of the formulation and finite element implementation of large strain isotropic elasticity. *International Journal for Numerical Methods in Engineering, 37*, 1981–2004.
59. Betsch, P., & Stein, E. (1999). Numerical implementation of multiplicative Elasto-plasticity into assumed strain elements with application to shells at large strains. *Computer Methods in Applied Mechanics and Engineering, 179*, 215–245.
60. Betsch, P., & Steinmann, P. (2000). Derivation of the fourth-order tangent operator based on a generalized eigenvalue problem. *International Journal of Solids and Structures, 37*, 1615–1628.
61. Simo, J. C., & Marsden, J. E. (1984). On the rotated stress tensor and the material version of the Doyle-Ericksen formula. *Archive for Rational Mechanics and Analysis, 86*, 213–231.
62. Sansour, C. (1993). On the spatial description in elasticity and the Doyle-Ericksen formula. *Computer Methods in Applied Mechanics and Engineering, 107*, 239–249.
63. Simo, J. C., & Pister, K. S. (1984). Remarks on rate constitutive equations for finite deformation problems: Computational implications. *Computer Methods in Applied Mechanics and Engineering, 46*, 201–215.

Chapter 7
Gradient and Related Operators

In a continuation of vector and tensor calculus, this chapter mainly studies the actions of **gradient**, **divergence** and **curl** operators on vectors and tensors. Needless to say that these differential operators are the workhorses of vector and tensor analysis. Recall that Chap. 6 dealt with the gradients of tensor functions which were non-constant tensorial variables. Here, any tensorial variable is assumed to be a function of **position** (and possibly **time**). Functions depending on space and time are ubiquitous in science and engineering. Examples of which include electric field in electromagnetism and velocity field in continuum mechanics.

Scalar, vector and tensor fields. The term *field* is used to designate a function defined over space and time. In general, one needs to distinguish between a

- scalar-valued function $\bar{h} : \mathbf{x} \in \mathscr{E}_{\mathrm{p}}^3 \times t \in \mathrm{R} \to \bar{h}(\mathbf{x}, t) \in \mathrm{R}$, called a **scalar field**,
- vector-valued function $\hat{\mathbf{h}} : \mathbf{x} \in \mathscr{E}_{\mathrm{p}}^3 \times t \in \mathrm{R} \to \hat{\mathbf{h}}(\mathbf{x}, t) \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}$, called a **vector field** and
- tensor-valued function $\tilde{\mathbf{H}} : \mathbf{x} \in \mathscr{E}_{\mathrm{p}}^3 \times t \in \mathrm{R} \to \tilde{\mathbf{H}}(\mathbf{x}, t) \in \mathscr{T}_{\mathrm{so}}\left(\mathscr{E}_{\mathrm{r}}^{\mathrm{o}3}\right)$, called a **tensor field**.

A field variable that is independent of time is referred to as a *stationary* or *steady-state* field variable. And, it is said to be *homogeneous* or *uniform* if it is only a function of time. Of special interest in this chapter is to only consider stationary fields.

7.1 Differentiation of Fields

Let $\bar{h}(\mathbf{x})$, $\hat{\mathbf{h}}(\mathbf{x})$ and $\tilde{\mathbf{H}}(\mathbf{x})$ be a nonlinear and sufficiently smooth scalar, vector and tensor field, respectively. The first-order Taylor series expansions of these three fields at $\mathbf{x}$ are given by

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_7

$$\bar{h}\left(\mathbf{x}+d\mathbf{x}\right)=\bar{h}\left(\mathbf{x}\right)+d\bar{h}+\bar{o}\left(d\mathbf{x}\right)\ , \tag{7.1a}$$

$$\hat{\mathbf{h}}\left(\mathbf{x}+d\mathbf{x}\right)=\hat{\mathbf{h}}\left(\mathbf{x}\right)+d\hat{\mathbf{h}}+\hat{\mathbf{o}}\left(d\mathbf{x}\right)\ , \tag{7.1b}$$

$$\tilde{\mathbf{H}}\left(\mathbf{x}+d\mathbf{x}\right)=\tilde{\mathbf{H}}\left(\mathbf{x}\right)+d\tilde{\mathbf{H}}+\tilde{\mathbf{o}}\left(d\mathbf{x}\right)\ , \tag{7.1c}$$

where the total differentials $d\bar{h}$, $d\hat{\mathbf{h}}$ and $d\tilde{\mathbf{H}}$ are

$$\boxed{d\bar{h}=\frac{\partial\bar{h}}{\partial\mathbf{x}}\cdot d\mathbf{x}\quad ,\quad d\hat{\mathbf{h}}=\frac{\partial\hat{\mathbf{h}}}{\partial\mathbf{x}}d\mathbf{x}\quad ,\quad d\tilde{\mathbf{H}}=\frac{\partial\tilde{\mathbf{H}}}{\partial\mathbf{x}}d\mathbf{x}\ ,} \tag{7.2}$$

and the Landau order symbols $\bar{o}\left(d\mathbf{x}\right)$, $\hat{\mathbf{o}}\left(d\mathbf{x}\right)$ and $\tilde{\mathbf{o}}\left(d\mathbf{x}\right)$ will tend to zero faster than $d\mathbf{x}\to\mathbf{0}$. This is indicated by

$$\lim_{d\mathbf{x}\to\mathbf{0}}\frac{\bar{o}\left(d\mathbf{x}\right)}{|d\mathbf{x}|}=0\quad ,\quad \lim_{d\mathbf{x}\to\mathbf{0}}\frac{\hat{\mathbf{o}}\left(d\mathbf{x}\right)}{|d\mathbf{x}|}=\mathbf{0}\quad ,\quad \lim_{d\mathbf{x}\to\mathbf{0}}\frac{\tilde{\mathbf{o}}\left(d\mathbf{x}\right)}{|d\mathbf{x}|}=\mathbf{O}\ . \tag{7.3}$$

In (7.2), the first-order tensor $\partial\bar{h}/\partial\mathbf{x}$ denotes the gradient (or derivative) of $\bar{h}$ at $\mathbf{x}$. And the second-order (third-order) tensor $\partial\hat{\mathbf{h}}/\partial\mathbf{x}$ ($\partial\tilde{\mathbf{H}}/\partial\mathbf{x}$) presents the gradient of $\hat{\mathbf{h}}$ ($\tilde{\mathbf{H}}$) at $\mathbf{x}$. Note that any of these objects can be determined from its directional derivative as follows:

$$\begin{aligned} D_{\mathbf{v}}\bar{h}\left(\mathbf{x}\right) &= \frac{d}{d\varepsilon}\bar{h}\left(\mathbf{x}+\varepsilon\mathbf{v}\right)\Big|_{\varepsilon=0}=\left[\frac{\partial\bar{h}}{\partial\left(\mathbf{x}+\varepsilon\mathbf{v}\right)}\cdot\frac{\partial\left(\mathbf{x}+\varepsilon\mathbf{v}\right)}{\partial\varepsilon}\right]_{\varepsilon=0} \\ &= \frac{\partial\bar{h}}{\partial\mathbf{x}}\cdot\mathbf{v}\ , \end{aligned} \tag{7.4a}$$

$$\begin{aligned} D_{\mathbf{v}}\hat{\mathbf{h}}\left(\mathbf{x}\right) &= \frac{d}{d\varepsilon}\hat{\mathbf{h}}\left(\mathbf{x}+\varepsilon\mathbf{v}\right)\Big|_{\varepsilon=0}=\left[\frac{\partial\hat{\mathbf{h}}}{\partial\left(\mathbf{x}+\varepsilon\mathbf{v}\right)}\frac{\partial\left(\mathbf{x}+\varepsilon\mathbf{v}\right)}{\partial\varepsilon}\right]_{\varepsilon=0} \\ &= \frac{\partial\hat{\mathbf{h}}}{\partial\mathbf{x}}\mathbf{v}\ , \end{aligned} \tag{7.4b}$$

$$\begin{aligned} D_{\mathbf{v}}\tilde{\mathbf{H}}\left(\mathbf{x}\right) &= \frac{d}{d\varepsilon}\tilde{\mathbf{H}}\left(\mathbf{x}+\varepsilon\mathbf{v}\right)\Big|_{\varepsilon=0}=\left[\frac{\partial\tilde{\mathbf{H}}}{\partial\left(\mathbf{x}+\varepsilon\mathbf{v}\right)}\frac{\partial\left(\mathbf{x}+\varepsilon\mathbf{v}\right)}{\partial\varepsilon}\right]_{\varepsilon=0} \\ &= \frac{\partial\tilde{\mathbf{H}}}{\partial\mathbf{x}}\mathbf{v}\ , \end{aligned} \tag{7.4c}$$

where the generic vector $\mathbf{v}$ is sometimes introduced as a unit vector in the literature. The ultimate goal here is to express the gradient of a tensorial field variable with respect to the standard and curvilinear bases. Regarding a Cartesian tensor field, this can be done in a straightforward manner. But, the gradient of a curvilinear one requires further consideration since the general basis vectors are no longer constant and their partial derivatives with respect to the curvilinear coordinates need to be naturally taken into account. The problem is thus treated by characterizing the nonzero objects $\partial\mathbf{g}_i/\partial\Theta^j$ that appear, for instance, in

$$\frac{\partial}{\partial \Theta^j} \left[\hat{\mathbf{h}} \right] = \frac{\partial}{\partial \Theta^j} \left[\underline{\hat{h}}^i \mathbf{g}_i \right] = \frac{\partial \underline{\hat{h}}^i}{\partial \Theta^j} \mathbf{g}_i + \underline{\hat{h}}^i \frac{\partial \mathbf{g}_i}{\partial \Theta^j} \ .$$

This leads to the introduction of *Christoffel symbols* described below.

7.1.1 Christoffel Symbols of First and Second Kind

Consider the following family of objects

$$\boldsymbol{\Gamma}_{ij} := \frac{\partial \mathbf{g}_i}{\partial \Theta^j} \ , \tag{7.5}$$

possessing the symmetry in the indices i and j due to

$$\boldsymbol{\Gamma}_{ij} = \frac{\partial \mathbf{g}_i}{\partial \Theta^j} = \frac{\partial^2 \mathbf{x}}{\partial \Theta^j \partial \Theta^i} = \frac{\partial^2 \mathbf{x}}{\partial \Theta^i \partial \Theta^j} = \frac{\partial \mathbf{g}_j}{\partial \Theta^i} = \boldsymbol{\Gamma}_{ji} \ . \tag{7.6}$$

Then, the quantities

$$\boxed{\Gamma^k_{ij} = \frac{\partial \mathbf{g}_i}{\partial \Theta^j} \cdot \mathbf{g}^k \quad \text{satisfying} \quad \Gamma^k_{ij} = \Gamma^k_{ji} \ ,} \tag{7.7}$$

are referred to as *Christoffel symbols of the second kind* or simply *Christoffel symbols*. These 18 independent quantities are strictly properties of the chosen coordinate system. The Christoffel symbols relate the sensitivity of the tangent vectors with respect to the curvilinear coordinates to the covariant basis vectors:

$$\boxed{\frac{\partial \mathbf{g}_i}{\partial \Theta^j} = \Gamma^k_{ij} \mathbf{g}_k \ .} \tag{7.8}$$

They also help connect the partial derivatives of dual vectors to the contravariant basis vectors:

$$\boxed{\frac{\partial \mathbf{g}^i}{\partial \Theta^j} = -\Gamma^i_{jk} \mathbf{g}^k \ .} \tag{7.9}$$

This result eventually relies on

$$\mathbf{g}^i \cdot \mathbf{g}_k = \delta^i_k \quad \Rightarrow \quad \frac{\partial \mathbf{g}^i}{\partial \Theta^j} \cdot \mathbf{g}_k + \mathbf{g}^i \cdot \frac{\partial \mathbf{g}_k}{\partial \Theta^j} = 0$$

$$\Rightarrow \quad \frac{\partial \mathbf{g}^i}{\partial \Theta^j} \cdot \mathbf{g}_k = -\mathbf{g}^i \cdot \Gamma^l_{kj} \mathbf{g}_l = -\Gamma^i_{kj} \quad \Rightarrow \quad \frac{\partial \mathbf{g}^i}{\partial \Theta^j} \cdot \mathbf{g}_k = -\Gamma^i_{jk} \ .$$

By use of (5.49), i.e. $\mathbf{g}^m = g^{ml}\mathbf{g}_l$, and $(5.51)_1$, i.e. $g_{km}g^{ml} = \delta^l_k$, the superscript index of Γ^m_{ij} can be lowered:

$$\boxed{\underbrace{\Gamma_{ijk} = \Gamma^m_{ij}\, g_{mk}}_{\text{note that } \Gamma^k_{ij} = \Gamma_{ijm} g^{mk}} \quad .} \tag{7.10}$$

The quantities Γ_{ijk} that also possess symmetry in the first two indices are termed *Christoffel symbols of the first kind*. These symbols

$$\boxed{\Gamma_{ijk} = \Gamma_{jik} = \frac{\partial \mathbf{g}_i}{\partial \Theta^j} \cdot \mathbf{g}_k \ ,} \tag{7.11}$$

basically relate the partial derivatives of tangent vectors to the dual vectors according to

$$\boxed{\frac{\partial \mathbf{g}_i}{\partial \Theta^j} = \Gamma_{ijk}\mathbf{g}^k \ .} \tag{7.12}$$

The structure of $(7.11)_2$ motivates to represent the partial derivatives of the covariant metric $(5.38)_1$, i.e. $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$. It follows that

$$\frac{\partial g_{ki}}{\partial \Theta^j} = \Gamma_{kji} + \Gamma_{ijk} \ . \tag{7.13}$$

By considering the symmetry of the Christoffel symbols, one then obtains

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ki}}{\partial \Theta^j} + \frac{\partial g_{kj}}{\partial \Theta^i} - \frac{\partial g_{ij}}{\partial \Theta^k} \right) \ . \tag{7.14}$$

The expressions (7.10) and (7.14), along with the identity $(5.51)_1$, help find out that

$$\boxed{\Gamma^k_{ij} = \frac{1}{2}\left(\frac{\partial g_{mi}}{\partial \Theta^j} + \frac{\partial g_{mj}}{\partial \Theta^i} - \frac{\partial g_{ij}}{\partial \Theta^m} \right) g^{mk} \ .} \tag{7.15}$$

The results (7.14)–(7.15) clearly show that the Christoffel symbols of the first and second kind only depend on the **metric coefficients**. In other words, if g_{ij} are known, then Γ_{ijk} and Γ^k_{ij} can consistently be computed.

Now, it is easy to see that the partial derivatives of the contravariant metric coefficients $g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j$, according to $(5.46)_1$, render

$$\boxed{\frac{\partial g^{ij}}{\partial \Theta^k} = -\Gamma^i_{kl}\, g^{lj} - \Gamma^j_{kl}\, g^{il} \ .} \tag{7.16}$$

Recall from $(5.30)_2$ that the Jacobian is $J = \det\begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 \end{bmatrix}$. The fact that the basis vectors vary from point to point in curvilinear coordinates implies that the Jacobian also varies with position. One then realizes that $J = J\left(\Theta^1, \Theta^2, \Theta^3\right)$. By using (5.47), (5.51)$_{1-2}$, (6.23)$_2$, (7.7)$_2$ and (7.16), along with the chain rule of differentiation, one can now establish the identity

$$\boxed{\underbrace{\Gamma^i_{ki} = \Gamma^i_{ik} = \frac{\partial \ln J}{\partial \Theta^k}}_{\text{or } \partial J/\partial\Theta^k = J\Gamma^i_{ki} = J\Gamma^i_{ik}} \,,} \tag{7.17}$$

owing to

$$\underbrace{\frac{\partial \ln J}{\partial \Theta^k}}_{= \frac{1}{J}\frac{\partial J}{\partial \Theta^k}} = \underbrace{\frac{1}{J}\frac{\partial J}{\partial g^{ij}}\frac{\partial g^{ij}}{\partial \Theta^k}}_{= \frac{1}{J}\left(-\frac{J}{2}g_{ij}\right)\left(-\Gamma^i_{kl}g^{lj} - \Gamma^j_{kl}g^{il}\right)} = \underbrace{\frac{1}{2}\Gamma^i_{kl}\delta^l_i + \frac{1}{2}\Gamma^j_{kl}\delta^l_j}_{= \frac{1}{2}\Gamma^i_{ki} + \frac{1}{2}\Gamma^j_{kj}} = \Gamma^i_{ki} \,.$$

Another useful identity is

$$\boxed{\frac{\partial\left(J\mathbf{g}^i\right)}{\partial \Theta^i} = \mathbf{0} \,,} \tag{7.18}$$

because

$$\begin{aligned}
\frac{\partial\left(J\mathbf{g}^i\right)}{\partial \Theta^i} &\xlongequal[\text{the product rule}]{\text{by using}} \frac{\partial J}{\partial \Theta^i}\mathbf{g}^i + J\frac{\partial \mathbf{g}^i}{\partial \Theta^i} \\
&\xlongequal[\text{(7.9) and (7.17)}]{\text{by using}} J\Gamma^k_{ik}\mathbf{g}^i - J\Gamma^i_{ik}\mathbf{g}^k \\
&\xlongequal[\text{and switching the names of dummy indices}]{\text{by using the symmetry of Christoffel symbols}} J\Gamma^k_{ki}\left[\mathbf{g}^i - \mathbf{g}^i\right] \\
&\xlongequal[\text{(1.4c), (1.4d), (1.5) and (a) in (1.76)}]{\text{by using}} \mathbf{0} \,.
\end{aligned}$$

As an example, consider the following nonzero Christoffel symbols for **cylindrical coordinates**:

$$\Gamma^1_{22} = -r \quad , \quad \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r} \,, \tag{7.19}$$

where the tangent vectors (5.7a)–(5.7c) along with the dual vectors (5.118a)–(5.118c) have been used.

Another example regards the nonzero Christoffel elements for **spherical coordinates**. Considering (5.11a)–(5.11c) and (5.121a)–(5.121c), they are given by

$$\Gamma^1_{22} = -r \quad , \quad \Gamma^1_{33} = -r\sin^2\theta \,, \tag{7.20a}$$

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r} \quad , \quad \Gamma^2_{33} = -\frac{1}{2}\sin 2\theta \,, \tag{7.20b}$$

$$\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r} \quad , \quad \Gamma^3_{23} = \Gamma^3_{32} = \cot\theta \ . \tag{7.20c}$$

Hint: The Christoffel symbols are **basis-dependent** quantities which carry information about the variation of a curvilinear basis in space. Although they are perfectly characterized by three indices, they should not be considered as the components of a third-order tensor. The reason is that they do not obey the tensor transformation laws explained in Sect. 2.9. This will be discussed in the following. ●

Given an old basis $\{\mathbf{g}_i\}$ and a new basis $\{\bar{\mathbf{g}}_i\}$ with the corresponding Christoffel symbols Γ^k_{ij} and $\bar{\Gamma}^k_{ij}$. Then,

$$\bar{\Gamma}^k_{ij} \neq \left(\mathbf{g}^l \cdot \bar{\mathbf{g}}_i\right)\left(\mathbf{g}^m \cdot \bar{\mathbf{g}}_j\right)\left(\bar{\mathbf{g}}^k \cdot \mathbf{g}_n\right)\Gamma^n_{lm} \ , \tag{7.21}$$

where, using $(5.58a)_2$ and $(5.58b)_2$,

$$\mathbf{g}^l \cdot \bar{\mathbf{g}}_i = \frac{\partial \Theta^l}{\partial \bar{\Theta}^i} \quad , \quad \mathbf{g}^m \cdot \bar{\mathbf{g}}_j = \frac{\partial \Theta^m}{\partial \bar{\Theta}^j} \quad , \quad \bar{\mathbf{g}}^k \cdot \mathbf{g}_n = \frac{\partial \bar{\Theta}^k}{\partial \Theta^n} \ . \tag{7.22}$$

Indeed, these Christoffel symbols are related by

$$\boxed{\bar{\Gamma}^k_{ij} = \underbrace{\frac{\partial \Theta^l}{\partial \bar{\Theta}^i}\frac{\partial \Theta^m}{\partial \bar{\Theta}^j}\frac{\partial \bar{\Theta}^k}{\partial \Theta^n}\Gamma^n_{lm}}_{\text{tensorial contribution}} + \underbrace{\frac{\partial^2 \Theta^n}{\partial \bar{\Theta}^i \bar{\Theta}^j}\frac{\partial \bar{\Theta}^k}{\partial \Theta^n}}_{\text{nontensorial part}} \ ,} \tag{7.23}$$

since, by means of the product and chain rules of differentiation along with (1.9a)–(1.9c), $(2.158)_1$, (5.14), $(5.27)_1$ $(7.7)_1$ and $(7.22)_{1-3}$,

$$\begin{aligned}
\bar{\Gamma}^k_{ij} &= \frac{\partial\left[\bar{\mathbf{g}}_i\right]}{\partial \bar{\Theta}^j} \cdot \left[\bar{\mathbf{g}}^k\right] \\
&= \frac{\partial\left[\left(\mathbf{g}^l \cdot \bar{\mathbf{g}}_i\right)\mathbf{g}_l\right]}{\partial \bar{\Theta}^j} \cdot \left[\left(\mathbf{g}_n \cdot \bar{\mathbf{g}}^k\right)\mathbf{g}^n\right] \\
&= \left(\mathbf{g}_n \cdot \bar{\mathbf{g}}^k\right)\left[\frac{\partial\left(\mathbf{g}^l \cdot \bar{\mathbf{g}}_i\right)}{\partial \bar{\Theta}^j}\delta^n_l + \left(\mathbf{g}^l \cdot \bar{\mathbf{g}}_i\right)\frac{\partial \mathbf{g}_l}{\partial \Theta^m}\frac{\partial \Theta^m}{\partial \bar{\Theta}^j} \cdot \mathbf{g}^n\right] \\
&= \frac{\partial \bar{\Theta}^k}{\partial \Theta^n}\left[\frac{\partial\left(\mathbf{g}^n \cdot \bar{\mathbf{g}}_i\right)}{\partial \bar{\Theta}^j} + \frac{\partial \Theta^l}{\partial \bar{\Theta}^i}\frac{\partial \Theta^m}{\partial \bar{\Theta}^j}\Gamma^n_{lm}\right] \\
&= \frac{\partial \bar{\Theta}^k}{\partial \Theta^n}\left[\frac{\partial \Theta^l}{\partial \bar{\Theta}^i}\frac{\partial \Theta^m}{\partial \bar{\Theta}^j}\Gamma^n_{lm} + \frac{\partial^2 \Theta^n}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i}\right] \ . \quad ●
\end{aligned}$$

The gradient of a tensor field is formulated by means of its partial derivatives with respect to the curvilinear coordinates, see (7.68a)–(7.69b). Representation of the partial derivatives of a tensorial field variable with respect to a basis is thus required. These derivatives are expressed in terms of a special derivative of components which crucially depends on the Christoffel symbols. This new mathematical object will be introduced in the following.

7.1.2 First Covariant Differentiation

Let $\hat{\mathbf{h}} = \hat{\mathbf{h}}\left(\Theta^1, \Theta^2, \Theta^3\right)$ be a generic vector field. Then, by using (5.64a)–(5.64b) and (7.8)–(7.9) together with the product rule of differentiation, the rate of change in this field variable can be expressed with respect to the curvilinear basis vectors as

$$\underbrace{\frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} = \underline{\hat{h}}^i\Big|_j \, \mathbf{g}_i}_{\text{in the literature, this is sometimes denoted by } \nabla_{\mathbf{g}_j}\hat{\mathbf{h}} = \nabla_j \hat{\mathbf{h}} = \left(\nabla_j \underline{\hat{h}}^i\right)\mathbf{g}_i} \quad , \tag{7.24a}$$

$$\underbrace{\frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} = \underline{\hat{h}}_i\Big|_j \, \mathbf{g}^i}_{\text{in the literature, this is sometimes denoted by } \nabla_{\mathbf{g}_j}\hat{\mathbf{h}} = \nabla_j \hat{\mathbf{h}} = \left(\nabla_j \underline{\hat{h}}_i\right)\mathbf{g}^i} \quad , \tag{7.24b}$$

where $\underline{\hat{h}}^i\Big|_j$ and $\underline{\hat{h}}_i\Big|_j$ present the *first-order covariant derivative* (or *differentiation*) of the contravariant and covariant components $\underline{\hat{h}}^i$ and $\underline{\hat{h}}_i$, respectively. These special components are given by

$$\underline{\hat{h}}^i\Big|_j = \frac{\partial \underline{\hat{h}}^i}{\partial \Theta^j} + \Gamma^i_{jm} \, \underline{\hat{h}}^m \ , \tag{7.25a}$$

$$\underline{\hat{h}}_i\Big|_j = \frac{\partial \underline{\hat{h}}_i}{\partial \Theta^j} - \Gamma^m_{ij} \, \underline{\hat{h}}_m \ . \tag{7.25b}$$

The covariant derivative consistently captures the change in a contravariant (covariant) component along with its companion covariant (contravariant) basis vector. In comparison with the partial derivative, it is thus a better measure of the rate of change in a field variable. For more details see, for example, Grinfeld [1].

The coordinate index on the covariant derivative can be raised to define the *contravariant derivative* of either contravariant or covariant components:

$$\underline{\hat{h}}^i\Big|^j = \underline{\hat{h}}^i\Big|_m \, g^{mj} \quad , \quad \underline{\hat{h}}_i\Big|^j = \underline{\hat{h}}_i\Big|_m \, g^{mj} \ . \tag{7.26}$$

By considering (5.73a)–(5.73d) and (7.8)–(7.9) along with the product rule of differentiation, the rate of change in the tensor field $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}\left(\Theta^1, \Theta^2, \Theta^3\right)$ can be decomposed as

$$\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} = \underline{\tilde{H}}^{ij}\Big|_k \mathbf{g}_i \otimes \mathbf{g}_j \quad \text{where} \quad \underline{\tilde{H}}^{ij}\Big|_k = \frac{\partial\, \underline{\tilde{H}}^{ij}}{\partial \Theta^k} + \Gamma^i_{km}\, \underline{\tilde{H}}^{mj} + \Gamma^j_{km}\, \underline{\tilde{H}}^{im} \,, \tag{7.27a}$$

$$\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} = \underline{\tilde{H}}^{i}_{\,.j}\Big|_k \mathbf{g}_i \otimes \mathbf{g}^j \quad \text{where} \quad \underline{\tilde{H}}^{i}_{\,.j}\Big|_k = \frac{\partial\, \underline{\tilde{H}}^{i}_{\,.j}}{\partial \Theta^k} + \Gamma^i_{km}\, \underline{\tilde{H}}^{m}_{\,.j} - \Gamma^m_{kj}\, \underline{\tilde{H}}^{i}_{\,.m} \,, \tag{7.27b}$$

$$\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} = \underline{\tilde{H}}_{i}^{\,.j}\Big|_k \mathbf{g}^i \otimes \mathbf{g}_j \quad \text{where} \quad \underline{\tilde{H}}_{i}^{\,.j}\Big|_k = \frac{\partial\, \underline{\tilde{H}}_{i}^{\,.j}}{\partial \Theta^k} - \Gamma^m_{ki}\, \underline{\tilde{H}}_{m}^{\,.j} + \Gamma^j_{km}\, \underline{\tilde{H}}_{i}^{\,.m} \,, \tag{7.27c}$$

$$\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} = \underline{\tilde{H}}_{ij}\Big|_k \mathbf{g}^i \otimes \mathbf{g}^j \quad \text{where} \quad \underline{\tilde{H}}_{ij}\Big|_k = \frac{\partial\, \underline{\tilde{H}}_{ij}}{\partial \Theta^k} - \Gamma^m_{ki}\, \underline{\tilde{H}}_{mj} - \Gamma^m_{kj}\, \underline{\tilde{H}}_{im} \,. \tag{7.27d}$$

Consistent with (7.26), the contravariant derivative of the curvilinear components of a tensor field can be written as

$$\left.\begin{aligned} \underline{\tilde{H}}^{ij}\Big|^k &= \underline{\tilde{H}}^{ij}\Big|_m g^{mk} \\ \underline{\tilde{H}}^{i}_{\,.j}\Big|^k &= \underline{\tilde{H}}^{i}_{\,.j}\Big|_m g^{mk} \end{aligned}\right\} \,, \quad \left.\begin{aligned} \underline{\tilde{H}}_{i}^{\,.j}\Big|^k &= \underline{\tilde{H}}_{i}^{\,.j}\Big|_m g^{mk} \\ \underline{\tilde{H}}_{ij}\Big|^k &= \underline{\tilde{H}}_{ij}\Big|_m g^{mk} \end{aligned}\right\} \,. \tag{7.28}$$

7.1.3 *Invariance of Covariant Derivative*

The partial derivative operator $\partial/\partial\Theta^i$ preserves the tensor property when applied to the **invariant** field variables such as $\bar{h}$ or $\hat{\mathbf{h}}$ or $\tilde{\mathbf{H}}$. But, this is not the case for variants. For instance, consider the old and new components of the object $\hat{\mathbf{h}}$ which are tensorially related by

$$\underline{\bar{\hat{h}}}^i = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\, \underline{\hat{h}}^j \,, \quad \leftarrow \text{note that } \hat{\mathbf{h}} = \underline{\bar{\hat{h}}}^i \bar{\mathbf{g}}_i = \underline{\hat{h}}^i \mathbf{g}_i = \underline{\hat{h}}^i \left(\mathbf{g}_i \cdot \bar{\mathbf{g}}^j\right) \bar{\mathbf{g}}_j = \underline{\hat{h}}^i \frac{\partial \bar{\Theta}^j}{\partial \Theta^i} \bar{\mathbf{g}}_j = \underline{\hat{h}}^j \frac{\partial \bar{\Theta}^i}{\partial \Theta^j} \bar{\mathbf{g}}_i \,, \tag{7.29a}$$

$$\underline{\bar{\hat{h}}}_i = \frac{\partial \Theta^j}{\partial \bar{\Theta}^i}\, \underline{\hat{h}}_j \,, \quad \leftarrow \text{note that } \hat{\mathbf{h}} = \underline{\bar{\hat{h}}}_i \bar{\mathbf{g}}^i = \underline{\hat{h}}_i \mathbf{g}^i = \underline{\hat{h}}_i \left(\mathbf{g}^i \cdot \bar{\mathbf{g}}_j\right) \bar{\mathbf{g}}^j = \underline{\hat{h}}_i \frac{\partial \Theta^i}{\partial \bar{\Theta}^j} \bar{\mathbf{g}}^j = \underline{\hat{h}}_j \frac{\partial \Theta^j}{\partial \bar{\Theta}^i} \bar{\mathbf{g}}^i \,, \tag{7.29b}$$

whereas their partial derivatives nontensorially transform according to

$$\frac{\partial\, \underline{\bar{\hat{h}}}^i}{\partial \bar{\Theta}^j} = \frac{\partial}{\partial \bar{\Theta}^j}\left[\frac{\partial \bar{\Theta}^i}{\partial \Theta^k}\, \underline{\hat{h}}^k\right]$$

$$= \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \underline{\hat{h}}^k}{\partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \frac{\partial^2 \bar{\Theta}^i}{\partial \Theta^l \partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \underline{\hat{h}}^k \,, \tag{7.30a}$$

$$\frac{\partial \bar{\underline{\hat{h}}}_i}{\partial \bar{\Theta}^j} = \frac{\partial}{\partial \bar{\Theta}^j} \left[\frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \underline{\hat{h}}_k \right]$$

$$= \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \underline{\hat{h}}_k}{\partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i} \underline{\hat{h}}_k \,. \tag{7.30b}$$

The covariant derivative now manifests itself as a powerful differential operator which overcomes this problem:

$$\bar{\underline{\hat{h}}}^i \Big|_j = \underbrace{\frac{\partial \bar{\underline{\hat{h}}}^i}{\partial \bar{\Theta}^j}}_{= \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \underline{\hat{h}}^k}{\partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \frac{\partial^2 \bar{\Theta}^i}{\partial \Theta^l \partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \underline{\hat{h}}^k} + \underbrace{\bar{\Gamma}^i_{jm} \bar{\underline{\hat{h}}}^m}_{= \left(\frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \frac{\partial \Theta^n}{\partial \bar{\Theta}^m} \Gamma^k_{ln} + \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^m \bar{\Theta}^j} \frac{\partial \bar{\Theta}^i}{\partial \Theta^l} \right) \frac{\partial \bar{\Theta}^m}{\partial \Theta^r} \underline{\hat{h}}^r}$$

$$= \frac{\partial \underline{\hat{h}}^k}{\partial \Theta^l} \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \underbrace{\frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \frac{\partial \Theta^n}{\partial \bar{\Theta}^m} \frac{\partial \bar{\Theta}^m}{\partial \Theta^r} \Gamma^k_{ln} \underline{\hat{h}}^r}_{= \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \delta^n_r \Gamma^k_{ln} \underline{\hat{h}}^r = \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \Gamma^k_{ln} \underline{\hat{h}}^n}$$

$$+ \frac{\partial^2 \bar{\Theta}^i}{\partial \Theta^l \partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \underline{\hat{h}}^k + \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^m \bar{\Theta}^j} \frac{\partial \bar{\Theta}^i}{\partial \Theta^l} \frac{\partial \bar{\Theta}^m}{\partial \Theta^r} \underline{\hat{h}}^r$$

$$= \underbrace{\left(\frac{\partial \underline{\hat{h}}^k}{\partial \Theta^l} + \Gamma^k_{ln} \underline{\hat{h}}^n \right)}_{= \underline{\hat{h}}^k \big|_l} \frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \underbrace{\left(\frac{\partial^2 \bar{\Theta}^i}{\partial \Theta^k \partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \frac{\partial \bar{\Theta}^i}{\partial \Theta^l} \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^m \bar{\Theta}^j} \frac{\partial \bar{\Theta}^m}{\partial \Theta^k} \right)}_{= \frac{\partial}{\partial \Theta^k} \left[\frac{\partial \bar{\Theta}^i}{\partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \right] = \frac{\partial}{\partial \Theta^k} \left[\frac{\partial \bar{\Theta}^i}{\partial \bar{\Theta}^j} \right] = \frac{\partial}{\partial \Theta^k} \left[\delta^i_j \right] = 0} \underline{\hat{h}}^k$$

$$= \boxed{\frac{\partial \bar{\Theta}^i}{\partial \Theta^k} \left(\underline{\hat{h}}^k \Big|_l \right) \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \,,} \tag{7.31a}$$

$$\bar{\underline{\hat{h}}}_i \Big|_j = \underbrace{\frac{\partial \bar{\underline{\hat{h}}}_i}{\partial \bar{\Theta}^j}}_{= \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \underline{\hat{h}}_k}{\partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i} \underline{\hat{h}}_k} - \underbrace{\bar{\Gamma}^m_{ij} \bar{\underline{\hat{h}}}_m}_{= \left(\frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \frac{\partial \bar{\Theta}^m}{\partial \Theta^n} \Gamma^n_{kl} + \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^i \bar{\Theta}^j} \frac{\partial \bar{\Theta}^m}{\partial \Theta^l} \right) \frac{\partial \Theta^r}{\partial \bar{\Theta}^m} \underline{\hat{h}}_r}$$

$$= \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \underline{\hat{h}}_k}{\partial \Theta^l} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} - \underbrace{\frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \frac{\partial \Theta^r}{\partial \bar{\Theta}^m} \frac{\partial \bar{\Theta}^m}{\partial \Theta^n} \Gamma^n_{kl} \underline{\hat{h}}_r}_{= \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \delta^r_n \Gamma^n_{kl} \underline{\hat{h}}_r = \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} \Gamma^n_{kl} \underline{\hat{h}}_n}$$

$$+ \frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i} \underline{\hat{h}}_k - \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^i \bar{\Theta}^j} \frac{\partial \Theta^r}{\partial \bar{\Theta}^m} \frac{\partial \bar{\Theta}^m}{\partial \Theta^l} \underline{\hat{h}}_r$$

$$= \underbrace{\left(\frac{\partial \underline{\hat{h}}_k}{\partial \Theta^l} - \Gamma^n_{kl} \underline{\hat{h}}_n \right)}_{= \underline{\hat{h}}_k \big|_l} \frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \frac{\partial \Theta^l}{\partial \bar{\Theta}^j} + \underbrace{\left(\frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i} - \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^j \bar{\Theta}^i} \frac{\partial \Theta^k}{\partial \bar{\Theta}^m} \frac{\partial \bar{\Theta}^m}{\partial \Theta^l} \right)}_{= \frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i} - \frac{\partial^2 \Theta^l}{\partial \bar{\Theta}^j \bar{\Theta}^i} \delta^k_l = \frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \partial \bar{\Theta}^i} - \frac{\partial^2 \Theta^k}{\partial \bar{\Theta}^j \bar{\Theta}^i} = 0} \underline{\hat{h}}_k$$

$$= \boxed{\frac{\partial \Theta^k}{\partial \bar{\Theta}^i} \left(\hat{\underline{h}}_k \Big|_l \right) \frac{\partial \Theta^l}{\partial \bar{\Theta}^j}} \,, \tag{7.31b}$$

where (5.14), (7.23), (7.25a)–(7.25b) and (7.29a)–(7.30b) have been used. It is not then difficult to verify that

$$\bar{\tilde{\underline{H}}}^{ij}\Big|_k = \frac{\partial \bar{\Theta}^i}{\partial \Theta^l} \frac{\partial \bar{\Theta}^j}{\partial \Theta^m} \left(\tilde{\underline{H}}^{lm}\Big|_n \right) \frac{\partial \Theta^n}{\partial \bar{\Theta}^k} \,, \tag{7.32a}$$

$$\bar{\tilde{\underline{H}}}^{i}{}_{.j}\Big|_k = \frac{\partial \bar{\Theta}^i}{\partial \Theta^l} \frac{\partial \Theta^m}{\partial \bar{\Theta}^j} \left(\tilde{\underline{H}}^{l}{}_{.m}\Big|_n \right) \frac{\partial \Theta^n}{\partial \bar{\Theta}^k} \,, \tag{7.32b}$$

$$\bar{\tilde{\underline{H}}}_{i}{}^{.j}\Big|_k = \frac{\partial \Theta^l}{\partial \bar{\Theta}^i} \frac{\partial \bar{\Theta}^j}{\partial \Theta^m} \left(\tilde{\underline{H}}_{l}{}^{.m}\Big|_n \right) \frac{\partial \Theta^n}{\partial \bar{\Theta}^k} \,, \tag{7.32c}$$

$$\bar{\tilde{\underline{H}}}_{ij}\Big|_k = \frac{\partial \Theta^l}{\partial \bar{\Theta}^i} \frac{\partial \Theta^m}{\partial \bar{\Theta}^j} \left(\tilde{\underline{H}}_{lm}\Big|_n \right) \frac{\partial \Theta^n}{\partial \bar{\Theta}^k} \,. \tag{7.32d}$$

One can finally conclude that:

> The covariant derivative, as a new differential operator, generates tensors out of tensors. And this is not the case for the partial derivative.

The goal is now to consider the covariant derivative of the metric coefficients and basis vectors. This is summarized in an important theorem described below.

7.1.4 Ricci's Theorem

Consider the identity tensor $\mathbf{I} = g^{ij} \mathbf{g}_i \otimes \mathbf{g}_j = g_{ij} \mathbf{g}^i \otimes \mathbf{g}^j$ according to (5.78). Although g^{ij} and g_{ij} vary in space, $\mathbf{I}$ apparently remains constant. This means that its rate of change and accordingly gradient must vanish:

$$\frac{\partial \mathbf{I}}{\partial \Theta^k} = \mathbf{O} \quad , \quad \frac{\partial \mathbf{I}}{\partial \mathbf{x}} = \mathsf{O} \,. \tag{7.33}$$

By virtue of (7.27a) and (7.27d), the left hand side of $(7.33)_1$ can be expressed as $\partial \mathbf{I}/\partial \Theta^k = g^{ij}\big|_k \mathbf{g}_i \otimes \mathbf{g}_j = g_{ij}\big|_k \mathbf{g}^i \otimes \mathbf{g}^j$. Consistent with this, the right hand side of $(7.33)_1$ renders $\mathbf{O} = 0\, \mathbf{g}_i \otimes \mathbf{g}_j = 0\, \mathbf{g}^i \otimes \mathbf{g}^j$ knowing that the zero tensor is a tensor with zero components in any coordinate system. One then concludes that

$$\boxed{g^{ij}\big|_k = 0 \quad , \quad g_{ij}\big|_k = 0 \,.} \tag{7.34}$$

← see (9.164) and (9.184)

As can be seen, the metric coefficients are **covariantly constant**. Note that theses results could also be easily obtained via

$$
\begin{aligned}
g^{ij}\big|_k &\xlongequal[\text{(7.27a)}]{\text{from}} \frac{\partial g^{ij}}{\partial \Theta^k} + \Gamma^i_{km} g^{mj} + \Gamma^j_{km} g^{im} \\
&\xlongequal[\text{(7.16)}]{\text{from}} 0 \,, \\
g_{ij}\big|_k &\xlongequal[\text{(5.39) and (7.27d)}]{\text{from}} \frac{\partial g_{ij}}{\partial \Theta^k} - \Gamma^m_{ik} g_{mj} - \Gamma^m_{jk} g_{im} \\
&\xlongequal[\text{(5.39) and (7.10)}]{\text{from}} \frac{\partial g_{ij}}{\partial \Theta^k} - \Gamma_{ikj} - \Gamma_{jki} \\
&\xlongequal[\text{(7.13)}]{\text{from}} 0 \,.
\end{aligned}
$$

In a similar fashion,

$$
\boxed{\delta^i_j\big|_k = 0 \,.} \tag{7.35}
$$

In accord with (7.34)–(7.35), the covariant basis vectors $\mathbf{g}_i$, $i = 1, 2, 3$, along with their companion dual vectors $\mathbf{g}^i$, $i = 1, 2, 3$, remain also covariantly constant. This can be shown by formally taking the covariant derivative of these basis vectors:

$$
\boxed{\mathbf{g}_i\big|_j \xlongequal[\text{(7.25b)}]{\text{from}} \frac{\partial \mathbf{g}_i}{\partial \Theta^j} - \Gamma^m_{ij} \mathbf{g}_m \xlongequal[\text{(7.8)}]{\text{from}} \Gamma^m_{ij} \mathbf{g}_m - \Gamma^m_{ij} \mathbf{g}_m \xlongequal[\text{(1.4c) and (1.5)}]{\text{from}} \mathbf{0} \,,} \tag{7.36a}
$$

$$
\boxed{\mathbf{g}^i\big|_j \xlongequal[\text{(7.25a)}]{\text{from}} \frac{\partial \mathbf{g}^i}{\partial \Theta^j} + \Gamma^i_{jm} \mathbf{g}^m \xlongequal[\text{(7.9)}]{\text{from}} -\Gamma^i_{jm} \mathbf{g}^m + \Gamma^i_{jm} \mathbf{g}^m \xlongequal[\text{(1.4c) and (1.5)}]{\text{from}} \mathbf{0} \,.} \tag{7.36b}
$$

The identities (7.34)–(7.36b) are known as the *metrinilic property*. They are also referred to as *Ricci's theorem*. As a result, one can establish the useful identities

$$
\underbrace{g_{jl}\left(\underline{\hat{h}}^l\Big|_k\right) = \underline{\hat{h}}_j\Big|_k \quad , \quad g^{jl}\left(\underline{\hat{h}}_l\Big|_k\right) = \underline{\hat{h}}^j\Big|_k}_{\text{note that } g_{jl}\left(\underline{\hat{h}}^l\big|_k\right) = g_{jl}\left(\frac{\partial \underline{\hat{h}}^l}{\partial \Theta^k} + \Gamma^l_{km}\underline{\hat{h}}^m\right) = \frac{\partial \underline{\hat{h}}_j}{\partial \Theta^k} - \underline{\hat{h}}^l\left(\Gamma_{jkl} + \Gamma_{lkj}\right) + \Gamma_{kmj}\underline{\hat{h}}^m = \frac{\partial \underline{\hat{h}}_j}{\partial \Theta^k} - \underline{\hat{h}}_l \Gamma^l_{jk} = \underline{\hat{h}}_j\big|_k} \,. \tag{7.37}
$$

In accord with the rules (6.4a) and (6.4c), the demand for satisfying the product rule implies that

$$
\boxed{\hat{\mathbf{h}}\Big|_j = \left[\underline{\hat{h}}^i \mathbf{g}_i\right]\Big|_j = \left[\underline{\hat{h}}^i\Big|_j\right]\mathbf{g}_i + \underline{\hat{h}}^i\left[\mathbf{g}_i\big|_j\right] = \underline{\hat{h}}^i\Big|_j \mathbf{g}_i \,,} \tag{7.38a}
$$

$$
\boxed{\hat{\mathbf{h}}\Big|_j = \left[\underline{\hat{h}}_i \mathbf{g}^i\right]\Big|_j = \left[\underline{\hat{h}}_i\Big|_j\right]\mathbf{g}^i + \underline{\hat{h}}_i\left[\mathbf{g}^i\big|_j\right] = \underline{\hat{h}}_i\Big|_j \mathbf{g}^i \,.} \tag{7.38b}
$$

As a result, the covariant derivative reduces to the ordinary partial derivative for invariant objects:

$$\boxed{\left.\hat{\mathbf{h}}\right|_j = \frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j}\ .} \qquad \leftarrow \text{see (7.24a)–(7.24b)} \tag{7.39}$$

This result also holds true for tensors of higher orders (see the similar result written in (7.117) for the scalar variables).

7.1.5 Second Covariant Differentiation

Some numerical methods are based on the second-order approximation of field variables. This basically requires the second-order derivatives of a vector and tensor field relative to the coordinates. Representation of these derivatives with respect to a basis will be demonstrated in the following.

Following similar procedures that led to (7.24a)–(7.25b) reveals

$$\frac{\partial^2 \hat{\mathbf{h}}}{\partial \Theta^k \partial \Theta^l} = \frac{\partial}{\partial \Theta^k} \frac{\partial \hat{\mathbf{h}}}{\partial \Theta^l} = \underline{\hat{h}}^i{}_{|lk} \mathbf{g}_i\ , \tag{7.40a}$$

$$\frac{\partial^2 \hat{\mathbf{h}}}{\partial \Theta^k \partial \Theta^l} = \frac{\partial}{\partial \Theta^k} \frac{\partial \hat{\mathbf{h}}}{\partial \Theta^l} = \underline{\hat{h}}_{i|lk} \mathbf{g}^i\ , \tag{7.40b}$$

where

$$\underline{\hat{h}}^i{}_{|lk} = \frac{\partial^2\, \underline{\hat{h}}^i}{\partial \Theta^k \partial \Theta^l} + \frac{\partial \Gamma^i_{lm}}{\partial \Theta^k}\, \underline{\hat{h}}^m + \Gamma^i_{kn} \Gamma^n_{lm}\, \underline{\hat{h}}^m + \Gamma^i_{lm} \frac{\partial\, \underline{\hat{h}}^m}{\partial \Theta^k} + \Gamma^i_{km} \frac{\partial\, \underline{\hat{h}}^m}{\partial \Theta^l}\ , \tag{7.41a}$$

$$\underline{\hat{h}}_{i|lk} = \frac{\partial^2\, \underline{\hat{h}}_i}{\partial \Theta^k \partial \Theta^l} + \Gamma^n_{ki} \Gamma^m_{nl}\, \underline{\hat{h}}_m - \frac{\partial \Gamma^m_{il}}{\partial \Theta^k}\, \underline{\hat{h}}_m - \Gamma^m_{li} \frac{\partial\, \underline{\hat{h}}_m}{\partial \Theta^k} - \Gamma^m_{ki} \frac{\partial\, \underline{\hat{h}}_m}{\partial \Theta^l}\ , \tag{7.41b}$$

may be called the pseudo second-order covariant derivatives. By means of $(7.7)_2$, (7.25a)–(7.25b), (7.27b), (7.27d) and (7.41a)–(7.41b), the true *second-order covariant derivatives* are given by

$$\begin{aligned}
\underline{\hat{h}}^i\Big|_{lk} &= \left(\underline{\hat{h}}^i\Big|_l\right)\Big|_k \\
&= \frac{\partial}{\partial\Theta^k}\left(\underline{\hat{h}}^i\Big|_l\right) + \Gamma^i_{km}\left(\underline{\hat{h}}^m\Big|_l\right) - \Gamma^m_{kl}\left(\underline{\hat{h}}^i\Big|_m\right) \\
&= \boxed{\underline{\hat{h}}^i{}_{|lk} - \Gamma^n_{kl}\left(\frac{\partial\,\underline{\hat{h}}^i}{\partial\Theta^n} + \Gamma^i_{nm}\,\underline{\hat{h}}^m\right)\,,}
\end{aligned} \tag{7.42a}$$

$$\begin{aligned}
\underline{\hat{h}}_i\Big|_{lk} &= \left(\underline{\hat{h}}_i\Big|_l\right)\Big|_k \\
&= \frac{\partial}{\partial\Theta^k}\left(\underline{\hat{h}}_i\Big|_l\right) - \Gamma^m_{ki}\left(\underline{\hat{h}}_m\Big|_l\right) - \Gamma^m_{kl}\left(\underline{\hat{h}}_i\Big|_m\right) \\
&= \boxed{\underline{\hat{h}}_{i|lk} - \Gamma^n_{kl}\left(\frac{\partial\,\underline{\hat{h}}_i}{\partial\Theta^n} - \Gamma^m_{ni}\,\underline{\hat{h}}_m\right)\,.}
\end{aligned} \tag{7.42b}$$

From (7.7)$_2$, (7.42a)$_3$ and (7.42b)$_3$, one immediately obtains

$$\underline{\hat{h}}^i\Big|_{lk} - \underline{\hat{h}}^i\Big|_{kl} = \underline{\hat{h}}^i{}_{|lk} - \underline{\hat{h}}^i{}_{|kl} \quad , \quad \underline{\hat{h}}_i\Big|_{lk} - \underline{\hat{h}}_i\Big|_{kl} = \underline{\hat{h}}_{i|lk} - \underline{\hat{h}}_{i|kl}\,. \tag{7.43}$$

For a tensor field that is known in its contravariant components, one will have

$$\frac{\partial^2\tilde{\mathbf{H}}}{\partial\Theta^k\partial\Theta^l} = \frac{\partial}{\partial\Theta^k}\frac{\partial\tilde{\mathbf{H}}}{\partial\Theta^l} = \underline{\tilde{H}}^{ij}{}_{|lk}\mathbf{g}_i\otimes\mathbf{g}_j\,, \tag{7.44}$$

where

$$\begin{aligned}
\underline{\tilde{H}}^{ij}{}_{|lk} = {} & \frac{\partial^2\,\underline{\tilde{H}}^{ij}}{\partial\Theta^k\partial\Theta^l} + \frac{\partial\Gamma^i_{lm}}{\partial\Theta^k}\,\underline{\tilde{H}}^{mj} + \Gamma^n_{lm}\Gamma^i_{nk}\,\underline{\tilde{H}}^{mj} + \frac{\partial\Gamma^j_{lm}}{\partial\Theta^k}\,\underline{\tilde{H}}^{im} + \Gamma^n_{lm}\Gamma^j_{nk}\,\underline{\tilde{H}}^{im} \\
& + \Gamma^i_{kn}\Gamma^j_{lm}\,\underline{\tilde{H}}^{nm} + \Gamma^i_{ln}\Gamma^j_{km}\,\underline{\tilde{H}}^{nm} \\
& + \Gamma^i_{lm}\frac{\partial\,\underline{\tilde{H}}^{mj}}{\partial\Theta^k} + \Gamma^i_{km}\frac{\partial\,\underline{\tilde{H}}^{mj}}{\partial\Theta^l} + \Gamma^j_{lm}\frac{\partial\,\underline{\tilde{H}}^{im}}{\partial\Theta^k} + \Gamma^j_{km}\frac{\partial\,\underline{\tilde{H}}^{im}}{\partial\Theta^l}\,.
\end{aligned} \tag{7.45}$$

It is then a simple exercise to show that the second covariant derivative of a contravariant tensor takes the following form

$$\begin{aligned}
\underline{\tilde{H}}^{ij}\Big|_{lk} &= \left(\underline{\tilde{H}}^{ij}\Big|_l\right)\Big|_k \\
&= \frac{\partial}{\partial\Theta^k}\left(\underline{\tilde{H}}^{ij}\Big|_l\right) + \Gamma^i_{km}\left(\underline{\tilde{H}}^{mj}\Big|_l\right) + \Gamma^j_{km}\left(\underline{\tilde{H}}^{im}\Big|_l\right) - \Gamma^m_{kl}\left(\underline{\tilde{H}}^{ij}\Big|_m\right) \\
&= \underline{\tilde{H}}^{ij}{}_{|lk} - \Gamma^n_{kl}\left(\frac{\partial\,\underline{\tilde{H}}^{ij}}{\partial\Theta^n} + \Gamma^i_{nm}\,\underline{\tilde{H}}^{mj} + \Gamma^j_{nm}\,\underline{\tilde{H}}^{im}\right)\,.
\end{aligned} \tag{7.46}$$

Consistent with (7.43)$_1$, one will also have

$$\underline{\tilde{H}}^{ij}\Big|_{lk} - \underline{\tilde{H}}^{ij}\Big|_{kl} = \underline{\tilde{H}}^{ij}{}_{|lk} - \underline{\tilde{H}}^{ij}{}_{|kl} \ . \tag{7.47}$$

It should be emphasized that the covariant derivative coincides with the ordinary partial derivative in the case of Cartesian coordinates. This also holds true for scalar field variables, see (7.117). Similarly to the partial derivative, the covariant derivative satisfies the well-known sum and product rules. Moreover, it is important to note that the covariant derivatives commute for **flat spaces**. For instance, let $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ be two vector fields that are known in their contravariant components. Accordingly, these rules are indicated by

$$\left(\underline{\hat{h}}^i_1 + \underline{\hat{h}}^i_2\right)\Big|_j = \underline{\hat{h}}^i_1\Big|_j + \underline{\hat{h}}^i_2\Big|_j \ , \quad \longleftarrow \text{the sum rule} \tag{7.48a}$$

$$\left(\underline{\hat{h}}^i_1\, \underline{\hat{h}}^j_2\right)\Big|_k = \left(\underline{\hat{h}}^i_1\Big|_k\right) \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 \left(\underline{\hat{h}}^j_2\Big|_k\right) \ , \quad \longleftarrow \text{the product rule} \tag{7.48b}$$

$$\left(\underline{\hat{h}}^i_1\right)\Big|_{jk} = \left(\underline{\hat{h}}^i_1\Big|_j\right)\Big|_k = \left(\underline{\hat{h}}^i_1\Big|_k\right)\Big|_j = \left(\underline{\hat{h}}^i_1\right)\Big|_{kj} \ , \quad \longleftarrow \text{the commutative law} \tag{7.48c}$$

which also hold true for tensors of higher orders. The properties (7.48a)–(7.48b) can be shown in a straightforward manner. But, the last property requires more consideration. It relies on the vanishing of an important mathematical object described below.

7.1.6 Riemann-Christoffel Curvature Tensor

By means of (7.41a)–(7.41b) and (7.43)$_{1-2}$, one can obtain (see Danielson [2] and Das [3])

$$\boxed{\underline{\hat{h}}^i\Big|_{lk} - \underline{\hat{h}}^i\Big|_{kl} = \underline{\mathbb{R}}^{i}_{\cdot\, jkl}\ \underline{\hat{h}}^j \ ,} \tag{7.49a}$$

$$\boxed{\underline{\hat{h}}_j\Big|_{kl} - \underline{\hat{h}}_j\Big|_{lk} = \underline{\mathbb{R}}^{i}_{\cdot\, jkl}\ \underline{\hat{h}}_i \ ,} \tag{7.49b}$$

where

$$\boxed{\underline{\mathbb{R}}^{i}_{\cdot\, jkl} = \frac{\partial \Gamma^i_{jl}}{\partial \Theta^k} - \frac{\partial \Gamma^i_{jk}}{\partial \Theta^l} + \Gamma^i_{km}\Gamma^m_{lj} - \Gamma^i_{lm}\Gamma^m_{kj} \ ,} \tag{7.50}$$

satisfying

$$
\begin{aligned}
\underline{\mathbb{R}}^{i\,\cdot\cdot\cdot}_{\cdot\, ikl} &\overset{\text{from}}{\underset{(7.7)\text{ and }(7.50)}{=\!=\!=\!=\!=}} \frac{\partial \Gamma^i_{li}}{\partial \Theta^k} - \frac{\partial \Gamma^i_{ki}}{\partial \Theta^l} + \Gamma^i_{km}\Gamma^m_{li} - \Gamma^i_{lm}\Gamma^m_{ki} \\
&\overset{\text{from}}{\underset{(7.17)}{=\!=\!=}} \cancel{\frac{\partial^2 \ln \mathcal{J}}{\partial \Theta^k \partial \Theta^l}} - \cancel{\frac{\partial^2 \ln \mathcal{J}}{\partial \Theta^l \partial \Theta^k}} + \Gamma^i_{km}\Gamma^m_{li} - \Gamma^i_{lm}\Gamma^m_{ki} \\
&\overset{\text{by renaming}}{\underset{\text{the dummy indices}}{=\!=\!=\!=\!=\!=}} \cancel{\Gamma^m_{ki}\Gamma^i_{lm}} - \cancel{\Gamma^i_{lm}\Gamma^m_{ki}} = 0 \,,
\end{aligned} \tag{7.51}
$$

present the mixed components of the fourth-order tensor

$$
\mathbb{R} = \underline{\mathbb{R}}^{i\,\cdot\cdot\cdot}_{\cdot\, jkl}\ \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \,, \tag{7.52}
$$

known as the *Riemann-Christoffel curvature tensor* (or simply *Riemann-Christoffel tensor*). The first contravariant component can be lowered to provide the fully covariant form

$$
\mathbb{R} = \underline{\mathbb{R}}_{ijkl}\,\mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \quad \text{with} \quad \underline{\mathbb{R}}_{ijkl} = g_{ip}\,\underline{\mathbb{R}}^{p\,\cdot\cdot\cdot}_{\cdot\, jkl} \,. \tag{7.53}
$$

From (7.10), (7.13), (7.14) and (7.50), these fully covariant components may be represented by

$$
\underline{\mathbb{R}}_{ijkl} = \frac{\partial \Gamma_{jli}}{\partial \Theta^k} - \frac{\partial \Gamma_{jki}}{\partial \Theta^l} + \Gamma_{ilp}\Gamma^p_{jk} - \Gamma_{ikp}\Gamma^p_{jl} \,, \tag{7.54}
$$

or

$$
\begin{aligned}
\underline{\mathbb{R}}_{ijkl} = \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial\Theta^k\partial\Theta^j} + \frac{\partial^2 g_{jk}}{\partial\Theta^l\partial\Theta^i} - \frac{\partial^2 g_{jl}}{\partial\Theta^k\partial\Theta^i} - \frac{\partial^2 g_{ik}}{\partial\Theta^l\partial\Theta^j}\right) \\
+ g^{pq}\left(\Gamma_{ilp}\Gamma_{jkq} - \Gamma_{ikp}\Gamma_{jlq}\right) \,.
\end{aligned} \tag{7.55}
$$

It is not then difficult to deduce that

$$
\boxed{\underline{\mathbb{R}}_{ijkl} = \underline{\mathbb{R}}_{klij} \quad \text{and} \quad \underline{\mathbb{R}}_{ijkl} = -\underline{\mathbb{R}}_{ijlk} \,,} \tag{7.56}
$$

since, for instance,

$$
\begin{aligned}
\underline{\mathbb{R}}_{klij} &= \frac{1}{2}\left(\frac{\partial^2 g_{kj}}{\partial\Theta^i\partial\Theta^l} + \frac{\partial^2 g_{li}}{\partial\Theta^j\partial\Theta^k} - \frac{\partial^2 g_{lj}}{\partial\Theta^i\partial\Theta^k} - \frac{\partial^2 g_{ki}}{\partial\Theta^j\partial\Theta^l}\right) \\
&\quad + g^{pq}\left(\Gamma_{kjp}\Gamma_{liq} - \Gamma_{kip}\Gamma_{ljq}\right) \\
&= \frac{1}{2}\left(\frac{\partial^2 g_{kj}}{\partial\Theta^l\partial\Theta^i} + \frac{\partial^2 g_{li}}{\partial\Theta^k\partial\Theta^j} - \frac{\partial^2 g_{lj}}{\partial\Theta^k\partial\Theta^i} - \frac{\partial^2 g_{ki}}{\partial\Theta^l\partial\Theta^j}\right) \\
&\quad + g^{qp}\Gamma_{jkq}\Gamma_{ilp} - g^{pq}\Gamma_{ikp}\Gamma_{jlq} \\
&= \frac{1}{2}\left(\frac{\partial^2 g_{jk}}{\partial\Theta^l\partial\Theta^i} + \frac{\partial^2 g_{il}}{\partial\Theta^k\partial\Theta^j} - \frac{\partial^2 g_{jl}}{\partial\Theta^k\partial\Theta^i} - \frac{\partial^2 g_{ik}}{\partial\Theta^l\partial\Theta^j}\right) \\
&\quad + g^{pq}\left(\Gamma_{jkq}\Gamma_{ilp} - \Gamma_{ikp}\Gamma_{jlq}\right) = \underline{\mathbb{R}}_{ijkl} \,,
\end{aligned}
$$

where (5.39), (5.47), $(7.11)_1$ and (7.55) have been used (along with switching the order of the partial derivatives and the names of the dummy indices). The antisymmetric property $(7.56)_2$ can also be shown in an analogous manner. This is left to be undertaken by the ambitious reader. The relations $(7.56)_{1-2}$ then imply that

$$\boxed{\underline{\mathbb{R}}_{ijkl} = -\,\underline{\mathbb{R}}_{jikl}\ .} \tag{7.57}$$

This property also holds true for the mixed components given in (7.50). For instance,

$$\boxed{\underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,jkl} = -\,\underline{\mathbb{R}}^{\,.\,i\,.\,.}_{j\,.\,kl}\ ,} \tag{7.58}$$

owing to

$$\begin{aligned}
&\xrightarrow[\text{(7.53) and (7.57)}]{\text{from}} \quad g_{ip}\,\underline{\mathbb{R}}^{p\,.\,.\,.}_{\,.\,jkl} = -g_{jp}\,\underline{\mathbb{R}}^{p\,.\,.\,.}_{\,.\,ikl} \\
&\xrightarrow[\text{(5.14) and (5.51)}]{\text{from}} \quad \underline{\mathbb{R}}^{q\,.\,.\,.}_{\,.\,jkl} = -g^{qi}g_{jp}\,\underline{\mathbb{R}}^{p\,.\,.\,.}_{\,.\,ikl} \\
&\xrightarrow[\text{index juggling}]{\text{by}} \quad \underline{\mathbb{R}}^{q\,.\,.\,.}_{\,.\,jkl} = -\,\underline{\mathbb{R}}^{\,.\,q\,.\,.}_{j\,.\,kl} \\
&\xrightarrow[\text{the free index } q \text{ to } i]{\text{by renaming}} \quad \underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,jkl} = -\,\underline{\mathbb{R}}^{\,.\,i\,.\,.}_{j\,.\,kl}\ .
\end{aligned}$$

The result (7.58) helps directly obtain (7.49b) from (7.49a).[1]

The expression (7.49a) measures noncommutativity of the covariant derivative for an object with one index. Using (7.45) and (7.47), this can be generalized to a quantity with two indices:

$$\boxed{\underline{\tilde{H}}^{ij}\Big|_{lk} - \underline{\tilde{H}}^{ij}\Big|_{kl} = \underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,mkl}\ \underline{\tilde{H}}^{mj} + \underline{\mathbb{R}}^{j\,.\,.\,.}_{\,.\,mkl}\ \underline{\tilde{H}}^{im}\ .} \tag{7.59}$$

---

[1] The proof is not difficult. First, by using the relations $\underline{\hat{h}}^i = g^{ir}\,\underline{\hat{h}}_r$ and $\underline{\hat{h}}^j = g^{js}\,\underline{\hat{h}}_s$, taking into account that the metric coefficients are **covariantly constant** according to (7.34), the equation (7.49a) can be written as

$$g^{ir}\left[\,\underline{\hat{h}}_r\Big|_{lk} - \underline{\hat{h}}_r\Big|_{kl}\right] = g^{js}\,\underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,jkl}\ \underline{\hat{h}}_s\ .$$

Multiplying both sides of this equation with g_{ti} and considering the identity $g_{ti}\,g^{ir} = \delta^r_t$ then leads to

$$\delta^r_t\left[\,\underline{\hat{h}}_r\Big|_{lk} - \underline{\hat{h}}_r\Big|_{kl}\right] = g_{ti}\,g^{js}\,\underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,jkl}\ \underline{\hat{h}}_s\ .$$

Now, by considering the replacement property of the mixed Kronecker delta and making use of index juggling, one obtains

$$\underline{\hat{h}}_t\Big|_{lk} - \underline{\hat{h}}_t\Big|_{kl} = \underline{\mathbb{R}}^{\,.\,s\,.\,.}_{t\,.\,kl}\ \underline{\hat{h}}_s\ .$$

By renaming $t \to j$, $s \to i$ and subsequently using $\underline{\mathbb{R}}^{\,.\,i\,.\,.}_{j\,.\,kl} = -\,\underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,jkl}$, one can finally arrive at the desired result

$$\underline{\hat{h}}_j\Big|_{kl} - \underline{\hat{h}}_j\Big|_{lk} = \underline{\mathbb{R}}^{i\,.\,.\,.}_{\,.\,jkl}\ \underline{\hat{h}}_i\ .$$

.

Notice that the free indices i and j in (7.59) can be lowered since the contravariant metric coefficients are covariantly constant. Moreover, the dummy index m on the right hand side of this equation can also be juggled. With the aid of (7.58), one can thus establish

$$\underline{\tilde{H}}_{ij}\Big|_{kl} - \underline{\tilde{H}}_{ij}\Big|_{lk} = \underline{\mathbb{R}}^{m}{}_{.\,ikl}^{\,\cdot\cdot\cdot}\,\underline{\tilde{H}}_{mj} + \underline{\mathbb{R}}^{m}{}_{.\,jkl}^{\,\cdot\cdot\cdot}\,\underline{\tilde{H}}_{im}\ . \tag{7.60}$$

It is then a simple exercise to show that

$$\underline{\tilde{H}}^{i}{}_{.j}\Big|_{lk} - \underline{\tilde{H}}^{i}{}_{.j}\Big|_{kl} = \underline{\mathbb{R}}^{i}{}_{.mkl}^{\,\cdot\cdot\cdot}\,\underline{\tilde{H}}^{m}{}_{.j} - \underline{\mathbb{R}}^{m}{}_{.\,jkl}^{\,\cdot\cdot\cdot}\,\underline{\tilde{H}}^{i}{}_{.m}\ , \tag{7.61a}$$

$$\underline{\tilde{H}}_{i}{}^{.j}\Big|_{kl} - \underline{\tilde{H}}_{i}{}^{.j}\Big|_{lk} = \underline{\mathbb{R}}^{m}{}_{.\,ikl}^{\,\cdot\cdot\cdot}\,\underline{\tilde{H}}_{m}{}^{.j} - \underline{\mathbb{R}}^{j}{}_{.mkl}^{\,\cdot\cdot\cdot}\,\underline{\tilde{H}}_{i}{}^{.m}\ . \tag{7.61b}$$

The Riemann-Christoffel tensor satisfies the so-called *first Bianchi identity*

$$\boxed{\left.\begin{aligned}\underline{\mathbb{R}}^{i}{}_{.\,jkl}^{\,\cdot\cdot\cdot} + \underline{\mathbb{R}}^{i}{}_{.klj}^{\,\cdot\cdot\cdot} + \underline{\mathbb{R}}^{i}{}_{.ljk}^{\,\cdot\cdot\cdot} &= 0\\ \underline{\mathbb{R}}_{ijkl} + \underline{\mathbb{R}}_{iklj} + \underline{\mathbb{R}}_{iljk} &= 0\end{aligned}\right\}\ ,} \quad \longleftarrow \text{the proof is given in Exercise 7.5} \tag{7.62}$$

which renders an **algebraic identity**. This fourth-order tensor also satisfies the **differential identity**

$$\boxed{\left.\begin{aligned}\underline{\mathbb{R}}^{i}{}_{.\,jkl}^{\,\cdot\cdot\cdot}\Big|_{m} + \underline{\mathbb{R}}^{i}{}_{.\,jlm}^{\,\cdot\cdot\cdot}\Big|_{k} + \underline{\mathbb{R}}^{i}{}_{.\,jmk}^{\,\cdot\cdot\cdot}\Big|_{l} &= 0\\ \underline{\mathbb{R}}_{ijkl}\Big|_{m} + \underline{\mathbb{R}}_{ijlm}\Big|_{k} + \underline{\mathbb{R}}_{ijmk}\Big|_{l} &= 0\end{aligned}\right\}\ ,} \quad \longleftarrow \text{the proof is given in Exercise 7.5} \tag{7.63}$$

known as the *second Bianchi identity*. In this expression, $\underline{\mathbb{R}}^{i}{}_{.\,jkl}^{\,\cdot\cdot\cdot}\Big|_{m}$ and $\underline{\mathbb{R}}_{ijkl}\Big|_{m}$ represent the components of partial differentiation of the fourth-order tensor $\underline{\mathbb{R}}$, i.e.

$$\frac{\partial\,\underline{\mathbb{R}}}{\partial\Theta^{m}} = \underline{\mathbb{R}}^{i}{}_{.\,jkl}^{\,\cdot\cdot\cdot}\Big|_{m}\,\mathbf{g}_{i}\otimes\mathbf{g}^{j}\otimes\mathbf{g}^{k}\otimes\mathbf{g}^{l} = \underline{\mathbb{R}}_{ijkl}\Big|_{m}\,\mathbf{g}^{i}\otimes\mathbf{g}^{j}\otimes\mathbf{g}^{k}\otimes\mathbf{g}^{l}\ .$$

They are given by

$$\underline{\mathbb{R}}^{i}{}_{.\,jkl}^{\,\cdot\cdot\cdot}\Big|_{m} = \frac{\partial\,\underline{\mathbb{R}}^{i}{}_{.\,jkl}^{\,\cdot\cdot\cdot}}{\partial\Theta^{m}} + \Gamma^{i}_{mn}\,\underline{\mathbb{R}}^{n}{}_{.\,jkl}^{\,\cdot\cdot\cdot} - \Gamma^{n}_{jm}\,\underline{\mathbb{R}}^{i}{}_{.nkl}^{\,\cdot\cdot\cdot} - \Gamma^{n}_{km}\,\underline{\mathbb{R}}^{i}{}_{.\,jnl}^{\,\cdot\cdot\cdot} - \Gamma^{n}_{lm}\,\underline{\mathbb{R}}^{i}{}_{.\,jkn}^{\,\cdot\cdot\cdot}\ , \tag{7.64a}$$

$$\underline{\mathbb{R}}_{ijkl}\Big|_{m} = \frac{\partial\,\underline{\mathbb{R}}_{ijkl}}{\partial\Theta^{m}} - \Gamma^{n}_{im}\,\underline{\mathbb{R}}_{njkl} - \Gamma^{n}_{jm}\,\underline{\mathbb{R}}_{inkl} - \Gamma^{n}_{km}\,\underline{\mathbb{R}}_{ijnl} - \Gamma^{n}_{lm}\,\underline{\mathbb{R}}_{ijkn}\ . \tag{7.64b}$$

Recall that the expressions $(7.27\text{a})_2$–$(7.27\text{d})_2$ were eventually a simple extension of the results (7.25a)–(7.25b). With regard to this, the relations (7.64a)–(7.64b) should now be considered as another example of extending the action of covariant differentiation on objects with one index to quantities with four indices.

Based on the above considerations, one can infer that the fourth-order covariant Riemann-Christoffel tensor with $3^4 = 81$ entries only has $3^2 \left(3^2 - 1\right)/12 = 6$ independent components in a three-dimensional space; namely, $\underline{\mathbb{R}}_{2323}$, $\underline{\mathbb{R}}_{2313}$, $\underline{\mathbb{R}}_{2312}$, $\underline{\mathbb{R}}_{1313}$, $\underline{\mathbb{R}}_{1312}$, $\underline{\mathbb{R}}_{1212}$ (see the Voigt notation (3.128)–(3.129)). In a similar manner, in a two-dimensional space, this tensor has $2^4 = 16$ components but the number of distinct entries is $2^2 \left(2^2 - 1\right)/12 = 1$; namely, $\underline{\mathbb{R}}_{1212}$. It is now easy to conclude that $\mathbb{R}$ vanishes identically in a one-dimensional space. This is equivalent to saying that the space curves embedded in a three-dimensional Euclidean space are intrinsically flat. For more details see, for example, Sochi [4].

The Riemann-Christoffel curvature tensor is of fundamental importance in tensor description of curved surfaces. This is due to the fact that all information regarding the curvature of space is embedded in this tensor as implied by its name. It measures the noncommutativity of covariant differentiation. For a flat Euclidean space, whose main characteristic is straightness, the order of covariant derivatives makes no distinctions. This implies that the curvature tensor vanishes for flat spaces. It can easily be shown that this tensor becomes null for the well-known coordinate systems that are commonly used to coordinate Euclidean spaces; examples of which include Cartesian, cylindrical and spherical coordinates. In other words, Euclidean spaces allow constructing a rectangular coordinate grid. Indeed, all three-dimensional curvilinear coordinate systems that are related to a Cartesian coordinate frame of the same dimension demonstrate flat spaces.

However, there are spaces which do not support a Cartesian grid. Such spaces with nonvanishing Riemann-Christoffel curvature tensor are called *Riemannian*. A well-known example in Riemannian geometry regards the surface of sphere which basically represents a two-dimensional object embedded in a three-dimensional Euclidean space. See Chap. 9 for an introduction to differential geometry of surfaces and curves.

7.2 Gradient of Scalar, Vector and Tensor Fields

In the literature, the gradient is often denoted by

$$\boxed{\operatorname{grad}(\bullet) := \frac{\partial(\bullet)}{\partial \mathbf{x}} \quad \text{representing} \quad \operatorname{grad}(\bullet) = \frac{\partial(\bullet)}{\partial x_i}\widehat{\mathbf{e}}_i = \frac{\partial(\bullet)}{\partial \Theta^i}\mathbf{g}^i \ .} \tag{7.65}$$

In a similar manner, the **differential vector** (or **Del operator** or **Nabla operator**) $\boldsymbol{\nabla}$ of vector calculus is given by[2]

[2] The scalars $\partial/\partial\Theta^j$, $j = 1, 2, 3$, in (7.66) represent the covariant components of the vector $\partial/\partial\mathbf{x}$. If they appear in a term of an equation, they should apply to all variables that are explicitly under the action of these differential operators except their own companion basis vectors because, in general, $\partial\mathbf{g}_i/\partial\Theta^j \neq \mathbf{0}$, see (7.8).

As an example, consider the equation

$$\boxed{\nabla := \frac{\partial}{\partial \mathbf{x}} \quad \text{representing} \quad \nabla = \frac{\partial}{\partial x_i}\widehat{\mathbf{e}}_i = \frac{\partial}{\partial \Theta^i}\mathbf{g}^i \ .} \tag{7.66}$$

Knowing that $\bar{h}$ can be expressed in terms of either the Cartesian coordinates (x_1, x_2, x_3) or the curvilinear coordinates $\left(\Theta^1, \Theta^2, \Theta^3\right)$, one can write

$$d\bar{h} = \frac{\partial \bar{h}}{\partial x_i} dx_i \ , \tag{7.67a}$$

$$d\bar{h} = \frac{\partial \bar{h}}{\partial \Theta^i} d\Theta^i \ . \tag{7.67b}$$

From (5.49), (5.107), (5.115a), $(7.2)_1$ and (7.65)–(7.67b), one can write

$$\text{grad}\bar{h} = \frac{\partial \bar{h}}{\partial \mathbf{x}} = \frac{\partial \bar{h}}{\partial x_i}\widehat{\mathbf{e}}_i \ , \quad \leftarrow \text{see (6.14a)} \tag{7.68a}$$

$$\text{grad}\bar{h} = \frac{\partial \bar{h}}{\partial \mathbf{x}} = g^{ij}\frac{\partial \bar{h}}{\partial \Theta^j}\mathbf{g}_i = \frac{\partial \bar{h}}{\partial \Theta^i}\mathbf{g}^i \ . \quad \leftarrow \text{see (6.14c)} \tag{7.68b}$$

Consider a scalar field which is basically a point function varying in a three-dimensional space. The equation $\bar{h}(\mathbf{x}) = \text{constant}$ describes a surface called *level surface* (or *isosurface* or *equiscalar surface*). This surface is constructed by points of a constant value such as temperature. The gradient grad$\bar{h}$ basically illustrates a vector **perpendicular** to the isosurface (Chavez [5]). The unit vector at an arbitrary point $\mathbf{x}$ in the direction where $\bar{h}$ has the greatest rate of increase in that point then renders $\widehat{\mathbf{n}} = \left(\text{grad}\bar{h}\right) / \left|\text{grad}\bar{h}\right|$. For a geometrical interpretation, see Fig. 7.1.

Consistent with (7.68a)–(7.68b), the gradient of vector and tensor fields is governed by

$$\text{grad}\hat{\mathbf{h}} = \frac{\partial \hat{\mathbf{h}}}{\partial x_j} \otimes \widehat{\mathbf{e}}_j \quad , \quad \text{grad}\tilde{\mathbf{H}} = \frac{\partial \tilde{\mathbf{H}}}{\partial x_k} \otimes \widehat{\mathbf{e}}_k \ , \tag{7.69a}$$

$$\text{grad}\hat{\mathbf{h}} = \frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} \otimes \mathbf{g}^j \quad , \quad \text{grad}\tilde{\mathbf{H}} = \frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} \otimes \mathbf{g}^k \ . \tag{7.69b}$$

Consider the fact that the standard basis $\{\widehat{\mathbf{e}}_i\}$ is fixed in space. The second- and third-order tensors in (7.69a) can then be expressed as

---

$$\alpha\left(\hat{\mathbf{h}} \cdot \nabla\right) + \beta\left(\hat{\mathbf{h}} \otimes \nabla\right) : \tilde{\mathbf{H}} = \gamma \ ,$$

where $\hat{\mathbf{h}}$ is a vector field that is known in its contravariant components, $\tilde{\mathbf{H}}$ ia a co-contravariant second-order tensor field and α, β, γ are constants. Then, one may rewrite this equation as

$$\frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} \cdot \left(\alpha \mathbf{I} + \beta \tilde{\mathbf{H}}\right) \mathbf{g}^j = \gamma \ ,$$

where the partial derivatives $\partial \hat{\mathbf{h}} / \partial \Theta^j$ have been represented in (7.24a) and (7.25a).

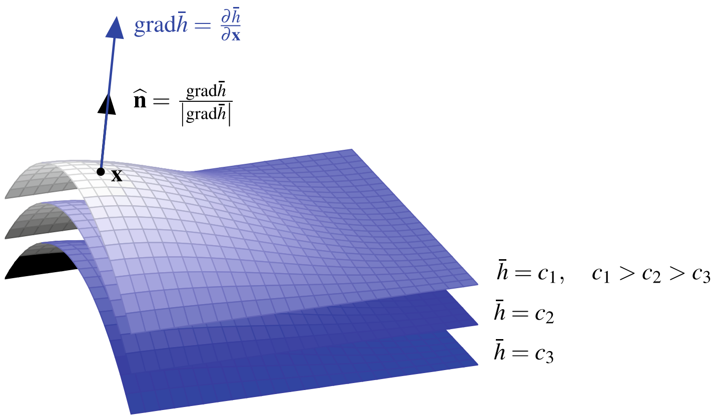


Fig. 7.1 Gradient of a scalar field along with its unit vector at an arbitrary point of a level surface

$$\underbrace{\mathrm{grad}\hat{\mathbf{h}} = \frac{\partial \hat{h}_i}{\partial x_j} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j}_{\text{with } \left(\mathrm{grad}\hat{\mathbf{h}}\right)_{ij} = \frac{\partial \hat{h}_i}{\partial x_j}} \quad , \quad \underbrace{\mathrm{grad}\tilde{\mathbf{H}} = \frac{\partial \tilde{H}_{ij}}{\partial x_k} \widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k}_{\text{with } \left(\mathrm{grad}\tilde{\mathbf{H}}\right)_{ijk} = \frac{\partial \tilde{H}_{ij}}{\partial x_k}} \; . \tag{7.70}$$

In a similar manner, with the aid of (7.24a)–(7.28), the gradients in (7.69b) can be decomposed with respect to the curvilinear basis vectors as

$$\begin{aligned} \mathrm{grad}\hat{\mathbf{h}} &= \underline{\hat{h}}^i \Big|_j \, \mathbf{g}_i \otimes \mathbf{g}^j = \underline{\hat{h}}_i \Big|_j \, \mathbf{g}^i \otimes \mathbf{g}^j \\ &= \underline{\hat{h}}^i \Big|^j \, \mathbf{g}_i \otimes \mathbf{g}_j = \underline{\hat{h}}_i \Big|^j \, \mathbf{g}^i \otimes \mathbf{g}_j \; , \end{aligned} \tag{7.71a}$$

$$\begin{aligned} \mathrm{grad}\tilde{\mathbf{H}} &= \underline{\tilde{H}}^{ij} \Big|_k \, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}^k = \underline{\tilde{H}}^i_{\cdot j} \Big|_k \, \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \\ &= \underline{\tilde{H}}_i^{\cdot j} \Big|_k \, \mathbf{g}^i \otimes \mathbf{g}_j \otimes \mathbf{g}^k = \underline{\tilde{H}}_{ij} \Big|_k \, \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \\ &= \underline{\tilde{H}}^{ij} \Big|^k \, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k = \underline{\tilde{H}}^i_{\cdot j} \Big|^k \, \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}_k \\ &= \underline{\tilde{H}}_i^{\cdot j} \Big|^k \, \mathbf{g}^i \otimes \mathbf{g}_j \otimes \mathbf{g}_k = \underline{\tilde{H}}_{ij} \Big|^k \, \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}_k \; . \end{aligned} \tag{7.71b}$$

The results (7.71a) and (7.71b) can consistently be extended to tensors of higher orders. This is left to be undertaken by the ambitious reader.

The gradient of a tensorial field variable may be denoted more explicitly by

$$\mathrm{grad}\bar{h} := \underbrace{\bar{h}\boldsymbol{\nabla}}_{\text{since, e.g., } \bar{h}\boldsymbol{\nabla} = \bar{h}\frac{\partial}{\partial x_i}\widehat{\mathbf{e}}_i = \frac{\partial \bar{h}}{\partial x_i}\widehat{\mathbf{e}}_i} \; , \tag{7.72a}$$

$$\text{grad}\hat{\mathbf{h}} := \underbrace{\hat{\mathbf{h}} \otimes \boldsymbol{\nabla}}_{\text{since, e.g., } \hat{\mathbf{h}} \otimes \boldsymbol{\nabla} = \hat{\mathbf{h}} \otimes \frac{\partial}{\partial x_j}\widehat{\mathbf{e}}_j = \frac{\partial \hat{\mathbf{h}}}{\partial x_j} \otimes \widehat{\mathbf{e}}_j} \quad , \tag{7.72b}$$

$$\text{grad}\tilde{\mathbf{H}} := \underbrace{\tilde{\mathbf{H}} \otimes \boldsymbol{\nabla}}_{\text{since, e.g., } \tilde{\mathbf{H}} \otimes \boldsymbol{\nabla} = \tilde{\mathbf{H}} \otimes \frac{\partial}{\partial \Theta^k}\mathbf{g}^k = \frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} \otimes \mathbf{g}^k} \quad . \tag{7.72c}$$

It is important to note that the gradient of a tensor field of order n delivers a tensor field of order $(n + 1)$.

As an example, consider the **cylindrical coordinates** (r, θ, z) with the basis vectors (5.7a)–(5.7c), the metric coefficients (5.116) and the Christoffel symbols (7.19). Referred to the orthonormal basis $\{\widehat{\mathbf{e}}_r, \widehat{\mathbf{e}}_\theta, \widehat{\mathbf{e}}_z\}$, a vector field $\hat{\mathbf{h}}$ admits the following decomposition

$$\hat{\mathbf{h}} = \underline{\hat{h}}^r \widehat{\mathbf{e}}_r + \underline{\hat{h}}^\theta \widehat{\mathbf{e}}_\theta + \underline{\hat{h}}^z \widehat{\mathbf{e}}_z \ . \tag{7.73}$$

Then, using $(7.71a)_1$, its gradient takes the following form (see the derivation in Chap. 2 of Lai et al. [6])

$$\begin{aligned}
\text{grad}\hat{\mathbf{h}} = &\frac{\partial \underline{\hat{h}}^r}{\partial r} \widehat{\mathbf{e}}_r \otimes \widehat{\mathbf{e}}_r + \left[\frac{\partial \underline{\hat{h}}^r}{r\partial \theta} - \frac{\underline{\hat{h}}^\theta}{r}\right] \widehat{\mathbf{e}}_r \otimes \widehat{\mathbf{e}}_\theta + \frac{\partial \underline{\hat{h}}^r}{\partial z} \widehat{\mathbf{e}}_r \otimes \widehat{\mathbf{e}}_z \quad \longleftarrow \text{see Exercise 7.7} \\
&+ \frac{\partial \underline{\hat{h}}^\theta}{\partial r} \widehat{\mathbf{e}}_\theta \otimes \widehat{\mathbf{e}}_r + \left[\frac{\partial \underline{\hat{h}}^\theta}{r\partial \theta} + \frac{\underline{\hat{h}}^r}{r}\right] \widehat{\mathbf{e}}_\theta \otimes \widehat{\mathbf{e}}_\theta + \frac{\partial \underline{\hat{h}}^\theta}{\partial z} \widehat{\mathbf{e}}_\theta \otimes \widehat{\mathbf{e}}_z \\
&+ \frac{\partial \underline{\hat{h}}^z}{\partial r} \widehat{\mathbf{e}}_z \otimes \widehat{\mathbf{e}}_r + \frac{\partial \underline{\hat{h}}^z}{r\partial \theta} \widehat{\mathbf{e}}_z \otimes \widehat{\mathbf{e}}_\theta + \frac{\partial \underline{\hat{h}}^z}{\partial z} \widehat{\mathbf{e}}_z \otimes \widehat{\mathbf{e}}_z \ .
\end{aligned} \tag{7.74}$$

As another example, consider the **spherical coordinates** (r, θ, ϕ) with the basis vectors (5.11a)–(5.11c), the metric coefficients (5.119) and the Christoffel symbols (7.20a)–(7.20c). Referred to the orthonormal basis $\left\{\widehat{\mathbf{e}}_r, \widehat{\mathbf{e}}_\theta, \widehat{\mathbf{e}}_\phi\right\}$, a vector field $\hat{\mathbf{h}}$ can be written as

$$\hat{\mathbf{h}} = \underline{\hat{h}}^r \widehat{\mathbf{e}}_r + \underline{\hat{h}}^\theta \widehat{\mathbf{e}}_\theta + \underline{\hat{h}}^\phi \widehat{\mathbf{e}}_\phi \ . \tag{7.75}$$

Then, using $(7.71a)_1$, its gradient represents

$$\begin{aligned}
\text{grad}\hat{\mathbf{h}} = &\frac{\partial \underline{\hat{h}}^r}{\partial r} \widehat{\mathbf{e}}_r \otimes \widehat{\mathbf{e}}_r + \left[\frac{\partial \underline{\hat{h}}^r}{r\partial \theta} - \frac{\underline{\hat{h}}^\theta}{r}\right] \widehat{\mathbf{e}}_r \otimes \widehat{\mathbf{e}}_\theta \quad \longleftarrow \text{see Exercise 7.7} \\
&+ \left[\frac{\partial \underline{\hat{h}}^r}{r \sin\theta \partial \phi} - \frac{\underline{\hat{h}}^\phi}{r}\right] \widehat{\mathbf{e}}_r \otimes \widehat{\mathbf{e}}_\phi + \frac{\partial \underline{\hat{h}}^\theta}{\partial r} \widehat{\mathbf{e}}_\theta \otimes \widehat{\mathbf{e}}_r + \left[\frac{\partial \underline{\hat{h}}^\theta}{r\partial \theta} + \frac{\underline{\hat{h}}^r}{r}\right] \widehat{\mathbf{e}}_\theta \otimes \widehat{\mathbf{e}}_\theta \\
&+ \left[\frac{\partial \underline{\hat{h}}^\theta}{r \sin\theta \partial \phi} - \cot\theta \frac{\underline{\hat{h}}^\phi}{r}\right] \widehat{\mathbf{e}}_\theta \otimes \widehat{\mathbf{e}}_\phi + \frac{\partial \underline{\hat{h}}^\phi}{\partial r} \widehat{\mathbf{e}}_\phi \otimes \widehat{\mathbf{e}}_r + \frac{\partial \underline{\hat{h}}^\phi}{r\partial \theta} \widehat{\mathbf{e}}_\phi \otimes \widehat{\mathbf{e}}_\theta \\
&+ \left[\frac{\partial \underline{\hat{h}}^\phi}{r \sin\theta \partial \phi} + \cot\theta \frac{\underline{\hat{h}}^\theta}{r} + \frac{\underline{\hat{h}}^r}{r}\right] \widehat{\mathbf{e}}_\phi \otimes \widehat{\mathbf{e}}_\phi \ .
\end{aligned} \tag{7.76}$$

7.3 Divergence of Vector and Tensor Fields

The divergence of a vector field $\hat{\mathbf{h}}(\mathbf{x})$ is a **scalar field** defined by

$$\boxed{\text{div}\hat{\mathbf{h}} = \text{grad}\hat{\mathbf{h}} : \mathbf{I} = \text{tr}\left(\text{grad}\hat{\mathbf{h}}\right)} \ . \tag{7.77}$$

By use of (2.89a) and $(7.70)_2$, this scalar field can be expressed in terms of the Cartesian components as

$$\text{div}\hat{\mathbf{h}} = \frac{\partial \hat{h}_i}{\partial x_i} \ . \tag{7.78}$$

And in terms of the curvilinear components, by using $(5.88)_{1-4}$ and $(7.71\text{a})_{1-4}$, it renders

$$\text{div}\hat{\mathbf{h}} = \underline{\hat{h}}^i\Big|_i = g^{ij}\ \underline{\hat{h}}_i\Big|_j = g_{ij}\ \underline{\hat{h}}^i\Big|^j = \underline{\hat{h}}_i\Big|^i \ . \tag{7.79}$$

Alternatively to (7.77), the divergence of a vector field may be introduced as

$$\boxed{\text{div}\hat{\mathbf{h}} = \frac{\partial \hat{\mathbf{h}}}{\partial \Theta^i} \cdot \mathbf{g}^i} \ , \tag{7.80}$$

since

$$\begin{aligned}
\text{div}\hat{\mathbf{h}} &\xlongequal[(7.77)]{\text{from}} \text{grad}\hat{\mathbf{h}} : \mathbf{I} \\
&\xlongequal[(5.78) \text{ and } (7.69\text{b})]{\text{from}} \left(\frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} \otimes \mathbf{g}^j\right) : \left(\mathbf{g}^i \otimes \mathbf{g}_i\right) \\
&\xlongequal[(2.73)]{\text{from}} \left(\frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} \cdot \mathbf{g}^i\right)\left(\mathbf{g}^j \cdot \mathbf{g}_i\right) \\
&\xlongequal[(5.27)]{\text{from}} \left(\frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} \cdot \mathbf{g}^i\right)\left(\delta_i^j\right) \\
&\xlongequal[(5.14)]{\text{from}} \frac{\partial \hat{\mathbf{h}}}{\partial \Theta^j} \cdot \mathbf{g}^j \ .
\end{aligned}$$

The divergence of a vector field can also be written as

$$\boxed{\text{div}\hat{\mathbf{h}} = \frac{1}{J}\frac{\partial}{\partial \Theta^i}\left[J\ \underline{\hat{h}}^i\right]} \ . \tag{7.81}$$

This expression is known as the *Voss-Weyl formula*. Note that

$$\frac{1}{J}\frac{\partial}{\partial \Theta^i}\left[J\,\underline{\hat{h}}^i\right] \xlongequal[\text{rule of differentiation}]{\text{by using the product}} \frac{1}{J}\frac{\partial J}{\partial \Theta^i}\,\underline{\hat{h}}^i + \frac{\partial\,\underline{\hat{h}}^i}{\partial \Theta^i}$$
$$\xlongequal[(7.17)]{\text{by using}} \Gamma^m_{mi}\,\underline{\hat{h}}^i + \frac{\partial\,\underline{\hat{h}}^i}{\partial \Theta^i}$$
$$\xlongequal[\text{renaming the the dummy indices}]{\text{by using (7.25a) and}} \underline{\hat{h}}^i\Big|_i$$
$$\xlongequal[(7.79)]{\text{by using}} \text{div}\hat{\mathbf{h}}\ .$$

The vector field $\hat{\mathbf{h}}$ is called *solenoidal* (or *divergence-free*) when $\text{div}\hat{\mathbf{h}} = 0$. Using the same recipes for introducing the divergence of a vector field, the divergence of a tensor field will be characterized in the following.

The divergence of a tensor field $\tilde{\mathbf{H}}(\mathbf{x})$ is a **vector field** defined by

$$\boxed{\text{div}\tilde{\mathbf{H}} = \text{grad}\tilde{\mathbf{H}} : \mathbf{I}\ .} \tag{7.82}$$

The third-order tensor $\text{grad}\tilde{\mathbf{H}}$ and the second-order tensor $\mathbf{I}$ are already given in $(7.70)_{3-4}$ and (2.23), respectively. Substituting these expressions into $(3.16\text{b})_4$ leads to the following decomposition of the first-order tensor $\text{div}\tilde{\mathbf{H}}$ with respect to the standard basis

$$\text{div}\tilde{\mathbf{H}} = \frac{\partial \tilde{H}_{ij}}{\partial x_j}\widehat{\mathbf{e}}_i \quad \text{with} \quad \left(\text{div}\tilde{\mathbf{H}}\right)_i = \frac{\partial \tilde{H}_{ij}}{\partial x_j}\ . \tag{7.83}$$

Guided by the derivation demonstrated in (3.16b) and by use of (5.14), (5.27), (5.78) and (7.71b), the vector field (7.82) can be expressed with respect to the curvilinear basis vectors as

$$\begin{aligned} \text{div}\tilde{\mathbf{H}} &= \underline{\tilde{H}}^{ij}\Big|_j \mathbf{g}_i = g^{jk}\ \underline{\tilde{H}}^{i}_{\ .j}\Big|_k \mathbf{g}_i \\ &= \underline{\tilde{H}}_i^{\ .j}\Big|_j \mathbf{g}^i = g^{jk}\ \underline{\tilde{H}}_{ij}\Big|_k \mathbf{g}^i \\ &= g_{jk}\ \underline{\tilde{H}}^{ij}\Big|^k \mathbf{g}_i = \underline{\tilde{H}}^{i}_{\ .j}\Big|^j \mathbf{g}_i \\ &= g_{jk}\ \underline{\tilde{H}}_i^{\ .j}\Big|^k \mathbf{g}^i = \underline{\tilde{H}}_{ij}\Big|^j \mathbf{g}^i\ . \end{aligned} \tag{7.84}$$

The divergence of a tensor field satisfies

$$\boxed{\left(\text{div}\tilde{\mathbf{H}}\right)\cdot\mathbf{w} = \text{div}\left(\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{w}\right) \quad \text{for any constant vector } \mathbf{w} \in \mathscr{E}_{\mathrm{r}}^{o3}\ ,} \tag{7.85}$$

since

$$\left(\operatorname{div}\tilde{\mathbf{H}}\right) \cdot \mathbf{w} \xlongequal[\text{(1.38) and (7.83)}]{\text{from}} \frac{\partial \tilde{H}_{ij}}{\partial x_j} w_i \xlongequal[\text{(2.49)}]{\text{from}} \frac{\partial \tilde{H}^{\mathrm{T}}_{ji}}{\partial x_j} w_i \xlongequal[\text{(2.22)}]{\text{from}} \frac{\partial}{\partial x_j}\left(\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{w}\right)_j \xlongequal[\text{(7.78)}]{\text{from}} \operatorname{div}\left(\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{w}\right) .$$

The expression (7.85) can be regarded as another definition for the divergence of a tensor field. Another form of this vector field is

$$\boxed{\operatorname{div}\tilde{\mathbf{H}} = \frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k}\mathbf{g}^k ,} \tag{7.86}$$

owing to

$$\operatorname{div}\tilde{\mathbf{H}} \xlongequal[\text{(7.82)}]{\text{from}} \left(\operatorname{grad}\tilde{\mathbf{H}}\right) : (\mathbf{I}) \xlongequal[\text{(5.78) and (7.69b)}]{\text{from}} \left(\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k} \otimes \mathbf{g}^k\right) : \left(\mathbf{g}^i \otimes \mathbf{g}_i\right) \xlongequal[\text{(3.16b)}]{\text{in view of}} \left(\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k}\mathbf{g}^i\right)\left(\mathbf{g}^k \cdot \mathbf{g}_i\right) \xlongequal[\text{(5.27)}]{\text{from}} \left(\frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k}\mathbf{g}^i\right)\left(\delta^k_i\right) \xlongequal[\text{(5.14)}]{\text{from}} \frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k}\mathbf{g}^k .$$

An extension of (7.81) is

$$\operatorname{div}\tilde{\mathbf{H}} = \left(\frac{1}{J}\frac{\partial}{\partial \Theta^j}\left[J\,\underline{\tilde{H}}^{ij}\right] + \Gamma^i_{jk}\,\underline{\tilde{H}}^{kj}\right)\mathbf{g}_i . \tag{7.87}$$

Using the Nabla operator of vector calculus, the divergence of a tensorial field variable may be written by

$$\operatorname{div}\hat{\mathbf{h}} := \underbrace{\hat{\mathbf{h}} \cdot \boldsymbol{\nabla}}_{\text{since, e.g., } \hat{\mathbf{h}} \cdot \boldsymbol{\nabla} = \hat{\mathbf{h}} \cdot \frac{\partial}{\partial x_j}\widehat{\mathbf{e}}_j = \frac{\partial \hat{\mathbf{h}}}{\partial x_j} \cdot \widehat{\mathbf{e}}_j} , \tag{7.88a}$$

$$\operatorname{div}\tilde{\mathbf{H}} := \underbrace{\tilde{\mathbf{H}}\,\boldsymbol{\nabla}}_{\text{since, e.g., } \tilde{\mathbf{H}}\,\boldsymbol{\nabla} = \tilde{\mathbf{H}}\frac{\partial}{\partial \Theta^k}\mathbf{g}^k = \frac{\partial \tilde{\mathbf{H}}}{\partial \Theta^k}\mathbf{g}^k} , \tag{7.88b}$$

in alignment with (7.72a)–(7.72c). At the end, it is important to note that the divergence of a tensor field of order n yields a tensor field of order $(n-1)$. And this differential operator is defined for tensorial field variables of at least order 1.

7.4 Curl of Vector and Tensor Fields

The *curl* (or *rotation*) of a vector field $\hat{\mathbf{h}}$ is again a **vector field** defined by

$$\boxed{\operatorname{curl}\hat{\mathbf{h}} := \boldsymbol{\nabla} \times \hat{\mathbf{h}} = -\hat{\mathbf{h}} \times \boldsymbol{\nabla}\ .} \tag{7.89}$$

For a Cartesian vector field, this definition leads to

$$\begin{aligned}
\operatorname{curl}\hat{\mathbf{h}} &\xlongequal[\text{(7.66) and (7.89)}]{\text{from}} -\hat{\mathbf{h}} \times \frac{\partial}{\partial x_k}\widehat{\mathbf{e}}_k \\
&\xlongequal[\text{(1.34), (1.49a) and (1.49b)}]{\text{from}} -\frac{\partial \hat{h}_j}{\partial x_k}\widehat{\mathbf{e}}_j \times \widehat{\mathbf{e}}_k \\
&\xlongequal[\text{(1.64)}]{\text{from}} -\frac{\partial \hat{h}_j}{\partial x_k}\left(\varepsilon_{jki}\widehat{\mathbf{e}}_i\right) \\
&\xlongequal[\text{(1.4f), (1.4g) and (1.54)}]{\text{from}} -\varepsilon_{ijk}\frac{\partial \hat{h}_j}{\partial x_k}\widehat{\mathbf{e}}_i\ ,
\end{aligned} \tag{7.90}$$

with

$$\left(\operatorname{curl}\hat{\mathbf{h}}\right)_i = -\varepsilon_{ijk}\frac{\partial \hat{h}_j}{\partial x_k}\ . \tag{7.91}$$

In matrix form,

$$\left[\operatorname{curl}\hat{\mathbf{h}}\right] = \begin{bmatrix} \partial\hat{h}_3/\partial x_2 - \partial\hat{h}_2/\partial x_3 \\ \partial\hat{h}_1/\partial x_3 - \partial\hat{h}_3/\partial x_1 \\ \partial\hat{h}_2/\partial x_1 - \partial\hat{h}_1/\partial x_2 \end{bmatrix}\ . \quad \leftarrow \text{see (2.66)} \tag{7.92}$$

The curl of a vector field can also be governed by

$$\boxed{\operatorname{curl}\hat{\mathbf{h}} = -\boldsymbol{\mathsf{E}} : \operatorname{grad}\hat{\mathbf{h}}\ .} \tag{7.93}$$

Note that by means of $(3.16\text{b})_4$, (3.17), $(7.70)_2$ and (7.93), one can again arrive at $(7.90)_4$. The above relation basically shows that the curl of a vector field is twice the axial vector of (the skew-symmetric portion of) its gradient.

The curl of a vector field is an object satisfying

$$\boxed{\left(\operatorname{curl}\hat{\mathbf{h}}\right)\cdot\mathbf{w} = \operatorname{div}\left(\hat{\mathbf{h}}\times\mathbf{w}\right) \quad \text{for any constant vector } \mathbf{w}\in\mathscr{E}_{\mathrm{r}}^{o3}\ ,} \tag{7.94}$$

since

$$\begin{aligned}\left(\operatorname{curl}\hat{\mathbf{h}}\right)\cdot\mathbf{w} &\xlongequal[\text{(1.38) and (7.90)}]{\text{from}} -\varepsilon_{ijk}\frac{\partial\hat{h}_j}{\partial x_k}w_i \\ &\xlongequal[\text{(1.67)}]{\text{from}} -\frac{\partial\left(\mathbf{w}\times\hat{\mathbf{h}}\right)_k}{\partial x_k} \\ &\xlongequal[\text{(1.49a)}]{\text{from}} \frac{\partial\left(\hat{\mathbf{h}}\times\mathbf{w}\right)_k}{\partial x_k} \\ &\xlongequal[\text{(7.78)}]{\text{from}} \operatorname{div}\left(\hat{\mathbf{h}}\times\mathbf{w}\right)\ .\end{aligned}$$

The curvilinear form of $(7.89)_2$ is

$$\boxed{\operatorname{curl}\hat{\mathbf{h}} = -\frac{\partial\hat{\mathbf{h}}}{\partial\Theta^i}\times\mathbf{g}^i\ ,} \tag{7.95}$$

because

$$\operatorname{curl}\hat{\mathbf{h}} \xlongequal[\text{(7.69b) and (7.93)}]{\text{from}} -\mathbf{E}:\frac{\partial\hat{\mathbf{h}}}{\partial\Theta^i}\otimes\mathbf{g}^i \xlongequal[\text{(3.22)}]{\text{from}} -\frac{\partial\hat{\mathbf{h}}}{\partial\Theta^i}\times\mathbf{g}^i\ .$$

Guided by (3.16b) and using (5.14), $(5.27)_{1-2}$, $(5.33)_1$, $(5.35)_1$, $(5.38)_1$, $(5.46)_1$, $(5.98)_{1-2}$, $(7.71a)_{1-4}$ and (7.93), the curl of a vector field admits the following forms

$$\begin{aligned}\operatorname{curl}\hat{\mathbf{h}} &= \underbrace{-J^{-1}\varepsilon^{ijk}g_{jl}\ \underline{\hat{h}}^l\Big|_k\ \mathbf{g}_i}_{=-g_{jl}\ \underline{\hat{h}}^l|_k\ \mathbf{g}^j\times\mathbf{g}^k} = \underbrace{-J^{-1}\varepsilon^{ijk}\ \underline{\hat{h}}_j\Big|_k\ \mathbf{g}_i}_{=-\ \underline{\hat{h}}_j|_k\ \mathbf{g}^j\times\mathbf{g}^k} \\ &= \underbrace{-J^{-1}\varepsilon^{ijk}g_{jl}g_{km}\ \underline{\hat{h}}^l\Big|^m\ \mathbf{g}_i}_{=-g_{jl}g_{km}\ \underline{\hat{h}}^l|^m\ \mathbf{g}^j\times\mathbf{g}^k} = \underbrace{-J^{-1}\varepsilon^{ijk}g_{km}\ \underline{\hat{h}}_j\Big|^m\ \mathbf{g}_i}_{=-g_{km}\ \underline{\hat{h}}_j|^m\ \mathbf{g}^j\times\mathbf{g}^k} \\ &= \underbrace{-J\varepsilon_{ijk}g^{km}\ \underline{\hat{h}}^j\Big|_m\ \mathbf{g}^i}_{=-g^{km}\ \underline{\hat{h}}^j|_m\ \mathbf{g}_j\times\mathbf{g}_k} = \underbrace{-J\varepsilon_{ijk}g^{jl}g^{km}\ \underline{\hat{h}}_l\Big|_m\ \mathbf{g}^i}_{=-g^{jl}g^{km}\ \underline{\hat{h}}_l|_m\ \mathbf{g}_j\times\mathbf{g}_k} \\ &= \underbrace{-J\varepsilon_{ijk}\ \underline{\hat{h}}^j\Big|^k\ \mathbf{g}^i}_{=-\ \underline{\hat{h}}^j|^k\ \mathbf{g}_j\times\mathbf{g}_k} = \underbrace{-J\varepsilon_{ijk}g^{jl}\ \underline{\hat{h}}_l\Big|^k\ \mathbf{g}^i}_{=-g^{jl}\ \underline{\hat{h}}_l|^k\ \mathbf{g}_j\times\mathbf{g}_k}\ .\end{aligned} \tag{7.96}$$

The vector field $\hat{\mathbf{h}}$ is called *irrotational* when $\text{curl}\hat{\mathbf{h}} = \mathbf{0}$. In what follows, the curl of a tensor field will be characterized.

The curl of a tensor field $\tilde{\mathbf{H}}$ is again a **tensor field** defined by

$$\boxed{\text{curl}\tilde{\mathbf{H}} := \boldsymbol{\nabla} \times \tilde{\mathbf{H}}^{\mathrm{T}} \ .} \tag{7.97}$$

This definition relies on the following rule

$$\boxed{\mathbf{u} \times (\mathbf{v} \otimes \mathbf{w}) := (\mathbf{u} \times \mathbf{v}) \otimes \mathbf{w} \ ,} \tag{7.98}$$

possessing linearity in each argument. When the curl operator operates on a Cartesian tensor, the resulting tensor can be expressed as

$$\begin{aligned} \text{curl}\tilde{\mathbf{H}} &= \frac{\partial}{\partial x_l}\widehat{\mathbf{e}}_l \times \tilde{\mathbf{H}}^{\mathrm{T}} = \widehat{\mathbf{e}}_l \times \frac{\partial \tilde{\mathbf{H}}^{\mathrm{T}}}{\partial x_l} \\ &= \frac{\partial \tilde{H}_{jk}}{\partial x_l}\widehat{\mathbf{e}}_l \times \left(\widehat{\mathbf{e}}_k \otimes \widehat{\mathbf{e}}_j\right) = \frac{\partial \tilde{H}_{jk}}{\partial x_l}\varepsilon_{lki}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \\ &= -\varepsilon_{ikl}\frac{\partial \tilde{H}_{jk}}{\partial x_l}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \ , \end{aligned} \tag{7.99}$$

with

$$\left(\text{curl}\tilde{\mathbf{H}}\right)_{ij} = -\varepsilon_{ikl}\frac{\partial \tilde{H}_{jk}}{\partial x_l} \ . \tag{7.100}$$

The curl of a tensor field satisfies

$$\boxed{\left(\text{curl}\tilde{\mathbf{H}}\right)\mathbf{w} = \text{curl}\left(\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{w}\right) \quad \text{for any constant vector } \mathbf{w} \in \mathscr{E}_{\mathrm{r}}^{\mathrm{o3}} \ ,} \tag{7.101}$$

because

$$\begin{aligned} \left(\text{curl}\tilde{\mathbf{H}}\right)_{ij}(\mathbf{w})_j &\overset{\text{from}}{\underset{(7.100)}{=\!=}} -\varepsilon_{ikl}\frac{\partial \tilde{H}_{jk}}{\partial x_l}w_j \\ &\overset{\text{from}}{\underset{(2.22)\text{ and }(2.49)}{=\!=}} -\varepsilon_{ikl}\frac{\partial\left(\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{w}\right)_k}{\partial x_l} \\ &\overset{\text{from}}{\underset{(7.91)}{=\!=}} \left(\text{curl}\left(\tilde{\mathbf{H}}^{\mathrm{T}}\mathbf{w}\right)\right)_i \ . \end{aligned}$$

Notice that the expressions (7.101) and (7.85) have the same structure.

By means of $(7.66)_3$, the curl of a tensor field, according to (7.97), can be represented by

$$\boxed{\operatorname{curl}\tilde{\mathbf{H}} = \mathbf{g}^k \times \frac{\partial \tilde{\mathbf{H}}^{\mathrm{T}}}{\partial \Theta^k}} \ . \tag{7.102}$$

It is then a simple exercise to decompose the curl of a tensor field with respect to the curvilinear basis tensors as

$$\begin{aligned}
\operatorname{curl}\tilde{\mathbf{H}} &= \underbrace{J^{-1}\varepsilon^{kli} \ \underline{\tilde{H}}^{j}_{\ .l}\Big|_k \mathbf{g}_i \otimes \mathbf{g}_j}_{\text{note that } \frac{\partial \tilde{\mathbf{H}}^{\mathrm{T}}}{\partial \Theta^k} = \underline{\tilde{H}}^{j}_{\ .l}\big|_k \mathbf{g}^l \otimes \mathbf{g}_j} = \underbrace{J^{-1}\varepsilon^{kli} \ \underline{\tilde{H}}_{jl}\Big|_k \mathbf{g}_i \otimes \mathbf{g}^j}_{\text{note that } \frac{\partial \tilde{\mathbf{H}}^{\mathrm{T}}}{\partial \Theta^k} = \underline{\tilde{H}}_{jl}\big|_k \mathbf{g}^l \otimes \mathbf{g}^j} \\
&= \underbrace{J\varepsilon_{kli} \ \underline{\tilde{H}}^{jl}\Big|^k \mathbf{g}^i \otimes \mathbf{g}_j}_{\text{note that } \frac{\partial \tilde{\mathbf{H}}^{\mathrm{T}}}{\partial \Theta^k} = \underline{\tilde{H}}^{jl}\big|_k \mathbf{g}_l \otimes \mathbf{g}_j} = \underbrace{J\varepsilon_{kli} \ \underline{\tilde{H}}^{\ .l}_{j}\Big|^k \mathbf{g}^i \otimes \mathbf{g}^j}_{\text{note that } \frac{\partial \tilde{\mathbf{H}}^{\mathrm{T}}}{\partial \Theta^k} = \underline{\tilde{H}}^{\ .l}_{j}\big|_k \mathbf{g}_l \otimes \mathbf{g}^j} \ .
\end{aligned} \tag{7.103a}$$

It is worthwhile to point out that the curl of a tensor field gives a tensor field with the same order. And the curl of a scalar field is not defined.

7.5 Laplacian and Hessian of Scalar, Vector and Tensor Fields

The *Laplace operator* (or *Laplacian*), denoted by ∇^2 or Δ, is defined as the dot product of the differential vector operator $\boldsymbol{\nabla}$ with itself. This second-order differential **scalar** operator, which does not change the order of a tensorial field variable, in Cartesian coordinates renders (Hildebrand [7])

$$\begin{aligned}
\nabla^2 &\xlongequal[\text{definition}]{\text{by}} \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \\
&\xlongequal[(7.66)]{\text{from}} \frac{\partial}{\partial x_i}\widehat{\mathbf{e}}_i \cdot \frac{\partial}{\partial x_j}\widehat{\mathbf{e}}_j \\
&\xlongequal[(1.9a)–(1.9c)]{\text{from}} \frac{\partial^2}{\partial x_i \partial x_j}\widehat{\mathbf{e}}_i \cdot \widehat{\mathbf{e}}_j \\
&\xlongequal[(1.35)]{\text{from}} \frac{\partial^2}{\partial x_i \partial x_j}\delta_{ij} \\
&\xlongequal[(1.36)]{\text{from}} \boxed{\frac{\partial^2}{\partial x_i \partial x_i}} \ .
\end{aligned} \tag{7.104}$$

Accordingly, by operating the Laplace operator upon a Cartesian tensor field, one will have

$$\nabla^2 \bar{h} = \frac{\partial^2 \bar{h}}{\partial x_i \partial x_i} \quad , \quad \nabla^2 \hat{\mathbf{h}} = \frac{\partial^2 \hat{h}_j}{\partial x_i \partial x_i}\widehat{\mathbf{e}}_j \quad , \quad \nabla^2 \tilde{\mathbf{H}} = \frac{\partial^2 \tilde{H}_{jk}}{\partial x_i \partial x_i}\widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k \ . \tag{7.105}$$

The vector field $(7.105)_2$ and the tensor field $(7.105)_3$ are objects satisfying

$$\boxed{\left(\nabla^2\hat{\mathbf{h}}\right)\cdot\mathbf{w}=\nabla^2\left(\hat{\mathbf{h}}\cdot\mathbf{w}\right)\quad\text{for any constant vector } \mathbf{w}\in\mathscr{E}_{\mathrm{r}}^{o3}}\ , \tag{7.106a}$$

$$\boxed{\left(\nabla^2\tilde{\mathbf{H}}\right)\cdot\mathbf{w}=\nabla^2\left(\tilde{\mathbf{H}}\mathbf{w}\right)\quad\text{for any constant vector } \mathbf{w}\in\mathscr{E}_{\mathrm{r}}^{o3}}\ . \tag{7.106b}$$

The conditions (7.106a) and (7.106b) may be used to define the tensor fields $\nabla^2\hat{\mathbf{h}}$ and $\nabla^2\tilde{\mathbf{H}}$, respectively.

The Laplace operator in curvilinear coordinates presents

$$\begin{aligned}
\nabla^2 &\xlongequal[\text{(7.66)}]{\text{by using}} \frac{\partial}{\partial\Theta^i}\mathbf{g}^i\cdot\frac{\partial}{\partial\Theta^j}\mathbf{g}^j \\
&\xlongequal[\text{and the product rule}]{\text{by using (1.9a)–(1.9c)}} \frac{\partial^2}{\partial\Theta^i\partial\Theta^j}\mathbf{g}^i\cdot\mathbf{g}^j+\frac{\partial}{\partial\Theta^j}\mathbf{g}^i\cdot\frac{\partial\mathbf{g}^j}{\partial\Theta^i} \\
&\xlongequal[\text{(5.46) and (7.9)}]{\text{by using}} g^{ij}\frac{\partial^2}{\partial\Theta^i\partial\Theta^j}-\Gamma^j_{ik}g^{ik}\frac{\partial}{\partial\Theta^j} \\
&\xlongequal[\text{the indices } j \text{ and } k \text{ in the last term}]{\text{by switching the names of}} \boxed{g^{ij}\left[\frac{\partial^2}{\partial\Theta^i\partial\Theta^j}-\Gamma^k_{ij}\frac{\partial}{\partial\Theta^k}\right]}\ .
\end{aligned} \tag{7.107}$$

The Laplacian of a tensor field is basically the divergence of its gradient. This can be demonstrated as follows:

$$\begin{aligned}
\nabla^2(\bullet) &= \operatorname{div}\left[\operatorname{grad}(\bullet)\right] \\
&= \operatorname{div}\left[\frac{\partial(\bullet)}{\partial\Theta^j}\otimes\mathbf{g}^j\right] \quad \longleftarrow \text{see, for instance, } (7.69\text{b})_1 \text{ and } (7.86) \\
&= \frac{\partial}{\partial\Theta^i}\left[\frac{\partial(\bullet)}{\partial\Theta^j}\otimes\mathbf{g}^j\right]\mathbf{g}^i \\
&= g^{ij}\left[\frac{\partial^2(\bullet)}{\partial\Theta^i\partial\Theta^j}-\Gamma^k_{ij}\frac{\partial(\bullet)}{\partial\Theta^k}\right]\ .
\end{aligned} \tag{7.108}$$

To take the Laplacian of a vector or tensor field in curvilinear coordinates, the corresponding first and second partial derivatives with respect to the coordinates are required. They are already given in Sects. 7.1.2 and 7.1.5. Of interest here is to only represent a scalar and vector field:

$$\nabla^2\bar{h}=g^{ij}\left[\frac{\partial^2\bar{h}}{\partial\Theta^i\partial\Theta^j}-\Gamma^k_{ij}\frac{\partial\bar{h}}{\partial\Theta^k}\right]\ , \tag{7.109a}$$

$$\nabla^2\hat{\mathbf{h}} = g^{ij}\left[\frac{\partial^2 \underline{\hat{h}}^m}{\partial\Theta^i\partial\Theta^j} + \left(\frac{\partial\Gamma^m_{jl}}{\partial\Theta^i} + \Gamma^m_{ik}\Gamma^k_{jl}\right)\underline{\hat{h}}^l \right.$$
$$\left. +2\Gamma^m_{jl}\frac{\partial\underline{\hat{h}}^l}{\partial\Theta^i} - \Gamma^k_{ij}\frac{\partial\underline{\hat{h}}^m}{\partial\Theta^k} - \Gamma^k_{ij}\Gamma^m_{kl}\underline{\hat{h}}^l\right]\mathbf{g}_m ,$$
$$= g^{ij}\left[\frac{\partial^2 \underline{\hat{h}}_m}{\partial\Theta^i\partial\Theta^j} + \left(\Gamma^k_{im}\Gamma^l_{jk} - \frac{\partial\Gamma^l_{mj}}{\partial\Theta^i}\right)\underline{\hat{h}}_l \right.$$
$$\left. -2\Gamma^l_{jm}\frac{\partial\underline{\hat{h}}_l}{\partial\Theta^i} - \Gamma^k_{ij}\frac{\partial\underline{\hat{h}}_m}{\partial\Theta^k} + \Gamma^k_{ij}\Gamma^l_{km}\underline{\hat{h}}_l\right]\mathbf{g}^m . \tag{7.109b}$$

Similarly to (7.109b), one can express $\nabla^2\tilde{\mathbf{H}}$ with respect to the curvilinear basis vectors in a lengthy but straightforward manner. This remains to be done by the interested reader.

Consider the Laplacian of a scalar field according to (7.109a). By means of (7.16) and (7.17), it can also be written as

$$\nabla^2\bar{h} = \frac{1}{J}\frac{\partial}{\partial\Theta^i}\left[J g^{ij}\frac{\partial\bar{h}}{\partial\Theta^j}\right] . \quad \longleftarrow \text{see (7.81)} \tag{7.110}$$

The Laplace operator appears in many branches of physics and engineering. Examples of which include the heat conduction equations of solids, convection-diffusion problems and equilibrium equations governing the fluid flow in porous media. Specifically, the so-called **Laplace equation** $\nabla^2\Phi = 0$ and, its inhomogeneous form, **Poisson equation** $\nabla^2\Phi = \Psi$ are frequently seen in the literature. If a scalar (vector) field $\bar{h}$ ($\hat{\mathbf{h}}$) satisfies the Laplace equation $\nabla^2\bar{h} = 0$ ($\nabla^2\hat{\mathbf{h}} = \mathbf{0}$), then it is called *harmonic*.

The *Hessian operator* (or simply *Hessian*) is defined as the tensor product of the Nabla operator $\boldsymbol{\nabla}$ with itself. This second-order differential **tensor** operator, which increases the order of a tensor field by 2, in Cartesian coordinates represents

$$\boxed{\boldsymbol{\nabla}\otimes\boldsymbol{\nabla} = \frac{\partial}{\partial\mathbf{x}}\otimes\frac{\partial}{\partial\mathbf{x}} = \frac{\partial}{\partial x_k}\widehat{\mathbf{e}}_k\otimes\frac{\partial}{\partial x_l}\widehat{\mathbf{e}}_l = \frac{\partial^2}{\partial x_k\partial x_l}\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l .} \tag{7.111}$$

Accordingly, the Hessian of Cartesian tensor fields will render

$$\bar{h}\boldsymbol{\nabla}\otimes\boldsymbol{\nabla} = \frac{\partial^2\bar{h}}{\partial x_k\partial x_l}\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l , \tag{7.112a}$$

$$\hat{\mathbf{h}}\otimes\boldsymbol{\nabla}\otimes\boldsymbol{\nabla} = \frac{\partial^2\hat{h}_j}{\partial x_k\partial x_l}\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l , \tag{7.112b}$$

$$\tilde{\mathbf{H}}\otimes\boldsymbol{\nabla}\otimes\boldsymbol{\nabla} = \frac{\partial^2\tilde{H}_{ij}}{\partial x_k\partial x_l}\widehat{\mathbf{e}}_i\otimes\widehat{\mathbf{e}}_j\otimes\widehat{\mathbf{e}}_k\otimes\widehat{\mathbf{e}}_l . \tag{7.112c}$$

By comparing (7.105)$_{1-3}$ with (7.112a)–(7.112c), one can simply write

$$\boxed{\left(\bar{h}\nabla \otimes \nabla\right) : \mathbf{I} = \nabla^2 \bar{h} \ , \ \left(\hat{\mathbf{h}} \otimes \nabla \otimes \nabla\right) : \mathbf{I} = \nabla^2 \hat{\mathbf{h}} \ , \ \left(\tilde{\mathbf{H}} \otimes \nabla \otimes \nabla\right) : \mathbf{I} = \nabla^2 \tilde{\mathbf{H}} \ .} \tag{7.113}$$

The result (7.113)$_1$ shows that the Laplacian of a scalar field is basically the trace of its Hessian.

The curvilinear form of (7.111)$_3$ is given by

$$\begin{aligned} \nabla \otimes \nabla &= \frac{\partial}{\partial \Theta^k} \mathbf{g}^k \otimes \frac{\partial}{\partial \Theta^l} \mathbf{g}^l \\ &= \frac{\partial^2}{\partial \Theta^k \partial \Theta^l} \mathbf{g}^k \otimes \mathbf{g}^l + \frac{\partial}{\partial \Theta^l} \mathbf{g}^k \otimes \frac{\partial \mathbf{g}^l}{\partial \Theta^k} \\ &= \frac{\partial^2}{\partial \Theta^k \partial \Theta^l} \mathbf{g}^k \otimes \mathbf{g}^l - \Gamma^l_{km} \frac{\partial}{\partial \Theta^l} \mathbf{g}^k \otimes \mathbf{g}^m \\ &= \boxed{\frac{\partial^2}{\partial \Theta^k \partial \Theta^l} \mathbf{g}^k \otimes \mathbf{g}^l - \Gamma^m_{kl} \frac{\partial}{\partial \Theta^m} \mathbf{g}^k \otimes \mathbf{g}^l \ .} \end{aligned} \tag{7.114}$$

Consistent with (7.109a)–(7.109b), one then arrives at

$$\bar{h}\nabla \otimes \nabla = \frac{\partial^2 \bar{h}}{\partial \Theta^k \partial \Theta^l} \mathbf{g}^k \otimes \mathbf{g}^l - \Gamma^m_{kl} \frac{\partial \bar{h}}{\partial \Theta^m} \mathbf{g}^k \otimes \mathbf{g}^l \ , \tag{7.115a}$$

$$\begin{aligned} \hat{\mathbf{h}} \otimes \nabla \otimes \nabla = &\left[\frac{\partial^2 \underline{\hat{h}}^j}{\partial \Theta^k \partial \Theta^l} + \left(\frac{\partial \Gamma^j_{ln}}{\partial \Theta^k} + \Gamma^j_{km} \Gamma^m_{ln} \right) \underline{\hat{h}}^n \right. \\ &\left. + \Gamma^j_{ln} \frac{\partial \underline{\hat{h}}^n}{\partial \Theta^k} + \Gamma^j_{kn} \frac{\partial \underline{\hat{h}}^n}{\partial \Theta^l} - \Gamma^m_{kl} \frac{\partial \underline{\hat{h}}^j}{\partial \Theta^m} - \Gamma^m_{kl} \Gamma^j_{mn} \underline{\hat{h}}^n \right] \mathbf{g}_j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \\ = &\left[\frac{\partial^2 \underline{\hat{h}}_j}{\partial \Theta^k \partial \Theta^l} + \left(\Gamma^m_{kj} \Gamma^n_{lm} - \frac{\partial \Gamma^n_{jl}}{\partial \Theta^k} \right) \underline{\hat{h}}_n \right. \\ &\left. - \Gamma^n_{lj} \frac{\partial \underline{\hat{h}}_n}{\partial \Theta^k} - \Gamma^n_{kj} \frac{\partial \underline{\hat{h}}_n}{\partial \Theta^l} - \Gamma^m_{kl} \frac{\partial \underline{\hat{h}}_j}{\partial \Theta^m} + \Gamma^m_{kl} \Gamma^n_{mj} \underline{\hat{h}}_n \right] \mathbf{g}^j \otimes \mathbf{g}^k \otimes \mathbf{g}^l \ . \end{aligned} \tag{7.115b}$$

It should not be difficult now to decompose the Hessian of a tensor field with respect to the curvilinear basis vectors. This is left to be undertaken by the ambitious reader.

Within a solution technique such as *Newton's method*, a tensor field is linearly approximated. This gives a basic information regarding the behavior of that tensorial field variable. However, the approximation using second-order derivatives provides more detailed description of a field variable. With regard to this, an application of the Hessian operator will be in the second-order approximations

$$\bar{h}(\mathbf{x}+d\mathbf{x}) \approx \underbrace{\bar{h}(\mathbf{x}) + \frac{\partial \bar{h}}{\partial \mathbf{x}} \cdot d\mathbf{x} + \frac{1}{2}\frac{\partial^2 \bar{h}}{\partial \mathbf{x}\partial \mathbf{x}} : d\mathbf{x} \otimes d\mathbf{x}}_{\text{or } \bar{h}(\mathbf{x}) + (\bar{h}\nabla)\cdot d\mathbf{x} + \frac{1}{2}(\bar{h}\nabla\otimes\nabla) : (d\mathbf{x}\otimes d\mathbf{x}) \approx \bar{h}(\mathbf{x}+d\mathbf{x})} , \tag{7.116a}$$

$$\hat{\mathbf{h}}(\mathbf{x}+d\mathbf{x}) \approx \underbrace{\hat{\mathbf{h}}(\mathbf{x}) + \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{x}} d\mathbf{x} + \frac{1}{2}\frac{\partial^2 \hat{\mathbf{h}}}{\partial \mathbf{x}\partial \mathbf{x}} : d\mathbf{x} \otimes d\mathbf{x}}_{\text{or } \hat{\mathbf{h}}(\mathbf{x}) + \left(\hat{\mathbf{h}}\otimes\nabla\right) d\mathbf{x} + \frac{1}{2}\left(\hat{\mathbf{h}}\otimes\nabla\otimes\nabla\right) : (d\mathbf{x}\otimes d\mathbf{x}) \approx \hat{\mathbf{h}}(\mathbf{x}+d\mathbf{x})} , \tag{7.116b}$$

$$\tilde{\mathbf{H}}(\mathbf{x}+d\mathbf{x}) \approx \underbrace{\tilde{\mathbf{H}}(\mathbf{x}) + \frac{\partial \tilde{\mathbf{H}}}{\partial \mathbf{x}} d\mathbf{x} + \frac{1}{2}\frac{\partial^2 \tilde{\mathbf{H}}}{\partial \mathbf{x}\partial \mathbf{x}} : d\mathbf{x} \otimes d\mathbf{x}}_{\text{or } \tilde{\mathbf{H}}(\mathbf{x}) + \left(\tilde{\mathbf{H}}\otimes\nabla\right) d\mathbf{x} + \frac{1}{2}\left(\tilde{\mathbf{H}}\otimes\nabla\otimes\nabla\right) : (d\mathbf{x}\otimes d\mathbf{x}) \approx \tilde{\mathbf{H}}(\mathbf{x}+d\mathbf{x})} . \tag{7.116c}$$

7.6 Exercises

Exercise 7.1

Suppose one is given a smooth scalar field $\bar{h} = x_1 x_2 x_3 - x_1$. Find a unit vector $\widehat{\mathbf{n}}$ perpendicular to the isosurface $\bar{h} = \text{constant}$ passing through (2, 0, 3).

Solution. First, using (7.68a)$_2$, the gradient of $\bar{h} = x_1 x_2 x_3 - x_1$ at $\mathbf{x} = (2, 0, 3)$ renders

$$\text{grad}\bar{h} = \frac{\partial \bar{h}}{\partial x_i}\widehat{\mathbf{e}}_i = (x_2 x_3 - 1)\,\widehat{\mathbf{e}}_1 + (x_1 x_3)\,\widehat{\mathbf{e}}_2 + (x_1 x_2)\,\widehat{\mathbf{e}}_3 = -\widehat{\mathbf{e}}_1 + 6\,\widehat{\mathbf{e}}_2 \ .$$

The fact that the gradient vector is normal to the level surface helps then obtain

$$\widehat{\mathbf{n}} = \frac{\text{grad}\bar{h}}{\left|\text{grad}\bar{h}\right|} = \frac{-\widehat{\mathbf{e}}_1 + 6\,\widehat{\mathbf{e}}_2}{\sqrt{37}} ,$$

where (1.39)$_3$ has been used.

Exercise 7.2

Consider a force $\mathbf{F}$ with the magnitude F acting in the direction radially away from the origin at $\mathbf{p} = \left(2a, 3a, 2\sqrt{3}c\right)$ on a hyperboloid of one sheet defined by $\bar{h} = x_1^2/a^2 + x_2^2/a^2 - x_3^2/c^2 = 1$. Compute the Cartesian components of $\mathbf{F}$ lying in the tangential plane to the surface at $\left(2a, 3a, 2\sqrt{3}c\right)$.

Solution. First, the vector $\mathbf{F}$ in the desired direction can be expressed as

$$\mathbf{F} = F\,\widehat{\mathbf{n}}_{\mathbf{F}} = \frac{(2Fa)\,\widehat{\mathbf{e}}_1 + (3Fa)\,\widehat{\mathbf{e}}_2 + \left(2\sqrt{3}Fc\right)\widehat{\mathbf{e}}_3}{\sqrt{13a^2 + 12c^2}} \,.$$

Using (7.68a)$_2$, the gradient of $\bar{h} = x_1^2/a^2 + x_2^2/a^2 - x_3^2/c^2$ at the point with the coordinates $\left(2a, 3a, 2\sqrt{3}c\right)$ then renders

$$\left.\frac{\partial \bar{h}}{\partial \mathbf{x}}\right|_{(2a,3a,2\sqrt{3}c)} = \left.\frac{\partial \bar{h}}{\partial x_i}\right|_{(2a,3a,2\sqrt{3}c)} \widehat{\mathbf{e}}_i = \frac{4}{a}\widehat{\mathbf{e}}_1 + \frac{6}{a}\widehat{\mathbf{e}}_2 - \frac{4\sqrt{3}}{c}\widehat{\mathbf{e}}_3 \,,$$

Subsequently, the unit normal to the surface becomes,

$$\widehat{\mathbf{n}} = \frac{\sqrt{a^2c^2}}{\sqrt{13c^2 + 12a^2}} \left[\frac{2}{a}\widehat{\mathbf{e}}_1 + \frac{3}{a}\widehat{\mathbf{e}}_2 - \frac{2\sqrt{3}}{c}\widehat{\mathbf{e}}_3\right] .$$

Guided by (2.140b), the force located in the tangent plane is basically the vector rejection of $\mathbf{F}$ from $\widehat{\mathbf{n}}$, that is,

$$\mathbf{F}_t = \mathbf{F} - \mathbf{F}_n = \mathbf{F} - (\widehat{\mathbf{n}} \cdot \mathbf{F})\,\widehat{\mathbf{n}} \,.$$

At the end, the Cartesian components of $\mathbf{F}_t$ are given by

$$(\mathbf{F}_t)_1 = \frac{2Fa}{\sqrt{13a^2 + 12c^2}} \left[1 - \frac{c^2}{13c^2 + 12a^2}\right] ,$$
$$(\mathbf{F}_t)_2 = \frac{3Fa}{\sqrt{13a^2 + 12c^2}} \left[1 - \frac{c^2}{13c^2 + 12a^2}\right] ,$$
$$(\mathbf{F}_t)_3 = \frac{2\sqrt{3}Fc}{\sqrt{13a^2 + 12c^2}} \left[1 - \frac{a^2}{13c^2 + 12a^2}\right] .$$

Exercise 7.3

Suppose one is given a scalar field $\bar{h}\,(x_1, x_2, x_3) = \mathrm{e}^{x_1} \cos(3x_1 - 2x_2 + x_3)$. Then, calculate the directional derivative of $\bar{h}$ at a point P corresponding to $\mathbf{x} = (1, 1, 1)$ in the direction of the line with the parametric equations $x_1 = 1 + 3t\,,\ x_2 = 2 - 2t\,,\ x_3 = 3 - t\ ;\ t \in \mathrm{R}$, for increasing values of x_1.

Solution. First, the unit vector in the desired direction renders

$$\widehat{\mathbf{n}} = \frac{3\,\widehat{\mathbf{e}}_1 - 2\,\widehat{\mathbf{e}}_2 - \widehat{\mathbf{e}}_3}{\sqrt{14}} \,.$$

Then, guided by (7.68a)$_2$, the gradient of $\bar{h} = \mathrm{e}^{x_1} \cos(3x_1 - 2x_2 + x_3)$ at P with $(1, 1, 1)$ becomes

$$\left.\frac{\partial \bar{h}}{\partial \mathbf{x}}\right|_{(1,1,1)} = \left.\frac{\partial \bar{h}}{\partial x_i}\right|_{(1,1,1)} \widehat{\mathbf{e}}_i = \mathrm{e}\,[\cos(2) - 3\sin(2)]\,\widehat{\mathbf{e}}_1 + 2\mathrm{e}\sin(2)\,\widehat{\mathbf{e}}_2 - \mathrm{e}\sin(2)\,\widehat{\mathbf{e}}_3 \ .$$

Finally, using (1.38)$_7$ and (7.4a)$_3$,

$$D_{\widehat{\mathbf{n}}}\bar{h}\,(\mathbf{x})\big|_{(1,1,1)} = \left.\frac{\partial \bar{h}}{\partial \mathbf{x}}\right|_{(1,1,1)} \cdot \widehat{\mathbf{n}} = \frac{3\mathrm{e}\cos(2) - 12\mathrm{e}\sin(2)}{\sqrt{14}} \ .$$

Exercise 7.4

Consider two covariant vector fields $\mathbf{u} = \underline{u}_i \mathbf{g}^i$, $\mathbf{v} = \underline{v}_j \mathbf{g}^j$ and a contravariant tensor field $\mathbf{A} = \underline{A}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j$. Let Φ be a scalar field that is constructed from these tensorial field variables according to $\Phi = \mathbf{u} \cdot \mathbf{A}\mathbf{v} = \underline{u}_i \, \underline{A}^{ij} \, \underline{v}_j$. Show that the covariant derivative of this scalar field is equal to its partial derivative, i.e.

$$\Phi|_k = \frac{\partial \Phi}{\partial \Theta^k} \ . \tag{7.117}$$

Solution. By means of (7.7)$_2$, (7.25b), (7.27a)$_2$ and in light of the product rule (7.48b), one will have

$$\begin{aligned}
\Phi|_k &= \left(\underline{u}_i \, \underline{A}^{ij} \, \underline{v}_j\right)\big|_k \\
&= \underline{u}_i\big|_k \, \underline{A}^{ij} \, \underline{v}_j + \underline{u}_i \, \underline{A}^{ij}\big|_k \, \underline{v}_j + \underline{u}_i \, \underline{A}^{ij} \, \underline{v}_j\big|_k \\
&= \frac{\partial \, \underline{u}_i}{\partial \Theta^k} \, \underline{A}^{ij} \, \underline{v}_j \cancel{- \Gamma^m_{ik} \, \underline{u}_m \, \underline{A}^{ij} \, \underline{v}_j} \\
&\quad + \underline{u}_i \frac{\partial \, \underline{A}^{ij}}{\partial \Theta^k} \, \underline{v}_j + \underbrace{\underline{u}_i \Gamma^i_{km} \, \underline{A}^{mj} \, \underline{v}_j}_{= \underline{u}_m \Gamma^m_{ki} \, \underline{A}^{ij} \, \underline{v}_j = \cancel{\underline{u}_m \Gamma^m_{ik} \, \underline{A}^{ij} \, \underline{v}_j}} + \underbrace{\underline{u}_i \Gamma^j_{km} \, \underline{A}^{im} \, \underline{v}_j}_{= \underline{u}_i \Gamma^m_{kj} \, \underline{A}^{ij} \, \underline{v}_m = \cancel{\underline{u}_i \Gamma^m_{jk} \, \underline{A}^{ij} \, \underline{v}_m}} \\
&\quad + \underline{u}_i \, \underline{A}^{ij} \frac{\partial \, \underline{v}_j}{\partial \Theta^k} \cancel{- \underline{u}_i \, \underline{A}^{ij} \Gamma^m_{jk} \, \underline{v}_m} \\
&= \frac{\partial \left(\underline{u}_i \, \underline{A}^{ij} \, \underline{v}_j\right)}{\partial \Theta^k} \\
&= \frac{\partial \Phi}{\partial \Theta^k} \ .
\end{aligned}$$

It should be emphasized that this identity generally holds true irrespective of the given types of components. The interested reader may thus want to verify this result for other forms of Φ such as $\underline{u}^i \, \underline{A}_{ij} \, \underline{v}^j$ or $\underline{u}^i g_{ij} \, \underline{A}^{jk} g_{kl} \, \underline{v}^l$.

Exercise 7.5

Verify the Bianchi identities (7.62) and (7.63).

Solution. The algebraic Bianchi identity $(7.62)_1$ can be shown by means of $(7.7)_2$ and (7.50) as follows:

$$\mathbb{R}^{i}_{.jkl} + \mathbb{R}^{i}_{.klj} + \mathbb{R}^{i}_{.ljk} = \frac{\partial \Gamma^i_{jl}}{\partial \Theta^k} - \frac{\partial \Gamma^i_{jk}}{\partial \Theta^l} + \Gamma^i_{km}\Gamma^m_{lj} - \Gamma^i_{lm}\Gamma^m_{kj}$$
$$+ \frac{\partial \Gamma^i_{kj}}{\partial \Theta^l} - \frac{\partial \Gamma^i_{kl}}{\partial \Theta^j} + \Gamma^i_{lm}\Gamma^m_{jk} - \Gamma^i_{jm}\Gamma^m_{lk}$$
$$+ \frac{\partial \Gamma^i_{lk}}{\partial \Theta^j} - \frac{\partial \Gamma^i_{lj}}{\partial \Theta^k} + \Gamma^i_{jm}\Gamma^m_{kl} - \Gamma^i_{km}\Gamma^m_{jl}$$
$$= 0 \, .$$

The fully covariant form $(7.62)_2$ can then be easily obtained from the mixed representation $(7.62)_1$ by index juggling.

The differential Bianchi identity requires more consideration. It can be shown in different ways. The proof here mainly relies on taking the covariant derivative of (7.49b) three times upon cyclic permutation of the indices k, l and m having in mind that the covariant derivative satisfies the product rule of differentiation.

To begin with, consider

$$\hat{\underline{h}}_j\Big|_{klm} - \hat{\underline{h}}_j\Big|_{lkm} = \mathbb{R}^{i}_{.jkl}\Big|_m \hat{\underline{h}}_i + \mathbb{R}^{i}_{.jkl}\ \hat{\underline{h}}_i\Big|_m \,, \tag{7.118a}$$

$$\hat{\underline{h}}_j\Big|_{lmk} - \hat{\underline{h}}_j\Big|_{mlk} = \mathbb{R}^{i}_{.jlm}\Big|_k \hat{\underline{h}}_i + \mathbb{R}^{i}_{.jlm}\ \hat{\underline{h}}_i\Big|_k \,, \tag{7.118b}$$

$$\hat{\underline{h}}_j\Big|_{mkl} - \hat{\underline{h}}_j\Big|_{kml} = \mathbb{R}^{i}_{.jmk}\Big|_l \hat{\underline{h}}_i + \mathbb{R}^{i}_{.jmk}\ \hat{\underline{h}}_i\Big|_l \,. \tag{7.118c}$$

Consider now the fact that $\hat{\underline{h}}_i\Big|_j$ is a second-order tensor. Consequently, using (7.60), the above relations represent

$$\hat{\underline{h}}_j\Big|_{klm} - \hat{\underline{h}}_j\Big|_{kml} = \mathbb{R}^{i}_{.jlm}\ \hat{\underline{h}}_i\Big|_k + \mathbb{R}^{i}_{.klm}\ \hat{\underline{h}}_j\Big|_i \,, \tag{7.119a}$$

$$\hat{\underline{h}}_j\Big|_{lmk} - \hat{\underline{h}}_j\Big|_{lkm} = \mathbb{R}^{i}_{.jmk}\ \hat{\underline{h}}_i\Big|_l + \mathbb{R}^{i}_{.lmk}\ \hat{\underline{h}}_j\Big|_i \,, \tag{7.119b}$$

$$\hat{\underline{h}}_j\Big|_{mkl} - \hat{\underline{h}}_j\Big|_{mlk} = \mathbb{R}^{i}_{.jkl}\ \hat{\underline{h}}_i\Big|_m + \mathbb{R}^{i}_{.mkl}\ \hat{\underline{h}}_j\Big|_i \,. \tag{7.119c}$$

From (7.118a)–(7.119c), it follows that

$$\underline{\mathbb{R}}^{i}_{(klm)}\ \underline{\hat{h}}_{j}\Big|_{i} = \underline{\mathbb{R}}^{i}_{j(kl\,|\,m)}\,\underline{\hat{h}}_{i}\ , \tag{7.120}$$

where

$$\underline{\mathbb{R}}^{i}_{(klm)} := \underline{\mathbb{R}}^{i}_{.\,klm}{}^{\cdot\cdot\cdot} + \underline{\mathbb{R}}^{i}_{.\,lmk}{}^{\cdot\cdot\cdot} + \underline{\mathbb{R}}^{i}_{.\,mkl}{}^{\cdot\cdot\cdot}\ , \tag{7.121a}$$

$$\underline{\mathbb{R}}^{i}_{j(kl\,|\,m)} := \underline{\mathbb{R}}^{i}_{.\,jkl}{}^{\cdot\cdot\cdot}\Big|_{m} + \underline{\mathbb{R}}^{i}_{.\,jlm}{}^{\cdot\cdot\cdot}\Big|_{k} + \underline{\mathbb{R}}^{i}_{.\,jmk}{}^{\cdot\cdot\cdot}\Big|_{l}\ . \tag{7.121b}$$

Notice that the algebraic Bianchi identity now takes the form

$$\underline{\mathbb{R}}^{i}_{(klm)} = 0\ .$$

The fact that $\underline{\hat{h}}_{i}$ is arbitrary in (7.120) then implies the differential Bianchi identity

$$\underline{\mathbb{R}}^{i}_{j(kl\,|\,m)} = 0\ .$$

This identity can also be obtained by use of (7.50) and (7.64a) in a lengthy but straightforward manner.

The desired relation $(7.63)_2$ finally follows from $(7.63)_1$ by index juggling.

Exercise 7.6

Consider the Cartesian vector field

$$\hat{\mathbf{h}} = \hat{\mathbf{h}}\,(\mathbf{x}) = (x_1 x_2 x_3)\,\widehat{\mathbf{e}}_1 + (x_1 x_2)\,\widehat{\mathbf{e}}_2 + (x_1)\,\widehat{\mathbf{e}}_3\ .$$

Determine grad$\hat{\mathbf{h}}$, div$\hat{\mathbf{h}}$ and curl$\hat{\mathbf{h}}$.

Moreover, consider the (linear) transformation equations

$$x_1 = \Theta^1\Theta^3 \quad , \quad x_2 = \Theta^2\Theta^3 \quad , \quad x_3 = \Theta^3 - 1\ .$$

First, compute the curvilinear bases $\{\mathbf{g}_i\}$ and $\{\mathbf{g}^i\}$. Then, express $\hat{\mathbf{h}}$ with respect to the covariant basis $\{\mathbf{g}_i\}$. Finally, calculate the **gradient**, **divergence** and **curl** of the resulting contravariant vector field in matrix form.

Solution. The desired tensor, scalar and vector fields are first computed with respect to the standard basis.

Using $(7.70)_2$,

$$\left[\,\mathrm{grad}\hat{\mathbf{h}}\,\right] = \left[\,\frac{\partial \hat{h}_i}{\partial x_j}\,\right] = \begin{bmatrix} x_2 x_3 & x_1 x_3 & x_1 x_2 \\ x_2 & x_1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\ .$$

From (7.78),

$$\text{div}\hat{\mathbf{h}} = \frac{\partial \hat{h}_i}{\partial x_i} = \frac{\partial \hat{h}_1}{\partial x_1} + \frac{\partial \hat{h}_2}{\partial x_2} + \frac{\partial \hat{h}_3}{\partial x_3} = x_2 x_3 + x_1 \ .$$

By means of (7.92),

$$\left[\text{curl}\hat{\mathbf{h}} \right] = \begin{bmatrix} \partial \hat{h}_3/\partial x_2 - \partial \hat{h}_2/\partial x_3 \\ \partial \hat{h}_1/\partial x_3 - \partial \hat{h}_3/\partial x_1 \\ \partial \hat{h}_2/\partial x_1 - \partial \hat{h}_1/\partial x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ x_1 x_2 - 1 \\ x_2 - x_1 x_3 \end{bmatrix} .$$

In the following, these tensorial field variables are represented with respect to the curvilinear basis vectors.

To begin with, one needs to have the covariant basis vectors. The tangent vectors $\mathbf{g}_i = \partial \mathbf{x}/\partial \Theta^i$, according to $(5.3)_1$, for the problem at hand become

$$\underbrace{\mathbf{g}_1 = \Theta^3 \widehat{\mathbf{e}}_1 \quad , \quad \mathbf{g}_2 = \Theta^3 \widehat{\mathbf{e}}_2 \quad , \quad \mathbf{g}_3 = \Theta^1 \widehat{\mathbf{e}}_1 + \Theta^2 \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3}_{\text{or} \quad \Theta^3 \widehat{\mathbf{e}}_1 = \mathbf{g}_1 \ , \ \Theta^3 \widehat{\mathbf{e}}_2 = \mathbf{g}_2 \ , \ \Theta^3 \widehat{\mathbf{e}}_3 = -\Theta^1 \mathbf{g}_1 - \Theta^2 \mathbf{g}_2 + \Theta^3 \mathbf{g}_3} \ .$$

From $(5.30)_2$ and $(5.38)_3$, one then obtains

$$J = \det \left[\ \mathbf{g}_1 \ \ \mathbf{g}_2 \ \ \mathbf{g}_3 \ \right] = \left(\Theta^3 \right)^2 \ ,$$

and

$$\left[g_{ij} \right] = \begin{bmatrix} \mathbf{g}_1 \cdot \mathbf{g}_1 & \mathbf{g}_1 \cdot \mathbf{g}_2 & \mathbf{g}_1 \cdot \mathbf{g}_3 \\ \mathbf{g}_2 \cdot \mathbf{g}_1 & \mathbf{g}_2 \cdot \mathbf{g}_2 & \mathbf{g}_2 \cdot \mathbf{g}_3 \\ \mathbf{g}_3 \cdot \mathbf{g}_1 & \mathbf{g}_3 \cdot \mathbf{g}_2 & \mathbf{g}_3 \cdot \mathbf{g}_3 \end{bmatrix} = \begin{bmatrix} \left(\Theta^3 \right)^2 & 0 & \Theta^1 \Theta^3 \\ 0 & \left(\Theta^3 \right)^2 & \Theta^2 \Theta^3 \\ \Theta^1 \Theta^3 & \Theta^2 \Theta^3 & \left(\Theta^1 \right)^2 + \left(\Theta^2 \right)^2 + 1 \end{bmatrix} .$$

One should also have

$$\begin{array}{lll} \dfrac{\partial \mathbf{g}_1}{\partial \Theta^1} = \mathbf{0} \quad , & \dfrac{\partial \mathbf{g}_1}{\partial \Theta^2} = \mathbf{0} \quad , & \dfrac{\partial \mathbf{g}_1}{\partial \Theta^3} = \widehat{\mathbf{e}}_1 \ , \\ \dfrac{\partial \mathbf{g}_2}{\partial \Theta^1} = \mathbf{0} \quad , & \dfrac{\partial \mathbf{g}_2}{\partial \Theta^2} = \mathbf{0} \quad , & \dfrac{\partial \mathbf{g}_2}{\partial \Theta^3} = \widehat{\mathbf{e}}_2 \ , \\ \dfrac{\partial \mathbf{g}_3}{\partial \Theta^1} = \widehat{\mathbf{e}}_1 \quad , & \dfrac{\partial \mathbf{g}_3}{\partial \Theta^2} = \widehat{\mathbf{e}}_2 \quad , & \dfrac{\partial \mathbf{g}_3}{\partial \Theta^3} = \mathbf{0} \ . \end{array}$$

It is easy to obtain the inverse of the given transformation equations:

$$\Theta^1 = \frac{x_1}{x_3 + 1} \quad , \quad \Theta^2 = \frac{x_2}{x_3 + 1} \quad , \quad \Theta^3 = x_3 + 1 \ .$$

Accordingly, the dual vectors $\mathbf{g}^i = \partial\Theta^i/\partial\mathbf{x}$ in (5.28) take the following form

$$\mathbf{g}^1 = \frac{1}{\Theta^3}\widehat{\mathbf{e}}_1 - \frac{\Theta^1}{\Theta^3}\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}^2 = \frac{1}{\Theta^3}\widehat{\mathbf{e}}_2 - \frac{\Theta^2}{\Theta^3}\widehat{\mathbf{e}}_3 \quad , \quad \mathbf{g}^3 = \widehat{\mathbf{e}}_3 \ .$$

Having obtained the partial derivatives of tangent vectors together with the dual vectors, the nonzero entries of the Christoffel symbols $\Gamma^k_{ij} = \left(\partial\mathbf{g}_i/\partial\Theta^j\right)\cdot\mathbf{g}^k$ in $(7.7)_1$ render

$$\Gamma^1_{13} = \Gamma^1_{31} = \frac{1}{\Theta^3} \quad , \quad \Gamma^1_{23} = \Gamma^1_{32} = \frac{1}{\Theta^3} \ .$$

Using the functional relationships between (x_1, x_2, x_3) and $\left(\Theta^1, \Theta^2, \Theta^3\right)$ together with $\{\widehat{\mathbf{e}}_i\}$ and $\{\mathbf{g}_i\}$, the given Cartesian vector field can now be expressed as

$$\hat{\mathbf{h}} = \left[\Theta^1\Theta^2\Theta^3\left(\Theta^3 - 1\right) - \left(\Theta^1\right)^2\right]\mathbf{g}_1 + \left[\Theta^1\Theta^2\Theta^3 - \Theta^1\Theta^2\right]\mathbf{g}_2 + \left[\Theta^1\Theta^3\right]\mathbf{g}_3 \ .$$

Considering $(7.71\text{a})_1$, the mixed contra-covariant components $\left[\underline{\hat{h}}^i\Big|_j\right]$ of the tensor field grad$\hat{\mathbf{h}}$ with respect to the co-contravariant basis $\left\{\mathbf{g}_i \otimes \mathbf{g}^j\right\}$ are given by

$$\begin{aligned}
\underline{\hat{h}}^1\Big|_1 &= \Theta^2\Theta^3\left(\Theta^3 - 1\right) - \Theta^1 \ , \\
\underline{\hat{h}}^1\Big|_2 &= \Theta^1\Theta^3\left(\Theta^3 - 1\right) \ , \\
\underline{\hat{h}}^1\Big|_3 &= -2\Theta^1\Theta^2 + 3\Theta^1\Theta^2\Theta^3 - \left(\Theta^1\right)^2\left(\Theta^3\right)^{-1} \ , \\
\underline{\hat{h}}^2\Big|_1 &= \Theta^2\Theta^3 - \Theta^2 \ , \\
\underline{\hat{h}}^2\Big|_2 &= \Theta^1\Theta^3 \ , \\
\underline{\hat{h}}^2\Big|_3 &= 2\Theta^1\Theta^2 - \Theta^1\Theta^2\left(\Theta^3\right)^{-1} \ , \\
\underline{\hat{h}}^3\Big|_1 &= \Theta^3 \ , \\
\underline{\hat{h}}^3\Big|_2 &= 0 \ , \\
\underline{\hat{h}}^3\Big|_3 &= \Theta^1 \ .
\end{aligned}$$

Accordingly, div$\hat{\mathbf{h}}$ in $(7.79)_1$ takes the form

$$\text{div}\hat{\mathbf{h}} = \underline{\hat{h}}^i\Big|_i = \underline{\hat{h}}^1\Big|_1 + \underline{\hat{h}}^2\Big|_2 + \underline{\hat{h}}^3\Big|_3 = \Theta^2\Theta^3\left(\Theta^3 - 1\right) + \Theta^1\Theta^3 \ .$$

Note that this result can easily be verified by introducing the given transformation equations into div$\hat{\mathbf{h}} = x_2x_3 + x_1$.

Guided by (7.96)$_1$, the contravariant components $\left(\text{curl}\hat{\mathbf{h}}\right)^i = -J^{-1}\varepsilon^{ijk} g_{jl} \left.\underline{\hat{h}}^l\right|_k$ of curl$\hat{\mathbf{h}}$ can finally be collected in the single-column matrix

$$\left[\text{curl}\hat{\mathbf{h}}\right]^{\text{con}} = \begin{bmatrix} \left(\Theta^1\Theta^3 - \Theta^1 - \Theta^2\right)\Theta^1 \\ \left[2\Theta^1\Theta^2\left(\Theta^3\right)^2 - \Theta^1\Theta^2\Theta^3 - \left(\Theta^2\right)^2\Theta^3 - 1\right]\left(\Theta^3\right)^{-1} \\ \left(\Theta^1 + \Theta^2 - \Theta^1\Theta^3\right)\Theta^3 \end{bmatrix} .$$

Exercise 7.7

Let $\bar{h}\,(r, \theta, z)$ be a scalar field and $\hat{\mathbf{h}}$ be a vector field with the decomposition (7.73) in **cylindrical** coordinates. Further, let $\bar{h}\,(r, \theta, \phi)$ be a scalar field and $\hat{\mathbf{h}}$ be a vector field of the form (7.75) in **spherical** coordinates. Then, write a **computer program** to symbolically compute grad$\hat{\mathbf{h}}$, div$\hat{\mathbf{h}}$, curl$\hat{\mathbf{h}}$ and $\nabla^2\bar{h}$ in each of these widely used coordinate systems, see Exercise 8.3.

Solution. The interested reader can download the desired code for free from https://data.uni-hannover.de/dataset/exercises-tensor-analysis.

Exercise 7.8

Applying the Nabla operator of vector calculus to products of smooth scalar fields ϕ, ψ, vector fields $\mathbf{u}$, $\mathbf{v}$ and tensor fields $\mathbf{A}$, $\mathbf{B}$ will provide numerous identities. Prove some important ones that are listed in the following

$$\text{grad}\,(\phi\psi) = (\text{grad}\phi)\,\psi + \phi\,\text{grad}\psi \ , \tag{7.122a}$$

$$\text{grad}\,(\phi\mathbf{u}) = \mathbf{u} \otimes \text{grad}\phi + \phi\,\text{grad}\mathbf{u} \ , \tag{7.122b}$$

$$\text{grad}\,(\mathbf{u} \cdot \mathbf{v}) = \left(\text{grad}^{\mathrm{T}}\mathbf{u}\right)\mathbf{v} + \left(\text{grad}^{\mathrm{T}}\mathbf{v}\right)\mathbf{u} \ , \tag{7.122c}$$

$$\text{div}\,(\phi\mathbf{u}) = \phi\,\text{div}\mathbf{u} + \mathbf{u} \cdot \text{grad}\phi \ , \tag{7.122d}$$

$$\text{div}\,(\phi\mathbf{A}) = \phi\,\text{div}\mathbf{A} + \mathbf{A}\,\text{grad}\phi \ , \tag{7.122e}$$

$$\text{div}\left(\mathbf{A}^{\mathrm{T}}\mathbf{u}\right) = (\text{div}\mathbf{A}) \cdot \mathbf{u} + \mathbf{A} : \text{grad}\mathbf{u} \ , \tag{7.122f}$$

$$\text{div}\,(\mathbf{u} \otimes \mathbf{v}) = (\text{grad}\mathbf{u})\,\mathbf{v} + \mathbf{u}\,\text{div}\mathbf{v} \ , \tag{7.122g}$$

$$\text{div}\,(\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \text{curl}\mathbf{u} - \mathbf{u} \cdot \text{curl}\mathbf{v} \ , \tag{7.122h}$$

$$\text{div}\,(\mathbf{A}\mathbf{B}) = \text{grad}\mathbf{A} : \mathbf{B} + \mathbf{A}\,\text{div}\mathbf{B} \ , \tag{7.122i}$$

$$\text{curl}\,(\phi\mathbf{u}) = \text{grad}\phi \times \mathbf{u} + \phi\,\text{curl}\mathbf{u} \ , \tag{7.122j}$$

$$\begin{aligned} \text{curl}\,(\mathbf{u} \times \mathbf{v}) &= \mathbf{u}\,\text{div}\mathbf{v} - \mathbf{v}\,\text{div}\mathbf{u} + (\text{grad}\mathbf{u})\,\mathbf{v} - (\text{grad}\mathbf{v})\,\mathbf{u} \\ &= \text{div}\,(\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u}) \ . \end{aligned} \tag{7.122k}$$

Moreover, verify that

$$\text{grad}\,\text{div}\mathbf{u} = \text{div}\,\text{grad}^{\text{T}}\mathbf{u}\ , \tag{7.123a}$$
$$\text{div}\,\text{curl}\mathbf{u} = 0\ , \tag{7.123b}$$
$$\text{curl}\,\text{grad}\phi = \mathbf{0}\ , \tag{7.123c}$$
$$\text{curl}\,\text{curl}\mathbf{u} = \text{grad}\,\text{div}\mathbf{u} - \nabla^2\mathbf{u}\ , \tag{7.123d}$$
$$\nabla^2\text{grad}\phi = \text{grad}\nabla^2\phi\ , \tag{7.123e}$$
$$\nabla^2\text{curl}\mathbf{u} = \text{curl}\nabla^2\mathbf{u}\ , \tag{7.123f}$$
$$\nabla^2\left(\mathbf{u}\cdot\mathbf{v}\right) = \left(\nabla^2\mathbf{u}\right)\cdot\mathbf{v} + 2\left(\text{grad}\mathbf{u}\right) : \left(\text{grad}\mathbf{v}\right) + \mathbf{u}\cdot\left(\nabla^2\mathbf{v}\right)\ , \tag{7.123g}$$

and

$$\text{curl}\,\text{grad}\mathbf{u} = \mathbf{O}\ , \tag{7.124a}$$
$$\begin{aligned}\text{curl}\,\text{curl}\mathbf{A} = &\left[\nabla^2\left(\text{tr}\mathbf{A}\right) - \text{div}\,\text{div}\mathbf{A}\right]\mathbf{I} - \nabla^2\mathbf{A}^{\text{T}} \\ &+ \text{grad}\,\text{div}\mathbf{A}^{\text{T}} + \text{grad}^{\text{T}}\,\text{div}\mathbf{A} - \text{grad}\,\text{grad}\left(\text{tr}\mathbf{A}\right)\ .\end{aligned} \tag{7.124b}$$

Solution. Here, all desired relations will be verified in indicial notation. Recall that the result of operating gradient - or any of its related operators - on a tensor field is another tensor field which is basically independent of any coordinate system. Thus, the desired identities generally remain valid irrespective of chosen coordinate system. This motivates to use the Cartesian form of components for convenience. However, the ambitious reader can use the curvilinear forms of components for any verification.

The expression (7.122a): By means of $(7.68a)_2$ along with the product rule of differentiation,

$$\begin{aligned}\left(\text{grad}\left(\phi\psi\right)\right)_i &= \frac{\partial\left(\phi\psi\right)}{\partial x_i} \\ &= \frac{\partial\phi}{\partial x_i}\psi + \phi\frac{\partial\psi}{\partial x_i} \\ &= \left(\text{grad}\phi\right)_i\psi + \phi\left(\text{grad}\psi\right)_i\ .\end{aligned}$$

The expression (7.122b): By means of $(2.24)_4$, $(7.68a)_2$ and $(7.70)_2$ along with the product rule of differentiation,

$$\begin{aligned}\left(\text{grad}\left(\phi\mathbf{u}\right)\right)_{ij} &= \frac{\partial\left(\phi u_i\right)}{\partial x_j} \\ &= u_i\frac{\partial\phi}{\partial x_j} + \phi\frac{\partial u_i}{\partial x_j} \\ &= \left(\mathbf{u}\otimes\text{grad}\phi\right)_{ij} + \phi\left(\text{grad}\mathbf{u}\right)_{ij}\ .\end{aligned}$$

The expression (7.122c): By means of (1.38)$_7$, (2.49), (7.68a)$_2$ and (7.70)$_2$ along with the product rule of differentiation,

$$\begin{aligned}
(\text{grad}\,(\mathbf{u}\cdot\mathbf{v}))_j &= \frac{\partial\,(u_i v_i)}{\partial x_j} \\
&= \frac{\partial u_i}{\partial x_j} v_i + u_i \frac{\partial v_i}{\partial x_j} \\
&= \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)_{ij} (\mathbf{v})_i + \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}}\right)_{ij} (\mathbf{u})_i \\
&= \left(\text{grad}^{\mathrm{T}}\mathbf{u}\right)_{ji} (\mathbf{v})_i + \left(\text{grad}^{\mathrm{T}}\mathbf{v}\right)_{ji} (\mathbf{u})_i \ .
\end{aligned}$$

The expression (7.122d): By means of (1.38)$_7$, (7.68a)$_2$ and (7.78) along with the product rule of differentiation,

$$\begin{aligned}
\text{div}\,(\phi\mathbf{u}) &= \frac{\partial\,(\phi u_i)}{\partial x_i} \\
&= \frac{\partial \phi}{\partial x_i} u_i + \phi \frac{\partial u_i}{\partial x_i} \\
&= (\mathbf{u})_i\,(\text{grad}\phi)_i + \phi\,\text{div}\mathbf{u} \\
&= \mathbf{u}\cdot\text{grad}\phi + \phi\,\text{div}\mathbf{u} \ .
\end{aligned}$$

The expression (7.122e): By means of (2.22)$_3$, (7.68a)$_2$ and (7.83)$_2$ along with the product rule of differentiation,

$$\begin{aligned}
(\text{div}\,(\phi\mathbf{A}))_i &= \frac{\partial\left(\phi A_{ij}\right)}{\partial x_j} \\
&= \frac{\partial \phi}{\partial x_j} A_{ij} + \phi \frac{\partial A_{ij}}{\partial x_j} \\
&= (\mathbf{A})_{ij}\,(\text{grad}\phi)_j + \phi\,(\text{div}\mathbf{A})_i \\
&= (\mathbf{A}\text{grad}\phi)_i + \phi\,(\text{div}\mathbf{A})_i \ .
\end{aligned}$$

The expression (7.122f): By means of (1.38)$_7$, (2.22)$_3$, (2.49), (2.75)$_4$, (7.70)$_2$, (7.78) and (7.83)$_2$ along with the product rule of differentiation,

$$\begin{aligned}
\text{div}\left(\mathbf{A}^{\mathrm{T}}\mathbf{u}\right) &= \frac{\partial\left(A^{\mathrm{T}}_{ij} u_j\right)}{\partial x_i} \\
&= \frac{\partial\left(A_{ji} u_j\right)}{\partial x_i} \\
&= \frac{\partial A_{ji}}{\partial x_i} u_j + A_{ji} \frac{\partial u_j}{\partial x_i}
\end{aligned}$$

$$= (\text{div}\mathbf{A})_j (\mathbf{u})_j + (\mathbf{A})_{ji} (\text{grad}\mathbf{u})_{ji}$$
$$= (\text{div}\mathbf{A}) \cdot \mathbf{u} + \mathbf{A} : \text{grad}\mathbf{u} \ .$$

The expression (7.122g): By means of $(2.22)_3$, $(2.24)_4$, $(7.70)_2$, (7.78) and $(7.83)_2$ along with the product rule of differentiation,

$$\begin{aligned}(\text{div}\,(\mathbf{u} \otimes \mathbf{v}))_i &= \frac{\partial \left(u_i v_j\right)}{\partial x_j} \\ &= \frac{\partial u_i}{\partial x_j} v_j + u_i \frac{\partial v_j}{\partial x_j} \\ &= (\text{grad}\mathbf{u})_{ij} (\mathbf{v})_j + (\mathbf{u})_i \,\text{div}\mathbf{v} \\ &= ((\text{grad}\mathbf{u})\,\mathbf{v})_i + (\mathbf{u}\text{div}\mathbf{v})_i \ .\end{aligned}$$

The expression (7.122h): By means of $(1.38)_7$, (1.54), $(1.67)_5$, (7.78) and (7.91) along with the product rule of differentiation,

$$\begin{aligned}\text{div}\,(\mathbf{u} \times \mathbf{v}) &= \frac{\partial \left(u_i v_j \varepsilon_{ijk}\right)}{\partial x_k} \\ &= \frac{\partial u_i}{\partial x_k} v_j \varepsilon_{ijk} + u_i \frac{\partial v_j}{\partial x_k} \varepsilon_{ijk} \\ &= v_j \left(-\varepsilon_{jik} \frac{\partial u_i}{\partial x_k}\right) - u_i \left(-\varepsilon_{ijk} \frac{\partial v_j}{\partial x_k}\right) \\ &= (\mathbf{v})_j (\text{curl}\mathbf{u})_j - (\mathbf{u})_i (\text{curl}\mathbf{v})_i \\ &= \mathbf{v} \cdot \text{curl}\mathbf{u} - \mathbf{u} \cdot \text{curl}\mathbf{v} \ .\end{aligned}$$

The expression (7.122i): By means of (2.26), $(3.16\text{b})_5$, $(7.70)_4$ and $(7.83)_2$ along with the product rule of differentiation,

$$\begin{aligned}(\text{div}\,(\mathbf{AB}))_i &= \frac{\partial \left(A_{ik} B_{kj}\right)}{\partial x_j} \\ &= \frac{\partial A_{ik}}{\partial x_j} B_{kj} + A_{ik} \frac{\partial B_{kj}}{\partial x_j} \\ &= (\text{grad}\mathbf{A})_{ikj} (\mathbf{B})_{kj} + (\mathbf{A})_{ik} (\text{div}\mathbf{B})_k \\ &= (\text{grad}\mathbf{A} : \mathbf{B})_i + (\mathbf{A}\text{div}\mathbf{B})_i \ .\end{aligned}$$

The expression (7.122j): By means of (1.54), $(1.67)_5$, $(7.68\text{a})_2$ and (7.91) along with the product rule of differentiation,

$$\begin{aligned}(\text{curl}\,(\phi\mathbf{u}))_i &= -\varepsilon_{ijk}\frac{\partial\left(\phi u_j\right)}{\partial x_k}\\ &= -\varepsilon_{ijk}\frac{\partial\phi}{\partial x_k}u_j - \phi\varepsilon_{ijk}\frac{\partial u_j}{\partial x_k}\\ &= (\text{grad}\phi)_k\,u_j\varepsilon_{kji} - \phi\varepsilon_{ijk}\frac{\partial u_j}{\partial x_k}\\ &= (\text{grad}\phi\times\mathbf{u})_i + (\phi\text{curl}\mathbf{u})_i\ .\end{aligned}$$

The expression (7.122k): By means of (1.36), (1.54), (1.58a), $(1.67)_5$, $(2.22)_3$, $(7.70)_2$, (7.78) and (7.91) along with the product rule of differentiation,

$$\begin{aligned}(\text{curl}\,(\mathbf{u}\times\mathbf{v}))_i &= -\varepsilon_{ijk}\frac{\partial\,(\mathbf{u}\times\mathbf{v})_j}{\partial x_k}\\ &= \varepsilon_{ikj}\frac{\partial\left(u_l v_m\varepsilon_{lmj}\right)}{\partial x_k}\\ &= (\delta_{il}\delta_{km}-\delta_{im}\delta_{lk})\left(\frac{\partial u_l}{\partial x_k}v_m + u_l\frac{\partial v_m}{\partial x_k}\right)\\ &= \underbrace{(\text{grad}\mathbf{u})_{ik}\,(\mathbf{v})_k + (\mathbf{u})_i\,(\text{div}\mathbf{v})}_{=\,(\text{div}\,(\mathbf{u}\otimes\mathbf{v}))_i\ ,\ \text{according to (7.122g)}}\\ &\quad \underbrace{-\,(\text{div}\mathbf{u})\,(\mathbf{v})_i - (\text{grad}\mathbf{v})_{ik}\,(\mathbf{u})_k}_{=\,-\,(\text{div}\,(\mathbf{v}\otimes\mathbf{u}))_i\ ,\ \text{according to (7.122g)}}\ .\end{aligned}$$

The expression (7.123a): Using (2.49), $(7.68\text{a})_2$, $(7.70)_2$, (7.78) and $(7.83)_2$,

$$\begin{aligned}(\text{grad}\,\text{div}\mathbf{u})_i &= \frac{\partial}{\partial x_i}\left(\frac{\partial u_j}{\partial x_j}\right)\\ &= \frac{\partial}{\partial x_j}\left(\frac{\partial u_j}{\partial x_i}\right)\\ &= \frac{\partial\,(\text{grad}\mathbf{u})_{ji}}{\partial x_j}\\ &= \frac{\partial\left(\text{grad}^{\text{T}}\mathbf{u}\right)_{ij}}{\partial x_j}\\ &= \left(\text{div}\,\text{grad}^{\text{T}}\mathbf{u}\right)_i\ .\end{aligned}$$

The expression (7.123b): Using (1.54), (7.78), (7.91) and in light of (2.79h),

$$\begin{aligned}
\operatorname{div}\operatorname{curl}\mathbf{u} &= \frac{\partial}{\partial x_i}\left(\operatorname{curl}\mathbf{u}\right)_i \\
&= \frac{\partial}{\partial x_i}\left(-\varepsilon_{ijk}\frac{\partial u_j}{\partial x_k}\right) \\
&= -\varepsilon_{ijk}\frac{\partial u_j}{\partial x_i \partial x_k} \\
&= \varepsilon_{jik}\frac{\partial u_j}{\partial x_i \partial x_k} \\
&= 0 \,.
\end{aligned}$$

The expression (7.123c): Using (1.54), (7.68a)$_2$, (7.91) and in light of (2.79h),

$$\begin{aligned}
\left(\operatorname{curl}\operatorname{grad}\phi\right)_i &= -\varepsilon_{ijk}\frac{\partial\left(\operatorname{grad}\phi\right)_j}{\partial x_k} \\
&= -\varepsilon_{ijk}\frac{\partial^2 \phi}{\partial x_k \partial x_j} \\
&= \varepsilon_{ikj}\frac{\partial^2 \phi}{\partial x_k \partial x_j} \\
&= 0 \,.
\end{aligned}$$

The expression (7.123d): Using (1.36), (1.54), (1.58a), (7.68a)$_2$, (7.78), (7.91) and (7.105)$_2$,

$$\begin{aligned}
\left(\operatorname{curl}\operatorname{curl}\mathbf{u}\right)_i &= -\varepsilon_{ijk}\frac{\partial\left(\operatorname{curl}\mathbf{u}\right)_j}{\partial x_k} \\
&= -\varepsilon_{ikj}\varepsilon_{lmj}\frac{\partial^2 u_l}{\partial x_k \partial x_m} \\
&= -\frac{\partial^2 u_i}{\partial x_k \partial x_k} + \frac{\partial^2 u_k}{\partial x_k \partial x_i} \\
&= -\left(\nabla^2 \mathbf{u}\right)_i + \frac{\partial}{\partial x_i}\frac{\partial u_k}{\partial x_k} \\
&= -\left(\nabla^2 \mathbf{u}\right)_i + \left(\operatorname{grad}\operatorname{div}\mathbf{u}\right)_i \,.
\end{aligned}$$

The expression (7.123e): Using (7.68a)$_2$ and (7.105)$_{1-2}$,

$$\begin{aligned}
\left(\nabla^2 \operatorname{grad}\phi\right)_i &= \frac{\partial^2}{\partial x_j \partial x_j}\left(\frac{\partial \phi}{\partial x_i}\right) \\
&= \frac{\partial}{\partial x_i}\left(\frac{\partial^2 \phi}{\partial x_j \partial x_j}\right)
\end{aligned}$$

$$= \frac{\partial}{\partial x_i} \left(\nabla^2 \phi\right)$$
$$= \left(\text{grad}\nabla^2 \phi\right)_i \ .$$

The expression (7.123f): Using (7.91) and (7.105)$_2$,

$$\begin{aligned}
\left(\nabla^2 \text{curl}\mathbf{u}\right)_i &= \frac{\partial^2}{\partial x_l \partial x_l} \left(-\varepsilon_{ijk} \frac{\partial u_j}{\partial x_k}\right) \\
&= -\varepsilon_{ijk} \frac{\partial}{\partial x_k} \left(\frac{\partial^2 u_j}{\partial x_l \partial x_l}\right) \\
&= -\varepsilon_{ijk} \frac{\partial \left(\nabla^2 \mathbf{u}\right)_j}{\partial x_k} \\
&= \left(\text{curl}\nabla^2 \mathbf{u}\right)_i \ .
\end{aligned}$$

The expression (7.123g): Using (1.38)$_7$, (2.75)$_4$, (7.70)$_2$ and (7.105)$_{1-2}$ along with the product rule of differentiation,

$$\begin{aligned}
\nabla^2 \left(\mathbf{u} \cdot \mathbf{v}\right) &= \frac{\partial}{\partial x_j} \left[\frac{\partial \left(u_i v_i\right)}{\partial x_j}\right] \\
&= \frac{\partial}{\partial x_j} \left[\frac{\partial u_i}{\partial x_j} v_i + u_i \frac{\partial v_i}{\partial x_j}\right] \\
&= \frac{\partial^2 u_i}{\partial x_j \partial x_j} v_i + \frac{\partial u_i}{\partial x_j} \frac{\partial v_i}{\partial x_j} + \frac{\partial u_i}{\partial x_j} \frac{\partial v_i}{\partial x_j} + u_i \frac{\partial^2 v_i}{\partial x_j \partial x_j} \\
&= \left(\nabla^2 \mathbf{u}\right) \cdot \mathbf{v} + 2 \left(\text{grad}\mathbf{u}\right) : \left(\text{grad}\mathbf{v}\right) + \mathbf{u} \cdot \left(\nabla^2 \mathbf{v}\right) \ .
\end{aligned}$$

The expression (7.124a): Using (1.54), (7.70)$_2$ and (7.100),

$$\begin{aligned}
\left(\text{curl grad}\mathbf{u}\right)_{ij} &= -\varepsilon_{ikl} \frac{\partial}{\partial x_l} \left[\left(\text{grad}\mathbf{u}\right)_{jk}\right] \\
&= -\varepsilon_{ikl} \frac{\partial}{\partial x_l} \left[\frac{\partial u_j}{\partial x_k}\right] \\
&= -\varepsilon_{ikl} \frac{\partial^2 u_j}{\partial x_l \partial x_k} \\
&= \varepsilon_{ilk} \frac{\partial^2 u_j}{\partial x_l \partial x_k} \\
&= \left(\mathbf{O}\right)_{ij} \ ,
\end{aligned}$$

taking into account $\varepsilon_{ilk} = -\varepsilon_{ikl}$ and $\partial^2 u_j / \partial x_l \partial x_k = \partial^2 u_j / \partial x_k \partial x_l$, see the expression (2.79h).

The expression (7.124b): With the aid of (1.36), (1.57)$_2$, (2.23), (2.49), (2.89a)$_2$, (7.68a)$_2$, (7.70)$_2$, (7.83)$_2$, (7.100), (7.105)$_1$ and (7.105)$_3$,

$$
\begin{aligned}
(\text{curl}\,\text{curl}\mathbf{A})_{ij} &= -\varepsilon_{ikl}\frac{\partial}{\partial x_l}\left[(\text{curl}\mathbf{A})_{jk}\right] \\
&= -\varepsilon_{ikl}\frac{\partial}{\partial x_l}\left[-\varepsilon_{jmn}\frac{\partial A_{km}}{\partial x_n}\right] \\
&= \underbrace{\varepsilon_{ikl}\varepsilon_{jmn}\frac{\partial^2 A_{km}}{\partial x_l \partial x_n}}_{=\left[\delta_{ij}\left(\delta_{km}\delta_{ln}-\delta_{kn}\delta_{lm}\right)-\delta_{im}\left(\delta_{kj}\delta_{ln}-\delta_{kn}\delta_{lj}\right)+\delta_{in}\left(\delta_{kj}\delta_{lm}-\delta_{km}\delta_{lj}\right)\right]\frac{\partial^2 A_{km}}{\partial x_l \partial x_n}} \\
&= \delta_{ij}\frac{\partial^2 A_{kk}}{\partial x_l \partial x_l} - \delta_{ij}\frac{\partial^2 A_{kl}}{\partial x_l \partial x_k} - \frac{\partial^2 A_{ji}}{\partial x_l \partial x_l} \\
&\quad + \frac{\partial^2 A_{ki}}{\partial x_j \partial x_k} + \frac{\partial^2 A_{jl}}{\partial x_l \partial x_i} - \frac{\partial^2 A_{kk}}{\partial x_j \partial x_i} \\
&= \underbrace{\delta_{ij}}_{=(\mathbf{I})_{ij}}\underbrace{\frac{\partial^2 A_{kk}}{\partial x_l \partial x_l}}_{=\nabla^2(\text{tr}\mathbf{A})} - \underbrace{\delta_{ij}}_{=(\mathbf{I})_{ij}}\frac{\partial}{\partial x_k}\underbrace{\frac{\partial A_{kl}}{\partial x_l}}_{=(\text{div}\mathbf{A})_k} - \underbrace{\frac{\partial^2 A^{\text{T}}_{ij}}{\partial x_l \partial x_l}}_{=\nabla^2\left(\mathbf{A}^{\text{T}}\right)_{ij}} \\
&\quad + \frac{\partial}{\partial x_j}\underbrace{\frac{\partial A^{\text{T}}_{ik}}{\partial x_k}}_{=\left(\text{div}\mathbf{A}^{\text{T}}\right)_i} + \frac{\partial}{\partial x_i}\underbrace{\frac{\partial A_{jl}}{\partial x_l}}_{=(\text{div}\mathbf{A})_j} - \frac{\partial}{\partial x_j}\underbrace{\frac{\partial A_{kk}}{\partial x_i}}_{=(\text{grad}\,(\text{tr}\mathbf{A}))_i} \\
&= (\mathbf{I})_{ij}\,\nabla^2(\text{tr}\mathbf{A}) - (\mathbf{I})_{ij}\,\text{div}\,\text{div}\mathbf{A} - \nabla^2\left(\mathbf{A}^{\text{T}}\right)_{ij} \\
&\quad + \left(\text{grad}\,\text{div}\mathbf{A}^{\text{T}}\right)_{ij} + \left(\text{grad}^{\text{T}}\,\text{div}\mathbf{A}\right)_{ij} - (\text{grad}\,\text{grad}\,(\text{tr}\mathbf{A}))_{ij} \ .
\end{aligned}
$$

Exercise 7.9

Consider the so-called *inverse square law*

$$\mathbf{u} = \alpha\frac{\mathbf{r}}{r^3} \quad \text{with} \quad \alpha > 0 \ . \tag{7.125}$$

In this expression, $\mathbf{r} = \mathbf{x} - \mathbf{x}_0$ where $\mathbf{x}_0$ denotes a fixed point in space and $r = |\mathbf{r}|$. The domain of $\mathbf{u}$ is the three-dimensional space, excluding $r = 0$. Note that the inverse square law appears in many branches of physics and engineering. Examples of which include Newton's law for gravitation, Coulomb's law for static electricity and transport problem of radiation mechanics.

First, prove the identities

$$\text{grad}\mathbf{r} = \mathbf{I}, \qquad \text{grad}r = \frac{\mathbf{r}}{r}\ , \tag{7.126a}$$

$$\text{div}\mathbf{r} = 3, \qquad \text{div}\left(\frac{\mathbf{r}\otimes\mathbf{r}}{r^2}\right) = \frac{2\mathbf{r}}{r^2}\ , \tag{7.126b}$$

$$\text{curl}\mathbf{r} = \mathbf{0}, \qquad \text{curl}\left(\frac{\mathbf{r}}{r}\right) = \mathbf{0}\ . \tag{7.126c}$$

Then, having in mind these identities, show that

$$\text{div}\mathbf{u} = 0\ , \tag{7.127a}$$

$$\nabla^2\mathbf{u} = \mathbf{0}\ , \tag{7.127b}$$

$$\text{curl}\,\text{curl}\mathbf{u} = \mathbf{0}\ . \tag{7.127c}$$

Solution. Similarly to the previous exercise, all desired relations will be verified in indicial notation using the Cartesian components of vectors.

The expression (7.126a)$_1$: With the aid of (2.23) and (7.70)$_2$,

$$\begin{aligned}(\text{grad}\mathbf{r})_{ij} &= \frac{\partial r_i}{\partial x_j} \\ &= \frac{\partial x_i}{\partial x_j} \\ &= \delta_{ij} \\ &= (\mathbf{I})_{ij}\ .\end{aligned}$$

The expression (7.126a)$_2$: With the aid of (1.36), (6.18)$_1$, (7.68a)$_2$ and (7.126a)$_1$ along with the chain rule of differentiation,

$$\begin{aligned}(\text{grad}r)_i &= \frac{\partial r}{\partial x_i} \\ &= \frac{\partial r}{\partial r_j}\frac{\partial r_j}{\partial x_i} \\ &= \frac{\partial r}{\partial r_j}\delta_{ji} \\ &= \frac{\partial r}{\partial r_i} \\ &= \frac{r_i}{r} \\ &= \left(\frac{\mathbf{r}}{r}\right)_i\ .\end{aligned}$$

The expression (7.126b)$_1$: With the aid of (1.37), (7.78) and in light of (7.126a)$_1$,

$$\begin{aligned}\text{div}\mathbf{r} &= \frac{\partial r_i}{\partial x_i} \\ &= \delta_{ii} \\ &= 3 \ .\end{aligned}$$

The expression (7.126b)$_2$: With the aid of (1.36), (7.83)$_2$, (7.126a)$_{1-2}$ and (7.126b)$_1$ along with the product rule of differentiation,

$$\begin{aligned}\left(\text{div}\left(r^{-2}\mathbf{r} \otimes \mathbf{r}\right)\right)_i &= \frac{\partial \left(r^{-2} r_i r_j\right)}{\partial x_j} \\ &= \frac{\partial r^{-2}}{\partial x_j} r_i r_j + r^{-2} \frac{\partial r_i}{\partial x_j} r_j + r^{-2} r_i \frac{\partial r_j}{\partial x_j} \\ &= -\frac{2}{r^3} \frac{r_j}{r} r_i r_j + r^{-2} \delta_{ij} r_j + 3r^{-2} r_i \\ &= \frac{2r_i}{r^2} \ .\end{aligned}$$

The expression (7.126c)$_1$: With the aid of (1.36), (1.52), (7.91) and (7.126a)$_1$,

$$\begin{aligned}(\text{curl}\mathbf{r})_i &= -\varepsilon_{ijk} \frac{\partial r_j}{\partial x_k} \\ &= -\varepsilon_{ijk} \delta_{jk} \\ &= -\varepsilon_{ijj} \\ &= 0 \ .\end{aligned}$$

The expression (7.126c)$_2$: With the aid of (1.36), (1.52), (7.91) and (7.126a)$_{1-2}$ along with the product rule of differentiation, one will have

$$\begin{aligned}\left(\text{curl}\frac{\mathbf{r}}{r}\right)_i &= -\varepsilon_{ijk} \frac{\partial \left(r^{-1} r_j\right)}{\partial x_k} \\ &= -\varepsilon_{ijk} \frac{\partial r^{-1}}{\partial x_k} r_j - \varepsilon_{ijk} r^{-1} \frac{\partial r_j}{\partial x_k} \\ &= -\varepsilon_{ijk} \frac{-1}{r^2} \frac{r_k}{r} r_j - \varepsilon_{ijk} r^{-1} \delta_{jk} \\ &= \frac{1}{r^3} \varepsilon_{ijk} r_j r_k - \frac{1}{r} \varepsilon_{ijj} \\ &= 0 \ ,\end{aligned}$$

taking into consideration $\varepsilon_{ijk} = -\varepsilon_{ikj}$ and $(\mathbf{r} \otimes \mathbf{r})_{jk} = (\mathbf{r} \otimes \mathbf{r})_{kj}$, see (2.79h).

The expression (7.127a): With the aid of (7.78), $(7.126a)_2$ and $(7.126b)_1$ along with the product rule of differentiation,

$$
\begin{aligned}
\operatorname{div}\mathbf{u} &= \frac{\partial u_i}{\partial x_i} \\
&= \alpha \frac{\partial \left(r^{-3} r_i\right)}{\partial x_i} \\
&= \alpha \frac{\partial r^{-3}}{\partial x_i} r_i + \alpha r^{-3} \frac{\partial r_i}{\partial x_i} \\
&= \alpha \frac{-3}{r^4} \frac{r_i}{r} r_i + \alpha r^{-3} \,(3) \\
&= \frac{-3\alpha r^2}{r^5} + 3\alpha r^{-3} \\
&= 0 \; .
\end{aligned}
$$

The expression (7.127b): With the aid of $(7.105)_2$, $(7.126a)_{1-2}$ and (7.127a) along with the product rule of differentiation,

$$
\begin{aligned}
\left(\nabla^2 \mathbf{u}\right)_j &= \frac{\partial^2 u_j}{\partial x_i \partial x_i} \\
&= \frac{\partial}{\partial x_i} \left[\alpha \frac{\partial \left(r^{-3} r_j\right)}{\partial x_i} \right] \\
&= \alpha \frac{\partial}{\partial x_i} \left[\frac{-3}{r^4} \frac{r_i}{r} r_j + r^{-3} \delta_{ji} \right] \\
&= \alpha \frac{\partial}{\partial x_i} \left[\frac{-3 r_i}{r^3} \frac{r_j}{r^2} + \frac{\delta_{ji}}{r^3} \right] \\
&= -3\alpha \left[\frac{r_i}{r^3} \frac{\partial}{\partial x_i} \left(\frac{r_j}{r^2} \right) + \frac{r_i}{r^5} \delta_{ji} \right] \\
&= -3\alpha \left[\frac{r_i}{r^3} \left(\frac{-2 r_i}{r^4} r_j + \frac{\delta_{ji}}{r^2} \right) + \frac{r_j}{r^5} \right] \\
&= -3\alpha \left[\frac{-2 r_j}{r^5} + \frac{r_j}{r^5} + \frac{r_j}{r^5} \right] \\
&= 0 \; .
\end{aligned}
$$

The expression (7.127c): At the end, this desired identity is obtained by substituting (7.127a) and (7.127b) into (7.123d).

References

1. Grinfeld, P. (2013). *Introduction to tensor analysis and the calculus of moving surfaces*. New York, NY: Springer.
2. Danielson, D. A. (1992). *Vectors and tensors in engineering and physics*. Addison-Wesley.
3. Das, A. (2007). *Tensors: The mathematics of relativity theory and continuum mechanics*. New York, NY: Springer.
4. Sochi, T. (2017). *Introduction to differential geometry of space curves and surfaces*. CreateSpace Independent Publishing Platform.
5. Chaves, E. W. V. (2013). *Notes on continuum mechanics*. Dordrecht: Springer.
6. Lai, W. M., Rubin, D., & Krempl, E. (2010). *Introduction to continuum mechanics* (4th ed.). Amsterdam: Elsevier Science.
7. Hildebrand, F. B. (1976). *Advanced calculus for applications* (2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall.

Chapter 8
Integral Theorems and Differential Forms

This chapter contains two sections. The first section deals with well-known theorems involving integrals of tensorial field variables to complete vector and tensor calculus started from Chap. 6. Specifically, the integral theorems of **Gauss** and **Stokes** are studied. They are of central importance in mathematics of physics because they eventually appear in many basic laws of physics such as conservation of linear momentum and conservation of electric charge. The beauty of these theorems is that they transform integration of field variables over closed domains into the integration over the boundaries of such domains.

The remaining part of this chapter aims at introducing what are known as *differential forms*. They are used in many physical theorems such as electromagnetic field theory. After some algebraic preliminaries, the calculus of these mathematical creatures will be studied. They are basically considered as a **complement** to vector analysis. The ultimate goal here is to introduce the so-called *generalized Stokes' theorem*. This elegant theorem unifies the four fundamental theorems of calculus (i.e. gradient theorem for line integrals, Green's theorem, Stokes' theorem and divergence theorem).

8.1 Integral Theorems

8.1.1 Divergence Theorem

Consider a closed surface[1] A enclosing a region V with the outward-pointing unit normal field $\widehat{\mathbf{n}}$ as illustrated in Fig. 8.1. The **divergence theorem** (or **Gauss's theorem**) states that the net flux of a vector field $\mathbf{u}$ out of a closed surface A oriented with

[1] A *closed surface* is a compact surface without boundary curve. But, it usually has an inner and outer side. Such a surface to be used in the divergence theorem must be at least piecewise smooth.

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_8

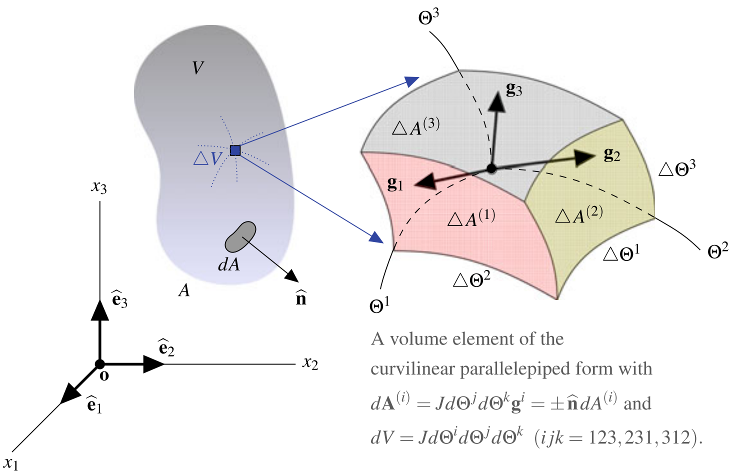


Fig. 8.1 Domain V (with the volume element $\triangle V$) and its bounded closed surface A (with the differential surface element dA and associated unit normal field $\widehat{\mathbf{n}}$)

$\widehat{\mathbf{n}}$ outward equals the volume integral of its divergence over the region V enclosed by that surface. This is indicated by

$$\int_A \mathbf{u} \cdot \widehat{\mathbf{n}}\, dA = \int_V \operatorname{div} \mathbf{u}\, dV \quad . \tag{8.1}$$

or, for instance in Cartesian coordinates, $\int_A u_i \widehat{n}_i \, dA = \int_V \frac{\partial u_i}{\partial x_i} \, dV$

<u>Proof</u>. Let the domain V be split into a finite number N of curvilinear hexahedra, i.e. $V \approx \bigcup_{e=1}^N V_e$. Further, let a generic hexahedron is denoted by $\triangle V$ and enclosed by $\triangle A$. Such an object is constructed by the coordinate surfaces Θ^1, $\Theta^1 + \triangle\Theta^1$, Θ^2, $\Theta^2 + \triangle\Theta^2$, Θ^3 and $\Theta^3 + \triangle\Theta^3$ as shown in Fig. 8.1. Recall from *Hint* on Sect. 5.8 that a surface vector $d\mathbf{A}$ admits two unit vectors oriented in opposite directions, i.e. $\pm\widehat{\mathbf{n}}$. With regard to this, a convention needs to be made for determining either $d\mathbf{A} = +\widehat{\mathbf{n}}\, dA$ or $d\mathbf{A} = -\widehat{\mathbf{n}}\, dA$. Regarding the surface vector of a closed surface, a normal vector pointing outward from the enclosed domain is conventionally granted. As an example, one will have $d\mathbf{A}^{(1)} = +\widehat{\mathbf{n}}^{(1)}\, dA^{(1)}$ at $\left(\Theta^1 + \triangle\Theta^1, \Theta^2, \Theta^3\right)$ and $d\mathbf{A}^{(1)} = -\widehat{\mathbf{n}}^{(1)}\, dA^{(1)}$ at $\left(\Theta^1, \Theta^2, \Theta^3\right)$ for $\triangle V$ in Fig. 8.1.

Now, guided by $(5.111)_2$ and $(5.113)_2$, the infinitesimal area elements $\widehat{\mathbf{n}}^{(i)}\, dA^{(i)}$ and the differential volume element dV of the very small curvilinear parallelepiped $\triangle V$, for even permutations of i, j and k (i.e. $ijk = 123, 231, 312$), are given by

$$\widehat{\mathbf{n}}^{(i)}\,dA^{(i)} = \begin{cases} -Jd\Theta^j d\Theta^k \mathbf{g}^i & \text{at } \left(\Theta^1, \Theta^2, \Theta^3\right) \\ +Jd\Theta^j d\Theta^k \mathbf{g}^i & \text{at } \left(\Theta^1 + \triangle\Theta^1\delta_{i1}, \Theta^2 + \triangle\Theta^2\delta_{i2}, \Theta^3 + \triangle\Theta^3\delta_{i3}\right) \end{cases}, \tag{8.2a}$$

$$dV = Jd\Theta^i d\Theta^j d\Theta^k \,. \tag{8.2b}$$

For the curvilinear parallelepiped, the **surface integral**, also known as the **flux integral**, on the left hand side of (8.1) then renders ($ijk = 123, 231, 312$)

$$\int_{\triangle A} \mathbf{u}\cdot\widehat{\mathbf{n}}\,dA = \sum_{i=1}^{3} \int_{\Theta^k}^{\Theta^k+\triangle\Theta^k} \int_{\Theta^j}^{\Theta^j+\triangle\Theta^j} \left[\left(\mathbf{u}\cdot J\mathbf{g}^i\right)\Big|_{\left(\Theta^1+\triangle\Theta^1\delta_{i1},\Theta^2+\triangle\Theta^2\delta_{i2},\Theta^3+\triangle\Theta^3\delta_{i3}\right)} - \left(\mathbf{u}\cdot J\mathbf{g}^i\right)\Big|_{\left(\Theta^1,\Theta^2,\Theta^3\right)} \right] d\Theta^j d\Theta^k \,.$$

It follows that

$$\int_{\triangle A} \mathbf{u}\cdot\widehat{\mathbf{n}}\,dA = \sum_{i=1}^{3} \int_{\Theta^k}^{\Theta^k+\triangle\Theta^k} \int_{\Theta^j}^{\Theta^j+\triangle\Theta^j} \int_{\Theta^i}^{\Theta^i+\triangle\Theta^i} \frac{\partial\left(\mathbf{u}\cdot J\mathbf{g}^i\right)}{\partial\Theta^i} d\Theta^i d\Theta^j d\Theta^k \,. \tag{8.3}$$

Using (7.18), (7.80) and (8.2b), the flux integral (8.3) represents

$$\begin{aligned} \int_{\triangle A} \mathbf{u}\cdot\widehat{\mathbf{n}}\,dA &= \sum_{i=1}^{3} \int_{\Theta^k}^{\Theta^k+\triangle\Theta^k} \int_{\Theta^j}^{\Theta^j+\triangle\Theta^j} \int_{\Theta^i}^{\Theta^i+\triangle\Theta^i} \left[\frac{\partial\mathbf{u}}{\partial\Theta^i}\cdot\left(J\mathbf{g}^i\right) + \mathbf{u}\cdot\frac{\partial\left(J\mathbf{g}^i\right)}{\partial\Theta^i} \right] d\Theta^i d\Theta^j d\Theta^k \\ &= \sum_{i=1}^{3} \int_{\triangle V} \frac{\partial\mathbf{u}}{\partial\Theta^i}\cdot\mathbf{g}^i\,dV \\ &= \int_{\triangle V} \mathrm{div}\mathbf{u}\,dV \,. \end{aligned} \tag{8.4}$$

At the end, the integral expression $(8.1)_1$ follows by summing over the entire region (noting that the sum of integrals over an inner surface shared between two volume elements vanishes).

The Gauss's theorem helps provide a new definition for the divergence of a tensorial field. Considering smallness of the curvilinear hexahedron, in the limit, the volume integral in (8.1) can be taken as the product of the integrand div**u** and the volume V. By assuming continuity of the integrand, one then has

$$\boxed{\mathrm{div}\mathbf{u} = \lim_{V\to 0}\frac{1}{V}\int_V \mathbf{g}^i\cdot\frac{\partial\mathbf{u}}{\partial\Theta^i}\,dV = \lim_{V\to 0}\frac{1}{V}\int_A \mathbf{u}\cdot\widehat{\mathbf{n}}\,dA \,.} \tag{8.5}$$

When the volume shrinks to zero in the limit, all points on the surface approach a point. Thus, the result will be independent of the actual shape of the volume element. Indeed, the definition (8.5) is independent of any coordinate system in alignment with (7.77). That is why in standard texts on calculus, the divergence theorem is often proved in a Cartesian frame by taking an incremental volume element of the rectangular parallelepiped form.

Notice that flux measures outward flow from (the closed surface of) an entire region whereas divergence measures the net outward flow per unit volume at the point under consideration. And if there is net flow out of (into) the closed surface, the flux integral is positive (negative).

As an example, consider a vector $\bar{\mathbf{r}} = \mathbf{x} - \mathbf{o}$ satisfying grad $\bar{\mathbf{r}} = \mathbf{I}$. Then, the divergence theorem helps establish

$$\int_A \bar{\mathbf{r}} \otimes \widehat{\mathbf{n}}\, dA = \int_V \operatorname{grad} \bar{\mathbf{r}}\, dV = \int_V \mathbf{I}\, dV \ , \qquad \longleftarrow \text{see (8.142f)}$$

or

$$\boxed{\mathbf{I} = \frac{1}{V} \int_A \bar{\mathbf{r}} \otimes \widehat{\mathbf{n}}\, dA \ .} \tag{8.6}$$

The relation (8.1) is a classical result. It may be utilized to transform a surface integral into a volume integral for a scalar and tensor field, that is,

$$\int_A \phi \widehat{\mathbf{n}}\, dA = \int_V \operatorname{grad}\phi\, dV \ , \tag{8.7}$$

or, for instance in Cartesian coordinates, $\int_A \phi \widehat{n}_i\, dA = \int_V \frac{\partial \phi}{\partial x_i}\, dV$

and

$$\int_A \mathbf{A}\widehat{\mathbf{n}}\, dA = \int_V \operatorname{div}\mathbf{A}\, dV \ . \tag{8.8}$$

or, for instance in Cartesian coordinates, $\int_A A_{ij} \widehat{n}_j\, dA = \int_V \frac{\partial A_{ij}}{\partial x_j}\, dV$

To prove (8.8), consider an arbitrary constant vector $\mathbf{v}$ (for which grad$\mathbf{v} = \mathbf{O}$). Then,

$$\begin{aligned} \mathbf{v} \cdot \int_A \mathbf{A}\widehat{\mathbf{n}}\, dA &\overset{\text{by assumption}}{=\!=} \int_A \mathbf{v} \cdot \mathbf{A}\widehat{\mathbf{n}}\, dA \\ &\overset{\text{from (2.51b)}}{=\!=} \int_A \left[\mathbf{A}^{\mathrm{T}}\mathbf{v}\right] \cdot \widehat{\mathbf{n}}\, dA \end{aligned}$$

$$\xlongequal[\text{(8.1)}]{\text{from}} \int\limits_V \operatorname{div}\left[\mathbf{A}^{\mathrm{T}}\mathbf{v}\right] dV$$

$$\xlongequal[\text{(1.9a) and (7.122f)}]{\text{from}} \int\limits_V \left[\mathbf{v}\cdot\operatorname{div}\mathbf{A} + \mathbf{A} : \mathbf{O}\right] dV$$

$$\xlongequal[\text{(2.78)}]{\text{from}} \int\limits_V \mathbf{v}\cdot\operatorname{div}\mathbf{A}\, dV$$

$$\xlongequal[\text{assumption}]{\text{by}} \mathbf{v}\cdot\int\limits_V \operatorname{div}\mathbf{A}\, dV\ ,$$

provides the desired result considering (1.14). From (8.8), one can establish (8.7) as follows:

$$\int\limits_A \mathbf{A}\widehat{\mathbf{n}}\, dA \xlongequal[\mathbf{A}=\phi\mathbf{I}]{\text{by setting}} \int\limits_A [\phi\mathbf{I}]\,\widehat{\mathbf{n}}\, dA$$

$$\xlongequal[\text{from (2.5) and (2.8b)}]{\text{on the one hand}} \int\limits_A \phi\widehat{\mathbf{n}}\, dA$$

$$\xlongequal[\text{from (8.8)}]{\text{on the other hand}} \int\limits_V \operatorname{div}[\phi\mathbf{I}]\ dV$$

$$\xlongequal[\text{(7.122e)}]{\text{from}} \int\limits_V \left[\phi\operatorname{div}\mathbf{I} + \mathbf{I}\operatorname{grad}\phi\right] dV$$

$$\xlongequal[\text{(1.4a), (1.4d), (a) in (1.76) and (2.5)}]{\text{from}} \int\limits_V \operatorname{grad}\phi\, dV\ ,$$

where the identity $\operatorname{div}\mathbf{I} = \mathbf{0}$ has been used, see (7.33). Following arguments similar to those which led to (8.5) now reveals

$$\boxed{\operatorname{grad}\phi = \lim_{V\to 0}\frac{1}{V}\int\limits_A \phi\widehat{\mathbf{n}}\, dA\ ,} \tag{8.9}$$

and

$$\boxed{\operatorname{div}\mathbf{A} = \lim_{V\to 0}\frac{1}{V}\int\limits_V \frac{\partial\mathbf{A}}{\partial\Theta^i}\mathbf{g}^i\, dV = \lim_{V\to 0}\frac{1}{V}\int\limits_A \mathbf{A}\widehat{\mathbf{n}}\, dA\ .} \tag{8.10}$$

By analogy with the procedure which led to (8.4) and considering (8.5), the curl of a vector field can also be written as

$$\boxed{\operatorname{curl}\mathbf{u} = \lim_{V \to 0} \frac{1}{V} \int_V \mathbf{g}^i \times \frac{\partial \mathbf{u}}{\partial \Theta^i} \, dV = \lim_{V \to 0} \frac{1}{V} \int_A \widehat{\mathbf{n}} \times \mathbf{u} \, dA \; .} \tag{8.11}$$

It is then a simple exercise to show that

$$\boxed{\operatorname{curl}\mathbf{A} = \lim_{V \to 0} \frac{1}{V} \int_V \mathbf{g}^i \times \frac{\partial \mathbf{A}^{\mathrm{T}}}{\partial \Theta^i} \, dV = \lim_{V \to 0} \frac{1}{V} \int_A \widehat{\mathbf{n}} \times \mathbf{A}^{\mathrm{T}} \, dA \; .} \tag{8.12}$$

In many branches of physics and engineering such as continuum mechanics, physical laws are often expressed in terms of integrals over the domain of a continuum body or its boundary.

As discussed, the divergence theorem transforms a surface integral into a volume integral. Another important integral theorem that transforms a line integral to a surface integral will be introduced in the following.

8.1.2 Stokes' Theorem

Consider a closed curve[2] $\mathcal{C}$ bounding an open surface A with positive orientation[3] as illustrated in Fig. 8.2. Recall that a convention helped determine either $d\mathbf{A} = +\widehat{\mathbf{n}} \, dA$ or $d\mathbf{A} = -\widehat{\mathbf{n}} \, dA$ for a closed surface. The appropriate choice for an open surface depends on assigning a direction of rotation to its boundary curve. The direction of unit vector perpendicular to an open surface is conventionally associated with the rotational direction of its bounding curve by the right-hand screw rule rendering a *positively oriented surface*. Therefore, one will have $d\mathbf{A} = +\widehat{\mathbf{n}} \, dA$ for the positively oriented surface shown in Fig. 8.2. Suppose that such a surface is parametrized by

$$\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}} \left(t^1, t^2 \right) \; , \tag{8.13}$$

for which the line element $d\mathbf{x}$ and the surface vector $d\mathbf{A}$ are given by

$$d\mathbf{x} = dt^1 \mathbf{a}_1 + dt^2 \mathbf{a}_2 , \, d\mathbf{A} = +\widehat{\mathbf{n}} \, dA = dt^1 dt^2 \mathbf{a}_1 \times \mathbf{a}_2 \; , \tag{8.14}$$

← see (9.56)-(9.57)

[2] A curve $\mathcal{C}$ with the parametrization

$$\mathbf{x} = \mathbf{x}^{\mathrm{c}} (t) \, , \quad \bar{t}_0 \leq t \leq \bar{t}_1 \; ,$$

is said to be *closed* when $\mathbf{x}^{\mathrm{c}} \left(\bar{t}_0 \right) = \mathbf{x}^{\mathrm{c}} \left(\bar{t}_1 \right)$.

[3] An *open surface* is a surface possessing a boundary curve. If the unit vector $\widehat{\mathbf{n}}$ normal to an open surface A and the tangent vector $\mathbf{a}_t$ to its bounding closed curve $\mathcal{C}$ are oriented in the right-handed sense, then A is called a *positively oriented surface*. Such a surface along with its bounding curve to be used in the Stokes' theorem must be at least piecewise smooth.

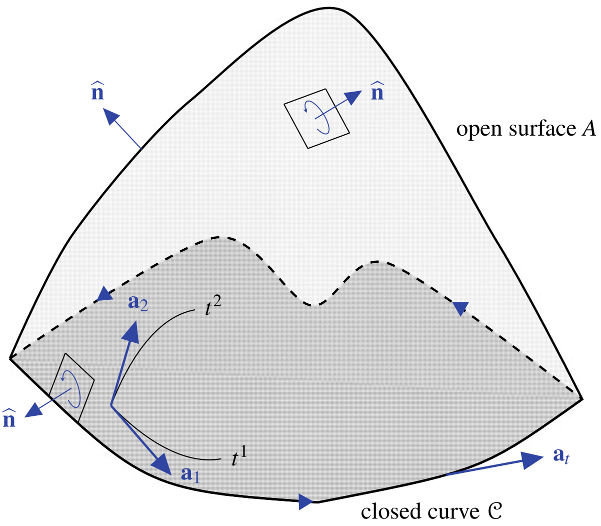


Fig. 8.2 Open surface with positive orientation

where

$$\mathbf{a}_1 = \frac{\partial \hat{\mathbf{x}}^{\,\mathrm{s}}}{\partial t^1} \quad , \quad \mathbf{a}_2 = \frac{\partial \hat{\mathbf{x}}^{\,\mathrm{s}}}{\partial t^2} \ . \tag{8.15}$$

Given the parametric equations $t^1 = t^1(t)$ and $t^2 = t^2(t)$, the bounding curve $\mathcal{C}$ can be described by the parametrization

$$\mathbf{x} = \mathbf{x}^{\mathrm{c}}(t) \quad \text{with the tangent vector} \quad \mathbf{a}_t = \frac{d\mathbf{x}}{dt} \ . \tag{8.16}$$

For a more detailed study of parametrization of surfaces and curves, see Chap. 9.

The **circulation** of a vector field $\mathbf{u}$ around an oriented closed curve $\mathcal{C}$ is defined by

$$\boxed{\varGamma = \int_{\mathcal{C}} \mathbf{u} \cdot d\mathbf{x} \ .} \tag{8.17}$$

This relation basically demonstrates the **line integral** of the tangential component of $\mathbf{u}$ along $\mathcal{C}$. And it measures how much this vector field tends to circulate around such curve.

Now, the *Stokes' theorem* states that the circulation of a vector field $\mathbf{u} = \mathbf{u}\left(t^1, t^2\right)$ around a closed curve $\mathcal{C}$ is equal to the normal component of its curl over an open surface A bounded by that curve. For a positively oriented surface shown in Fig. 8.2, this is indicated by

$$\boxed{\int_{\mathcal{C}} \mathbf{u} \cdot d\mathbf{x} = \int_{A} \text{curl}\mathbf{u} \cdot \widehat{\mathbf{n}}\, dA} \quad . \tag{8.18}$$

or, for instance in Cartesian coordinates, $\int_{\mathcal{C}} u_i dx_i = \int_A -\varepsilon_{ijk} \frac{\partial u_j}{\partial x_k} \widehat{n}_i \, dA$

Proof. First, the left hand side of (8.18)$_1$ with the aid of (8.14)$_1$ takes the following form

$$\int_{\mathcal{C}} \mathbf{u} \cdot d\mathbf{x} = \int_{\mathcal{C}} \left(\underline{u}_1 dt^1 + \underline{u}_2 dt^2 \right) \quad \text{where} \quad \underline{u}_1 = \mathbf{u} \cdot \mathbf{a}_1 \ , \ \ \underline{u}_2 = \mathbf{u} \cdot \mathbf{a}_2 \ . \tag{8.19}$$

Then, the right hand side of (8.18)$_1$ considering (8.14)$_2$ becomes

$$\int_A \text{curl}\mathbf{u} \cdot \widehat{\mathbf{n}}\, dA = \int_A \text{curl}\mathbf{u} \cdot (\mathbf{a}_1 \times \mathbf{a}_2)\, dt^1 dt^2 = \int_A \left(\frac{\partial\, \underline{u}_2}{\partial t^1} - \frac{\partial\, \underline{u}_1}{\partial t^2} \right) dt^1 dt^2 \ , \tag{8.20}$$

since

$$\begin{aligned}
\text{curl}\mathbf{u} \cdot (\mathbf{a}_1 \times \mathbf{a}_2) &\overset{\text{in light of (7.95)}}{=\!=} - \left(\frac{\partial \mathbf{u}}{\partial t^1} \times \mathbf{a}^1 + \frac{\partial \mathbf{u}}{\partial t^2} \times \mathbf{a}^2 \right) \cdot (\mathbf{a}_1 \times \mathbf{a}_2) \\
&\overset{\text{from (1.9b) and (1.78a)}}{=\!=} - \left(\frac{\partial \mathbf{u}}{\partial t^1} \cdot \mathbf{a}_1 \right) \left(\mathbf{a}^1 \cdot \mathbf{a}_2 \right) + \left(\frac{\partial \mathbf{u}}{\partial t^1} \cdot \mathbf{a}_2 \right) \left(\mathbf{a}^1 \cdot \mathbf{a}_1 \right) \\
&\quad - \left(\frac{\partial \mathbf{u}}{\partial t^2} \cdot \mathbf{a}_1 \right) \left(\mathbf{a}^2 \cdot \mathbf{a}_2 \right) + \left(\frac{\partial \mathbf{u}}{\partial t^2} \cdot \mathbf{a}_2 \right) \left(\mathbf{a}^2 \cdot \mathbf{a}_1 \right) \\
&\overset{\text{in light of (5.13) and (5.27)}}{=\!=} \frac{\partial \mathbf{u}}{\partial t^1} \cdot \mathbf{a}_2 - \frac{\partial \mathbf{u}}{\partial t^2} \cdot \mathbf{a}_1 \\
&\overset{\text{from (8.15)}}{=\!=} \frac{\partial \left(\mathbf{u} \cdot \mathbf{a}_2 \right)}{\partial t^1} - \frac{\partial \left(\mathbf{u} \cdot \mathbf{a}_1 \right)}{\partial t^2} \quad \leftarrow \text{note that } \frac{\partial \mathbf{a}_2}{\partial t^1} = \frac{\partial^2 \hat{\mathbf{x}}^s}{\partial t^1 \partial t^2} = \frac{\partial^2 \hat{\mathbf{x}}^s}{\partial t^2 \partial t^1} = \frac{\partial \mathbf{a}_1}{\partial t^2} \\
&\overset{\text{in light of (5.65b)}}{=\!=} \frac{\partial\, \underline{u}_2}{\partial t^1} - \frac{\partial\, \underline{u}_1}{\partial t^2} \ .
\end{aligned} \tag{8.21}$$

By means of **Green's theorem**, one can write

$$\boxed{\int_{\mathcal{C}} \left(\underline{u}_1 dt^1 + \underline{u}_2 dt^2 \right) = \int_A \left(\frac{\partial\, \underline{u}_2}{\partial t^1} - \frac{\partial\, \underline{u}_1}{\partial t^2} \right) dt^1 dt^2 \ .} \tag{8.22}$$

And this completes the proof of Stokes' theorem.

The relation (8.18) presents a classical result. It simply helps transform a line integral into a surface integral for a scalar and tensor field, that is,

$$\boxed{\underbrace{\int_{\mathcal{C}} \phi \, d\mathbf{x} = \int_{A} \widehat{\mathbf{n}} \times \text{grad}\phi \, dA}_{\text{or, for instance in Cartesian coordinates, } \int_{\mathcal{C}} \phi \, dx_i = \int_A \varepsilon_{ijk} \widehat{n}_j \frac{\partial \phi}{\partial x_k} dA} \quad ,} \tag{8.23}$$

and

$$\boxed{\underbrace{\int_{\mathcal{C}} \mathbf{A} \, d\mathbf{x} = \int_{A} (\text{curl}\mathbf{A})^{\text{T}} \, \widehat{\mathbf{n}} \, dA}_{\text{or, for instance in Cartesian coordinates, } \int_{\mathcal{C}} A_{ij} \, dx_j = \int_A \varepsilon_{jlk} \frac{\partial A_{ik}}{\partial x_l} \widehat{n}_j \, dA} \quad .} \tag{8.24}$$

To prove (8.24), suppose that $\mathbf{v}$ is an arbitrary constant vector. Then,

$$\begin{aligned}
\mathbf{v} \cdot \int_{A} (\text{curl}\mathbf{A})^{\text{T}} \, \widehat{\mathbf{n}} \, dA &\xlongequal[\text{assumption}]{\text{by}} \int_{A} \mathbf{v} \cdot (\text{curl}\mathbf{A})^{\text{T}} \, \widehat{\mathbf{n}} \, dA \\
&\xlongequal[\text{(2.51b) and (2.55b)}]{\text{from}} \int_{A} \widehat{\mathbf{n}} \cdot [(\text{curl}\mathbf{A}) \, \mathbf{v}] \, dA \\
&\xlongequal[\text{(7.101)}]{\text{from}} \int_{A} \left[\text{curl}\left(\mathbf{A}^{\text{T}}\mathbf{v}\right)\right] \cdot \widehat{\mathbf{n}} \, dA \\
&\xlongequal[\text{(8.18)}]{\text{from}} \int_{\mathcal{C}} \left(\mathbf{A}^{\text{T}}\mathbf{v}\right) \cdot d\mathbf{x} \\
&\xlongequal[\text{(2.51b)}]{\text{from}} \int_{\mathcal{C}} \mathbf{v} \cdot \mathbf{A} d\mathbf{x} \\
&\xlongequal[\text{assumption}]{\text{by}} \mathbf{v} \cdot \int_{\mathcal{C}} \mathbf{A} d\mathbf{x} \; ,
\end{aligned}$$

delivers the desired result taking into account (1.14). The relation (8.23) can be deduced from (8.24). This can be shown, for instance, in Cartesian coordinates as follows:

$$\begin{aligned}
\int_{\mathcal{C}} A_{ij} \, dx_j &\xlongequal[A_{ij} = \phi\delta_{ij}]{\text{by setting}} \int_{\mathcal{C}} \left[\phi\delta_{ij}\right] dx_j \\
&\xlongequal[\text{from (1.36)}]{\text{on the one hand}} \int_{\mathcal{C}} \phi \, dx_i \\
&\xlongequal[\text{from (8.24)}]{\text{on the other hand}} \int_{A} \varepsilon_{jlk} \frac{\partial \left(\phi\delta_{ik}\right)}{\partial x_l} \widehat{n}_j \, dA
\end{aligned}$$

$$\xrightarrow[\text{the product rule}]{\text{by using}} \int_A \varepsilon_{jlk}\delta_{ik}\,\widehat{n}_j \frac{\partial \phi}{\partial x_l}\,dA$$

$$\xrightarrow[\text{(1.36) and (1.54)}]{\text{by using}} \int_A \varepsilon_{ijl}\,\widehat{n}_j \frac{\partial \phi}{\partial x_l}\,dA$$

$$\xrightarrow[\text{of the dummy index } l \text{ to } k]{\text{by changing the name}} \int_A \varepsilon_{ijk}\,\widehat{n}_j \frac{\partial \phi}{\partial x_k}\,dA\ ,$$

where the identity $\partial \mathbf{I}/\partial \mathbf{x} = \mathsf{O}$, according to $(7.33)_2$, has been used.

As an example, consider a vector field of the form $\mathbf{u} = \text{grad}\phi$ where ϕ is said to be its *potential*. Then, using (1.40), (7.123c) and (8.18), the circulation of this vector field vanishes:

$$\boxed{\int_C (\text{grad}\phi) \cdot d\mathbf{x} = \int_A (\text{curl}\,\text{grad}\phi) \cdot \widehat{\mathbf{n}}\,dA = \int_A \mathbf{0} \cdot \widehat{\mathbf{n}}\,dA = 0\ .} \tag{8.25}$$

Another example regards a tensor field of the form $\mathbf{A} = \text{grad}\mathbf{u}$ satisfying (7.124a). By means of (2.4) and (8.24), one then infers that

$$\boxed{\int_C (\text{grad}\mathbf{u})\,d\mathbf{x} = \int_A (\text{curl}\,\text{grad}\mathbf{u})^{\mathrm{T}}\,\widehat{\mathbf{n}}\,dA = \int_A (\mathbf{O})^{\mathrm{T}}\,\widehat{\mathbf{n}}\,dA = \int_A \mathbf{O}\,\widehat{\mathbf{n}}\,dA = \mathbf{0}\ .} \tag{8.26}$$

8.2 Differential Forms

Differential forms (or simply *forms*) appear in many physical contexts such as electromagnetic field theory. They are also used to numerically treat boundary value problems in solid as well as fluid mechanics. In differential geometry, their most important application is integration on manifolds. Examples of which include the divergence and Stokes' theorems.

For these mathematical creatures, differentiation can be made in a specific manner even without defining a metric. In particular, a great advantage of the differential forms is that they behave much like vectors. And they can thus be considered as a **complement** to vector analysis. A differential form is simply an integrand, i.e. an object which can be integrated over some **oriented** region.

Assigning an orientation to curves (one-dimensional manifolds) can be done in a straightforward manner. One can also orient many surfaces (two-dimensional manifolds) properly as demonstrated in the previous section for integral theorems (note that the **Möbius strip** and the **Klein bottle** ([1]) are two well-known examples of nonorientable surfaces). But, orientation cannot conveniently be specified for objects in higher dimensions. This means that vector calculus is not an appropriate tool for carrying out integration on objects in higher dimensions. Instead, differential forms and their derivatives, called *exterior derivatives*, are utilized. The reason is that they are capable of generalizing integral theorems to any arbitrary dimension. They help express the widely used integral theorems in a more sophisticated and unified format. It is known as the *generalized Stokes' theorem* which holds true for all dimensions. This great advantage motivates to entirely devote this section to the study of differential forms and their calculus. For a detailed account on differential forms, see Spivak [2], Lovelock and Rund [3], do Carmo [4], Bachman [5], Renteln [6], Hubbard and Hubbard [7], Nguyen-Schäfer and Schmidt [8] and Fortney [9] among many others.

Denoting by n the space dimension, a differential k-form, where the integer k belongs to the interval $0 \leq k \leq n$, is a **linear transformation** that takes k vectors as inputs and provides a scalar as an output (for reasons that become clear later, a differential k-form vanishes when $k > n$). Here, the number k is known as the *degree* of form. For the sake of consistency with the dimension of vector spaces introduced so far, the main focus here will be on objects belonging to the three-dimensional real vector space $\mathcal{E}_{\mathrm{r}}^3$. Note that for subsequent developments, a special vector space, called the *Minkowski spacetime*, will be introduced which has one extra dimension, see Sect. 8.2.4.1.

Denoting by $\overset{k}{\omega}$ a differential k-form, the linear mapping

$$\overset{k}{\omega} : \underbrace{\mathcal{E}_{\mathrm{r}}^3 \times \cdots \times \mathcal{E}_{\mathrm{r}}^3}_{k \text{ times}} \to \mathrm{R} , \tag{8.27}$$

helps construct the space of differential forms on $\mathcal{E}_{\mathrm{r}}^3$. That will be a new vector space usually denoted by

$$\bigwedge{}^k \mathcal{E}_{\mathrm{r}}^3 . \quad \longleftarrow \text{see (8.30), (8.37), (8.44) and (8.55) for } 0 \leq k \leq 3 \tag{8.28}$$

8.2.1 Differential 1-Form

A differential 0-form presents the simplest case of differential forms. Without taking any vector, it represents a scalar function

$$\boxed{\overset{0}{\omega} = \overset{0}{\omega}\left(\Theta^1, \Theta^2, \Theta^3\right) ,} \tag{8.29}$$

which belongs to the following space

$$\bigwedge{}^0 \mathcal{E}_{\mathrm{r}}^3 = \mathrm{Span}\left\{1\right\} . \tag{8.30}$$

And its dimension is 1. A differential 1-form is then constructed by the sum

$$\boxed{\overset{1}{\omega}(\mathbf{u}) = \overset{1}{\omega}_i \left(\Theta^1, \Theta^2, \Theta^3\right) d\Theta^i(\mathbf{u}) \ ,} \tag{8.31}$$

where $\mathbf{u} = \underline{u}^j \mathbf{g}_j$ is a vector field. It is important to note that $d\Theta^i(\mathbf{u})$ represents the **directional derivative** of Θ^i in the direction of $\mathbf{u}$. It is known as the *elementary* (or *simple*) *1-form* (having in mind that the elementary 0-form is 1). Using (5.27)$_1$, (5.28) and (7.4a)$_3$, it takes the following form

$$\boxed{d\Theta^i(\mathbf{u}) = \frac{\partial \Theta^i}{\partial \mathbf{x}} \cdot \mathbf{u} = \left(\mathbf{g}^i\right) \cdot \left(\underline{u}^j \mathbf{g}_j\right) = \delta^i_j \, \underline{u}^j = \underline{u}^i \ ,} \tag{8.32}$$

where the replacement property of the Kronecker delta along with the bilinearity of the dot product have been used. Trivially,

$$\boxed{d\Theta^i\left(\mathbf{g}_j\right) = \delta^i_j \ .} \tag{8.33}$$

One can now rephrase (8.31) in the form

$$\boxed{\overset{1}{\omega}(\mathbf{u}) = \overset{1}{\omega}_i \, \underline{u}^i \ .} \tag{8.34}$$

It is worthwhile to point out that the direction specified by $\mathbf{u}$ is fixed at a given point. As a result,

$$d\left(d\Theta^i\right) = d\,\underline{u}^i = 0 \ . \tag{8.35}$$

A quick example here is $\overset{1}{\omega} = \sin\Theta^2 d\Theta^1 + \cos\Theta^1 d\Theta^2 + \sin^2\Theta^3 d\Theta^3$ to be evaluated on $\mathbf{u} = \cos\Theta^1 \mathbf{g}_1 - \sin\Theta^2 \mathbf{g}_2 + \mathbf{g}_3$. The result is $\overset{1}{\omega}(\mathbf{u}) = \sin^2\Theta^3$.

Let $\overset{1}{\omega} = \overset{1}{\omega}_i d\Theta^i$ and $\overset{1}{\varsigma} = \overset{1}{\varsigma}_i d\Theta^i$ be two 1-forms. Further, let f be a scalar function of the local curvilinear coordinates. Then, the sum $\overset{1}{\omega} + \overset{1}{\varsigma}$ and the product $f\overset{1}{\omega}$ obey

$$\overset{1}{\omega} + \overset{1}{\varsigma} = \underbrace{\left(\overset{1}{\omega}_i + \overset{1}{\varsigma}_i\right) d\Theta^i}_{\text{in general, this result holds true for any } \overset{k}{\omega} \text{ and } \overset{k}{\varsigma}} \ , \tag{8.36a}$$

$$f\overset{1}{\omega} = \underbrace{\left(f\overset{1}{\omega}_i\right) d\Theta^i}_{\text{in general, this result holds true for any } \overset{k}{\omega}} \ . \tag{8.36b}$$

As can be seen, the result of any of these mathematical operations, i.e. **addition** and **scalar multiplication**, is again a differential 1-form. This motivates to construct the space of differential 1-forms on $\mathcal{E}_{\mathrm{r}}^3$. It is a new vector space; given by,

$$\bigwedge{}^1 \mathcal{E}_{\mathrm{r}}^3 = \mathrm{Span}\left\{d\Theta^1, d\Theta^2, d\Theta^3\right\} . \tag{8.37}$$

And its dimension is 3.
Hint: A differential 1-form $\overset{1}{\omega} = \overset{1}{\omega}_i d\Theta^i$, according to (8.31), may be thought of as an object $\mathbf{w} = \underline{w}_i \mathbf{g}^i$ by setting up the following correspondence

$$d\Theta^i \longleftrightarrow \mathbf{g}^i .$$

Then,

$$\overset{1}{\omega}_1 d\Theta^1 + \overset{1}{\omega}_2 d\Theta^2 + \overset{1}{\omega}_3 d\Theta^3 \longleftrightarrow \underline{w}_1 \mathbf{g}^1 + \underline{w}_2 \mathbf{g}^2 + \underline{w}_3 \mathbf{g}^3 .$$

With regards to this, $\overset{1}{\omega}$ may be identified as a covariant first-order tensor decomposed with respect to a (differential 1-form) basis $\left\{d\Theta^i\right\}$ by means of the corresponding (differential 0-form) components $\overset{1}{\omega}_i$. That is why $\overset{1}{\omega}$ is often introduced as an object behaving very similar to $\mathbf{w}$.

Recall from the expression (5.66a) that a covector $\underline{u}_i$ is related to its companion vector $\underline{u}^j$ via $\underline{u}_i = g_{ij}\, \underline{u}^j$. This reveals the fact that a covector is basically a linear combination of a vector. In this context, the differential 1-form (8.34) should thus be realized as a **covector**.

8.2.2 *Differential 2-Form*

A differential 2-form is constructed by the so-called *wedge* (or *exterior*) *product* of two differential 1-forms. The wedge product of a k-form and a l-form generates a $(k+l)$-form. Similarly to the cross product, the wedge product possesses the **skew-symmetric** and **bilinearity** properties. For all 1-forms $\overset{1}{\omega}, \overset{1}{\varsigma}, \overset{1}{\pi} \in \bigwedge{}^1 \mathcal{E}_{\mathrm{r}}^3$ and all functions $\alpha, \beta \in \bigwedge{}^0 \mathcal{E}_{\mathrm{r}}^3$, it satisfies.[4]

[4] Note that (8.38a) is a special case of the **skew-commutative property**

$$\overset{k}{\varsigma} \wedge \overset{l}{\pi} = (-1)^{kl}\, \overset{l}{\pi} \wedge \overset{k}{\varsigma} , \quad \text{for any } \overset{k}{\varsigma} = \overset{k}{\varsigma}_{i_1 \cdots i_k} d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k} \ \text{ and } \ \overset{l}{\pi} = \overset{l}{\pi}_{j_1 \cdots j_l} d\Theta^{j_1} \wedge \cdots \wedge d\Theta^{j_l} ,$$

owing to

$$\begin{aligned}
\overset{k}{\varsigma} \wedge \overset{l}{\pi} &= \overset{k}{\varsigma}_{i_1 \cdots i_k} \overset{l}{\pi}_{j_1 \cdots j_l} \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) \wedge \left(d\Theta^{j_1} \wedge \cdots \wedge d\Theta^{j_l}\right) \\
&= (-1)^l\, \overset{k}{\varsigma}_{i_1 \cdots i_k} \overset{l}{\pi}_{j_1 \cdots j_l} \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_{k-1}}\right) \wedge \left(d\Theta^{j_1} \wedge \cdots \wedge d\Theta^{j_l}\right) \wedge d\Theta^{i_k} \\
&= (-1)^{kl}\, \overset{l}{\pi}_{j_1 \cdots j_l} \overset{k}{\varsigma}_{i_1 \cdots i_k} \left(d\Theta^{j_1} \wedge \cdots \wedge d\Theta^{j_l}\right) \wedge \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) = (-1)^{kl}\, \overset{l}{\pi} \wedge \overset{k}{\varsigma} .
\end{aligned}$$

$$\overset{1}{\varsigma} \wedge \overset{1}{\pi} = -\overset{1}{\pi} \wedge \overset{1}{\varsigma} , \tag{8.38a}$$

$$\left(\alpha\overset{1}{\omega} + \beta\overset{1}{\varsigma}\right) \wedge \overset{1}{\pi} = \alpha\left(\overset{1}{\omega} \wedge \overset{1}{\pi}\right) + \beta\left(\overset{1}{\varsigma} \wedge \overset{1}{\pi}\right) , \tag{8.38b}$$

$$\overset{1}{\omega} \wedge \overset{1}{\omega} = 0 . \tag{8.38c}$$

$\longleftarrow$ this property is an immediate consequence of (8.38a)

Suppose one is given $\overset{1}{\varsigma} = \overset{1}{\varsigma}_i d\Theta^i$ and $\overset{1}{\pi} = \overset{1}{\pi}_j d\Theta^j$. Then, these rules help represent

$$\begin{aligned}\overset{1}{\varsigma} \wedge \overset{1}{\pi} &= \overset{1}{\varsigma}_i \overset{1}{\pi}_j d\Theta^i \wedge d\Theta^j \\ &= \sum_{\substack{r,s=1 \\ r<s}}^{n=3} \left(\overset{1}{\varsigma}_r \overset{1}{\pi}_s - \overset{1}{\varsigma}_s \overset{1}{\pi}_r\right) d\Theta^r \wedge d\Theta^s ,\end{aligned} \tag{8.39}$$

or

$$\overset{1}{\varsigma} \wedge \overset{1}{\pi} = \sum_{\substack{r,s=1 \\ r<s}}^{n=3} \det\begin{bmatrix} \overset{1}{\varsigma}_r & \overset{1}{\pi}_r \\ \overset{1}{\varsigma}_s & \overset{1}{\pi}_s \end{bmatrix} d\Theta^r \wedge d\Theta^s . \tag{8.40}$$

As a quick example, consider a differential 2-form

$$\overset{2}{\omega} = d\Theta^1 \wedge d\Theta^2 + d\Theta^2 \wedge d\Theta^3 + d\Theta^1 \wedge d\Theta^3 ,$$

which may be rephrased as

$$\begin{aligned}\overset{2}{\omega} &= d\Theta^2 \wedge \left(-d\Theta^1 + d\Theta^3\right) - d\Theta^3 \wedge d\Theta^1 + d\Theta^3 \wedge d\Theta^3 \\ &= d\Theta^2 \wedge \left(-d\Theta^1 + d\Theta^3\right) + d\Theta^3 \wedge \left(-d\Theta^1 + d\Theta^3\right) \\ &= \left(d\Theta^2 + d\Theta^3\right) \wedge \left(-d\Theta^1 + d\Theta^3\right) \\ &= d\left(\Theta^3 + \Theta^2\right) \wedge d\left(\Theta^3 - \Theta^1\right) .\end{aligned}$$

The most general differential 2-form is now given by

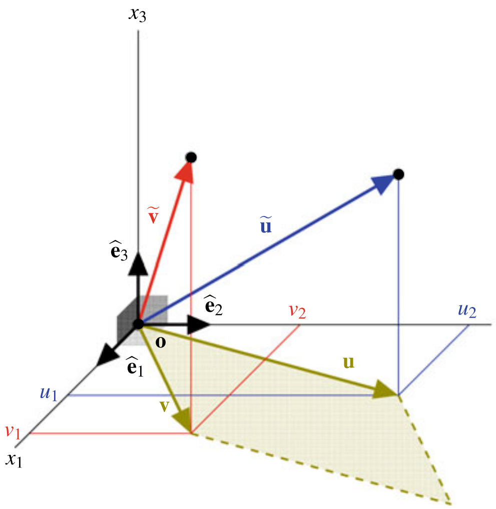
$$\boxed{\overset{2}{\omega}(\mathbf{u}, \mathbf{v}) = \overset{2}{\omega}_{ij}\left(\Theta^1, \Theta^2, \Theta^3\right)\left(d\Theta^i \wedge d\Theta^j\right)(\mathbf{u}, \mathbf{v}) ,} \tag{8.41}$$

where

$$\boxed{\left(d\Theta^i \wedge d\Theta^j\right)(\mathbf{u}, \mathbf{v}) = \det\begin{bmatrix} d\Theta^i(\mathbf{u}) & d\Theta^i(\mathbf{v}) \\ d\Theta^j(\mathbf{u}) & d\Theta^j(\mathbf{v}) \end{bmatrix} .} \tag{8.42}$$

or $\left(d\Theta^i \wedge d\Theta^j\right)(\mathbf{u}, \mathbf{v}) = \underline{u}^i \, \underline{v}^j - \underline{u}^j \, \underline{v}^i$

Also note that (8.38c) can be resulted from $\overset{k}{\varsigma} \wedge \overset{l}{\pi} = (-1)^{kl} \overset{l}{\pi} \wedge \overset{k}{\varsigma}$ by choosing $k = l = 1$ and $\overset{k}{\varsigma} = \overset{k}{\pi}$.



Consider a Cartesian coordinate frame with the following arbitrary vectors

$$\widetilde{\mathbf{u}} = \sum_{i=1}^{3} u_i \widehat{\mathbf{e}}_i \quad , \quad \widetilde{\mathbf{v}} = \sum_{i=1}^{3} v_i \widehat{\mathbf{e}}_i \; .$$

Further, consider the projection of these vectors onto the x_1x_2-plane, that is,

$$\mathbf{u} = \sum_{i=1}^{2} u_i \widehat{\mathbf{e}}_i \quad , \quad \mathbf{v} = \sum_{i=1}^{2} v_i \widehat{\mathbf{e}}_i \; .$$

The area of the parallelogram spanned by **u** and **v** then renders $A = |u_1 v_2 - u_2 v_1|$.
Note that

$$(dx_1 \wedge dx_2)(\mathbf{u}, \mathbf{v}) = \det \begin{bmatrix} dx_1(\mathbf{u}) & dx_1(\mathbf{v}) \\ dx_2(\mathbf{u}) & dx_2(\mathbf{v}) \end{bmatrix} .$$

This represents the **signed area**. Finally, the geometric and signed areas are related by $A = |(dx_1 \wedge dx_2)(\mathbf{u}, \mathbf{v})|$.

Fig. 8.3 Area and signed area

Hint: In Cartesian coordinates, consider a parallelogram defined by the projection of **u** and **v**, for instance, onto the plane spanned by $\widehat{\mathbf{e}}_1$ and $\widehat{\mathbf{e}}_2$. Then, $(dx_1 \wedge dx_2)(\mathbf{u}, \mathbf{v})$ evaluates the *signed area* of that parallelogram (note that the signed area of a region may be negative in contrast to its geometric area which is always positive). This provides the geometric meaning of an elementary 2-form (see Fig. 8.3).

From (8.41) and (8.42), it follows that

$$\begin{aligned} \overset{2}{\omega}(\mathbf{u}, \mathbf{v}) &= \left(\overset{2}{\omega}_{ij} - \overset{2}{\omega}_{ji}\right) \underline{u}^i \, \underline{v}^j \\ &= \sum_{\substack{r,\, s=1 \\ r<s}}^{n=3} \left(\overset{2}{\omega}_{rs} - \overset{2}{\omega}_{sr}\right) \left(\underline{u}^r \, \underline{v}^s - \underline{u}^s \, \underline{v}^r\right) \; . \end{aligned} \tag{8.43}$$

As a simple example, consider

$$\left.\begin{aligned} \overset{1}{\varsigma} &= 3d\Theta^1 - d\Theta^2 + 2d\Theta^3 \\ \overset{1}{\pi} &= d\Theta^1 + 2d\Theta^2 - d\Theta^3 \end{aligned}\right\} ,$$

to be evaluated on

$$\left.\begin{aligned} \mathbf{u} &= 2\mathbf{g}_1 + \mathbf{g}_2 - \mathbf{g}_3 \\ \mathbf{v} &= \mathbf{g}_1 - 3\mathbf{g}_2 + 2\mathbf{g}_3 \end{aligned}\right\} .$$

Then, using $(8.43)_2$, the wedge product $\overset{2}{\omega} = \overset{1}{\varsigma} \wedge \overset{1}{\pi}$ delivers

$$\overset{2}{\omega} = \sum_{\substack{r,s=1 \\ r<s}}^{n=3} \left(\overset{1}{\varsigma}_r \overset{1}{\pi}_s - \overset{1}{\varsigma}_s \overset{1}{\pi}_r \right) \left(\underline{u}^r \, \underline{v}^s - \underline{u}^s \, \underline{v}^r \right)$$

$$= \underbrace{\left(\overset{1}{\varsigma}_1 \overset{1}{\pi}_2 - \overset{1}{\varsigma}_2 \overset{1}{\pi}_1 \right)}_{= (3)\,(2) - (-1)\,(1) = 7} \underbrace{\left(\underline{u}^1 \, \underline{v}^2 - \underline{u}^2 \, \underline{v}^1 \right)}_{= (2)\,(-3) - (1)\,(1) = -7}$$

$$+ \underbrace{\left(\overset{1}{\varsigma}_1 \overset{1}{\pi}_3 - \overset{1}{\varsigma}_3 \overset{1}{\pi}_1 \right)}_{= (3)\,(-1) - (2)\,(1) = -5} \underbrace{\left(\underline{u}^1 \, \underline{v}^3 - \underline{u}^3 \, \underline{v}^1 \right)}_{= (2)\,(2) - (-1)\,(1) = 5}$$

$$+ \underbrace{\left(\overset{1}{\varsigma}_2 \overset{1}{\pi}_3 - \overset{1}{\varsigma}_3 \overset{1}{\pi}_2 \right)}_{= (-1)\,(-1) - (2)\,(2) = -3} \underbrace{\left(\underline{u}^2 \, \underline{v}^3 - \underline{u}^3 \, \underline{v}^2 \right)}_{= (1)\,(2) - (-1)\,(-3) = -1}$$

$$= -49 - 25 + 3 = -71 \; .$$

It is now easy to realize that the space of differential 2-forms on $\mathcal{E}_{\mathrm{r}}^3$ is of the following form

$$\bigwedge{}^2 \mathcal{E}_{\mathrm{r}}^3 = \mathrm{Span} \left\{ d\Theta^1 \wedge d\Theta^2, d\Theta^2 \wedge d\Theta^3, d\Theta^3 \wedge d\Theta^1 \right\} \; , \tag{8.44}$$

whose dimension, similarly to (8.37), is 3.

8.2.3 *Differential 3-Form*

Three differential 1-forms $\overset{1}{\varsigma}$, $\overset{1}{\pi}$ and $\overset{1}{\xi}$ can be used to construct a differential 3-form, $\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi}$, satisfying the **associative rule**

$$\underbrace{\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} = \left(\overset{1}{\varsigma} \wedge \overset{1}{\pi} \right) \wedge \overset{1}{\xi} = \overset{1}{\varsigma} \wedge \left(\overset{1}{\pi} \wedge \overset{1}{\xi} \right)}_{\substack{\text{note that the wedge product is much like the cross product} \\ \text{having in mind that the cross product is not associative}}} \; . \tag{8.45}$$

The skew-symmetric and bilinearity properties of the wedge product then help establish

$$\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} = \underbrace{-\overset{1}{\varsigma} \wedge \overset{1}{\xi} \wedge \overset{1}{\pi}}_{= \overset{1}{\pi} \wedge \overset{1}{\xi} \wedge \overset{1}{\varsigma} = -\overset{1}{\pi} \wedge \overset{1}{\varsigma} \wedge \overset{1}{\xi} = \overset{1}{\xi} \wedge \overset{1}{\varsigma} \wedge \overset{1}{\pi} = -\overset{1}{\xi} \wedge \overset{1}{\pi} \wedge \overset{1}{\varsigma}} \; , \tag{8.46}$$

and

$$\underbrace{\left(\alpha\overset{1}{\omega} + \beta\overset{1}{\varsigma}\right) \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} = \alpha\overset{1}{\omega} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} + \beta\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi}}_{\substack{\text{or } \overset{1}{\varsigma} \wedge \left(\alpha\overset{1}{\omega} + \beta\overset{1}{\pi}\right) \wedge \overset{1}{\xi} = \alpha\overset{1}{\varsigma} \wedge \overset{1}{\omega} \wedge \overset{1}{\xi} + \beta\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} \\ \text{or } \overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \left(\alpha\overset{1}{\omega} + \beta\overset{1}{\xi}\right) = \alpha\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\omega} + \beta\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi}}} \ . \tag{8.47}$$

If any two of differential 1-forms are equal, one will have

$$\overset{1}{\varsigma} \wedge \overset{1}{\xi} \wedge \overset{1}{\xi} = \overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\varsigma} = \overset{1}{\pi} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} = 0 \ . \tag{8.48}$$

Let $\overset{1}{\varsigma} = \overset{1}{\varsigma}_i d\Theta^i$, $\overset{1}{\pi} = \overset{1}{\pi}_j d\Theta^j$ and $\overset{1}{\xi} = \overset{1}{\xi}_k d\Theta^k$. It is then easy to see that

$$\begin{aligned} \overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} &= \overset{1}{\varsigma}_i \overset{1}{\pi}_j \overset{1}{\xi}_k d\Theta^i \wedge d\Theta^j \wedge d\Theta^k \\ &= \left[\overset{1}{\varsigma}_1 \left(\overset{1}{\pi}_2 \overset{1}{\xi}_3 - \overset{1}{\pi}_3 \overset{1}{\xi}_2\right) - \overset{1}{\pi}_1 \left(\overset{1}{\varsigma}_2 \overset{1}{\xi}_3 - \overset{1}{\varsigma}_3 \overset{1}{\xi}_2\right) \right. \\ &\quad \left. + \overset{1}{\xi}_1 \left(\overset{1}{\varsigma}_2 \overset{1}{\pi}_3 - \overset{1}{\varsigma}_3 \overset{1}{\pi}_2\right)\right] d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \\ &= \varepsilon^{ijk} \overset{1}{\varsigma}_i \overset{1}{\pi}_j \overset{1}{\xi}_k d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \ , \end{aligned} \tag{8.49}$$

or

$$\overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi} = \det \begin{bmatrix} \overset{1}{\varsigma}_1 & \overset{1}{\pi}_1 & \overset{1}{\xi}_1 \\ \overset{1}{\varsigma}_2 & \overset{1}{\pi}_2 & \overset{1}{\xi}_2 \\ \overset{1}{\varsigma}_3 & \overset{1}{\pi}_3 & \overset{1}{\xi}_3 \end{bmatrix} d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \ . \tag{8.50}$$

A quick example here is

$$\begin{aligned} \overset{3}{\omega} &= \left(3d\Theta^1 + d\Theta^2\right) \wedge \left(d\Theta^1 - d\Theta^3\right) \wedge \left(d\Theta^1 + d\Theta^2 + d\Theta^3\right) \\ &= \left(-3d\Theta^1 \wedge d\Theta^3 + d\Theta^2 \wedge d\Theta^1 - d\Theta^2 \wedge d\Theta^3\right) \wedge \left(d\Theta^1 + d\Theta^2 + d\Theta^3\right) \\ &= -d\Theta^2 \wedge d\Theta^3 \wedge d\Theta^1 - 3d\Theta^1 \wedge d\Theta^3 \wedge d\Theta^2 + d\Theta^2 \wedge d\Theta^1 \wedge d\Theta^3 \\ &= -d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 + 3d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 - d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \\ &= d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \ . \end{aligned}$$

In alignment with (8.31) and (8.41), the most general form of a differential 3-form is given by

$$\boxed{\overset{3}{\omega}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \overset{3}{\omega}_{ijk}\left(\Theta^1, \Theta^2, \Theta^3\right)\left(d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\right)(\mathbf{u}, \mathbf{v}, \mathbf{w}) \ ,} \tag{8.51}$$

where

$$\boxed{\underbrace{\left(d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\right)(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \det \begin{bmatrix} d\Theta^i(\mathbf{u}) & d\Theta^i(\mathbf{v}) & d\Theta^i(\mathbf{w}) \\ d\Theta^j(\mathbf{u}) & d\Theta^j(\mathbf{v}) & d\Theta^j(\mathbf{w}) \\ d\Theta^k(\mathbf{u}) & d\Theta^k(\mathbf{v}) & d\Theta^k(\mathbf{w}) \end{bmatrix}}_{\text{or } \left(d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\right)(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \underline{u}^i \left(\underline{v}^j \underline{w}^k - \underline{v}^k \underline{w}^j\right) - \underline{v}^i \left(\underline{u}^j \underline{w}^k - \underline{u}^k \underline{w}^j\right) + \underline{w}^i \left(\underline{u}^j \underline{v}^k - \underline{u}^k \underline{v}^j\right)} .} \tag{8.52}$$

The most general form of a differential 3-form, according to (8.51), may also be written in the form

$$\overset{3}{\omega}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \varXi \, \varepsilon^{ijk} \overset{3}{\omega}_{ijk} , \tag{8.53}$$

where, knowing that $\varepsilon^{ijk} \overset{3}{\omega}_{ijk} = \overset{3}{\omega}_{123} - \overset{3}{\omega}_{132} - \overset{3}{\omega}_{213} + \overset{3}{\omega}_{231} + \overset{3}{\omega}_{312} - \overset{3}{\omega}_{321}$,

$$\varXi = \det \begin{bmatrix} \underline{u}^1 & \underline{v}^1 & \underline{w}^1 \\ \underline{u}^2 & \underline{v}^2 & \underline{w}^2 \\ \underline{u}^3 & \underline{v}^3 & \underline{w}^3 \end{bmatrix} . \tag{8.54}$$

As an example, suppose one is given

$$\left.\begin{aligned} \overset{1}{\varsigma} &= -d\Theta^1 + d\Theta^2 + d\Theta^3 \\ \overset{1}{\pi} &= 2d\Theta^1 - d\Theta^2 - d\Theta^3 \\ \overset{1}{\xi} &= d\Theta^1 + d\Theta^2 - 2d\Theta^3 \end{aligned}\right\} \quad \text{and} \quad \left.\begin{aligned} \mathbf{u} &= \mathbf{g}_1 + \mathbf{g}_2 - \mathbf{g}_3 \\ \mathbf{v} &= 3\mathbf{g}_1 + \mathbf{g}_2 + 2\mathbf{g}_3 \\ \mathbf{w} &= \mathbf{g}_1 - \mathbf{g}_2 - \mathbf{g}_3 \end{aligned}\right\} .$$

Then, by means of (8.50) and (8.52), the covector $\overset{3}{\omega} = \overset{1}{\varsigma} \wedge \overset{1}{\pi} \wedge \overset{1}{\xi}$ renders

$$\overset{3}{\omega} = \underbrace{\det \begin{bmatrix} -1 & 2 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -2 \end{bmatrix}}_{=3} \underbrace{\det \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & -1 \\ -1 & 2 & -1 \end{bmatrix}}_{=10} = 30 .$$

One should now realize that the space of differential 3-forms on $\mathcal{E}_{\mathrm{r}}^3$ takes the following form

$$\bigwedge{}^3 \mathcal{E}_{\mathrm{r}}^3 = \operatorname{Span}\left\{d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3\right\} , \tag{8.55}$$

whose dimension, similarly to (8.30), is 1.
Hint: Notice that there will be no (nonzero) differential forms of degree greater than 3 on $\mathcal{E}_{\mathrm{r}}^3$. Because one should choose more than 3 distinct integer values ranging from 1 and 3 to construct elementary differential forms.

At the end, a differential k-form is given by

$$\boxed{\overset{k}{\omega}(\mathbf{u}_1, \ldots, \mathbf{u}_k) = \overset{k}{\omega}_{i_1 \cdots i_k}\left(\Theta^1, \ldots, \Theta^n\right)\left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right)(\mathbf{u}_1, \ldots, \mathbf{u}_k) ,} \tag{8.56}$$

where $\overset{k}{\omega} \in \bigwedge^k \mathcal{E}_{\mathrm{r}}^n$ for which n denotes the space dimension and $\mathcal{E}_{\mathrm{r}}^n$ presents a real vector space of dimension n. In the above expression,

$$\boxed{\left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right)(\mathbf{u}_1, \ldots, \mathbf{u}_k) = \det \begin{bmatrix} d\Theta^{i_1}(\mathbf{u}_1) & \cdots & d\Theta^{i_1}(\mathbf{u}_k) \\ \vdots & \vdots & \vdots \\ d\Theta^{i_k}(\mathbf{u}_1) & \cdots & d\Theta^{i_k}(\mathbf{u}_k) \end{bmatrix}} \ . \qquad (8.57)$$

8.2.4 Hodge Star Operator

The *Hodge star operator* (or *Hodge dual* or *star operator*) converts a differential form into its so-called *dual form*. This useful operator helps express the differential operators of vector calculus (such as gradient and divergence) in the language of **exterior calculus**, see Sect. 8.2.6.

The star operator $*$ is a linear map from $\bigwedge^k \mathcal{E}_{\mathrm{r}}^n$ to $\bigwedge^{n-k} \mathcal{E}_{\mathrm{r}}^n$, that is,

$$* : \bigwedge^k \mathcal{E}_{\mathrm{r}}^n \to \bigwedge^{n-k} \mathcal{E}_{\mathrm{r}}^n \ . \qquad (8.58)$$

In general, the Hodge star of a k-form $\overset{k}{\omega} = \overset{k}{\omega}_{i_1 \cdots i_k} d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k} \in \bigwedge^k \mathcal{E}_{\mathrm{r}}^n$ is obtained by

$$\boxed{*\overset{k}{\omega} = \overset{k}{\omega}_{i_1 \cdots i_k} * \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) \in \bigwedge^{n-k} \mathcal{E}_{\mathrm{r}}^n \ ,} \qquad (8.59)$$

where the action of the star operator on elementary differential forms is defined as

$$\boxed{* \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) = \frac{J\, g^{i_1 j_1} \cdots g^{i_k j_k} \varepsilon_{j_1 \cdots j_k l_1 \cdots l_{n-k}}}{(n-k)!} d\Theta^{l_1} \wedge \cdots \wedge d\Theta^{l_{n-k}} \ ,} \qquad (8.60)$$

where

$$\boxed{J = \sqrt{\left|\det\left[\, g_{ij} \,\right]\right|} \ .} \qquad (8.61)$$

← see (5.41)

An elementary differential form, up to a sign, remains unchanged under the action of **double-star operator**:

$$\begin{aligned} * * \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) &= * \left(* \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right)\right) \\ &= (-1)^{k(n-k)}\, s \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) \ , \end{aligned} \qquad (8.62)$$

where s denotes the sign of the determinant of the matrix of the inner product on the vector space.[5] For a three-dimensional Euclidean space, one will have

$$* (1) = \frac{J}{3!} \varepsilon_{ijk} d\Theta^i \wedge d\Theta^j \wedge d\Theta^k$$

$$= J d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \,, \tag{8.63a}$$

$$* \left(d\Theta^i\right) = \frac{J}{2} g^{il} \varepsilon_{ljk} d\Theta^j \wedge d\Theta^k \,, \tag{8.63b}$$

$$* \left(d\Theta^i \wedge d\Theta^j\right) = J g^{il} g^{jm} \varepsilon_{lmk} d\Theta^k \,, \tag{8.63c}$$

$$* \left(d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\right) = \underbrace{J g^{il} g^{jm} g^{kn} \varepsilon_{lmn} = J^{-1} \varepsilon^{ijk}}_{\text{or simply } J^{-1} = * \left(d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3\right)} \quad . \quad \longleftarrow \text{see (5.53)} \tag{8.63d}$$

These relations truly satisfy $(8.62)_2$ as follows:

$$* * (1) = J * \left(d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3\right) = J \, J^{-1}$$

$$= \boxed{1 \,,} \tag{8.64a}$$

$$* * \left(d\Theta^i\right) = \frac{J}{2} g^{il} \varepsilon_{ljk} * \left(d\Theta^j \wedge d\Theta^k\right) = \frac{J}{2} g^{il} \varepsilon_{ljk} J g^{jr} g^{ks} \varepsilon_{rst} d\Theta^t$$

$$= \frac{J^2}{2} g^{li} g^{jr} g^{ks} \varepsilon_{ljk} \varepsilon_{rst} d\Theta^t = \frac{J^2}{2} J^{-2} \varepsilon^{irs} \varepsilon_{rst} d\Theta^t$$

$$= \frac{1}{2} \varepsilon^{irs} \varepsilon_{trs} d\Theta^t = \frac{1}{2} 2 \delta^i_t d\Theta^t$$

$$= \boxed{d\Theta^i \,,} \tag{8.64b}$$

$$* * \left(d\Theta^i \wedge d\Theta^j\right) = J g^{il} g^{jm} \varepsilon_{lmk} * \left(d\Theta^k\right) = J g^{il} g^{jm} \varepsilon_{lmk} \frac{J}{2} g^{kr} \varepsilon_{rst} d\Theta^s \wedge d\Theta^t$$

$$= \frac{J^2}{2} g^{li} g^{mj} g^{kr} \varepsilon_{lmk} \varepsilon_{rst} d\Theta^s \wedge d\Theta^t = \frac{J^2}{2} J^{-2} \varepsilon^{ijr} \varepsilon_{rst} d\Theta^s \wedge d\Theta^t$$

$$= \frac{1}{2} \varepsilon^{ijr} \varepsilon_{str} d\Theta^s \wedge d\Theta^t = \frac{1}{2} \left(\delta^i_s \delta^j_t - \delta^i_t \delta^j_s\right) d\Theta^s \wedge d\Theta^t$$

$$= \frac{1}{2} d\Theta^i \wedge d\Theta^j - \frac{1}{2} d\Theta^j \wedge d\Theta^i$$

$$= \boxed{d\Theta^i \wedge d\Theta^j \,,} \tag{8.64c}$$

$$* * \left(d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3\right) = J^{-1} * (1)$$

$$= \boxed{d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \,,} \tag{8.64d}$$

where (1.54), (1.58a), (1.58b), (5.14), (5.53), (8.38a) and $(8.62)_1$ have been used. Thus, in three-dimensional Euclidean spaces, one will have

[5] Note that the parity of the signature of the inner product on the vector space is always +1 for Euclidean spaces while it is −1 in, for instance, the four-dimensional Minkowski space with the Cartesian metric (8.70).

$$\boxed{* * \overset{k}{\omega} = \overset{k}{\omega} \in \bigwedge{}^k \mathcal{E}_{\mathrm{r}}^3 \ .} \tag{8.65}$$

In the following, the goal is to express the Hodge star of elementary differential forms, according to (8.63a)-(8.63d), in the well-known coordinate systems introduced in this text, see Exercise 5.1.

First, consider the **Cartesian coordinates** $\left(\Theta^1 = x, \Theta^2 = y, \Theta^3 = z\right)$ for which simply $J = 1$ and $g^{ij} = \delta^{ij}$. In this case,

$$* (1) = dx \wedge dy \wedge dz \ , \tag{8.66a}$$

$$* (dx) = dy \wedge dz \quad , \quad \left.\begin{aligned} * (dy) = dz \wedge dx \\ * (dz) = dx \wedge dy \end{aligned}\right\} , \tag{8.66b}$$

$$* (dx \wedge dy) = dz \quad , \quad \left.\begin{aligned} * (dy \wedge dz) = dx \\ * (dz \wedge dx) = dy \end{aligned}\right\} , \tag{8.66c}$$

$$* (dx \wedge dy \wedge dz) = 1 \ . \tag{8.66d}$$

Then, consider the **cylindrical coordinates** $\left(\Theta^1 = r, \Theta^2 = \theta, \Theta^3 = z\right)$ for which $J = r$ and the only nonzero elements of the contravariant metric coefficients are $g^{11} = 1$, $g^{22} = r^{-2}$ and $g^{33} = 1$. After some algebraic manipulations, one can arrive at

$$* (1) = r \, dr \wedge d\theta \wedge dz \ , \tag{8.67a}$$

$$* (dr) = r \, d\theta \wedge dz \quad , \quad \left.\begin{aligned} * (d\theta) = r^{-1} dz \wedge dr \\ * (dz) = r \, dr \wedge d\theta \end{aligned}\right\} , \tag{8.67b}$$

$$* (dr \wedge d\theta) = r^{-1} dz \quad , \quad \left.\begin{aligned} * (d\theta \wedge dz) = r^{-1} dr \\ * (dz \wedge dr) = r \, d\theta \end{aligned}\right\} , \tag{8.67c}$$

$$* (dr \wedge d\theta \wedge dz) = r^{-1} \ . \tag{8.67d}$$

In the **spherical coordinates** $\left(\Theta^1 = r, \Theta^2 = \theta, \Theta^3 = \phi\right)$ for which $J = r^2 \sin\theta$ and the only nonzero elements of $\left[\, g^{ij} \,\right]$ are $g^{11} = 1$, $g^{22} = r^{-2}$, $g^{33} = r^{-2} \sin^{-2}\theta$, one will finally have

$$* (1) = r^2 \sin\theta dr \wedge d\theta \wedge d\phi \ , \tag{8.68a}$$

$$* (dr) = r^2 \sin\theta d\theta \wedge d\phi \quad , \quad \left.\begin{aligned} * (d\theta) = \sin\theta d\phi \wedge dr \\ * (d\phi) = \sin^{-1}\theta dr \wedge d\theta \end{aligned}\right\} , \tag{8.68b}$$

$$* (dr \wedge d\theta) = \sin\theta d\phi \quad , \quad \left.\begin{aligned} * (d\theta \wedge d\phi) = r^{-2} \sin^{-1}\theta dr \\ * (d\phi \wedge dr) = \sin^{-1}\theta d\theta \end{aligned}\right\} , \tag{8.68c}$$

$$*(dr \wedge d\theta \wedge d\phi) = r^{-2} \sin^{-1}\theta \ . \tag{8.68d}$$

8.2.4.1 Hodge Star Operator in Minkowski Spacetime

Space and time are inextricably linked in the greatest 20th century development in science, i.e. Einstein's theory of relativity, which represents the world with four dimensions (recall that space and time are separate entities in the Newtonian picture of the world). The *Minkowski space* (or *Minkowski spacetime*) is a combination of three ordinary Euclidean space dimensions and one time dimension into a four-dimensional body. It is a mathematical setting on which the **special theory of relativity** is properly formulated. The goal here is to study the application of Hodge star operator in this (four-dimensional real vector) space. To this end, consider Cartesian coordinates for consistency.

To begin with, the spacetime coordinates are labeled according to

$$\boxed{x^0 = ct \quad , \quad x^1 = x \quad , \quad x^2 = y \quad , \quad x^3 = z \ ,} \tag{8.69}$$

where c denotes the **speed of light**. One can then denote the spacetime coordinates of a point by $x^\mu = (ct, x, y, z)$ and its space coordinates by $x^i = (x, y, z)$. Note that in this context, Greek indices assume the values 0, 1, 2, 3 while Latin index letters range over 1, 2, 3.

The Minkowski covariant metric coefficients is given by

$$[g_{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} . \tag{8.70}$$

Accordingly, the spacetime interval between two infinitesimally close points of the coordinates $x^\mu = (ct, x, y, z)$ and $x^\mu + dx^\mu = (c\,t + c\,dt, x + dx, y + dy, z + dz)$ takes the following form

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - dx^2 - dy^2 - dz^2 \ . \tag{8.71}$$

Notice that the **metric signature** $(+ - --)$ is adopted here, although this convention is not unique and the opposite signature $(- + ++)$ is also seen in the literature (for both cases, s in (8.62) takes the value -1). Moreover, the introduced metric is not positive definite and, therefore, an arbitrary vector in Minkowski spacetime may have zero length even if its components are all nonzero.

A differential k-form in this space is a function of $(c\,t, x, y, z)$ and an elementary 4-form is of the following format

$$dx^I = c\,dt \wedge dx \wedge dy \wedge dz \ . \tag{8.72}$$

In the following, the goal is to represent the Hodge dual of elementary differential forms in this four-dimensional space.

First, the **four** elementary 1-forms under the action of the Hodge star operator represent (see Exercise 8.4)

$$*c\,dt = dx \wedge dy \wedge dz \ , \tag{8.73a}$$

$$*dx = c\,dt \wedge dy \wedge dz \ , \tag{8.73b}$$

$$*dy = c\,dt \wedge dz \wedge dx \ , \tag{8.73c}$$

$$*dz = c\,dt \wedge dx \wedge dy \ . \tag{8.73d}$$

By applying the Hodge star operator to (**six** independent) elementary 2-forms in the Minkowski space, one then arrives at

$$*\,(c\,dt \wedge dx) = -dy \wedge dz \ , \tag{8.74a}$$

$$*\,(c\,dt \wedge dy) = -dz \wedge dx \ , \tag{8.74b}$$

$$*\,(c\,dt \wedge dz) = -dx \wedge dy \ , \tag{8.74c}$$

$$*\,(dx \wedge dy) = -dz \wedge c\,dt \ , \tag{8.74d}$$

$$*\,(dx \wedge dz) = -c\,dt \wedge dy \ , \tag{8.74e}$$

$$*\,(dy \wedge dz) = -dx \wedge c\,dt \ . \tag{8.74f}$$

From (8.62) and (8.73a)-(8.73d), one immediately obtains

$$*\,(dx \wedge dy \wedge dz) = c\,dt \ , \tag{8.75a}$$

$$*\,(c\,dt \wedge dy \wedge dz) = dx \ , \tag{8.75b}$$

$$*\,(c\,dt \wedge dz \wedge dx) = dy \ , \tag{8.75c}$$

$$*\,(c\,dt \wedge dx \wedge dy) = dz \ . \tag{8.75d}$$

Finally, the Hodge dual of the elementary 4-form (8.72) renders

$$*\,(c\,dt \wedge dx \wedge dy \wedge dz) = -1 \ . \tag{8.76}$$

8.2.5 Exterior Derivatives

So far, algebraic properties of differential forms have been introduced. The goal here is to study *differential calculus* of these mathematical entities. This relies on the powerful concept of *exterior derivative* described below.

In general, the exterior derivative of a k-form will produce a $(k+1)$-form. Let $\overset{k}{\varsigma} \in \bigwedge^k \mathcal{E}^3_{\mathrm{r}}$ be a differential k-form and $\overset{l}{\pi} \in \bigwedge^l \mathcal{E}^3_{\mathrm{r}}$ be a differential l-form. Further, let $f, g \in \bigwedge^0 \mathcal{E}_{\mathrm{r}}^3$ be two functions of the local curvilinear coordinates and $\alpha, \beta \in \mathrm{R}$ be two constants. The exterior derivative, denoted by d, then satisfies the following properties

$$d\left(\alpha \overset{k}{\varsigma} + \beta \overset{l}{\pi}\right) = \alpha d\overset{k}{\varsigma} + \beta d\overset{l}{\pi} \ , \tag{8.77a}$$

$$d\left(f \overset{k}{\varsigma}\right) = df \wedge \overset{k}{\varsigma} + f\, d\overset{k}{\varsigma} \ , \tag{8.77b}$$

$$d\,(f\,g) = (df)\,g + f\,dg \ . \tag{8.77c}$$

8.2.5.1 Exterior Derivative of Differential 0-Form

The exterior derivative of a differential 0-form, according to (8.29), is given by

$$\boxed{d\overset{0}{\omega} = \frac{\partial \overset{0}{\omega}}{\partial \Theta^i} d\Theta^i \ .} \tag{8.78}$$

In this expression, the **differential 1-form** $d\overset{0}{\omega}$ may be identified as a covariant first-order tensor decomposed with respect to the (differential 1-form) basis $\{d\Theta^i\}$ by means of the corresponding (differential 0-form) components $\partial\overset{0}{\omega}/\partial\Theta^i$. Indeed, this differential 1-form takes a vector $\mathbf{u}$ and returns a scalar. This is indicated by

$$\left(d\overset{0}{\omega}\right)(\mathbf{u}) = \frac{\partial \overset{0}{\omega}}{\partial \Theta^i} d\Theta^i\,(\mathbf{u}) = \frac{\partial \overset{0}{\omega}}{\partial \Theta^i}\, \underline{u}^i \ . \tag{8.79}$$

As an example, consider $\overset{0}{\omega} = \Theta^1\left(\Theta^1 + \Theta^3 + 1\right) - \Theta^2$ whose exterior derivative should be evaluated on $\mathbf{u} = \mathbf{g}_1 + \Theta^3\mathbf{g}_2 - 2\mathbf{g}_3$. The result is

$$\begin{aligned}\left(d\overset{0}{\omega}\right)(\mathbf{u}) &= \frac{\partial \overset{0}{\omega}}{\partial \Theta^i}\, \underline{u}^i \\ &= \left(\cancel{2\Theta^1} + \cancel{\Theta^3} + 1\right)(1) + \cancel{(-1)\left(\Theta^3\right)} + \cancel{\left(\Theta^1\right)(-2)} = 1 \ .\end{aligned}$$

8.2.5.2 Exterior Derivative of Differential 1-Form

The exterior derivative of a differential 1-form, according to (8.31), is defined by

$$\boxed{d\overset{1}{\omega} = d\overset{1}{\omega}_i \wedge d\Theta^i = \frac{\partial \overset{1}{\omega}_i}{\partial \Theta^j} d\Theta^j \wedge d\Theta^i = -\frac{\partial \overset{1}{\omega}_i}{\partial \Theta^j} d\Theta^i \wedge d\Theta^j \ ,} \tag{8.80}$$

in alignment with (8.35) and (8.77b) taking into account (8.38a) and (8.78). This **differential 2-form** may be realized as a covariant second-order tensor whose components with respect to the basis $\left\{d\Theta^i \wedge d\Theta^j\right\}$ are $-\partial\overset{1}{\omega}_i/\partial\Theta^j$.

As a quick example, suppose one is given

$$\overset{1}{\omega} = 2\Theta^1\Theta^2 d\Theta^1 + \Theta^1\Theta^1 d\Theta^3 \ .$$

Its exterior derivative then renders

$$\begin{aligned} d\overset{1}{\omega} &= d\left(2\Theta^1\Theta^2\right) \wedge d\Theta^1 + d\left(\Theta^1\Theta^1\right) \wedge d\Theta^3 \\ &= \cancelto{0}{2\Theta^2 d\Theta^1 \wedge d\Theta^1} + 2\Theta^1 d\Theta^2 \wedge d\Theta^1 + 2\Theta^1 d\Theta^1 \wedge d\Theta^3 \\ &= -2\Theta^1\left(d\Theta^1 \wedge d\Theta^2 - d\Theta^1 \wedge d\Theta^3\right) \ . \end{aligned}$$

Using (8.41) , $(8.42)_2$ and $(8.43)_2$, the action of the introduced differential 2-form $(8.80)_3$ on the vectors $\mathbf{u}$ and $\mathbf{v}$ represents

$$\begin{aligned} \left(d\overset{1}{\omega}\right)(\mathbf{u}, \mathbf{v}) &= -\frac{\partial\overset{1}{\omega}_i}{\partial\Theta^j}\left(d\Theta^i \wedge d\Theta^j\right)(\mathbf{u}, \mathbf{v}) \\ &= -\frac{\partial\overset{1}{\omega}_i}{\partial\Theta^j}\left(\underline{u}^i\,\underline{v}^j - \underline{u}^j\,\underline{v}^i\right) \\ &= \sum_{\substack{r,\,s=1 \\ r<s}}^{n=3} -\left(\frac{\partial\overset{1}{\omega}_r}{\partial\Theta^s} - \frac{\partial\overset{1}{\omega}_s}{\partial\Theta^r}\right)\left(\underline{u}^r\,\underline{v}^s - \underline{u}^s\,\underline{v}^r\right) \ . \end{aligned} \tag{8.81}$$

As an example, consider

$$\overset{1}{\omega} = \Theta^1\Theta^2\Theta^3 d\Theta^1 + \Theta^3 d\Theta^2 + 2\Theta^2 d\Theta^3 \ ,$$

whose exterior derivative is supposed to be calculated on

$$\mathbf{u} = -\mathbf{g}_1 + \mathbf{g}_2 + \mathbf{g}_3 \quad \text{and} \quad \mathbf{v} = \mathbf{g}_1 + 2\mathbf{g}_2 - \mathbf{g}_3 \ ,$$

at a point P corresponding to $\left(\Theta^1, \Theta^2, \Theta^3\right) = (0.5, -0.5, 0.5)$. The result will be

$$
\begin{aligned}
\left[\left(d\overset{1}{\omega}\right)(\mathbf{u},\mathbf{v})\right]\Big|_P &= \sum_{\substack{r,s=1\\ r<s}}^{n=3} -\left[\left(\frac{\partial \overset{1}{\omega}_r}{\partial \Theta^s} - \frac{\partial \overset{1}{\omega}_s}{\partial \Theta^r}\right)\left(\underline{u}^r\,\underline{v}^s - \underline{u}^s\,\underline{v}^r\right)\right]\Big|_P \\
&= -\underbrace{\left[\left(\frac{\partial \overset{1}{\omega}_1}{\partial \Theta^2} - \frac{\partial \overset{1}{\omega}_2}{\partial \Theta^1}\right)\left(\underline{u}^1\,\underline{v}^2 - \underline{u}^2\,\underline{v}^1\right)\right]\Big|_P}_{=\left(\Theta^1\Theta^3 - 0\right)\big|_P\,[(-1)(2)-(1)(1)] = (0.25)[-3] = -0.75} \\
&\quad -\underbrace{\left[\left(\frac{\partial \overset{1}{\omega}_1}{\partial \Theta^3} - \frac{\partial \overset{1}{\omega}_3}{\partial \Theta^1}\right)\left(\underline{u}^1\,\underline{v}^3 - \underline{u}^3\,\underline{v}^1\right)\right]\Big|_P}_{=\left(\Theta^1\Theta^2 - 0\right)\big|_P\,[(-1)(-1)-(1)(1)] = (-0.25)[0] = 0} \\
&\quad -\underbrace{\left[\left(\frac{\partial \overset{1}{\omega}_2}{\partial \Theta^3} - \frac{\partial \overset{1}{\omega}_3}{\partial \Theta^2}\right)\left(\underline{u}^2\,\underline{v}^3 - \underline{u}^3\,\underline{v}^2\right)\right]\Big|_P}_{=(1-2)\,[(1)(-1)-(1)(2)] = (-1)[-3] = +3} \\
&= 0.75 - 0 - 3 = -2.25\ .
\end{aligned}
$$

The exterior derivative of the differential 1-form (8.78) vanishes:

$$
\underbrace{d\left(d\overset{0}{\omega}\right) = \frac{\partial^2 \overset{0}{\omega}}{\partial \Theta^j \partial \Theta^i} d\Theta^j \wedge d\Theta^i = 0}_{\text{since } \partial^2\overset{0}{\omega}/\left(\partial\Theta^j\partial\Theta^i\right) = \partial^2\overset{0}{\omega}/\left(\partial\Theta^i\partial\Theta^j\right) \text{ and } d\Theta^j \wedge d\Theta^i = -d\Theta^i \wedge d\Theta^j,\ \text{see (2.79h)}} \quad . \tag{8.82}
$$

This is a special case of what is known as *Poincaré's Lemma* given in (8.88).

8.2.5.3 Exterior Derivative of Differential 2-Form

Guided by (8.78) and (8.80), the exterior derivative of a differential 2-form, according to (8.41), is given by

$$
\begin{aligned}
d\overset{2}{\omega} &= d\overset{2}{\omega}_{ij} \wedge d\Theta^i \wedge d\Theta^j \\
&= \frac{\partial \overset{2}{\omega}_{ij}}{\partial \Theta^k} d\Theta^k \wedge d\Theta^i \wedge d\Theta^j \\
&= \boxed{\frac{\partial \overset{2}{\omega}_{ij}}{\partial \Theta^k} d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\ .}
\end{aligned} \tag{8.83}
$$

Guided by (8.51)-(8.54), this **differential 3-form** takes the three generic vectors **u**,**v** and **w** and returns the scalar

$$\left(d\overset{2}{\omega}\right)(\mathbf{u},\mathbf{v},\mathbf{w}) = \frac{\partial \overset{2}{\omega}_{ij}}{\partial \Theta^k}\left(d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\right)(\mathbf{u},\mathbf{v},\mathbf{w}) = \Xi\, \varepsilon^{ijk} \frac{\partial \overset{2}{\omega}_{ij}}{\partial \Theta^k} \ . \tag{8.84}$$

As an example, the exterior derivative of

$$\overset{2}{\omega} = \left[-\Theta^2 + \left(\Theta^3\right)^3\right] d\Theta^1 \wedge d\Theta^2 + 2\Theta^1 \left(\Theta^2\right)^2 d\Theta^2 \wedge d\Theta^3 + \left(\Theta^1\right)^2 \Theta^2 d\Theta^3 \wedge d\Theta^1 \ ,$$

is

$$\begin{aligned} d\overset{2}{\omega} &= \left[-d\Theta^2 + 3\left(\Theta^3\right)^2 d\Theta^3\right] \wedge d\Theta^1 \wedge d\Theta^2 \\ &\quad + \left[2\left(\Theta^2\right)^2 d\Theta^1 + 4\Theta^1\Theta^2 d\Theta^2\right] \wedge d\Theta^2 \wedge d\Theta^3 \\ &\quad + \left[2\Theta^1\Theta^2 d\Theta^1 + \left(\Theta^1\right)^2 d\Theta^2\right] \wedge d\Theta^3 \wedge d\Theta^1 \\ &= \left[3\left(\Theta^3\right)^2 + 2\left(\Theta^2\right)^2 + \left(\Theta^1\right)^2\right] d\Theta^1 \wedge d\Theta^2 \wedge d\Theta^3 \ , \end{aligned}$$

whose action on the given vectors

$$\left.\begin{aligned} \mathbf{u} &= 2\mathbf{g}_1 + \mathbf{g}_2 + 2\mathbf{g}_3 \\ \mathbf{v} &= \mathbf{g}_1 - 2\mathbf{g}_2 - \mathbf{g}_3 \\ \mathbf{w} &= 3\mathbf{g}_1 - \mathbf{g}_2 - \mathbf{g}_3 \end{aligned}\right\} \ ,$$

at a point P corresponding to $\left(\Theta^1, \Theta^2, \Theta^3\right) = (0.75, 0.5, 0.25)$ renders

$$\left[\left(d\overset{2}{\omega}\right)(\mathbf{u},\mathbf{v},\mathbf{w})\right]\Big|_P = \underbrace{\left[3\,(0.25)^2 + 2\,(0.5)^2 + (0.75)^2\right]}_{=\,1.25} \underbrace{\det\begin{bmatrix} 2 & 1 & 3 \\ 1 & -2 & -1 \\ 2 & -1 & -1 \end{bmatrix}}_{=\,10} = 12.5 \ .$$

In general, the exterior derivative of a k-form $\overset{k}{\varsigma} = \overset{k}{\varsigma}_{i_1\cdots i_k} d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}$ is given by

$$\begin{aligned} d\overset{k}{\varsigma} &= d\overset{k}{\varsigma}_{i_1\cdots i_k}\left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) \\ &= \frac{\partial \overset{k}{\varsigma}_{i_1\cdots i_k}}{\partial \Theta^m} d\Theta^m \wedge \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) \\ &= (-1)^k \frac{\partial \overset{k}{\varsigma}_{i_1\cdots i_k}}{\partial \Theta^m}\left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) \wedge d\Theta^m \ . \end{aligned} \tag{8.85}$$

In a similar manner,

$$d\overset{l}{\pi} = \frac{\partial \overset{l}{\pi}_{j_1\cdots j_l}}{\partial \Theta^m} d\Theta^m \wedge \left(d\Theta^{j_1} \wedge \cdots \wedge d\Theta^{j_l}\right)$$
$$= (-1)^l \frac{\partial \overset{l}{\pi}_{j_1\cdots j_l}}{\partial \Theta^m} \left(d\Theta^{j_1} \wedge \cdots \wedge d\Theta^{j_l}\right) \wedge d\Theta^m \,. \tag{8.86}$$

It is then a simple exercise to show that

$$\boxed{\underbrace{d\left(\overset{k}{\varsigma} \wedge \overset{l}{\pi}\right) = d\overset{k}{\varsigma} \wedge \overset{l}{\pi} + (-1)^k \overset{k}{\varsigma} \wedge d\overset{l}{\pi}}_{\text{note that (8.77b) is a special case of this general identity}} \,.} \tag{8.87}$$

One can finally extend the identity (8.82) to

$$\underbrace{d\left(d\overset{k}{\varsigma}\right) = \frac{\partial^2 \overset{k}{\varsigma}_{i_1\cdots i_k}}{\partial\Theta^n \partial\Theta^m} d\Theta^n \wedge d\Theta^m \wedge \left(d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k}\right) = 0}_{\text{note that } \partial^2 \overset{k}{\varsigma}_{i_1\cdots i_k} / \left(\partial\Theta^n \partial\Theta^m\right) = \partial^2 \overset{k}{\varsigma}_{i_1\cdots i_k} / \left(\partial\Theta^m \partial\Theta^n\right) \text{ whereas } d\Theta^n \wedge d\Theta^m = -d\Theta^m \wedge d\Theta^n} \quad , \tag{8.88}$$

called the *Poincaré's Lemma*.

8.2.6 Hodge Star Operator in Exterior Calculus

The goal here is to write down the widely used differential operators of vector calculus such as gradient and divergence in the language of exterior calculus. A key feature of the relations developed here is their **coordinate free** representation.

To begin with, consider a vector field $\mathbf{u} \in \mathcal{E}_r^3$ which is associated with a differential 1-form $\overset{1}{\omega}_{\mathbf{u}} \in \bigwedge^1 \mathcal{E}_r^3$ in the following way[6]:

$$\underbrace{\mathbf{u} = \underline{u}^i \mathbf{g}_i \quad \longleftrightarrow \quad \overset{1}{\omega}_{\mathbf{u}} = g_{ij}\, \underline{u}^j d\Theta^i}_{\text{or } \mathbf{u} = g^{ij}\, \underline{u}_j \mathbf{g}_i \quad \longleftrightarrow \quad \overset{1}{\omega}_{\mathbf{u}} = \underline{u}_i d\Theta^i} \,. \tag{8.89}$$

[6] It is important to point out that in differential geometry the basis $\{\mathbf{g}_i\}$ belongs to the so-called *tangent space* while the basis $\{\mathbf{g}^i\}$ is inhabitant of what is known as the *cotangent space*, see the next chapter for more details. With regard to this, the object $\underline{u}^i$ is referred to as a vector and $\underline{u}_i = g_{ij}\, \underline{u}^j$ renders its companion covector.

In a similar manner, $(\text{grad} f)^i = g^{ij}\left(\partial f / \partial\Theta^j\right)$ should be realized as a vector and consequently $(\text{grad} f)_i = \left(\partial f / \partial\Theta^i\right)$ demonstrates its companion covector, see (8.99). One should also consider $(\text{curl}\mathbf{u})^i = -J^{-1} \varepsilon^{ijk} g_{jl}\, \underline{u}^l\big|_k$ as a vector and therefore $(\text{curl}\mathbf{u})_i = -J \varepsilon_{ijk} g^{km}\, \underline{u}^j\big|_m$ presents its dual vector, see (8.107) and (8.109).

Indeed,

$$\boxed{\overset{1}{\omega}_{\mathbf{u}} = \mathbf{u} \cdot d\mathbf{x} \quad \text{where} \quad d\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \Theta^i} d\Theta^i = d\Theta^i \mathbf{g}_i \ .} \tag{8.90}$$

The expressions in (8.89) is usually demonstrated by using the *musical* operators $\flat$ and $\sharp$. In this regard, the **flat operator** $\flat$ acts on a vector field $\mathbf{u}$ to provide its associated differential 1-form $\overset{1}{\omega}_{\mathbf{u}}$, that is,

$$\boxed{\overset{1}{\omega}_{\mathbf{u}} = (\mathbf{u})^{\flat} = \underline{u}_i d\Theta^i \ .} \tag{8.91}$$

In a similar manner, the result of acting the **sharp operator** $\sharp$ on a differential 1-form $\overset{1}{\omega}_{\mathbf{u}}$ is

$$\boxed{\left(\overset{1}{\omega}_{\mathbf{u}}\right)^{\sharp} = \mathbf{u} = \underline{u}^i \mathbf{g}_i \ .} \tag{8.92}$$

Consider now two vectors $\mathbf{u} = \underline{u}^i \mathbf{g}_i$ and $\mathbf{v} = \underline{v}^j \mathbf{g}_j$ whose dot product and cross product are

$$\boxed{\mathbf{u} \cdot \mathbf{v} = \underline{u}^i g_{ij} \underline{v}^j \quad , \quad \mathbf{u} \times \mathbf{v} = (\mathbf{u} \times \mathbf{v})^k \mathbf{g}_k = J \, \underline{u}^i \, \underline{v}^j \varepsilon_{ijm} g^{mk} \mathbf{g}_k \ .} \tag{8.93}$$

In the context of exterior calculus, these **algebraic** relations are represented by

$$\boxed{\mathbf{u} \cdot \mathbf{v} = * \left[\overset{1}{\omega}_{\mathbf{u}} \wedge \left(* \overset{1}{\omega}_{\mathbf{v}} \right) \right] \quad , \quad \mathbf{u} \times \mathbf{v} = \left[* \left(\overset{1}{\omega}_{\mathbf{u}} \wedge \overset{1}{\omega}_{\mathbf{v}} \right) \right]^{\sharp} \ ,} \tag{8.94}$$

because

$$\begin{aligned} \overset{1}{\omega}_{\mathbf{u}} \wedge \left(* \overset{1}{\omega}_{\mathbf{v}} \right) &= \underline{u}_i d\Theta^i \wedge \left(\underline{v}_m * \left(d\Theta^m \right) \right) \\ &= \frac{J}{2} \, \underline{u}_i \, \underline{v}_m g^{mn} \varepsilon_{njk} d\Theta^i \wedge d\Theta^j \wedge d\Theta^k \ , \end{aligned} \tag{8.95a}$$

$$\begin{aligned} * \left[\overset{1}{\omega}_{\mathbf{u}} \wedge \left(* \overset{1}{\omega}_{\mathbf{v}} \right) \right] &= \frac{1}{2} \, \underline{u}_i \, \underline{v}_m g^{mn} \varepsilon_{njk} \varepsilon^{ijk} \\ &= \underline{u}_i \, \underline{v}_m g^{mn} \delta^i_n \\ &= \underline{u}_i \, \underline{v}_m g^{mi} \\ &= \underline{u}^i g_{ij} \, \underline{v}^j \ , \end{aligned} \tag{8.95b}$$

and

$$\overset{1}{\omega}_{\mathbf{u}} \wedge \overset{1}{\omega}_{\mathbf{v}} = \underline{u}_i \, \underline{v}_j d\Theta^i \wedge d\Theta^j \ , \tag{8.96a}$$

$$\begin{aligned} * \left(\overset{1}{\omega}_{\mathbf{u}} \wedge \overset{1}{\omega}_{\mathbf{v}} \right) &= J \, \underline{u}_i \, \underline{v}_j g^{il} g^{jm} \varepsilon_{lmk} d\Theta^k \\ &= J^{-1} \underline{u}_i \, \underline{v}_j \varepsilon^{ijm} g_{mk} d\Theta^k \ , \end{aligned} \tag{8.96b}$$

$$\begin{aligned} \left[* \left(\overset{1}{\omega}_{\mathbf{u}} \wedge \overset{1}{\omega}_{\mathbf{v}} \right) \right]^{\sharp} &= J^{-1} \underline{u}_i \, \underline{v}_j \varepsilon^{ijm} \delta^k_m \mathbf{g}_k \\ &= J^{-1} \underline{u}_i \, \underline{v}_j \varepsilon^{ijk} \mathbf{g}_k \\ &= J \, \underline{u}^i \, \underline{v}^j \varepsilon_{ijm} g^{mk} \mathbf{g}_k \ , \end{aligned} \tag{8.96c}$$

where the expressions (1.58b)$_3$, (5.14), (5.47), (5.53)$_{1-2}$, (5.66a)–(5.66b), (8.59), (8.63b)-(8.63d) and (8.91)-(8.93) have been used.

Consider now the scalar triple product of the three vectors **u**, **v** and **w** according to

$$\boxed{(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = J \varepsilon_{ijk} \, \underline{u}^i \, \underline{v}^j \, \underline{w}^k \ .} \tag{8.97}$$

It is then a simple exercise to see that

$$\boxed{(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = * \left(\overset{1}{\omega}_{\mathbf{u}} \wedge \overset{1}{\omega}_{\mathbf{v}} \wedge \overset{1}{\omega}_{\mathbf{w}} \right) \ .} \tag{8.98}$$

8.2.6.1 Gradient

Recall from (7.68b)$_2$ that the gradient of a scalar function f is

$$\operatorname{grad} f = g^{ij} \frac{\partial f}{\partial \Theta^j} \mathbf{g}_i \ , \tag{8.99}$$

while its exterior derivative, according to (8.78), renders

$$df = \frac{\partial f}{\partial \Theta^i} d\Theta^i \ . \tag{8.100}$$

Thus, the vector field $\operatorname{grad} f$ and its corresponding differential 1-form $\overset{1}{\omega}_{\operatorname{grad} f}$ are related by

$$\boxed{\overset{1}{\omega}_{\operatorname{grad} f} = df = (\operatorname{grad} f)^{\flat} \quad , \quad \operatorname{grad} f = (df)^{\sharp} \ .} \tag{8.101}$$

Indeed, one can establish the identity

$$\boxed{df = \operatorname{grad} f \cdot d\mathbf{x} \quad \text{where} \quad d\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \Theta^i} d\Theta^i = d\Theta^i \mathbf{g}_i \ .} \tag{8.102}$$

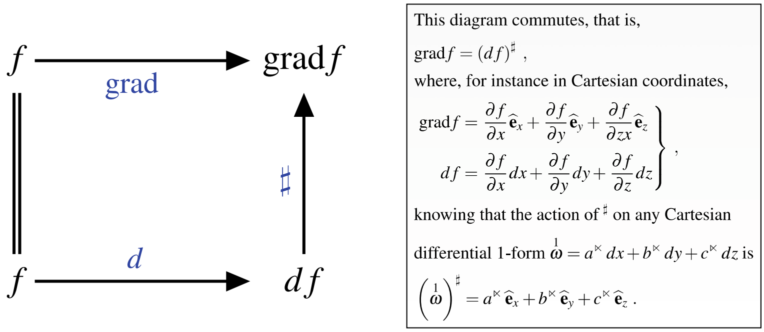


Fig. 8.4 Gradient of a scalar function in the language of exterior calculus

See Fig. 8.4 for a geometrical interpretation. It is important to note that any scalar function f of the local coordinates (or tensors) belongs to both R and $\bigwedge^0 \mathcal{E}^3_{\text{r}}$.

8.2.6.2 Divergence

Consider a vector field $\mathbf{u} = \underline{u}^i \mathbf{g}_i$ whose divergence, according to (7.81), is

$$\text{div}\,\mathbf{u} = \frac{1}{J}\frac{\partial}{\partial \Theta^i}\left[J\,\underline{u}^i\right] .$$

One can then associate this vector field with the differential 1-form $\overset{1}{\omega}_{\mathbf{u}} = \underline{u}_i d\Theta^i$. The goal here is to show that

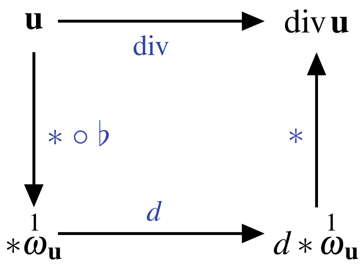
$$\boxed{\text{div}\,\mathbf{u} = *d * \overset{1}{\omega}_{\mathbf{u}}\ .} \tag{8.103}$$

For a geometrical interpretation of this expression, see Fig. 8.5. It can be verified in three steps as follows. First,

$$\begin{aligned} *\overset{1}{\omega}_{\mathbf{u}} &\xlongequal[\text{(8.59) and (8.91)}]{\text{from}} \underline{u}_i * \left(d\Theta^i\right) \\ &\xlongequal[\text{(8.63b)}]{\text{from}} \frac{J}{2}\left(\underline{u}_i g^{il}\right)\varepsilon_{ljk} d\Theta^j \wedge d\Theta^k \\ &\xlongequal[\text{(5.47) and (5.66b)}]{\text{from}} \frac{J}{2}\,\underline{u}^l \varepsilon_{ljk} d\Theta^j \wedge d\Theta^k\ . \end{aligned} \tag{8.104}$$

Then,

$$d * \overset{1}{\omega}_{\mathbf{u}} \xlongequal[\text{(8.83) and (8.104)}]{\text{from}} \frac{1}{2}\frac{\partial}{\partial \Theta^i}\left[J\,\underline{u}^l\right]\varepsilon_{ljk} d\Theta^i \wedge d\Theta^j \wedge d\Theta^k\ . \tag{8.105}$$



This diagram commutes, that is,

$\operatorname{div}\mathbf{u} = *d*\overset{1}{\omega}_{\mathbf{u}}$,

where, for instance in Cartesian coordinates,

$\operatorname{div}\mathbf{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}$,

and, knowing that $\overset{1}{\omega}_{\mathbf{u}} = u_x dx + u_y dy + u_z dz$,

$$\left.\begin{aligned} *\overset{1}{\omega}_{\mathbf{u}} &= u_x dy\wedge dz + u_y dz\wedge dx + u_z dx\wedge dy \\ d*\overset{1}{\omega}_{\mathbf{u}} &= \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}\right) dx\wedge dy\wedge dz \\ *d*\overset{1}{\omega}_{\mathbf{u}} &= \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} \end{aligned}\right\} .$$

Fig. 8.5 Divergence of a vector field in the language of exterior calculus

Finally,

$$\begin{aligned} *d*\overset{1}{\omega}_{\mathbf{u}} &\underset{\text{(8.63d) and (8.105)}}{\overset{\text{from}}{=\!=}} \frac{1}{2J}\frac{\partial}{\partial\Theta^i}\left[J\,\underline{u}^l\right]\varepsilon_{ljk}\varepsilon^{ijk} \\ &\underset{\text{(1.58b)}}{\overset{\text{from}}{=\!=}} \frac{1}{2J}\frac{\partial}{\partial\Theta^i}\left[J\,\underline{u}^l\right]2\delta^i_l \\ &\underset{\text{(5.14)}}{\overset{\text{from}}{=\!=}} \frac{1}{J}\frac{\partial}{\partial\Theta^i}\left[J\,\underline{u}^i\right] . \end{aligned} \tag{8.106}$$

8.2.6.3 Curl

Consider the curl of a vector field **u** according to

$$\operatorname{curl}\mathbf{u} = (\operatorname{curl}\mathbf{u})^i\,\mathbf{g}_i \ , \tag{8.107}$$

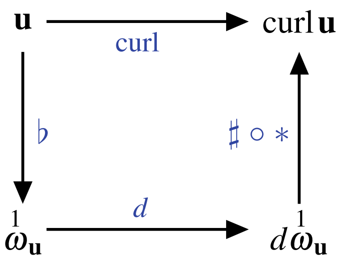
with its corresponding differential 1-form

$$\overset{1}{\omega}_{\operatorname{curl}\mathbf{u}} = (\operatorname{curl}\mathbf{u})^\flat = (\operatorname{curl}\mathbf{u})_i\,d\Theta^i \ , \tag{8.108}$$

where, using (7.96)$_1$ and (7.96)$_9$,

$$\underbrace{(\operatorname{curl}\mathbf{u})^i = -J^{-1}\varepsilon^{ijk}g_{jl}\;\underline{u}^l\big|_k \quad , \quad (\operatorname{curl}\mathbf{u})_i = -J\varepsilon_{ijk}g^{km}\;\underline{u}^j\big|_m}_{\text{note that } \underline{u}^j\big|_m = \partial\,\underline{u}^j/\partial\Theta^m + \Gamma^j_{mr}\,\underline{u}^r\text{, according to (7.25a)}} . \tag{8.109}$$

The object curl**u** from vector calculus can be expressed in the language of exterior calculus via the following expression



This diagram commutes, that is,

$$\text{curl}\,\mathbf{u} = \left(*d\overset{1}{\omega}_{\mathbf{u}}\right)^{\sharp} ,$$

where, for instance in Cartesian coordinates,

$$\text{curl}\,\mathbf{u} = a^{\rtimes}\,\widehat{\mathbf{e}}_x + b^{\rtimes}\,\widehat{\mathbf{e}}_y + c^{\rtimes}\,\widehat{\mathbf{e}}_z ,$$

and, knowing that $\overset{1}{\omega}_{\mathbf{u}} = u_x dx + u_y dy + u_z dz$,

$$\left.\begin{aligned} d\overset{1}{\omega}_{\mathbf{u}} &= a^{\rtimes}\,dy\wedge dz + b^{\rtimes}\,dz\wedge dx + c^{\rtimes}\,dx\wedge dy \\ *d\overset{1}{\omega}_{\mathbf{u}} &= a^{\rtimes}\,dx + b^{\rtimes}\,dy + c^{\rtimes}\,dz \\ \left(*d\overset{1}{\omega}_{\mathbf{u}}\right)^{\sharp} &= a^{\rtimes}\,\widehat{\mathbf{e}}_x + b^{\rtimes}\,\widehat{\mathbf{e}}_y + c^{\rtimes}\,\widehat{\mathbf{e}}_z \end{aligned}\right\} .$$

In these expressions,

$$a^{\rtimes} = \frac{\partial u_z}{\partial y} - \frac{\partial u_y}{\partial z} ,\; b^{\rtimes} = \frac{\partial u_x}{\partial z} - \frac{\partial u_z}{\partial x} ,\; c^{\rtimes} = \frac{\partial u_y}{\partial x} - \frac{\partial u_x}{\partial y} .$$

Fig. 8.6 Curl of a vector field in the language of exterior calculus

$$\boxed{\underbrace{\text{curl}\,\mathbf{u} = \left(*d\overset{1}{\omega}_{\mathbf{u}}\right)^{\sharp}}_{\text{or } \overset{1}{\omega}_{\text{curl}\,\mathbf{u}} = (\text{curl}\,\mathbf{u})^{\flat} = *d\overset{1}{\omega}_{\mathbf{u}}}} \quad . \tag{8.110}$$

This result, shown schematically in Fig. 8.6, can be verified in three steps as follows. First,

$$d\overset{1}{\omega}_{\mathbf{u}} \xlongequal[\text{(8.80) and (8.91)}]{\text{from}} \frac{\partial\,\underline{u}_n}{\partial\Theta^m} d\Theta^m \wedge d\Theta^n \; . \tag{8.111}$$

Then,

$$\begin{aligned} *d\overset{1}{\omega}_{\mathbf{u}} &\xlongequal[\text{(8.59) and (8.111)}]{\text{from}} \frac{\partial\,\underline{u}_n}{\partial\Theta^m} * \left(d\Theta^m \wedge d\Theta^n\right) \\ &\xlongequal[\text{(5.66a) and (8.63c)}]{\text{from}} \frac{\partial\left(g_{nr}\,\underline{u}^r\right)}{\partial\Theta^m} J g^{mk} g^{nj} \varepsilon_{kji} d\Theta^i \\ &\xlongequal[\text{of differentiation and (7.13)}]{\text{from the product rule}} \left[\Gamma_{nmr}\,\underline{u}^r + \Gamma_{rmn}\,\underline{u}^r + g_{nr}\frac{\partial\,\underline{u}^r}{\partial\Theta^m}\right] J g^{mk} g^{nj} \varepsilon_{kji} d\Theta^i \\ &\xlongequal[\text{(5.39), (5.51) and (7.10)}]{\text{from}} 0 + J\Gamma^j_{rm}\,\underline{u}^r g^{mk} \varepsilon_{kji} d\Theta^i + J\frac{\partial\,\underline{u}^r}{\partial\Theta^m}\delta^j_r g^{mk} \varepsilon_{kji} d\Theta^i \\ &\xlongequal[\text{(1.54), (5.14) and (7.7)}]{\text{from}} -J\varepsilon_{ijk} g^{mk}\left(\frac{\partial\,\underline{u}^j}{\partial\Theta^m} + \Gamma^j_{mr}\,\underline{u}^r\right) d\Theta^i \\ &\xlongequal[\text{(5.47) and (7.25a)}]{\text{from}} -J\varepsilon_{ijk} g^{km}\;\; \underline{u}^j\big|_m \, d\Theta^i \; , \end{aligned} \tag{8.112}$$

where the identity

$$\Gamma_{nmr} g^{mk} g^{nj} \varepsilon_{kji} = 0 \; , \tag{8.113}$$

has been used owing to

$$\begin{aligned} \Gamma_{nmr} g^{mk} g^{nj} \varepsilon_{kji} &\xlongequal[\text{the dummy indices}]{\text{by renaming}} \Gamma_{mnr} g^{nk} g^{mj} \varepsilon_{kji} \\ &\xlongequal[\text{the dummy indices}]{\text{by renaming}} \Gamma_{mnr} g^{nj} g^{mk} \varepsilon_{jki} \\ &\xlongequal[\text{(1.54) and (7.11)}]{\text{from}} -\Gamma_{nmr} g^{nj} g^{mk} \varepsilon_{kji} \ . \end{aligned}$$

Finally,

$$\begin{aligned} \left(* d \overset{1}{\omega}_{\mathbf{u}} \right)^{\sharp} &\xlongequal[\text{(8.92) and (8.112)}]{\text{from}} -J \left(g^{is} \varepsilon_{sjk} g^{km} \right) \ \underline{u}^j \big|_m \, \mathbf{g}_i \\ &\xlongequal[\text{(5.47) and (5.53)}]{\text{from}} -J \left(J^{-2} \varepsilon^{itm} g_{tj} \right) \ \underline{u}^j \big|_m \, \mathbf{g}_i \\ &\xlongequal[\text{the dummy indices}]{\text{by renaming}} -J^{-1} \varepsilon^{ijk} g_{jl} \ \underline{u}^l \big|_k \, \mathbf{g}_i \ . \end{aligned} \tag{8.114}$$

It is not then difficult to see that

$$\boxed{ * d \overset{1}{\omega}_{\mathbf{u}} = \text{curl}\mathbf{u} \cdot d\mathbf{x} \quad \text{where} \quad d\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \Theta^i} d\Theta^i = d\Theta^i \mathbf{g}_i \ . } \tag{8.115}$$

It is worthwhile to point out that the identities $\text{div}\,\text{curl}\mathbf{u} = 0$ and $\text{curl}\,\text{grad} f = \mathbf{0}$, respectively given in (7.123b) and (7.123c), can also be inferred by using the powerful tool of exterior calculus, see Figs. 8.7 and 8.8. Indeed, $\text{div}\,\text{curl}\mathbf{u} = 0$ corresponds to $dd\overset{1}{\omega}_{\mathbf{u}} = 0$ because

$$\begin{aligned} \text{div}\,\text{curl}\mathbf{u} &\xlongequal[\text{(8.103)}]{\text{from}} *d * \overset{1}{\omega}_{\text{curl}\mathbf{u}} \\ &\xlongequal[\text{(8.108)}]{\text{from}} *d * (\text{curl}\mathbf{u})^{\flat} \end{aligned}$$

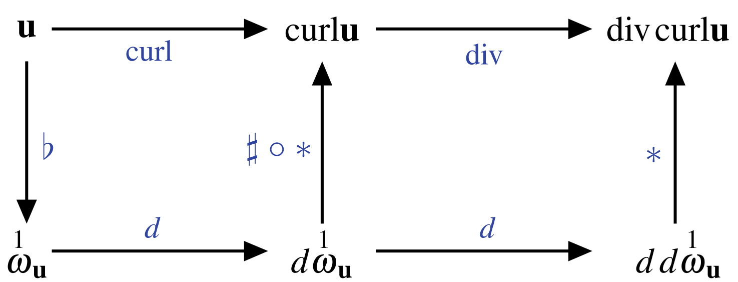


Fig. 8.7 Divergence of curl of a vector field in the language of exterior calculus

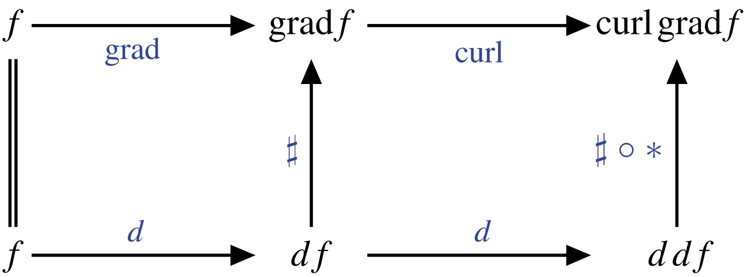


Fig. 8.8 Curl of gradient of a scalar field in the language of exterior calculus

$$
\begin{aligned}
&\xlongequal[\text{(8.110)}]{\text{from}} *d**d\overset{1}{\omega}_{\mathbf{u}} \\
&\xlongequal[\text{light of (8.62)}]{\text{in}} *dd\overset{1}{\omega}_{\mathbf{u}} \\
&\xlongequal[\text{light of (8.88)}]{\text{in}} *0 \\
&\xlongequal{\text{trivially}} 0 \,. \qquad (8.116)
\end{aligned}
$$

And, curl grad $f = \mathbf{0}$ corresponds to $ddf = 0$ since

$$
\begin{aligned}
\operatorname{curl}\operatorname{grad} f &\xlongequal[\text{(8.110)}]{\text{from}} \left(*d\overset{1}{\omega}_{\operatorname{grad} f}\right)^{\sharp} \\
&\xlongequal[\text{(8.101)}]{\text{from}} (*ddf)^{\sharp} \\
&\xlongequal[\text{(8.82)}]{\text{from}} (*0)^{\sharp} \\
&\xlongequal{\text{trivially}} \mathbf{0} \,. \qquad (8.117)
\end{aligned}
$$

8.2.6.4 Laplacian

The Laplacian of a scalar function f, according to (7.109a), is given by

$$
\nabla^2 f = g^{ij}\left[\frac{\partial^2 f}{\partial\Theta^i\,\partial\Theta^j} - \Gamma^k_{ij}\frac{\partial f}{\partial\Theta^k}\right] .
$$

As schematically illustrated in Fig. 8.9, the object $\nabla^2 f$ in the language of exterior calculus represents

$$
\boxed{\nabla^2 f = *d*df \ .} \qquad (8.118)
$$

This can be verified in three steps as follows. First,

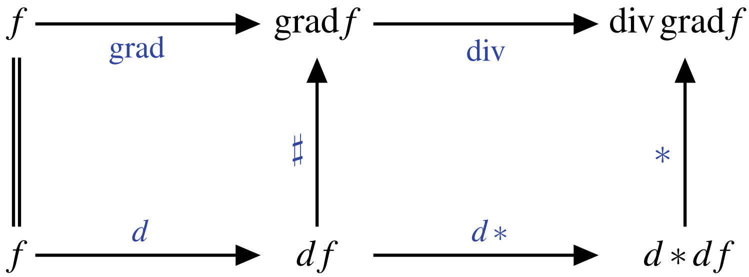


Fig. 8.9 Laplacian of a scalar field in the language of exterior calculus

$$
\begin{aligned}
*df &\xlongequal[\text{(8.59) and (8.78)}]{\text{from}} \frac{\partial f}{\partial \Theta^m} * \left(d\Theta^m\right) \\
&\xlongequal[\text{(8.63b)}]{\text{from}} \frac{J}{2} \frac{\partial f}{\partial \Theta^m} g^{mn} \varepsilon_{njk} d\Theta^j \wedge d\Theta^k \ .
\end{aligned}
\tag{8.119}
$$

Then,

$$
d * df \xlongequal[\text{(8.83) and (8.119)}]{\text{from}} \frac{1}{2} \frac{\partial}{\partial \Theta^i} \left[J \frac{\partial f}{\partial \Theta^m} g^{mn} \right] \varepsilon_{njk} d\Theta^i \wedge d\Theta^j \wedge d\Theta^k \ . \tag{8.120}
$$

Finally,

$$
\begin{aligned}
*d * df &\xlongequal[\text{(8.59) and (8.120)}]{\text{from}} \frac{1}{2} \frac{\partial}{\partial \Theta^i} \left[J \frac{\partial f}{\partial \Theta^m} g^{mn} \right] \varepsilon_{njk} * \left(d\Theta^i \wedge d\Theta^j \wedge d\Theta^k \right) \\
&\xlongequal[\text{(8.63d)}]{\text{from}} \frac{1}{2J} \frac{\partial}{\partial \Theta^i} \left[J \frac{\partial f}{\partial \Theta^m} g^{mn} \right] \varepsilon_{njk} \varepsilon^{ijk} \\
&\xlongequal[\text{of differentiation and (1.58b)}]{\text{from the product rule}} \left[\frac{\partial J}{J \partial \Theta^i} \frac{\partial f}{\partial \Theta^m} g^{mn} + \frac{\partial^2 f}{\partial \Theta^i \partial \Theta^m} g^{mn} + \frac{\partial f}{\partial \Theta^m} \frac{\partial g^{mn}}{\partial \Theta^i} \right] \delta^i_n \\
&\xlongequal[\text{and (7.16)-(7.17)}]{\text{from (5.14)}} \cancel{\Gamma^k_{ki} \frac{\partial f}{\partial \Theta^m} g^{mi}} + \frac{\partial^2 f}{\partial \Theta^i \partial \Theta^m} g^{mi} - \Gamma^m_{il} \frac{\partial f}{\partial \Theta^m} g^{li} - \cancel{\Gamma^i_{il} \frac{\partial f}{\partial \Theta^m} g^{ml}} \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} g^{ij} \left[\frac{\partial^2 f}{\partial \Theta^i \partial \Theta^j} - \Gamma^k_{ij} \frac{\partial f}{\partial \Theta^k} \right] .
\end{aligned}
\tag{8.121}
$$

8.2.7 Differential Forms Integration

The basic idea here is to integrate a differential k-form $\overset{k}{\omega}$ over an oriented domain Ω_k which is basically a k-dimensional subspace of the n-dimensional space. The goal is thus to consistently define the differential form integral

$$\int_{\Omega_k} \overset{k}{\omega} = \int_{\Omega_k} \overset{k}{\omega}_{i_1 \cdots i_k} \left(\Theta^1, \ldots, \Theta^n\right) d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k} \ .$$

This can be done by using the parametrization

$$\left.\begin{array}{l} \Theta^1 = \Theta^1 \left(t^1, \ldots, t^k\right) \\ \vdots \qquad\quad \vdots \\ \Theta^n = \Theta^n \left(t^1, \ldots, t^k\right) \end{array}\right\} , \tag{8.122}$$

which helps write

$$\mathbf{x} = \hat{\mathbf{x}} \left(\Theta^1, \ldots, \Theta^n\right) = \bar{\mathbf{x}} \left(t^1, \ldots, t^k\right) \ , \tag{8.123}$$

and, accordingly,

$$\left.\begin{array}{l} \mathbf{a}_1 = \dfrac{\partial \bar{\mathbf{x}}}{\partial t^1} = \displaystyle\sum_{s=1}^{n} \dfrac{\partial \Theta^s}{\partial t^1} \dfrac{\partial \hat{\mathbf{x}}}{\partial \Theta^s} = \displaystyle\sum_{s=1}^{n} \dfrac{\partial \Theta^s}{\partial t^1} \mathbf{g}_s \\ \vdots \qquad\qquad \vdots \\ \mathbf{a}_k = \dfrac{\partial \bar{\mathbf{x}}}{\partial t^k} = \displaystyle\sum_{s=1}^{n} \dfrac{\partial \Theta^s}{\partial t^k} \dfrac{\partial \hat{\mathbf{x}}}{\partial \Theta^s} = \displaystyle\sum_{s=1}^{n} \dfrac{\partial \Theta^s}{\partial t^k} \mathbf{g}_s \end{array}\right\} . \tag{8.124}$$

Hint: It is assumed that (8.122) is an orientation-preserving parametrization. Orientation of the object Ω_k along with its consistent parametrization is an enrich topic which is well studied within the context of differential geometry of manifolds.

Now, the integral of $\overset{k}{\omega}$ over Ω_k is defined by

$$\int_{\Omega_k} \overset{k}{\omega} = \int dt^1 \cdots \int dt^k \overset{k}{\omega}_{i_1 \cdots i_k} \left(\Theta^1, \ldots, \Theta^n\right) d\Theta^{i_1} \wedge \cdots \wedge d\Theta^{i_k} \left(\mathbf{a}_1, \ldots, \mathbf{a}_k\right) \ . \tag{8.125}$$

Using $(8.32)_4$, (8.57) and (8.124), this takes the form

$$\boxed{\int_{\Omega_k} \overset{k}{\omega} = \int dt^1 \cdots \int dt^k \overset{k}{\omega}_{i_1 \cdots i_k} \left(\Theta^1, \ldots, \Theta^n\right) \det \begin{bmatrix} \frac{\partial \Theta^{i_1}}{\partial t^1} & \cdots & \frac{\partial \Theta^{i_1}}{\partial t^k} \\ \vdots & \vdots & \vdots \\ \frac{\partial \Theta^{i_k}}{\partial t^1} & \cdots & \frac{\partial \Theta^{i_k}}{\partial t^k} \end{bmatrix}} \ . \tag{8.126}$$

In the following, some examples are provided in order to illustrate this expression.

The first example regards a 1-form $\overset{1}{\omega} = y dx + x dy + dz$ to be integrated over a helix parametrized by

$$\mathbf{x} = \bar{\mathbf{x}}\,(t) = \cos t\,\widehat{\mathbf{e}}_x + \sin t\,\widehat{\mathbf{e}}_y + t\,\widehat{\mathbf{e}}_z \;, \quad 0 \le t \le 2\pi \;, \qquad \leftarrow \text{see Fig. 9.17}$$

and oriented by the tangent vector $[\,\mathbf{t}\,(x, y, z)\,] = \begin{bmatrix} -y & x & 1 \end{bmatrix}^{\mathrm{T}}$. The tangent vector to this **space curve**,

$$\mathbf{a}_t = \frac{d\bar{\mathbf{x}}}{dt} = -\sin t\,\widehat{\mathbf{e}}_x + \cos t\,\widehat{\mathbf{e}}_y + \widehat{\mathbf{e}}_z \;,$$

helps obtain the integrand

$$\begin{aligned} \overset{1}{\omega}\,(\mathbf{a}_t) &= y dx\,(\mathbf{a}_t) + x dy\,(\mathbf{a}_t) + dz\,(\mathbf{a}_t) \\ &= \sin t\,(-\sin t) + \cos t\,(\cos t) + 1 \\ &= 2 - 2\sin^2 t \;, \end{aligned}$$

and, consequently,

$$\int_{\Omega_1} \overset{1}{\omega} = \int_0^{2\pi} \left(2 - 2\sin^2 t\right) dt = 2\pi \;.$$

It is important to note that the above parametrization is **orientation-preserving** due to

$$\mathbf{t} \cdot \mathbf{a}_t > 0 \;.$$

As another example, consider a 2-form $\overset{2}{\omega} = -dx \wedge (2dy + dz)$ to be integrated over the parametrized surface

$$\begin{aligned} \mathbf{x} = \bar{\mathbf{x}}\left(t^1, t^2\right) &= 2t^2\cos t^1\,\widehat{\mathbf{e}}_x + 2t^2\sin t^1\,\widehat{\mathbf{e}}_y \\ &\quad + \left[\left(2t^2\cos t^1\right)^2 + 2\left(2t^2\sin t^1\right)^2\right]\widehat{\mathbf{e}}_z \;, \quad 0 \le t^1 \le 2\pi \;,\; 0 \le t^2 \le 1 \;, \end{aligned}$$

with the unit normal oriented outwards (see Figs. 8.2 and 9.8). By using the tangent vectors to this **curved surface**,

$$\begin{aligned} \mathbf{a}_1 &= \frac{\partial\bar{\mathbf{x}}}{\partial t^1} = -2t^2\sin t^1\,\widehat{\mathbf{e}}_x + 2t^2\cos t^1\,\widehat{\mathbf{e}}_y + 8\left(t^2\right)^2\cos t^1\sin t^1\,\widehat{\mathbf{e}}_z \;, \\ \mathbf{a}_2 &= \frac{\partial\bar{\mathbf{x}}}{\partial t^2} = 2\cos t^1\,\widehat{\mathbf{e}}_x + 2\sin t^1\,\widehat{\mathbf{e}}_y + 8t^2\left(1 + \sin^2 t^1\right)\widehat{\mathbf{e}}_z \;, \end{aligned}$$

one can evaluate the integrand

$$\overset{2}{\omega}(\mathbf{a}_1, \mathbf{a}_2) = -2dx \wedge dy(\mathbf{a}_1, \mathbf{a}_2) - dx \wedge dz(\mathbf{a}_1, \mathbf{a}_2)$$
$$= 8t^2 - \left[-32\left(t^2\right)^2 \sin t^1\right]$$
$$= 8t^2 + 32\left(t^2\right)^2 \sin t^1 ,$$

and, accordingly,

$$\int_{\Omega_2} \overset{2}{\omega} = \int_0^{2\pi} dt^1 \left\{ \int_0^1 \left[8t^2 + 32\left(t^2\right)^2 \sin t^1\right] dt^2 \right\} = 8\pi .$$

Notice that the above parametrization is **orientation-preserving** owing to

$$\det\left[\,\widehat{\mathbf{n}} \;\; \mathbf{a}_1 \;\; \mathbf{a}_2\,\right] > 0 \quad \text{where} \quad \widehat{\mathbf{n}} = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{|\mathbf{a}_1 \times \mathbf{a}_2|} .$$

The last example here regards a 3-form $\overset{3}{\omega} = 2z^2 dx \wedge dy \wedge dz$ to be integrated over a three-dimensional region bounded by the solid torus

$$\mathbf{x} = \bar{\mathbf{x}}\left(t^1, t^2, t^3\right) = \left(2 + t^1 \cos t^3\right) \cos t^2 \,\widehat{\mathbf{e}}_x$$
$$+ \left(2 + t^1 \cos t^3\right) \sin t^2 \,\widehat{\mathbf{e}}_y$$
$$+ t^1 \sin t^3 \,\widehat{\mathbf{e}}_z , \quad 0 \le t^1 \le 1 , \; 0 \le t^2, t^3 \le 2\pi ,$$

← see Fig. 9.3

with the standard orientation. Introducing the tangent vectors

$$\mathbf{a}_1 = \frac{\partial \bar{\mathbf{x}}}{\partial t^1} = \cos t^2 \cos t^3 \,\widehat{\mathbf{e}}_x + \sin t^2 \cos t^3 \,\widehat{\mathbf{e}}_y + \sin t^3 \,\widehat{\mathbf{e}}_z ,$$
$$\mathbf{a}_2 = \frac{\partial \bar{\mathbf{x}}}{\partial t^2} = -\left(2 + t^1 \cos t^3\right) \sin t^2 \,\widehat{\mathbf{e}}_x + \left(2 + t^1 \cos t^3\right) \cos t^2 \,\widehat{\mathbf{e}}_y ,$$
$$\mathbf{a}_3 = \frac{\partial \bar{\mathbf{x}}}{\partial t^2} = -t^1 \cos t^2 \sin t^3 \,\widehat{\mathbf{e}}_x - t^1 \sin t^2 \sin t^3 \,\widehat{\mathbf{e}}_y + t^1 \cos t^3 \,\widehat{\mathbf{e}}_z ,$$

in the integrand

$$\overset{3}{\omega}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_2) = 2z^2 dx \wedge dy \wedge dz(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$$
$$= 2\left[t^1 \sin t^3\right]^2 \det\left[\,\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \mathbf{a}_3\,\right]$$
$$= 2\left(t^1\right)^3 \left(2 + t^1 \cos t^3\right) \left(\sin t^3\right)^2 ,$$

helps compute

$$\int_{\Omega_3} \overset{3}{\omega} = \int_0^1 dt^1 \left(\int_0^{2\pi} dt^2 \left\{ \int_0^{2\pi} \left[2 \left(t^1\right)^3 \left(2 + t^1 \cos t^3\right) \left(\sin t^3\right)^2 \right] dt^3 \right\} \right)$$
$$= \boxed{+2\pi^2 \ .}$$

Note that the introduced parametrization **preserves orientation** because

$$\det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} > 0 \ .$$

Consider again that solid torus with parametrization of the form

$$\begin{aligned} \mathbf{x} = \bar{\mathbf{x}} \left(t^1, t^2, t^3\right) = & \left(2 + t^1 \cos t^2\right) \cos t^3 \, \widehat{\mathbf{e}}_x \\ & + \left(2 + t^1 \cos t^2\right) \sin t^3 \, \widehat{\mathbf{e}}_y \\ & + t^1 \sin t^2 \, \widehat{\mathbf{e}}_z \ , \quad 0 \le t^1 \le 1 \ , \ 0 \le t^2, t^3 \le 2\pi \ , \end{aligned}$$

for which

$$\begin{aligned} \mathbf{a}_1 &= \frac{\partial \bar{\mathbf{x}}}{\partial t^1} = \cos t^2 \cos t^3 \, \widehat{\mathbf{e}}_x + \cos t^2 \sin t^3 \, \widehat{\mathbf{e}}_y + \sin t^2 \, \widehat{\mathbf{e}}_z \ , \\ \mathbf{a}_2 &= \frac{\partial \bar{\mathbf{x}}}{\partial t^2} = -t^1 \sin t^2 \cos t^3 \, \widehat{\mathbf{e}}_x - t^1 \sin t^2 \sin t^3 \, \widehat{\mathbf{e}}_y + t^1 \cos t^2 \, \widehat{\mathbf{e}}_z \ , \\ \mathbf{a}_3 &= \frac{\partial \bar{\mathbf{x}}}{\partial t^3} = - \left(2 + t^1 \cos t^2\right) \sin t^3 \, \widehat{\mathbf{e}}_x + \left(2 + t^1 \cos t^2\right) \cos t^3 \, \widehat{\mathbf{e}}_y \ . \end{aligned}$$

Now,

$$\begin{aligned} \overset{3}{\omega} \left(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_2\right) &= 2z^2 dx \wedge dy \wedge dz \left(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\right) \\ &= 2 \left[t^1 \sin t^2\right]^2 \det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \\ &= -2 \left(t^1\right)^3 \left(2 + t^1 \cos t^2\right) \left(\sin t^2\right)^2 \ , \end{aligned}$$

and, consequently,

$$\int_{\Omega_3} \overset{3}{\omega} = \int_0^1 dt^1 \left(\int_0^{2\pi} dt^2 \left\{ \int_0^{2\pi} \left[-2 \left(t^1\right)^3 \left(2 + t^1 \cos t^2\right) \left(\sin t^2\right)^2 \right] dt^3 \right\} \right)$$
$$= \boxed{-2\pi^2 \ .}$$

Notice that the above parametrization is **orientation-reversing**, that is,

$$\det \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} < 0 \ .$$

This example shows that the parametrization of any object in this context should be chosen with care. The results $+2\pi^2$ and $-2\pi^2$ here are in alignment with the basic property

$$\int_a^b f(x)dx = -\int_b^a f(x)dx \ ,$$

of the definite integral in one dimension.

8.2.8 Generalized Stokes' Theorem

The generalized Stokes' theorem needs to be recognized as one of the most profound theorem in modern mathematics. This powerful theorem manifests itself as the beauty, elegance and utility of mathematics. It is basically the essence of calculus which holds true for any dimension of interest. The fundamental theorem of calculus in one dimension, Green's theorem in two dimensions and divergence theorem in three dimensions are just the well-known examples of this sophisticated theorem in lower dimensions.

The generalized Stokes' theorem reads

$$\boxed{\int_{\Omega_k} d\overset{k-1}{\omega} = \int_{\partial\Omega_k} \overset{k-1}{\omega} \ ,} \tag{8.127}$$

(the proof can be found, e.g., in Spivak [2]) where $\partial\Omega_k$ denotes the boundary of Ω_k. Note that d and ∂ are opposite to one another. This means that removing the exterior derivative is the same as taking the boundary and removing the boundary amounts to taking the exterior derivative. One should also note that d is an operator which characterizes how an object changes **locally** while ∂ is a **global** operator which identifies the exterior of a shape.

In the following, it will be demonstrated that how the well-known integral theorems in a three-dimensional space can be written in the unified format of the generalized Stokes' theorem.

To begin with, consider a space curve $\mathcal{C}$ parametrized by

$$\mathbf{x} = \hat{\mathbf{x}}\left(\Theta^1(t), \Theta^2(t), \Theta^3(t)\right) = \mathbf{x}^{\mathrm{c}}(t) \ , \quad a \le t \le b \ . \quad \leftarrow \text{see (9.278)}$$

Consequently, the tangent vector to this curve takes the form

$$\mathbf{a}_t = \frac{d\mathbf{x}^{\mathrm{c}}}{dt} = \frac{d\Theta^j}{dt}\mathbf{g}_j \ .$$

The **gradient theorem for line integrals** is now given by

$$\underbrace{\int_{\mathcal{C}} \operatorname{grad} f \cdot d\mathbf{x} = f(b) - f(a)}_{\text{this is an extension of the \textbf{fundamental theorem of calculus} to any plane or space curve}} \quad , \tag{8.128}$$

or

$$\underbrace{\int_{\mathcal{C}} \operatorname{grad} f \cdot d\mathbf{x}}_{\text{i.e. sum of infinitesimal changes on inside}} = \underbrace{\int_{\partial\mathcal{C}} f}_{\text{i.e. total change on outside}} \quad . \tag{8.129}$$

Let $\mathbf{u} = \underline{u}^i \mathbf{g}_i$ be a vector and $\overset{1}{\omega}_{\mathbf{u}} = \underline{u}^i g_{ij} d\Theta^j$ be its corresponding covector. Then, the component of $\mathbf{u}$ in the direction of $\mathbf{a}_t$ renders

$$\mathbf{u} \cdot \mathbf{a}_t = \left(\underline{u}^i \mathbf{g}_i\right) \cdot \left(\frac{d\Theta^j}{dt} \mathbf{g}_j\right) = \underline{u}^i g_{ij} \frac{d\Theta^j}{dt} \ .$$

As a result, one can establish the identity

$$\begin{aligned} \int_{\mathcal{C}} \mathbf{u} \cdot d\mathbf{x} = \int_{\mathcal{C}} \mathbf{u} \cdot \mathbf{a}_t dt = \int_{\mathcal{C}} \underline{u}^i g_{ij} \frac{d\Theta^j}{dt} dt = \int_{\mathcal{C}} \underline{u}^i g_{ij} d\Theta^j \\ = \int_{\mathcal{C}} \overset{1}{\omega}_{\mathbf{u}} \ , \end{aligned} \tag{8.130}$$

in alignment with (8.90). Let $\mathbf{u} = \operatorname{grad} f$. From $(8.101)_1$ and $(8.130)_4$, it then follows that

$$\int_{\mathcal{C}} \operatorname{grad} f \cdot d\mathbf{x} = \int_{\mathcal{C}} \overset{1}{\omega}_{\operatorname{grad} f} = \int_{\mathcal{C}} df \ . \tag{8.131}$$

Using (8.129) and $(8.131)_2$, one immediately obtains

$$\int_{\mathcal{C}} df = \int_{\partial\mathcal{C}} f \ , \tag{8.132}$$

or, by setting $\mathcal{C} = \Omega_1$ and $f = \overset{0}{\omega}$,

$$\boxed{\int_{\Omega_1} d\overset{0}{\omega} = \int_{\partial\Omega_1} \overset{0}{\omega} \ .} \tag{8.133}$$

Next, suppose one is given a parametrized surface A according to

$$\mathbf{x} = \hat{\mathbf{x}}\left(\Theta^1\left(t^1, t^2\right), \Theta^2\left(t^1, t^2\right), \Theta^3\left(t^1, t^2\right)\right) = \hat{\mathbf{x}}^s\left(t^1, t^2\right) , \quad \leftarrow \text{see (9.1)}$$

with the following tangent vectors

$$\mathbf{a}_1 = \frac{\partial \hat{\mathbf{x}}^s}{\partial t^1} = \frac{\partial \Theta^j}{\partial t^1}\mathbf{g}_j \quad , \quad \mathbf{a}_2 = \frac{\partial \hat{\mathbf{x}}^s}{\partial t^2} = \frac{\partial \Theta^k}{\partial t^2}\mathbf{g}_k ,$$

and the surface element

$$\begin{aligned} \widehat{\mathbf{n}}\, dA &= (\mathbf{a}_1 \times \mathbf{a}_2)\, dt^1 dt^2 \quad \leftarrow \text{see (9.56) and (9.57)} \\ &= \left(\mathbf{g}_j \times \mathbf{g}_k\right) \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} dt^1 dt^2 \\ &= J \varepsilon_{jkl} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} dt^1 dt^2 \mathbf{g}^l . \quad \leftarrow \text{see (5.33)} \end{aligned}$$

Let $\mathbf{u} = \underline{u}^i \mathbf{g}_i$ be a vector corresponding to the covector $\overset{1}{\omega}_{\mathbf{u}} = \underline{u}^i g_{ij} d\Theta^j$. Then, introducing $\mathbf{a}_1$ and $\mathbf{a}_2$ into $*\overset{1}{\omega}_{\mathbf{u}}$ yields

$$\begin{aligned} *\overset{1}{\omega}_{\mathbf{u}}(\mathbf{a}_1, \mathbf{a}_2) &= \underbrace{\frac{1}{2} J\, \underline{u}^i \varepsilon_{ijk} d\Theta^j \wedge d\Theta^k (\mathbf{a}_1, \mathbf{a}_2)}_{= \frac{1}{2} J\, \underline{u}^i \varepsilon_{ijk} \left(\frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} - \frac{\partial \Theta^k}{\partial t^1} \frac{\partial \Theta^j}{\partial t^2} \right) = \frac{1}{2} J\, \underline{u}^i \varepsilon_{ijk} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} - \frac{1}{2} J\, \underline{u}^i \varepsilon_{ikj} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2}} \\ &= J\, \underline{u}^i \varepsilon_{ijk} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} , \end{aligned}$$

where (1.54), (8.41)-(8.42) and $(8.104)_3$ have been used. With the aid of (8.125), the integral of this differential 2-form takes the following form

$$\int_A *\overset{1}{\omega}_{\mathbf{u}} = \int \int J\, \underline{u}^i \varepsilon_{ijk} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} dt^1 dt^2 .$$

And this helps establish

$$\underbrace{\int_A \mathbf{u} \cdot \widehat{\mathbf{n}}\, dA = \int_A *\overset{1}{\omega}_{\mathbf{u}}}_{\text{knowing that } \mathbf{u} \cdot \widehat{\mathbf{n}}\, dA = \left(\underline{u}^i \mathbf{g}_i\right) \cdot \left(J \varepsilon_{jkl} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} dt^1 dt^2 \mathbf{g}^l\right) = J\, \underline{u}^i \varepsilon_{ijk} \frac{\partial \Theta^j}{\partial t^1} \frac{\partial \Theta^k}{\partial t^2} dt^1 dt^2} . \tag{8.134}$$

Let $\mathbf{u} = \text{curl}\mathbf{v}$. Making use of (8.65), $(8.110)_3$ and $(8.134)_1$, one then obtains

$$\int_A \text{curl}\mathbf{v} \cdot \widehat{\mathbf{n}} \, dA = \int_A * \overset{1}{\omega}_{\text{curl}\mathbf{v}}$$
$$= \int_A * * \, d\overset{1}{\omega}_{\mathbf{v}}$$
$$= \int_A d\overset{1}{\omega}_{\mathbf{v}} \,. \tag{8.135}$$

By virtue of $(8.130)_4$ and $(8.135)_3$, the Stokes' theorem $(8.18)_1$ can now be rewritten as

$$\int_A d\overset{1}{\omega}_{\mathbf{v}} = \int_A \text{curl}\mathbf{v} \cdot \widehat{\mathbf{n}} \, dA$$
$$= \int_{\partial A} \mathbf{v} \cdot d\mathbf{x}$$
$$= \int_{\partial A} \overset{1}{\omega}_{\mathbf{v}} \,, \tag{8.136}$$

or, by setting $A = \Omega_2$ and $\overset{1}{\omega}_{\mathbf{v}} = \overset{1}{\omega}$,

$$\boxed{\int_{\Omega_2} d\overset{1}{\omega} = \int_{\partial\Omega_2} \overset{1}{\omega} \,.} \tag{8.137}$$

At the end, the goal is to translate the divergence theorem $(8.1)_1$ in the language of differential forms. Using $(5.113)_2$, $(8.63a)_2$, (8.65), (8.103) and (8.126), one can arrive at

$$\int_V \text{div}\mathbf{u} \, dV = \int_V * \text{div}\mathbf{u} = \int_V d * \overset{1}{\omega}_{\mathbf{u}} \,. \tag{8.138}$$

By means of $(8.134)_1$ and $(8.138)_2$, the divergence theorem $\int_V \text{div}\mathbf{u} \, dV = \int_A \mathbf{u} \cdot \widehat{\mathbf{n}} \, dA$ can then be rephrased as

$$\int_V d * \overset{1}{\omega}_{\mathbf{u}} = \int_{\partial V} * \overset{1}{\omega}_{\mathbf{u}} \,, \tag{8.139}$$

or, by setting $V = \Omega_3$ and $*\overset{1}{\omega}_{\mathbf{u}} = \overset{2}{\omega}$,

$$\boxed{\int_{\Omega_3} d\overset{2}{\omega} = \int_{\partial\Omega_3} \overset{2}{\omega} \, .} \tag{8.140}$$

8.3 Exercises

Exercise 8.1

Show that the integral over a closed surface A of a surface vector $d\mathbf{A}$ vanishes, that is,

$$\int_A d\mathbf{A} = \mathbf{0} \, . \tag{8.141}$$

Solution. First, this important relation will be verified for an infinitesimal tetrahedron as illustrated in Fig. 8.10. Following a simple argument then reveals that this result also holds true for an arbitrary closed surface. According to Fig. 8.10 and guided by (5.115b), the surface vector $d\mathbf{A} = \widehat{\mathbf{n}}\, dA$ acting along the inclined plane renders

$$\widehat{\mathbf{n}}\, dA = \frac{1}{2} d\mathbf{x}^{\bullet} \times d\mathbf{x}^{\star} \, ,$$

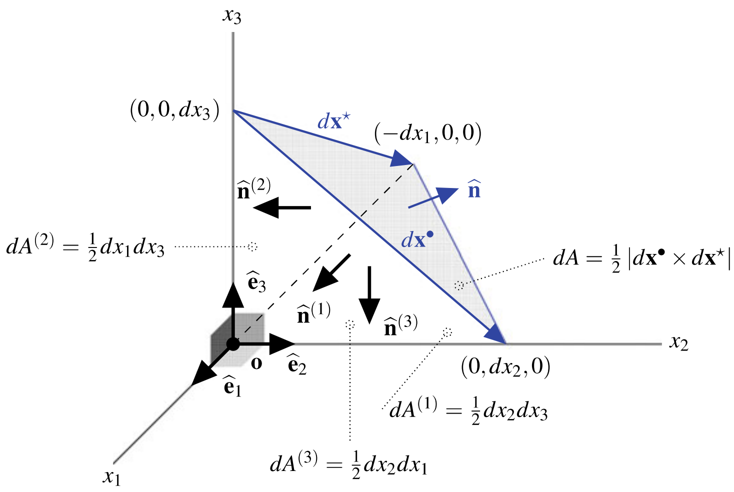


Fig. 8.10 Infinitesimal tetrahedron with edge lengths dx_1, dx_2, dx_3 and outwardly oriented unit vectors on its faces in a right-handed Cartesian coordinate frame

and its value equals the half of the area of the parallelogram constructed by the infinitesimal vectors $d\mathbf{x}^{\bullet}$ and $d\mathbf{x}^{\star}$. Having in mind the bilinearity of the cross product, substituting $d\mathbf{x}^{\bullet} = -dx_3\widehat{\mathbf{e}}_3 + dx_2\widehat{\mathbf{e}}_2$ and $d\mathbf{x}^{\star} = -dx_3\widehat{\mathbf{e}}_3 - dx_1\widehat{\mathbf{e}}_1$ into this relation provides

$$\widehat{\mathbf{n}}\, dA = \underbrace{\frac{1}{2}dx_1dx_3}_{= dA^{(2)}}\widehat{\mathbf{e}}_2 - \underbrace{\frac{1}{2}dx_2dx_3}_{= dA^{(1)}}\widehat{\mathbf{e}}_1 + \underbrace{\frac{1}{2}dx_2dx_1}_{= dA^{(3)}}\widehat{\mathbf{e}}_3 \ ,$$

where the expressions (1.52) and (1.64) have been used. Finally, considering $\widehat{\mathbf{n}}^{(1)} = +\widehat{\mathbf{e}}_1$, $\widehat{\mathbf{n}}^{(2)} = -\widehat{\mathbf{e}}_2$, $\widehat{\mathbf{n}}^{(3)} = -\widehat{\mathbf{e}}_3$ in the above relation yields

$$\widehat{\mathbf{n}}\, dA + dA^{(1)}\widehat{\mathbf{n}}^{(1)} + dA^{(2)}\widehat{\mathbf{n}}^{(2)} + dA^{(3)}\widehat{\mathbf{n}}^{(3)} = \mathbf{0} \ .$$

This result states that the sum of all surface vectors over the entire surface of an infinitesimal tetrahedron equals the zero vector.

Given a closed surface A. Suppose that the region enclosed by A is split into infinitesimal tetrahedrons. The fact that the contribution of all inner surfaces that are shared between infinitesimal volume elements is canceled during integration implies (8.141).

Notice that by choosing $\phi = 1$ in (8.7), one can simply arrive at the same result.

Exercise 8.2

Let ϕ, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{A}$ respectively be smooth scalar, vector, vector and tenor fields defined on a closed surface A bounding a domain V. Further, let $\widehat{\mathbf{n}}$ be the outward unit normal field acting along that surface. Verify that

$$\int_A \phi\mathbf{u} \cdot \widehat{\mathbf{n}}\, dA = \int_V \operatorname{div}(\phi\mathbf{u})\, dV \ , \tag{8.142a}$$

$$\int_A \mathbf{v}\,(\mathbf{u} \cdot \widehat{\mathbf{n}})\ dA = \int_V \operatorname{div}(\mathbf{v} \otimes \mathbf{u})\, dV \ , \tag{8.142b}$$

$$\int_A \phi\mathbf{A}\widehat{\mathbf{n}}\, dA = \int_V \operatorname{div}(\phi\mathbf{A})\, dV \ , \tag{8.142c}$$

$$\int_A \widehat{\mathbf{n}} \times \mathbf{u}\, dA = \int_V \operatorname{curl}\mathbf{u}\, dV \ , \tag{8.142d}$$

$$\int_A \widehat{\mathbf{n}} \cdot \operatorname{curl}\mathbf{u}\, dA = 0 \ , \tag{8.142e}$$

$$\int_A \mathbf{u} \otimes \widehat{\mathbf{n}}\, dA = \int_V \text{grad}\mathbf{u}\, dV \ , \tag{8.142f}$$

$$\int_A \mathbf{u} \cdot \mathbf{A}\widehat{\mathbf{n}}\, dA = \int_V \text{div}\left(\mathbf{A}^{\mathrm{T}}\mathbf{u}\right) dV \ , \tag{8.142g}$$

$$\int_A \mathbf{A}\widehat{\mathbf{n}} \otimes \mathbf{u}\, dA = \int_V \left[\text{div}\mathbf{A} \otimes \mathbf{u} + \mathbf{A}\text{grad}^{\mathrm{T}}\mathbf{u}\right] dV \ , \tag{8.142h}$$

$$\int_A \mathbf{u} \times \mathbf{A}\widehat{\mathbf{n}}\, dA = \int_V \left[\mathsf{E} : (\text{grad}\mathbf{u})\, \mathbf{A}^{\mathrm{T}} + \mathbf{u} \times \text{div}\mathbf{A}\right] dV \ , \tag{8.142i}$$

where $\mathsf{E} = \varepsilon_{ijk}\widehat{\mathbf{e}}_i \otimes \widehat{\mathbf{e}}_j \otimes \widehat{\mathbf{e}}_k$ in the last relation is the permutation tensor introduced in (3.17).

Solution. In this exercise, all desired identities will be verified in indicial notation. Notice that the integral theorems, introduced in (8.1), are entirely **coordinate free**. Thus, the Cartesian form of components will be utilized here for convenience. However, the interested reader can prove any of these expressions by means of the curvilinear forms of components.

The expression (8.142a): Using $(1.38)_7$, (7.78) and $(8.1)_2$,

$$\begin{aligned}
\int_A \phi\mathbf{u} \cdot \widehat{\mathbf{n}}\, dA &= \int_A \phi u_i \widehat{n}_i\, dA \\
&= \int_V \frac{\partial\left(\phi u_i\right)}{\partial x_i} dV \\
&= \int_V \text{div}\left(\phi\mathbf{u}\right) dV \ .
\end{aligned}$$

The expression (8.142b): Using $(1.38)_7$, $(2.24)_4$, $(7.83)_2$ and $(8.8)_2$,

$$\begin{aligned}
\int_A \left(\mathbf{v}\left(\mathbf{u} \cdot \widehat{\mathbf{n}}\right)\right)_i\, dA &= \int_A v_i \left(u_j \widehat{n}_j\right)\, dA \\
&= \int_V \frac{\partial\left(v_i u_j\right)}{\partial x_j} dV \\
&= \int_V \frac{\partial\left(\mathbf{v} \otimes \mathbf{u}\right)_{ij}}{\partial x_j} dV \\
&= \int_V \left(\text{div}\left(\mathbf{v} \otimes \mathbf{u}\right)\right)_i dV \ .
\end{aligned}$$

The expression (8.142c): Using $(2.22)_3$, $(7.83)_2$ and $(8.8)_2$,

$$\begin{aligned}\int_A (\phi\mathbf{A}\widehat{\mathbf{n}})_i\; dA &= \int_A \phi A_{ij}\widehat{n}_j\, dA \\ &= \int_V \frac{\partial\left(\phi A_{ij}\right)}{\partial x_j} dV \\ &= \int_V (\mathrm{div}\,(\phi\mathbf{A}))_i\, dV\ .\end{aligned}$$

The expression (8.142d): Using (1.54), $(1.67)_6$, (7.91) and $(8.1)_2$,

$$\begin{aligned}\int_A (\widehat{\mathbf{n}}\times\mathbf{u})_i\, dA &= \int_A \varepsilon_{ijk}\widehat{n}_j u_k\, dA \\ &= \int_V \frac{\partial\left(\varepsilon_{ijk}u_k\right)}{\partial x_j}\, dV \\ &= \int_V -\varepsilon_{ikj}\frac{\partial u_k}{\partial x_j}\, dV \\ &= \int_V (\mathrm{curl}\mathbf{u})_i\, dV\ .\end{aligned}$$

The expression (8.142e): Using $(1.38)_7$, (1.54), (7.91) and $(8.1)_2$,

$$\begin{aligned}\int_A \widehat{\mathbf{n}}\cdot\mathrm{curl}\,\mathbf{u}\, dA &= \int_A \widehat{n}_i\,(\mathrm{curl}\mathbf{u})_i\; dA \\ &= \int_A \left(-\varepsilon_{ijk}\frac{\partial u_j}{\partial x_k}\right)\widehat{n}_i\, dA \\ &= \int_V \frac{\partial}{\partial x_i}\left(-\varepsilon_{ijk}\frac{\partial u_j}{\partial x_k}\right)\, dV \\ &= \int_V \varepsilon_{jik}\frac{\partial^2 u_j}{\partial x_i\partial x_k}\, dV \\ &= 0\ ,\end{aligned}$$

taking into consideration the fact that $\varepsilon_{jik} = -\varepsilon_{jki}$ and that $\partial^2 u_j/\partial x_i\partial x_k = \partial^2 u_j/\partial x_k\partial x_i$, see (2.79h).
The expression (8.142f): Using $(2.24)_4$, $(7.70)_2$ and $(8.7)_2$,

$$\int_A (\mathbf{u} \otimes \widehat{\mathbf{n}})_{ij} \, dA = \int_A u_i \widehat{n}_j \, dA$$
$$= \int_V \frac{\partial u_i}{\partial x_j} \, dV$$
$$= \int_V (\mathrm{grad}\mathbf{u})_{ij} \, dV \, .$$

The expression (8.142g): Using $(1.38)_7$, $(2.22)_3$, (2.49), (7.78) and $(8.1)_2$,

$$\int_A \mathbf{u} \cdot \mathbf{A}\widehat{\mathbf{n}} \, dA = \int_A (\mathbf{u})_i \, (\mathbf{A}\widehat{\mathbf{n}})_i \, dA$$
$$= \int_A u_i A_{ij} \widehat{n}_j \, dA$$
$$= \int_V \frac{\partial \left(u_i A_{ij}\right)}{\partial x_j} \, dV$$
$$= \int_V \frac{\partial \left(A^{\mathrm{T}}_{ji} u_i\right)}{\partial x_j} \, dV$$
$$= \int_V \frac{\partial \left(\mathbf{A}^{\mathrm{T}} \mathbf{u}\right)_j}{\partial x_j} \, dV$$
$$= \int_V \mathrm{div} \left(\mathbf{A}^{\mathrm{T}} \mathbf{u}\right) dV \, .$$

The expression (8.142h): Using $(2.22)_3$, $(2.24)_4$, (2.26), (2.49), $(7.70)_2$, $(7.83)_2$ and $(8.8)_2$ along with the product rule of differentiation,

$$\int_A (\mathbf{A}\widehat{\mathbf{n}} \otimes \mathbf{u})_{ij} \, dA = \int_A (\mathbf{A}\widehat{\mathbf{n}})_i \, (\mathbf{u})_j \, dA$$
$$= \int_A A_{ik} \widehat{n}_k u_j \, dA$$
$$= \int_V \frac{\partial \left(A_{ik} u_j\right)}{\partial x_k} \, dV$$
$$= \int_V \left[\frac{\partial A_{ik}}{\partial x_k} u_j + A_{ik} \frac{\partial u_j}{\partial x_k} \right] dV$$

$$= \int_V \left[(\text{div}\mathbf{A})_i \, (\mathbf{u})_j + (\mathbf{A})_{ik} \, (\text{grad}\mathbf{u})_{jk} \right] dV$$

$$= \int_V \left(\text{div}\mathbf{A} \otimes \mathbf{u} + \mathbf{A}\text{grad}^{\text{T}}\mathbf{u} \right)_{ij} dV \; .$$

The expression (8.142i): Using $(1.67)_6$, $(2.22)_3$, (2.26), (2.49), (3.16b), (3.17), $(7.70)_2$, $(7.83)_2$ and $(8.8)_2$ along with the product rule of differentiation,

$$\begin{aligned}
\int_A (\mathbf{u} \times \mathbf{A}\widehat{\mathbf{n}})_i \, dA &= \int_A \varepsilon_{ijk} \, (\mathbf{u})_j \, (\mathbf{A}\widehat{\mathbf{n}})_k \, dA \\
&= \int_A \varepsilon_{ijk} u_j A_{kl} \widehat{n}_l \, dA \\
&= \int_V \frac{\partial \left(\varepsilon_{ijk} u_j A_{kl} \right)}{\partial x_l} \, dV \\
&= \int_V \left[\varepsilon_{ijk} \frac{\partial u_j}{\partial x_l} A_{kl} + \varepsilon_{ijk} u_j \frac{\partial A_{kl}}{\partial x_l} \right] dV \\
&= \int_V \left[\varepsilon_{ijk} \, (\text{grad}\mathbf{u})_{jl} \left(\mathbf{A}^{\text{T}} \right)_{lk} + \varepsilon_{ijk} \, (\mathbf{u})_j \, (\text{div}\mathbf{A})_k \right] dV \\
&= \int_V \left(\mathbf{E} : (\text{grad}\mathbf{u}) \, \mathbf{A}^{\text{T}} + \mathbf{u} \times \text{div}\mathbf{A} \right)_i dV \; .
\end{aligned}$$

Exercise 8.3

Let A be an **open** surface bounded by the closed curve $\mathcal{C}$ and positively oriented with the outward unit normal field $\widehat{\mathbf{n}}$. Further, let $\mathbf{u}$, $\mathbf{v}$ be two smooth vector fields and $\mathbf{A}$ be a smooth tensor field. Then, use the Stokes' theorem to verify the following identities

$$\int_{\mathcal{C}} (\mathbf{u} \cdot \mathbf{v}) \, d\mathbf{x} = \int_A \left[\widehat{\mathbf{n}} \times \left(\text{grad}^{\text{T}}\mathbf{u} \right) \mathbf{v} + \widehat{\mathbf{n}} \times \left(\text{grad}^{\text{T}}\mathbf{v} \right) \mathbf{u} \right] dA \; , \tag{8.143a}$$

$$\int_{\mathcal{C}} (\mathbf{u} \otimes \mathbf{v}) \, d\mathbf{x} = \int_A \left[(\text{grad}\mathbf{u}) \, (\mathbf{v} \times \widehat{\mathbf{n}}) + (\widehat{\mathbf{n}} \cdot \text{curl}\mathbf{v}) \, \mathbf{u} \right] dA \; , \tag{8.143b}$$

$$\int_{\mathcal{C}} \mathbf{u} \times d\mathbf{x} = \int_A \left[(\text{div}\mathbf{u}) \, \widehat{\mathbf{n}} - \left(\text{grad}^{\text{T}}\mathbf{u} \right) \widehat{\mathbf{n}} \right] dA \; , \tag{8.143c}$$

$$\int_{\mathcal{C}} (\widehat{\mathbf{n}} \times \mathbf{u}) \cdot d\mathbf{x} = \int_{A} \left[(\mathbf{I} - \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}) : \text{grad}\mathbf{u} - (\mathbf{u} \cdot \widehat{\mathbf{n}}) \left(\mathbf{I} : \text{grad}\widehat{\mathbf{n}} \right) \right] dA \ . \qquad (8.143\text{d})$$

Next, let A be a **closed** surface without any boundary curve. Then, use (8.143d) to verify that

$$\int_{A} \bar{\kappa} \, \widehat{\mathbf{n}} \, dA = \mathbf{0} \ , \qquad (8.144\text{a})$$

$$\int_{A} \bar{\kappa} \ (\bar{\mathbf{r}} \cdot \widehat{\mathbf{n}}) \, dA = 2A \ , \qquad (8.144\text{b})$$

$$\int_{A} \bar{\kappa} \ (\bar{\mathbf{r}} \times \widehat{\mathbf{n}}) \, dA = \mathbf{0} \ , \qquad (8.144\text{c})$$

$$\int_{A} \bar{\kappa} \ (\bar{\mathbf{r}} \cdot \bar{\mathbf{r}}) \, \widehat{\mathbf{n}} \, dA = \int_{A} 2 \left[\mathbf{I} - \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} \right] \bar{\mathbf{r}} \, dA \ , \qquad (8.144\text{d})$$

$$\int_{A} \bar{\kappa} \ (\bar{\mathbf{r}} \cdot \widehat{\mathbf{n}}) \, \bar{\mathbf{r}} \, dA = \int_{A} \left[3\mathbf{I} - \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} \right] \bar{\mathbf{r}} \, dA \ , \qquad (8.144\text{e})$$

where

$$\bar{\kappa} = \text{div}\, \widehat{\mathbf{n}} \quad , \quad \bar{\mathbf{r}} = \mathbf{x} - \mathbf{o} \ . \qquad (8.145)$$

Solution. Here, similarly to the previous exercise, the identities (8.143a)-(8.143d) are verified in indicial notation using the Cartesian components of tensorial field variables for simplicity. It will be shown that the results (8.144a)-(8.144e) can be deduced from (8.143d) by taking $\mathbf{u}$ as some appropriate linear functions of $\widehat{\mathbf{n}}$ when the surface is closed.

The expression (8.143a): By means of $(1.38)_7$, $(1.67)_6$, $(2.22)_3$, (2.49), $(7.70)_2$ and $(8.23)_2$ along with the product rule of differentiation,

$$\begin{aligned}
\int_{\mathcal{C}} (\mathbf{u} \cdot \mathbf{v}) \, (d\mathbf{x})_i &= \int_{\mathcal{C}} u_l v_l dx_i \\
&= \int_{A} \varepsilon_{ijk} \, \widehat{n}_j \, \frac{\partial \left(u_l v_l \right)}{\partial x_k} dA \\
&= \int_{A} \left[\varepsilon_{ijk} \, (\widehat{\mathbf{n}})_j \, (\text{grad}\mathbf{u})_{lk} \, (\mathbf{v})_l + \varepsilon_{ijk} \, (\widehat{\mathbf{n}})_j \, (\mathbf{u})_l \, (\text{grad}\mathbf{v})_{lk} \right] dA \\
&= \int_{A} \left(\widehat{\mathbf{n}} \times \left[\text{grad}^{\text{T}} \mathbf{u} \right] \mathbf{v} + \widehat{\mathbf{n}} \times \left[\text{grad}^{\text{T}} \mathbf{v} \right] \mathbf{u} \right)_i dA \ .
\end{aligned}$$

The expression (8.143b): By means of (1.54), $(1.67)_6$, $(2.22)_3$, $(2.24)_4$, $(7.70)_2$, (7.91) and $(8.24)_2$ along with the product rule of differentiation,

$$\begin{aligned}
\int_{\mathcal{C}} (\mathbf{u} \otimes \mathbf{v})_{ij} \, (d\mathbf{x})_j &= \int_{\mathcal{C}} u_i v_j dx_j \\
&= \int_A \varepsilon_{jlk} \frac{\partial \left(u_i v_k\right)}{\partial x_l} \widehat{n}_j \, dA \\
&= \int_A \left[\frac{\partial u_i}{\partial x_l} \left(\varepsilon_{lkj} v_k \widehat{n}_j\right) + \widehat{n}_j \left(-\varepsilon_{jkl} \frac{\partial v_k}{\partial x_l}\right) u_i \right] dA \\
&= \int_A \left\{ (\mathrm{grad}\mathbf{u})_{il} \, (\mathbf{v} \times \widehat{\mathbf{n}})_l + \left[(\widehat{\mathbf{n}})_j \, (\mathrm{curl}\mathbf{v})_j \right] (\mathbf{u})_i \right\} dA \ .
\end{aligned}$$

The expression (8.143c): By means of (1.36), (1.58a), $(1.67)_6$, $(2.22)_3$, (2.49), $(7.70)_2$, (7.78) and $(8.24)_2$,

$$\begin{aligned}
\int_{\mathcal{C}} (\mathbf{u} \times d\mathbf{x})_i &= \int_{\mathcal{C}} \varepsilon_{imj} u_m dx_j \\
&= \int_A \varepsilon_{jlk} \frac{\partial \left(\varepsilon_{imk} u_m\right)}{\partial x_l} \widehat{n}_j \, dA \\
&= \int_A \left[\left(\delta_{ji}\delta_{lm} - \delta_{jm}\delta_{li}\right) \frac{\partial u_m}{\partial x_l} \widehat{n}_j \right] dA \\
&= \int_A \left[\frac{\partial u_m}{\partial x_m} \widehat{n}_i - \frac{\partial u_j}{\partial x_i} \widehat{n}_j \right] dA \\
&= \int_A \left[(\mathrm{div}\mathbf{u}) \, (\widehat{\mathbf{n}})_i - (\mathrm{grad}\mathbf{u})_{ji} \, (\widehat{\mathbf{n}})_j \right] dA \\
&= \int_A \left[(\mathrm{div}\mathbf{u}) \, (\widehat{\mathbf{n}})_i - \left(\mathrm{grad}^{\mathrm{T}}\mathbf{u}\right)_{ij} (\widehat{\mathbf{n}})_j \right] dA \ .
\end{aligned}$$

The expression (8.143d): By means of $(1.38)_7$, $(2.24)_4$, $(2.75)_4$, $(7.77)_1$, $(7.122\mathrm{k})_1$ and $(8.18)_1$,

$$\begin{aligned}
\int_{\mathcal{C}} (\widehat{\mathbf{n}} \times \mathbf{u}) \cdot d\mathbf{x} &= \int_{A} \operatorname{curl}(\widehat{\mathbf{n}} \times \mathbf{u}) \cdot \widehat{\mathbf{n}}\, dA \\
&= \int_{A} (\operatorname{curl}(\widehat{\mathbf{n}} \times \mathbf{u}))_i\, (\widehat{\mathbf{n}})_i\, dA \\
&= \int_{A} \big[(\widehat{\mathbf{n}})_i\, (\widehat{\mathbf{n}})_i \operatorname{div}\mathbf{u} - (\mathbf{u})_i\, (\widehat{\mathbf{n}})_i \operatorname{div}\widehat{\mathbf{n}} \\
&\quad + (\operatorname{grad}\widehat{\mathbf{n}})_{ij}\, (\mathbf{u})_j\, (\widehat{\mathbf{n}})_i - (\operatorname{grad}\mathbf{u})_{ij}\, (\widehat{\mathbf{n}})_j\, (\widehat{\mathbf{n}})_i \big]\, dA \\
&= \int_{A} \big[(\mathbf{I} - \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}})_{ij}\, (\operatorname{grad}\mathbf{u})_{ij} - (\mathbf{u})_k\, (\widehat{\mathbf{n}})_k\, (\mathbf{I})_{ij}\, (\operatorname{grad}\widehat{\mathbf{n}})_{ij} \big]\, dA \ ,
\end{aligned}$$

taking into consideration $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$ and, consequently,

$$\operatorname{grad}(\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}}) = 2\widehat{\mathbf{n}}\,(\operatorname{grad}\widehat{\mathbf{n}}) = \mathbf{0} \ \Rightarrow \ \boxed{\widehat{\mathbf{n}}\,(\operatorname{grad}\widehat{\mathbf{n}}) = \mathbf{0} \quad \text{or} \quad \left(\operatorname{grad}^{\mathrm{T}}\widehat{\mathbf{n}}\right)\widehat{\mathbf{n}} = \mathbf{0}\,.} \tag{8.146}$$

The expressions (8.144a)-(8.144e): Let $\mathbf{u}$ be a linear function of $\widehat{\mathbf{n}}$ according to $\mathbf{u} = \widehat{\mathbf{n}} \times \mathbf{w}$. The Stokes' theorem (8.18) then takes the following form

$$\int_{\mathcal{C}} (\widehat{\mathbf{n}} \times \mathbf{w}) \cdot d\mathbf{x} = \int_{A} \operatorname{curl}(\widehat{\mathbf{n}} \times \mathbf{w}) \cdot \widehat{\mathbf{n}}\, dA \ . \tag{8.147}$$

Consider $(\widehat{\mathbf{n}} \times \mathbf{w}) \cdot d\mathbf{x} = \mathbf{w} \cdot (d\mathbf{x} \times \widehat{\mathbf{n}})$ according to $(1.73)_2$. Then, using (8.143d), the above expression can be rewritten as

$$\int_{\mathcal{C}} \mathbf{w} \cdot (d\mathbf{x} \times \widehat{\mathbf{n}}) = \int_{A} \big[(\mathbf{I} - \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}) : \operatorname{grad}\mathbf{w} - (\mathbf{w} \cdot \widehat{\mathbf{n}})\, (\mathbf{I} : \operatorname{grad}\widehat{\mathbf{n}}) \big]\, dA \ . \tag{8.148}$$

Now, let $\mathbf{w}$ be an arbitrary constant vector $\mathbf{c}$ with $\operatorname{grad}\mathbf{c} = \mathbf{O}$. For a closed surface, one then arrives at the desired result (8.144a), knowing that $\mathbf{I} : \operatorname{grad}\widehat{\mathbf{n}} = \operatorname{div}\widehat{\mathbf{n}}$ and $(\mathbf{I} - \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}) : \mathbf{O} = 0$.

Finally, the desired identities (8.144b)-(8.144e) are followed from (8.148) in a straightforward manner by choosing

$$\mathbf{w} = \bar{\mathbf{r}} \quad , \quad \mathbf{w} = \mathbf{c} \times \bar{\mathbf{r}} \quad , \quad \mathbf{w} = (\bar{\mathbf{r}} \cdot \bar{\mathbf{r}})\, \mathbf{c} \quad , \quad \mathbf{w} = (\bar{\mathbf{r}} \otimes \bar{\mathbf{r}})\, \mathbf{c} \ ,$$

respectively.

Exercise 8.4

Verify (8.73c), (8.74a), (8.75b) and (8.76).

Solution. First, consider (8.69) and (8.70). Then, in (8.60), consider

$$J = \sqrt{\left|\det\left[\, g_{\mu\nu}\,\right]\right|} = +1 \quad , \quad \varepsilon_{0123} = +1 \ ,$$

and

$$\left[\, g^{\mu\nu}\,\right] = \left[\, g_{\mu\nu}\,\right]^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} . \tag{8.149}$$

The desired relations are now verified step by step in the following.

The expression (8.73c):

$$\begin{aligned}
*dx^2 &= \frac{g^{22}\varepsilon_{2ijk}}{(4-1)!} dx^i \wedge dx^j \wedge dx^k \\
&= -\frac{1}{6}\varepsilon_{2013} dx^0 \wedge dx^1 \wedge dx^3 - \frac{1}{6}\varepsilon_{2031} dx^0 \wedge dx^3 \wedge dx^1 - \frac{1}{6}\varepsilon_{2103} dx^1 \wedge dx^0 \wedge dx^3 \\
&\quad - \frac{1}{6}\varepsilon_{2130} dx^1 \wedge dx^3 \wedge dx^0 - \frac{1}{6}\varepsilon_{2301} dx^3 \wedge dx^0 \wedge dx^1 - \frac{1}{6}\varepsilon_{2310} dx^3 \wedge dx^1 \wedge dx^0 \\
&= -\frac{1}{6}\varepsilon_{0123} dx^0 \wedge dx^1 \wedge dx^3 + \frac{1}{6}\varepsilon_{0123} dx^0 \wedge dx^3 \wedge dx^1 + \frac{1}{6}\varepsilon_{0123} dx^1 \wedge dx^0 \wedge dx^3 \\
&\quad - \frac{1}{6}\varepsilon_{0123} dx^1 \wedge dx^3 \wedge dx^0 - \frac{1}{6}\varepsilon_{0123} dx^3 \wedge dx^0 \wedge dx^1 + \frac{1}{6}\varepsilon_{0123} dx^3 \wedge dx^1 \wedge dx^0 \\
&= dx^0 \wedge dx^3 \wedge dx^1 \ .
\end{aligned}$$

The expression (8.74a):

$$\begin{aligned}
*\left(dx^0 \wedge dx^1\right) &= \frac{g^{00}g^{11}}{(4-2)!}\varepsilon_{01ij} dx^i \wedge dx^j \\
&= -\frac{1}{2}\varepsilon_{0123} dx^2 \wedge dx^3 - \frac{1}{2}\varepsilon_{0132} dx^3 \wedge dx^2 \\
&= -\frac{1}{2}\varepsilon_{0123} dx^2 \wedge dx^3 + \frac{1}{2}\varepsilon_{0132} dx^2 \wedge dx^3 \\
&= -\frac{1}{2}\varepsilon_{0123} dx^2 \wedge dx^3 - \frac{1}{2}\varepsilon_{0123} dx^2 \wedge dx^3 \\
&= -dx^2 \wedge dx^3 \ .
\end{aligned}$$

The expression (8.75b):

$$*\left(dx^0 \wedge dx^2 \wedge dx^3\right) = \frac{g^{00}g^{22}g^{33}}{(4-3)!}\varepsilon_{023i} dx^i = \varepsilon_{0231} dx^1 = \varepsilon_{0123} dx^1 = dx^1 \ .$$

The expression (8.76):

$$*\left(dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3\right) = \frac{g^{00} g^{11} g^{22} g^{33}}{(4-4)!} \varepsilon_{0123} = -1 \ .$$

At the end, note that

$$*(1) = \frac{1}{(4-0)!} \varepsilon_{ijkl} dx^i \wedge dx^j \wedge dx^k \wedge dx^l = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3 \ .$$

Exercise 8.5

Let **u** be a smooth vector field and f be a smooth scalar field. Then, derive div **u**, curl**u** and grad f, $\nabla^2 f$ in **cylindrical** and **spherical coordinates** using the powerful tool of exterior calculus.

Solution. To solve this exercise, a vector field along with its corresponding 1-form needs to be defined consistently for each coordinate system (having in mind that a scalar field is basically a 0-form). Then, the Hodge star operator $*$ and the exterior derivative d should appropriately be applied. Note that the results obtained in terms of differential forms should finally be translated to the language of vector calculus. The procedure will be shown step by step in the following.

*Divergence of a vector field in **cylindrical** coordinates:* By virtue of the relations (5.7a)-(5.7c), (5.64a), (5.116), (8.46), (8.48), (8.59), (8.67b), (8.67d), (8.83), (8.89) and (8.103),

$$\begin{aligned} \mathbf{u} &= \underline{u}^r \widehat{\mathbf{e}}_r + \underline{u}^\theta \widehat{\mathbf{e}}_\theta + \underline{u}^z \widehat{\mathbf{e}}_z \\ &= \underline{u}^r \mathbf{g}_r + \frac{\underline{u}^\theta}{r} \mathbf{g}_\theta + \underline{u}^z \mathbf{g}_z \ , \end{aligned} \tag{8.150a}$$

$$\overset{1}{\omega}_{\mathbf{u}} = \underline{u}^r dr + r\, \underline{u}^\theta d\theta + \underline{u}^z dz \ , \tag{8.150b}$$

$$*\overset{1}{\omega}_{\mathbf{u}} = r\, \underline{u}^r d\theta \wedge dz + \underline{u}^\theta dz \wedge dr + r\, \underline{u}^z dr \wedge d\theta \ , \tag{8.150c}$$

$$d * \overset{1}{\omega}_{\mathbf{u}} = \left[\frac{\partial \left(r\, \underline{u}^r\right)}{\partial r} + \frac{\partial\, \underline{u}^\theta}{\partial \theta} + \frac{\partial \left(r\, \underline{u}^z\right)}{\partial z} \right] dr \wedge d\theta \wedge dz \ , \tag{8.150d}$$

$$*d * \overset{1}{\omega}_{\mathbf{u}} = \frac{\partial \left(r\, \underline{u}^r\right)}{r \partial r} + \frac{\partial\, \underline{u}^\theta}{r \partial \theta} + \frac{\partial\, \underline{u}^z}{\partial z} = \boxed{\operatorname{div} \mathbf{u} \ .} \tag{8.150e}$$

*Curl of a vector field in **cylindrical** coordinates:* By means of the expressions (5.7a)–(5.7c), (5.117), (8.38a), (8.38c), (8.59), (8.67c), (8.80), (8.92), (8.110) and (8.150b),

$$d\overset{1}{\omega}_{\mathbf{u}} = \frac{\partial \underline{u}^r}{\partial \theta} d\theta \wedge dr + \frac{\partial \underline{u}^r}{\partial z} dz \wedge dr + \frac{\partial \left(r\, \underline{u}^\theta\right)}{\partial r} dr \wedge d\theta + \frac{\partial \left(r\, \underline{u}^\theta\right)}{\partial z} dz \wedge d\theta + \frac{\partial \underline{u}^z}{\partial r} dr \wedge dz + \frac{\partial \underline{u}^z}{\partial \theta} d\theta \wedge dz \,, \tag{8.151a}$$

$$*d\overset{1}{\omega}_{\mathbf{u}} = \left[\frac{\partial \underline{u}^z}{r\partial \theta} - \frac{\partial \underline{u}^\theta}{\partial z}\right] dr + \left[r \frac{\partial \underline{u}^r}{\partial z} - r \frac{\partial \underline{u}^z}{\partial r}\right] d\theta + \left[\frac{\partial \left(r\, \underline{u}^\theta\right)}{r\partial r} - \frac{\partial \underline{u}^r}{r\partial \theta}\right] dz \,, \tag{8.151b}$$

$$\begin{aligned}\left(*d\overset{1}{\omega}_{\mathbf{u}}\right)^\sharp &= \left[\frac{\partial \underline{u}^z}{r\partial \theta} - \frac{\partial \underline{u}^\theta}{\partial z}\right] \mathbf{g}_r + \left[\frac{\partial \underline{u}^r}{r\partial z} - \frac{\partial \underline{u}^z}{r\partial r}\right] \mathbf{g}_\theta + \left[\frac{\partial \left(r\, \underline{u}^\theta\right)}{r\partial r} - \frac{\partial \underline{u}^r}{r\partial \theta}\right] \mathbf{g}_z \\ &= \left[\frac{\partial \underline{u}^z}{r\partial \theta} - \frac{\partial \underline{u}^\theta}{\partial z}\right] \widehat{\mathbf{e}}_r + \left[\frac{\partial \underline{u}^r}{\partial z} - \frac{\partial \underline{u}^z}{\partial r}\right] \widehat{\mathbf{e}}_\theta \\ &\quad + \left[\frac{\partial \left(r\, \underline{u}^\theta\right)}{r\partial r} - \frac{\partial \underline{u}^r}{r\partial \theta}\right] \widehat{\mathbf{e}}_z = \boxed{\operatorname{curl} \mathbf{u} \,.}\end{aligned} \tag{8.151c}$$

Gradient and Laplacian of a scalar field in ***cylindrical*** *coordinates:* By using the equations (5.7a)-(5.7c), (5.117), (8.46), (8.48), (8.59), (8.67b), (8.67d), (8.83), $(8.92)_2$, $(8.101)_3$ and (8.118),

$$df = \frac{\partial f}{\partial r} dr + \frac{\partial f}{\partial \theta} d\theta + \frac{\partial f}{\partial z} dz \,, \tag{8.152a}$$

$$\begin{aligned}(df)^\sharp &= \frac{\partial f}{\partial r} \mathbf{g}_r + \frac{\partial f}{r^2 \partial \theta} \mathbf{g}_\theta + \frac{\partial f}{\partial z} \mathbf{g}_z \\ &= \frac{\partial f}{\partial r} \widehat{\mathbf{e}}_r + \frac{\partial f}{r\partial \theta} \widehat{\mathbf{e}}_\theta + \frac{\partial f}{\partial z} \widehat{\mathbf{e}}_z = \boxed{\operatorname{grad} f \,,}\end{aligned} \tag{8.152b}$$

$$*df = r \frac{\partial f}{\partial r} d\theta \wedge dz + \frac{\partial f}{r\partial \theta} dz \wedge dr + r \frac{\partial f}{\partial z} dr \wedge d\theta \,, \tag{8.152c}$$

$$d * df = \left[\frac{\partial}{\partial r}\left(r \frac{\partial f}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{\partial f}{r\partial \theta}\right) + \frac{\partial}{\partial z}\left(r \frac{\partial f}{\partial z}\right)\right] dr \wedge d\theta \wedge dz \,, \tag{8.152d}$$

$$*d * df = \frac{\partial}{r\partial r}\left(r \frac{\partial f}{\partial r}\right) + \frac{\partial^2 f}{r^2 \partial \theta^2} + \frac{\partial^2 f}{\partial z^2} = \boxed{\nabla^2 f \,.} \tag{8.152e}$$

Divergence of a vector field in ***spherical*** *coordinates:* Making use of the relations (5.11a)-(5.11c), (5.64a), (5.119), (8.46), (8.48), (8.59), (8.68b), (8.68d), (8.83), (8.89) and (8.103),

$$\begin{aligned}\mathbf{u} &= \underline{u}^r \widehat{\mathbf{e}}_r + \underline{u}^\theta \widehat{\mathbf{e}}_\theta + \underline{u}^\phi \widehat{\mathbf{e}}_\phi \\ &= \underline{u}^r \mathbf{g}_r + \frac{\underline{u}^\theta}{r} \mathbf{g}_\theta + \frac{\underline{u}^\phi}{r \sin\theta} \mathbf{g}_\phi \ , \end{aligned} \tag{8.153a}$$

$$\overset{1}{\omega}_{\mathbf{u}} = \underline{u}^r dr + r\, \underline{u}^\theta d\theta + r \sin\theta\, \underline{u}^\phi d\phi \ , \tag{8.153b}$$

$$*\overset{1}{\omega}_{\mathbf{u}} = r^2 \sin\theta\, \underline{u}^r d\theta \wedge d\phi + r \sin\theta\, \underline{u}^\theta d\phi \wedge dr + r\, \underline{u}^\phi dr \wedge d\theta \ , \tag{8.153c}$$

$$\begin{aligned} d * \overset{1}{\omega}_{\mathbf{u}} = \Bigg[&\frac{\partial \left(r^2 \sin\theta\, \underline{u}^r \right)}{\partial r} + \frac{\partial \left(r \sin\theta\, \underline{u}^\theta \right)}{\partial \theta} \\ &+ \frac{\partial \left(r\, \underline{u}^\phi \right)}{\partial \phi} \Bigg] dr \wedge d\theta \wedge d\phi \ , \end{aligned} \tag{8.153d}$$

$$*d * \overset{1}{\omega}_{\mathbf{u}} = \frac{\partial \left(r^2\, \underline{u}^r \right)}{r^2 \partial r} + \frac{\partial \left(\sin\theta\, \underline{u}^\theta \right)}{r \sin\theta \partial \theta} + \frac{\partial\, \underline{u}^\phi}{r \sin\theta \partial \phi} = \boxed{\operatorname{div} \mathbf{u} \ .} \tag{8.153e}$$

*Curl of a vector field in **spherical** coordinates:* With the aid of the expressions (5.11a)–(5.11c), (5.120), (8.38a), (8.38c), (8.59), (8.68c), (8.80), (8.92), (8.110) and (8.153b),

$$\begin{aligned} d\overset{1}{\omega}_{\mathbf{u}} = &\frac{\partial\, \underline{u}^r}{\partial \theta} d\theta \wedge dr + \frac{\partial\, \underline{u}^r}{\partial \phi} d\phi \wedge dr \\ &+ \frac{\partial \left(r\, \underline{u}^\theta \right)}{\partial r} dr \wedge d\theta + \frac{\partial \left(r\, \underline{u}^\theta \right)}{\partial \phi} d\phi \wedge d\theta \\ &+ \frac{\partial \left(r \sin\theta\, \underline{u}^\phi \right)}{\partial r} dr \wedge d\phi + \frac{\partial \left(r \sin\theta\, \underline{u}^\phi \right)}{\partial \theta} d\theta \wedge d\phi \ , \end{aligned} \tag{8.154a}$$

$$\begin{aligned} *d\overset{1}{\omega}_{\mathbf{u}} = &\left[\frac{\partial \left(r \sin\theta\, \underline{u}^\phi \right)}{r^2 \sin\theta \partial \theta} - \frac{\partial\, \underline{u}^\theta}{r \sin\theta \partial \phi} \right] dr + \left[\frac{\partial\, \underline{u}^r}{\sin\theta \partial \phi} - \frac{\partial \left(r\, \underline{u}^\phi \right)}{\partial r} \right] d\theta \\ &+ \sin\theta \left[\frac{\partial \left(r\, \underline{u}^\theta \right)}{\partial r} - \frac{\partial\, \underline{u}^r}{\partial \theta} \right] d\phi \ , \end{aligned} \tag{8.154b}$$

$$\begin{aligned} \left(*d\overset{1}{\omega}_{\mathbf{u}} \right)^\sharp = &\left[\frac{\partial \left(\sin\theta\, \underline{u}^\phi \right)}{r \sin\theta \partial \theta} - \frac{\partial\, \underline{u}^\theta}{r \sin\theta \partial \phi} \right] \mathbf{g}_r + \left[\frac{\partial\, \underline{u}^r}{r^2 \sin\theta \partial \phi} - \frac{\partial \left(r\, \underline{u}^\phi \right)}{r^2 \partial r} \right] \mathbf{g}_\theta \\ &+ \left[\frac{\partial \left(r\, \underline{u}^\theta \right)}{r^2 \sin\theta \partial r} - \frac{\partial\, \underline{u}^r}{r^2 \sin\theta \partial \theta} \right] \mathbf{g}_\phi \ , \\ = &\left[\frac{\partial \left(\sin\theta\, \underline{u}^\phi \right)}{r \sin\theta \partial \theta} - \frac{\partial\, \underline{u}^\theta}{r \sin\theta \partial \phi} \right] \widehat{\mathbf{e}}_r + \left[\frac{\partial\, \underline{u}^r}{r \sin\theta \partial \phi} - \frac{\partial \left(r\, \underline{u}^\phi \right)}{r \partial r} \right] \widehat{\mathbf{e}}_\theta \\ &+ \left[\frac{\partial \left(r\, \underline{u}^\theta \right)}{r \partial r} - \frac{\partial\, \underline{u}^r}{r \partial \theta} \right] \widehat{\mathbf{e}}_\phi = \boxed{\operatorname{curl} \mathbf{u} \ .} \end{aligned} \tag{8.154c}$$

*Gradient and Laplacian of a scalar field in **spherical** coordinates:* Using the equations (5.11a)-(5.11c), (5.120), (8.46), (8.48), (8.59), (8.68b), (8.68d), (8.83), $(8.92)_2$, $(8.101)_3$ and (8.118),

$$df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \theta}d\theta + \frac{\partial f}{\partial \phi}d\phi \ , \tag{8.155a}$$

$$\begin{aligned}(df)^\sharp &= \frac{\partial f}{\partial r}\mathbf{g}_r + \frac{\partial f}{r^2\partial \theta}\mathbf{g}_\theta + \frac{\partial f}{r^2\sin^2\theta\partial\phi}\mathbf{g}_\phi \\ &= \frac{\partial f}{\partial r}\widehat{\mathbf{e}}_r + \frac{\partial f}{r\partial\theta}\widehat{\mathbf{e}}_\theta + \frac{\partial f}{r\sin\theta\partial\phi}\widehat{\mathbf{e}}_\phi = \boxed{\operatorname{grad} f \ ,}\end{aligned} \tag{8.155b}$$

$$\begin{aligned}*df &= r^2\sin\theta\frac{\partial f}{\partial r}d\theta\wedge d\phi + \sin\theta\frac{\partial f}{\partial\theta}d\phi\wedge dr \\ &\quad + \sin^{-1}\theta\frac{\partial f}{\partial\phi}dr\wedge d\theta \ ,\end{aligned} \tag{8.155c}$$

$$\begin{aligned}d*df &= \left[\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial f}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right)\right. \\ &\quad \left.+\frac{\partial}{\partial\phi}\left(\sin^{-1}\theta\frac{\partial f}{\partial\phi}\right)\right]dr\wedge d\theta\wedge d\phi \ ,\end{aligned} \tag{8.155d}$$

$$\begin{aligned}*d*df &= \frac{\partial}{r^2\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{\partial}{r^2\sin\theta\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) \\ &\quad + \frac{\partial^2 f}{r^2\sin^2\theta\partial\phi^2} = \boxed{\nabla^2 f \ .}\end{aligned} \tag{8.155e}$$

Exercise 8.6

The goal of this exercise is to translate the **Maxwell's equations of electromagnetism** from the language of vector calculus to the language of differential forms. To do so, consider the following four Maxwell's partial differential equations

$$\operatorname{div}\mathbf{E} = 4\pi\rho \ , \qquad \leftarrow \text{Gauss's law} \tag{8.156a}$$

$$\operatorname{div}\mathbf{B} = 0 \ , \qquad \leftarrow \text{Gauss's law for magnetism} \tag{8.156b}$$

$$\operatorname{curl}\mathbf{E} + \frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} = \mathbf{0} \ , \qquad \leftarrow \text{Faraday's law of induction} \tag{8.156c}$$

$$\operatorname{curl}\mathbf{B} - \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j} \ , \qquad \leftarrow \text{Ampere's circuital law} \tag{8.156d}$$

where **E** denotes the *electric field*, **B** presents the *magnetic field*, **j** is the *current density*, ρ represents the *charge density* and c stands for the *light speed* (note the abuse of notation in this exercise where the vector fields are represented by uppercase boldface Latin letters). To this end, suppose that the space is coordinated by a

Cartesian coordinate system for convenience. It is worthwhile to point out that the smooth vector fields **E**, **B**, **j** and the scalar field ρ are time-dependent quantities. And this means that their entries are three space coordinates and time. Consequently, the introduced vector fields can be represented by

$$\left.\begin{aligned} \mathbf{E} &= E_x\,\widehat{\mathbf{e}}_x + E_y\,\widehat{\mathbf{e}}_y + E_z\,\widehat{\mathbf{e}}_z \\ \mathbf{B} &= B_x\,\widehat{\mathbf{e}}_x + B_y\,\widehat{\mathbf{e}}_y + B_z\,\widehat{\mathbf{e}}_z \\ \mathbf{j} &= j_x\,\widehat{\mathbf{e}}_x + j_y\,\widehat{\mathbf{e}}_y + j_z\,\widehat{\mathbf{e}}_z \end{aligned}\right\} . \tag{8.157}$$

They are associated with

$$\left.\begin{aligned} \overset{1}{\omega}_{\mathbf{E}} &= E_x\,dx + E_y\,dy + E_z\,dz \\ \overset{1}{\omega}_{\mathbf{B}} &= B_x\,dx + B_y\,dy + B_z\,dz \\ \overset{1}{\omega}_{\mathbf{j}} &= j_x\,dx + j_y\,dy + j_z\,dz \end{aligned}\right\} . \tag{8.158}$$

First, consider the *Faraday 2-form*

$$\overset{2}{F} = \overset{1}{\omega}_{\mathbf{E}} \wedge c\,dt + *_s\overset{1}{\omega}_{\mathbf{B}} \ , \tag{8.159}$$

where

$$*_s\,\overset{1}{\omega}_{\mathbf{B}} = B_x\,dy \wedge dz + B_y\,dz \wedge dx + B_z\,dx \wedge dy \ , \tag{8.160}$$

and the *Maxwell 2-form*

$$\overset{2}{M} = \overset{1}{\omega}_{\mathbf{B}} \wedge c\,dt - *_s\overset{1}{\omega}_{\mathbf{E}} \ , \tag{8.161}$$

where

$$*_s\,\overset{1}{\omega}_{\mathbf{E}} = E_x\,dy \wedge dz + E_y\,dz \wedge dx + E_z\,dx \wedge dy \ . \tag{8.162}$$

Note that in these relations, $*_s$ denotes the **space** Hodge star operator according to (8.66a)-(8.66d). Then, consider the *charge-current 1-form*

$$\begin{aligned} \overset{1}{J} &= \rho c\,dt - \frac{1}{c}\overset{1}{\omega}_{\mathbf{j}} \\ &= \rho c\,dt - \frac{j_x}{c}\,dx - \frac{j_y}{c}\,dy - \frac{j_z}{c}\,dz \ , \end{aligned} \tag{8.163}$$

and its dual *charge-current 3-form* which admits the following alternative forms

$$\overset{3}{J} = *\overset{1}{J}$$
$$= \rho *_s (1) - \frac{1}{c}\left(*_s \overset{1}{\omega}_{\mathbf{j}}\right) \wedge c\,dt$$
$$= \rho dx \wedge dy \wedge dz$$
$$- \frac{j_x}{c} c\,dt \wedge dy \wedge dz - \frac{j_y}{c} c\,dt \wedge dz \wedge dx - \frac{j_z}{c} c\,dt \wedge dx \wedge dy \,, \tag{8.164}$$

where $*$ is the **spacetime** Hodge star operator according to (8.73a)-(8.76).

Now, show that the Maxwell's equations can be rewritten in the language of exterior calculus as

$$d\overset{2}{F} = 0 \,, \tag{8.165a}$$
$$d\overset{2}{M} = -4\pi \overset{3}{J} \quad \text{or} \quad * d * \overset{2}{F} = 4\pi \overset{1}{J} \,. \tag{8.165b}$$

Moreover, verify that the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} = 0 \,, \tag{8.166}$$

is equivalent to

$$d\overset{3}{J} = d * \overset{1}{J} = 0 \,. \tag{8.167}$$

Finally, denoting by ϕ the *electric potential* and by $\mathbf{A}$ the *magnetic vector potential*, translate the following Maxwell's celebrated equations in potential formulation

$$-4\pi\rho = \operatorname{div}\operatorname{grad}\phi + \frac{1}{c^2}\operatorname{div}\frac{\partial \mathbf{A}}{\partial t} \,, \tag{8.168a}$$
$$4\pi\mathbf{j} = \operatorname{curl}\operatorname{curl}\mathbf{A} + \operatorname{grad}\frac{\partial \phi}{\partial t} + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} \,, \tag{8.168b}$$

to the language of differential forms.[7]

[7] Recall from (7.123b) that the curl of a vector field was divergence-free. One can then deduce from (8.156b) that

$$\mathbf{B} = \frac{1}{c}\operatorname{curl}\mathbf{A} \,,$$

where $\mathbf{A}$ is referred to as the *magnetic vector potential*. This result helps rewrite (8.156c) as $\operatorname{curl}\left(\mathbf{E} + c^{-2}\partial\mathbf{A}/\partial t\right) = \mathbf{0}$. Now, recall from (7.123c) that the gradient of a scalar field was curl-free. Consequently,

$$\mathbf{E} = -\operatorname{grad}\phi - \frac{1}{c^2}\frac{\partial \mathbf{A}}{\partial t} \,,$$

Solution. Here, the spacetime differential operator d is split into the timelike part d_t and spacelike portion d_s for the sake of clarity. For instance, the spacetime exterior derivative of the charge density is written as

$$d\rho = d_t\rho + d_s\rho \ , \tag{8.169}$$

where

$$d_t\rho = \frac{\partial\rho}{c\partial t}cdt \quad , \quad d_s\rho = \frac{\partial\rho}{\partial x}dx + \frac{\partial\rho}{\partial y}dy + \frac{\partial\rho}{\partial z}dz \ . \tag{8.170}$$

Some useful identities need to be established for subsequent developments. For this reason, consider a generic vector $\mathbf{P} = P_x\widehat{\mathbf{e}}_x + P_y\widehat{\mathbf{e}}_y + P_z\widehat{\mathbf{e}}_z$ with its corresponding 1-form $\overset{1}{\omega}_{\mathbf{P}} = P_x dx + P_y dy + P_z dz$ and

$$\frac{\partial\mathbf{P}}{\partial t} = \frac{\partial P_x}{\partial t}\widehat{\mathbf{e}}_x + \frac{\partial P_y}{\partial t}\widehat{\mathbf{e}}_y + \frac{\partial P_z}{\partial t}\widehat{\mathbf{e}}_z \ \leftrightarrow \ \overset{1}{\omega}_{\frac{\partial\mathbf{P}}{\partial t}} = \frac{\partial P_x}{\partial t}dx + \frac{\partial P_y}{\partial t}dy + \frac{\partial P_z}{\partial t}dz \ . \tag{8.171}$$

Then,

$$\frac{\partial\overset{1}{\omega}_{\mathbf{P}}}{\partial t} = \overset{1}{\omega}_{\frac{\partial\mathbf{P}}{\partial t}} \ . \tag{8.172}$$

Moreover,

$$d_t\overset{1}{\omega}_{\mathbf{P}} = -\frac{1}{c}\overset{1}{\omega}_{\frac{\partial\mathbf{P}}{\partial t}} \wedge c\,dt \quad , \quad d_t *_s \overset{1}{\omega}_{\mathbf{P}} = \frac{1}{c}\left(*_s\overset{1}{\omega}_{\frac{\partial\mathbf{P}}{\partial t}}\right) \wedge c\,dt \ . \tag{8.173}$$

To begin with, consider the following Maxwell's equations in terms of differential forms

$$*_s d_s *_s \overset{1}{\omega}_{\mathbf{E}} = 4\pi\rho \ , \tag{8.174a}$$

$$*_s d_s *_s \overset{1}{\omega}_{\mathbf{B}} = 0 \ , \tag{8.174b}$$

$$*_s d_s\overset{1}{\omega}_{\mathbf{E}} + \frac{1}{c}\overset{1}{\omega}_{\frac{\partial\mathbf{B}}{\partial t}} = 0 \ , \tag{8.174c}$$

$$*_s d_s\overset{1}{\omega}_{\mathbf{B}} - \frac{1}{c}\overset{1}{\omega}_{\frac{\partial\mathbf{E}}{\partial t}} = \frac{4\pi}{c}\overset{1}{\omega}_{\mathbf{j}} \ , \tag{8.174d}$$

where $(8.91)_1$, (8.103), $(8.110)_3$ and (8.172) have been utilized. These four relations can elegantly be presented in two equations by using the Faraday and Maxwell 2-forms (8.159)-(8.162) along with the charge-current forms (8.163)-(8.164) which

---

where ϕ is known as the *electric potential*. Substituting these results back into (8.156a) and (8.156d) leads to the elegant expressions (8.168a)-(8.168b) representing the Maxwell's equations in terms of potentials.

are all the inhabitants of the four-dimensional Minkowski spacetime. To show this, one needs to compute the exterior derivative of the introduced Faraday differential 2-form. This is given by

$$\begin{aligned}
d\overset{2}{F} &= \frac{\partial E_x}{\partial y}\, dy \wedge dx \wedge c\,dt + \frac{\partial E_x}{\partial z}\, dz \wedge dx \wedge c\,dt + \frac{\partial E_y}{\partial x}\, dx \wedge dy \wedge c\,dt \\
&\quad + \frac{\partial E_y}{\partial z}\, dz \wedge dy \wedge c\,dt + \frac{\partial E_z}{\partial x}\, dx \wedge dz \wedge c\,dt + \frac{\partial E_z}{\partial y}\, dy \wedge dz \wedge c\,dt \\
&\quad + \frac{\partial B_x}{c\partial t}\, c\,dt \wedge dy \wedge dz + \frac{\partial B_x}{\partial x}\, dx \wedge dy \wedge dz + \frac{\partial B_y}{c\partial t}\, c\,dt \wedge dz \wedge dx \\
&\quad + \frac{\partial B_y}{\partial y}\, dy \wedge dz \wedge dx + \frac{\partial B_z}{c\partial t}\, c\,dt \wedge dx \wedge dy + \frac{\partial B_z}{\partial z}\, dz \wedge dx \wedge dy \\
&= \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} + \frac{\partial B_z}{c\partial t} \right) c\,dt \wedge dx \wedge dy \\
&\quad + \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} + \frac{\partial B_y}{c\partial t} \right) c\,dt \wedge dz \wedge dx \\
&\quad + \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} + \frac{\partial B_x}{c\partial t} \right) c\,dt \wedge dy \wedge dz \\
&\quad + \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) dx \wedge dy \wedge dz \ ,
\end{aligned} \tag{8.175}$$

or

$$\begin{aligned}
d\overset{2}{F} &= (\operatorname{div} \mathbf{B})\, dx \wedge dy \wedge dz \\
&\quad + \left(\operatorname{curl}\mathbf{E} + \frac{\partial \mathbf{B}}{c\partial t} \right)_x c\,dt \wedge dy \wedge dz + \left(\operatorname{curl}\mathbf{E} + \frac{\partial \mathbf{B}}{c\partial t} \right)_y c\,dt \wedge dz \wedge dx \\
&\quad + \left(\operatorname{curl}\mathbf{E} + \frac{\partial \mathbf{B}}{c\partial t} \right)_z c\,dt \wedge dx \wedge dy \ ,
\end{aligned} \tag{8.176}$$

where (7.78), (7.92), (8.46), (8.48) and $(8.85)_2$ have been used. As can be seen, the **homogeneous** Maxwell's equations $\operatorname{div} \mathbf{B} = 0$ and $c \operatorname{curl}\mathbf{E} + \partial \mathbf{B}/\partial t = \mathbf{0}$ are equivalent to the single expression $d\overset{2}{F} = 0$. Indeed,

$$\operatorname{div} \mathbf{B} = 0 \ , \quad c \operatorname{curl}\mathbf{E} + \partial \mathbf{B}/\partial t = \mathbf{0} \qquad \Longleftrightarrow \qquad d\overset{2}{F} = 0 \ .$$

Following similar procedures which led to (8.176), one can obtain the exterior derivative of the Maxwell differential 2-form according to

$$\overset{2}{dM} = -(\text{div}\,\mathbf{E})\; dx\wedge dy\wedge dz + \left(\text{curl}\mathbf{B} - \frac{\partial\mathbf{E}}{c\partial t}\right)_x c\,dt\wedge dy\wedge dz + \left(\text{curl}\mathbf{B} - \frac{\partial\mathbf{E}}{c\partial t}\right)_y c\,dt\wedge dz\wedge dx + \left(\text{curl}\mathbf{B} - \frac{\partial\mathbf{E}}{c\partial t}\right)_z c\,dt\wedge dx\wedge dy\;. \tag{8.177}$$

Making use of (8.156a), (8.156d) and $(8.164)_3$, the differential 3-form (8.177) yields the desired relation

$$\overset{2}{dM} = -4\pi\,[\,\rho\,dx\wedge dy\wedge dz - \frac{j_x}{c}\,c\,dt\wedge dy\wedge dz - \frac{j_y}{c}\,c\,dt\wedge dz\wedge dx - \frac{j_z}{c}\,c\,dt\wedge dx\wedge dy\,] = -4\pi\overset{3}{J}\;. \tag{8.178}$$

Interestingly enough, $*d*\overset{2}{F} = 4\pi\overset{1}{J}$ also corresponds to the **non-homogeneous** Maxwell's equations $\text{div}\,\mathbf{E} = 4\pi\rho$ and $c\,\text{curl}\,\mathbf{B} - \partial\mathbf{E}/\partial t = 4\pi\mathbf{j}$ because from

$$*\overset{2}{F} = E_x\,dy\wedge dz + E_y\,dz\wedge dx + E_z\,dx\wedge dy - B_x\,dx\wedge c\,dt - B_y\,dy\wedge c\,dt - B_z\,dz\wedge c\,dt\;, \tag{8.179}$$

one can obtain

$$\begin{aligned} d*\overset{2}{F} = {} & \underline{\tfrac{\partial E_x}{c\partial t}\,c\,dt\wedge dy\wedge dz} + \underline{\tfrac{\partial E_x}{\partial x}\,dx\wedge dy\wedge dz} + \underline{\tfrac{\partial E_y}{c\partial t}\,c\,dt\wedge dz\wedge dx} \\ & + \underline{\tfrac{\partial E_y}{\partial y}\,dy\wedge dz\wedge dx} + \underline{\tfrac{\partial E_z}{c\partial t}\,c\,dt\wedge dx\wedge dy} + \underline{\tfrac{\partial E_z}{\partial z}\,dz\wedge dx\wedge dy} \\ & - \underline{\tfrac{\partial B_x}{\partial y}\,dy\wedge dx\wedge c\,dt} - \underline{\tfrac{\partial B_x}{\partial z}\,dz\wedge dx\wedge c\,dt} - \underline{\tfrac{\partial B_y}{\partial x}\,dx\wedge dy\wedge c\,dt} \\ & - \underline{\tfrac{\partial B_y}{\partial z}\,dz\wedge dy\wedge c\,dt} - \underline{\tfrac{\partial B_z}{\partial x}\,dx\wedge dz\wedge c\,dt} - \underline{\tfrac{\partial B_z}{\partial y}\,dy\wedge dz\wedge c\,dt} \\ = {} & \underline{\left(\tfrac{\partial E_x}{\partial x} + \tfrac{\partial E_y}{\partial y} + \tfrac{\partial E_z}{\partial z}\right) dx\wedge dy\wedge dz} \\ & + \underline{\left(\tfrac{\partial E_x}{c\partial t} + \tfrac{\partial B_y}{\partial z} - \tfrac{\partial B_z}{\partial y}\right) c\,dt\wedge dy\wedge dz} \\ & + \underline{\left(\tfrac{\partial E_y}{c\partial t} + \tfrac{\partial B_z}{\partial x} - \tfrac{\partial B_x}{\partial z}\right) c\,dt\wedge dz\wedge dx} \\ & + \underline{\left(\tfrac{\partial E_z}{c\partial t} + \tfrac{\partial B_x}{\partial y} - \tfrac{\partial B_y}{\partial x}\right) c\,dt\wedge dx\wedge dy}\;, \end{aligned} \tag{8.180}$$

and, subsequently,

$$
\begin{aligned}
* d * \overset{2}{F} &= (\operatorname{div} \mathbf{E})\, c\, dt \\
&\quad - \left(\operatorname{curl}\mathbf{B} - \frac{\partial \mathbf{E}}{c \partial t}\right)_x dx - \left(\operatorname{curl}\mathbf{B} - \frac{\partial \mathbf{E}}{c \partial t}\right)_y dy - \left(\operatorname{curl}\mathbf{B} - \frac{\partial \mathbf{E}}{c \partial t}\right)_z dz \\
&= 4\pi \left[\rho c\, dt - \frac{j_x}{c} dx - \frac{j_y}{c} dy - \frac{j_z}{c} dz\right] \\
&= 4\pi \overset{1}{J} \ , \qquad (8.181)
\end{aligned}
$$

where the expressions (7.78), (7.92), (8.46), (8.48), (8.74a)-(8.75d), $(8.85)_1$, (8.156a), (8.156d), $(8.158)_1$, (8.159)-(8.160) and $(8.163)_2$ have been used.

Next, the newfangled continuity equation (8.167) is verified. At this stage, one should realize that continuity equation (8.166) is established by the divergence of (8.156d) and the partial derivative of (8.156a) with respect to time taking into account that the divergence of curl of a vector field, according to (7.123b), identically vanishes. With the aid of (8.103), the relation (8.166) in vector calculus simply finds its equivalent form in exterior calculus. This is given by

$$
\frac{\partial \rho}{\partial t} + *_s d_s *_s \overset{1}{\omega}_{\mathbf{j}} = 0 \ . \qquad (8.182)
$$

This expression has alternative forms. It is usually demonstrated in two more elegant formats by using the charge-current differential forms, see $(8.164)_1$. That will be a simple exercise to see that

$$
\begin{aligned}
d\overset{3}{J} &= \frac{1}{c}\left(\frac{\partial \rho}{\partial t} + \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z}\right) c\, dt \wedge dx \wedge dy \wedge dz \\
&= \frac{1}{c}\left(\frac{\partial \rho}{\partial t} + \operatorname{div}\mathbf{j}\right) c\, dt \wedge dx \wedge dy \wedge dz \ . \qquad (8.183)
\end{aligned}
$$

One thus infers that $\partial\rho/\partial t + \operatorname{div}\mathbf{j} = 0$ amounts to writing $d\overset{3}{J} = 0$ and vice versa. Guided by $(8.88)_2$ and using $(8.165\text{b})_1$, one will have

$$
d\overset{3}{J} = \frac{-1}{4\pi} dd\overset{2}{M} = 0 \ . \qquad (8.184)
$$

The continuity equation (8.166) is thus a consequence of this result. By (8.59), (8.76) and (8.183),

$$
\frac{\partial \rho}{\partial t} + \operatorname{div}\mathbf{j} = -c * d\overset{3}{J} \ . \qquad (8.185)
$$

It is then easy to deduce that the vanishing of exterior derivative of Hodge dual of the charge-current 1-form is equivalent to the continuity equation. Notice that

$$\frac{\partial \rho}{\partial t} + \text{div}\mathbf{j} = -c * d * \overset{1}{J} . \tag{8.186}$$

Finally, the Maxwell's celebrated equations in terms of potentials are rewritten in the language of exterior calculus.

Making use of (8.65), (8.88), $(8.110)_3$ and (8.174b), one will have

$$*_s \overset{1}{\omega}_{\mathbf{B}} = \frac{1}{c} d_s \overset{1}{\omega}_{\mathbf{A}} ,$$

or

$$\overset{1}{\omega}_{\mathbf{B}} = \frac{1}{c} *_s d_s \overset{1}{\omega}_{\mathbf{A}} = \frac{1}{c} \overset{1}{\omega}_{\text{curl}\mathbf{A}} , \tag{8.187}$$

where $\overset{1}{\omega}_{\mathbf{A}}$ may be called the *magnetic 1-form potential* which corresponds to the *magnetic vector potential* **A**. By means of (8.65), $(8.101)_1$, (8.88), (8.174c) and $(8.187)_1$ (along with interchanging derivatives and having in mind that $*$, d are linear operators), one can get

$$*_s d_s \overset{1}{\omega}_{\mathbf{E}} + \frac{1}{c} \overset{1}{\omega}_{\frac{\partial \mathbf{B}}{\partial t}} = 0 \quad \Rightarrow \quad d_s \overset{1}{\omega}_{\mathbf{E}} + \frac{1}{c^2} d_s \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}} = 0 ,$$

or

$$\overset{1}{\omega}_{\mathbf{E}} = -d_s \phi - \frac{1}{c^2} \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}} = -\overset{1}{\omega}_{\text{grad}\phi} - \frac{1}{c^2} \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}} , \tag{8.188}$$

where the scalar field ϕ is known as the *electric potential*. That will be a simple exercise now to substitute $(8.187)_2$ and $(8.188)_2$ back into (8.174a) and (8.174d) in order to provide the potential formulation of Maxwell's equations in the language of exterior calculus:

$$-4\pi\rho = *_s d_s *_s \overset{1}{\omega}_{\text{grad}\phi} + \frac{1}{c^2} *_s d_s *_s \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}} , \tag{8.189a}$$

$$4\pi \overset{1}{\omega}_{\mathbf{j}} = \overset{1}{\omega}_{\text{curl curl}\mathbf{A}} + \overset{1}{\omega}_{\text{grad}\frac{\partial \phi}{\partial t}} + \frac{1}{c^2} \overset{1}{\omega}_{\frac{\partial^2 \mathbf{A}}{\partial t^2}} . \tag{8.189b}$$

One can alternatively use (8.165a) and $(8.165\text{b})_1$ to arrive at the above results. This will be shown in the following.

The relation $d\overset{2}{F} = 0$, considering the converse to the Poincaré's Lemma, implies that

$$\overset{2}{F} = d\overset{1}{K} .$$

And any additive decomposition of the new spacetime 1-form $\overset{1}{K}$ into the space and time portions can be written as

$$\overset{1}{K} = -\phi c\, dt + \frac{1}{c} \overset{1}{\omega}_{\mathbf{A}} \, . \tag{8.190}$$

Guided by (8.169) and using (8.65), (8.77b), (8.82), $(8.101)_1$, $(8.110)_3$, and $(8.173)_1$, the exterior derivative of (8.190) gives

$$\begin{aligned} d\overset{1}{K} &= -\left(d_t\phi \wedge cdt + d_s\phi \wedge cdt\right) + \frac{1}{c}\left[d_t\overset{1}{\omega}_{\mathbf{A}} + d_s\overset{1}{\omega}_{\mathbf{A}}\right] \\ &= -\left(\overset{1}{\omega}_{\text{grad}\phi} \wedge cdt\right) + \frac{1}{c}\left[-\frac{1}{c}\overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}} \wedge cdt + *_s\overset{1}{\omega}_{\text{curl}\mathbf{A}}\right] \\ &= \left(-\overset{1}{\omega}_{\text{grad}\phi} - \frac{1}{c^2}\overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}}\right) \wedge cdt + *_s\left(\frac{1}{c}\overset{1}{\omega}_{\text{curl}\mathbf{A}}\right) \, . \end{aligned} \tag{8.191}$$

Comparing (8.191) with (8.159), i.e. $\overset{2}{F} = \overset{1}{\omega}_{\mathbf{E}} \wedge c\, dt + *_s\overset{1}{\omega}_{\mathbf{B}}$, leads to the desired results $(8.187)_2$ and $(8.188)_2$. These differential forms help express the Maxwell 2-form (8.161), i.e. $\overset{2}{M} = \overset{1}{\omega}_{\mathbf{B}} \wedge c\, dt - *_s\overset{1}{\omega}_{\mathbf{E}}$, as

$$\overset{2}{M} = \frac{1}{c}\overset{1}{\omega}_{\text{curl}\mathbf{A}} \wedge c\, dt + *_s\overset{1}{\omega}_{\text{grad}\phi} + \frac{1}{c^2} *_s \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}} \, . \tag{8.192}$$

It is not then difficult to represent the spacetime exterior derivative of this differential 2-form according to

$$\begin{aligned} d\overset{2}{M} &= \underbrace{\frac{1}{c}d_s\overset{1}{\omega}_{\text{curl}\mathbf{A}} \wedge c\, dt}_{= \frac{1}{c} *_s \overset{1}{\omega}_{\text{curl}\,\text{curl}\mathbf{A}} \wedge c\, dt} + \underbrace{d_t *_s \overset{1}{\omega}_{\text{grad}\phi}}_{= \frac{1}{c} *_s \overset{1}{\omega}_{\text{grad}\frac{\partial\phi}{\partial t}} \wedge c\, dt} + \underbrace{d_s *_s \overset{1}{\omega}_{\text{grad}\phi}}_{= *_s \text{div}\,(\text{grad}\phi)} \\ &\quad + \underbrace{\frac{1}{c^2}d_t *_s \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}}}_{= \frac{1}{c^3} *_s \overset{1}{\omega}_{\frac{\partial^2 \mathbf{A}}{\partial t^2}} \wedge c\, dt} + \underbrace{\frac{1}{c^2}d_s *_s \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}}}_{= \frac{1}{c^2} *_s \text{div}\left(\frac{\partial \mathbf{A}}{\partial t}\right)} \\ &= \frac{1}{c}\left[*_s\left(\overset{1}{\omega}_{\text{curl}\,\text{curl}\mathbf{A}} + \overset{1}{\omega}_{\text{grad}\frac{\partial\phi}{\partial t}} + \frac{1}{c^2}\overset{1}{\omega}_{\frac{\partial^2 \mathbf{A}}{\partial t^2}}\right)\right] \wedge c\, dt \\ &\quad + \left[*_s d_s *_s \overset{1}{\omega}_{\text{grad}\phi} + \frac{1}{c^2} *_s d_s *_s \overset{1}{\omega}_{\frac{\partial \mathbf{A}}{\partial t}}\right] *_s (1) \, . \end{aligned} \tag{8.193}$$

On the other hand, from $(8.164)_2$ and $(8.165\text{b})_1$, it follows that

$$d\overset{2}{M} = \frac{1}{c}\left[*_s\left(4\pi\overset{1}{\omega}_{\mathbf{j}}\right)\right] \wedge c\, dt + \left[-4\pi\rho\right] *_s (1) \, . \tag{8.194}$$

By comparing (8.193) with (8.194), one can finally arrive at the desired equations (8.189a) and (8.189b).

References

1. Klein, F. (1874). Bemerkungen über den Zusammenhang der Flächen. *Mathematische Annalen, 7*, 549–557.
2. Spivak, M. (1971). *Calculus on manifolds: A modern approach to classical theorems of advanced calculus*. CRC Press.
3. Lovelock, D., & Rund, H. (1975). *Tensors, differential forms, and variational principles*. New York: Dover Publications.
4. do Carmo, M.P. (1994). *Differential forms and applications*. Berlin, Heidelberg: Springer.
5. Bachman, D. (2012). *A geometric approach to differential forms* (2nd ed.). Boston, MA: Birkhäuser.
6. Renteln, P. (2014). *Manifolds, tensors, and forms: An introduction for mathematicians and physicists*. New York: Cambridge University Press.
7. Hubbard, J. H., & Hubbard, B. B. (2015). *Vector calculus, linear algebra, and differential forms: A unified approach* (5th ed.). Matrix Editions.
8. Nguyen-Schäfer, H., & Schmidt, J. P. (2017). *Tensor analysis and elementary differential geometry for physicists and engineers* (2nd ed.). Berlin, Heidelberg: Springer.
9. Fortney, J. P. (2018). *A visual introduction to differential forms and calculus on manifolds*. Cham: Birkhäuser.

Chapter 9
Differential Geometry of Surfaces and Curves

This chapter provides an introduction to differential geometry of embedded surfaces and curves in the three-dimensional Euclidean space. The geometry of manifolds[1] of dimension two or **curved surfaces** is extensively used in many branches of physics and engineering. Examples of which include general relativity and structural mechanics. The geometry of manifolds of dimension one or **space curves** also plays an important role in many scientific applications such as beams or coordinated drone swarms. The goal of this chapter is therefore to study embedded two-dimensional surfaces as well as one-dimensional curves in the three-dimensional Euclidean flat space and introduce their fundamental properties. A topic of great importance in this context regards *curvature* which has a very long history. Many different objects are used in the literature to characterize curvature, each of which has its own application. This motivates to introduce *second-order surface mixed curvature tensor* along with its principal values (representing *mean curvature* and *Gaussian curvature*), *fourth-order surface Riemann-Christoffel curvature tensor*, *second-order Ricci curvature tensor* and *Ricci scalar*. *Normal curvature* and *geodesic curvature* of a curve embedded in a surface are also addressed. It is also the goal of this chapter to describe some crucially important derivative operators such as *surface covariant differentiation*, *Lie derivative* and *invariant time differentiation* and study their basic properties.[2]

[1] There is no unique definition for a manifold. Its widely accepted definition is provided within the realm of topology which is outside the scope of this text. Here and elsewhere in this text, this useful term means a subspace of the three-dimensional Euclidean space which can be parametrized with sufficiently smooth functions defined over a region of such a subspace.

[2] *Tangent space* of a manifold varies from point to point due to curvature. Consequently, the vectors (tensors) of a vector (tensor) field in two infinitesimally nearby points cannot simply be added and/or subtracted. Thus, the rate of change in an object with respect to a coordinate line, vector field or time-like variable should be treated with care. In this regard, some special techniques need to be established. A well-known example regards the surface covariant differentiation.

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*,
https://doi.org/10.1007/978-3-031-33953-0_9

Recall that the fourth-order Riemann-Christoffel curvature tensor acted as a decisive criterion for **flatness** of space. The Euclidean space represents a flat space with this fourth-order tensor being vanished all over the space. Indeed, the Euclidean space is one for which all well-known postulates of Euclidean geometry - expressing relationships between primitive quantities such as points and lines - perfectly applies. Straightness is its key characteristic and it is usually coordinated with the popular Cartesian coordinate system. For such a popular coordinate system, the matrix form of metric coefficients becomes diagonal with all diagonal elements being +1. Accordingly, all of its Christoffel elements vanish. When the space is Euclidean and referred to a Cartesian coordinate frame, any general basis that may be resulted from embedding a curvilinear coordinate system can be expressed as a linear combination of the standard Cartesian basis. Thus, the metric coefficients and subsequently the Christoffel elements can consistently be computed. This procedure was followed so far.

Over two thousand years geometers and/or mathematicians worked with Euclidean geometry and tried to prove his fifth axiom, i.e. the so-called *parallel postulate*, from the other four. Unfortunately, the attempts to directly prove this debatable postulate from the others were not successful. And this motivated some mathematicians to use the proof by contradiction. In 19th century, these attempts led to the foundation of what is called the *non-Euclidean* (or *Gauss-Bolyai-Lobachevsky*) *geometry*. A geometry that the most important 20th century development in science, i.e. Einstein's general theory of relativity, relies on. It may be classified into *hyperbolic geometry* and *elliptic geometry*. These geometries are called non-Euclidean because they at least violate Euclid's fifth postulate. For instance, there are infinitely many lines parallel to a given line or the sum of the angles of a triangle is less than 180° for the saddle shape surfaces. Interestingly, there are no lines parallel to a given line or the sum of the angles of a triangle is greater than 180° for a sphere.[3]

The study of geometric objects on a *differentiable* (or *differential*) *manifold* (i.e. a space which allows one to locally apply the tools of calculus) relies on two basic quantities; namely, *connection* and *metric*. The former connects different tangent spaces of a manifold as implied by its name. One can thus compare the vectors sitting at two infinitesimally nearby points which helps define covariant derivative and *parallel transport*.[4] But the latter, which acts as the central tool in measuring lengths and areas, defines the dot product on a manifold. One may now classify the differential manifolds into (i) *connected manifold* (i.e. space equipped with a connection) and (ii) *metric manifold* (i.e. space equipped with both connection and metric).

Torsion, *curvature* and *non-metricity* are three crucially important tensorial field variables - of order 3,4 and 3, respectively, - in describing differential geometry of

[3] In spherical geometry, lines are defined as *great circles* (i.e. the largest circles on the surface of a sphere). And parallel lines are lines that do not meet.

[4] The idea of parallel transport geometrically means keeping a vector as constant as possible when it is moving around on a curved space. Technically, the parallel transport of a vector along a curve represents a condition which states that the covariant derivative of such a vector must vanish.

manifolds. The torsion (curvature) tensor is defined in terms of the connection (connection and its partial derivatives) while the non-metricity tensor is introduced as the covariant differentiation of the metric. The torsion (curvature) of space measures the closure gap that evolves when two vectors are *parallel transported* along each other (a vector is parallel transported along an infinitesimal closed loop) while the non-metricity measures the change in the length of and the angle between parallel transported vectors. Based on these notions, one may define (i) *flat* (*non-flat*) *manifold* (i.e. space with zero (nonzero) curvature), (ii) *symmetric* (*non-symmetric*) *manifold* (i.e. space with vanishing (non-vanishing) torsion) and (iii) *metrically* (*non-metrically*) *connected manifold* (i.e. metric space with zero (nonzero) non-metricity).

The basic idea regarding differential geometry of a two-dimensional manifold dates back to the pioneering work of Gauss. In particular, he established the use of curvilinear coordinates for describing embedded surfaces. Moreover, he proposed that a curved surface can be studied either as a three-dimensional object or analyzed as a two-dimensional one (two completely different perspectives). This classifies the properties of surfaces or curves - in relation to a larger space embracing them - into *extrinsic* and *intrinsic objects*. Extrinsic quantities are those that depend on external embedding space; examples of which include the second-order curvature tensor and the surface normal vectors. Whereas intrinsic quantities are those which can be attained by measuring distances along the surface; examples of which include the metric coefficients and the Gaussian curvature.[5] The Gaussian curvature is basically a mathematical entity connecting extrinsic and intrinsic points of view.

A two-dimensional manifold is called *extrinsically flat* if the second-order curvature tensor vanishes identically for such a space. And it is referred to as *intrinsically flat* if the fourth-order Riemann-Christoffel curvature tensor (or equivalently the Gaussian curvature in this context) vanishes all over that space. For instance, a plane sheet is extrinsically and intrinsically flat while the surface of a cylinder is extrinsically curved but intrinsically flat. Moreover, the surface of a sphere is an example of extrinsically and intrinsically curved space. The reason for the surface of a cylinder to be intrinsically flat is that it can simply be constructed from a thin sheet of paper without any distortion implied by compression, stretching or shearing (only bending type deformations are required). Actually, these two surfaces are said to be *isometric*.[6]

Of interest here is to consider the two-dimensional surfaces embedded in the three-dimensional Euclidean space. From now on, this enveloping three-dimensional Euclidean space is referred to as the *ambient space*. Any surface in this text is

---

[5] To better understand the concept of intrinsic properties, consider the surface as a planet with entities living therein. Now, suppose that these entities have no knowledge of the external embedding space and the planet is all they know. In other words, these inhabitants only have a two-dimensional perception. Then, any quantity that can be measured or detected by these two-dimensional inhabitants - such as lengths of planar curves passing through a point or vectors lying in the local tangent planes - is referred to as intrinsic object.

[6] An *isometry* (or *isometric mapping*) is a mapping that preserves lengths of curves. This important mapping also preserves the angles and the surface areas. Intrinsic property is sometimes introduced as a property that is preserved under isometric transformations.

assumed to be a **sufficiently smooth** and *regular* two-dimensional subspace of the three-dimensional ambient space, if not otherwise stated. It is also assumed that there exist **tangent** and **normal** spaces at each point of the surface.[7] Moreover, any two-dimensional space in this chapter should be regarded as a *Riemannian manifold* which represents a non-flat, symmetric and metrically connected manifold. Finally, suppose that the standard properties of the real vector spaces introduced in Chaps. 1 and 3 hold true here (note that such properties will not be referenced in the upcoming developments for convenience).

For a detailed exposition on the rich area of differential geometry, the reader is referred to Weatherburn [1], Stoker [2], Kreyszig [3], Spivak [4], Gray et al. [5], O'Neill [6], Pressley [7], do Carmo [8] and Banchoff and Lovett [9] among many others.

9.1 Representation of Tensorial Variables on Surfaces

9.1.1 Surfaces in Three-Dimensional Euclidean Space

A two-dimensional surface $\mathbb{S}$ in the ambient space may be defined by the point function

$$\underbrace{\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right)}_{\text{note that } \Theta^i = \Theta^i\left(t^\alpha\right) \text{ which helps write } \mathbf{x} = \hat{\mathbf{x}}\left(\Theta^i\right) = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^\alpha\right), \text{ see (5.1)}} \quad , \tag{9.1}$$

where $\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right) : U \subset \mathrm{R}^2 \to \mathscr{E}_{\mathrm{p}}^3$ presents an arbitrary point on the surface and the real numbers t^α, $\alpha = 1, 2$, are called the *surface* (or *Gaussian*) *coordinates*. The points on a curve $\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2 = \text{constant}\right)$ is called a t^1-curve (or t^1-line). Similarly, the points on a curve $\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1 = \text{constant}, t^2\right)$ is called a t^2-curve (or t^2-line). The two families of t^1-curves and t^2-curves constitute the *coordinate* (or *parametric*) *curves* of a surface. Note that all Greek indices in this chapter range over $\{1, 2\}$ and the summation convention will also be implied when they are repeated. It is assumed that the function in (9.1) is sufficiently differentiable in order to have a *sufficiently smooth surface*.

Now, suppose that the ambient space is referred to Cartesian coordinates. Then, the functional relation (9.1), that parametrically describes the surface, represents

$$x_1 = \hat{x}_1^{\mathrm{s}}\left(t^1, t^2\right) \quad , \quad x_2 = \hat{x}_2^{\mathrm{s}}\left(t^1, t^2\right) \quad , \quad x_3 = \hat{x}_3^{\mathrm{s}}\left(t^1, t^2\right) \ . \tag{9.2}$$

In the following, some well-known surfaces will be introduced having in mind that the surface parametrization is not **unique**. ✷

[7] Note that the two-dimensional tangent space (or simply tangent plane) is spanned by the surface covariant basis. Whereas the one-dimensional normal space is defined by a unit vector normal to the surface which itself is constructed from the covariant basis vectors.

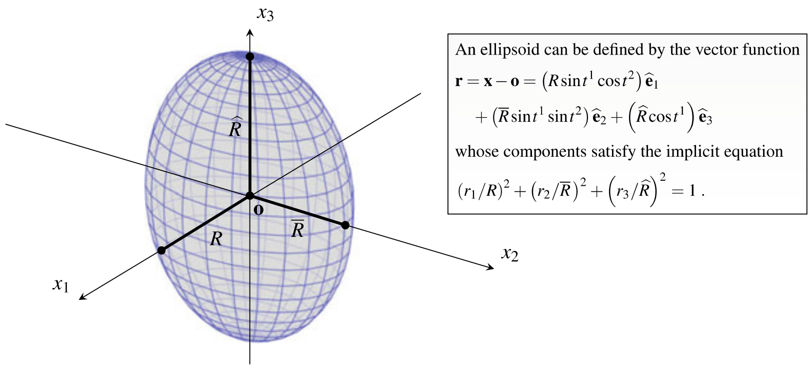


Fig. 9.1 Ellipsoid

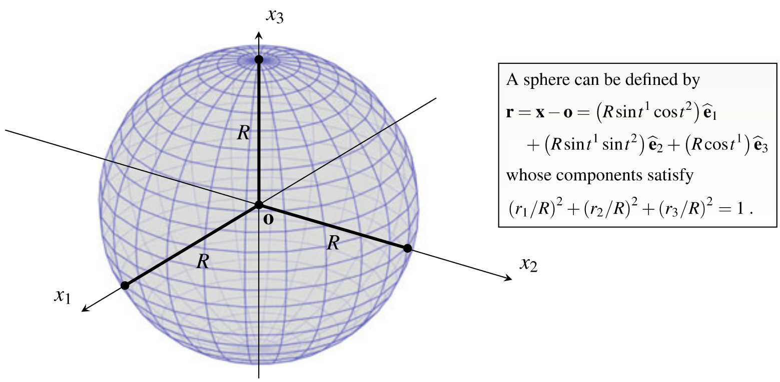


Fig. 9.2 Sphere

The first example presents an **ellipsoid** (Fig. 9.1). This *surface of revolution* is described by

$$\boxed{\underbrace{x_1 = R \sin t^1 \cos t^2 \quad , \quad x_2 = \overline{R} \sin t^1 \sin t^2 \quad , \quad x_3 = \widehat{R} \cos t^1}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + \left(x_2/\overline{R}\right)^2 + \left(x_3/\widehat{R}\right)^2 = 1} ,} \tag{9.3}$$

where the positive real numbers R, $\overline{R}$, $\widehat{R}$ denote the semiaxes and $0 \leq t^1 \leq \pi$, $0 \leq t^2 < 2\pi$.

A very practical example of elliptic geometry is **sphere** (Fig. 9.2). It can be provided by substituting $R = \overline{R} = \widehat{R}$ into (9.3), that is,

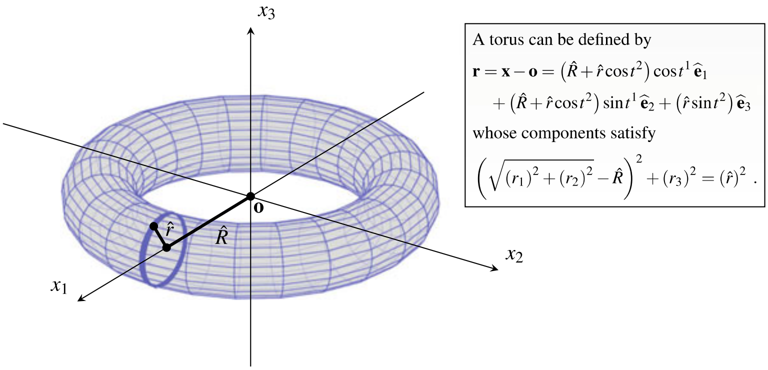


Fig. 9.3 Torus

$$\boxed{\underbrace{x_1 = R\sin t^1\cos t^2 \quad , \quad x_2 = R\sin t^1\sin t^2 \quad , \quad x_3 = R\cos t^1}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + (x_2/R)^2 + (x_3/R)^2 = 1} \ .} \tag{9.4}$$

Another example regards a **torus** (Fig. 9.3) which is also a surface of revolution generated by rotating a circle around a coplanar axis (note that such an axis is assumed to lie outside of that circle). It is defined by the following equation

$$\boxed{\underbrace{x_1 = \left(\hat{R} + \hat{r}\cos t^2\right)\cos t^1 \quad , \quad x_2 = \left(\hat{R} + \hat{r}\cos t^2\right)\sin t^1 \quad , \quad x_3 = \hat{r}\sin t^2}_{\text{these coordinates satisfy the implicit relation } \left(\sqrt{(x_1)^2 + (x_2)^2} - \hat{R}\right)^2 + (x_3)^2 = (\hat{r})^2} \ ,} \tag{9.5}$$

where $\hat{r} > 0$ is the radius of generating circle (minor radius), $\hat{R} > 0$ presents the distance between its center and the axis of revolution (major radius) and $0 \leq t^1 < 2\pi \ , \ 0 \leq t^2 < 2\pi$.

As a further example, consider an **elliptic paraboloid** (Fig. 9.4). It is parametrically defined by

$$\boxed{\underbrace{x_1 = Rt^1\cos t^2 \quad , \quad x_2 = \overline{R}t^1\sin t^2 \quad , \quad x_3 = \widehat{R}\left(t^1\right)^2}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + \left(x_2/\overline{R}\right)^2 = x_3/\widehat{R}} \ ,} \tag{9.6}$$

where $R > 0$, $\overline{R} > 0$, $\widehat{R} \neq 0$ are real numbers, $0 \leq t^1 < \infty$ and $0 \leq t^2 < 2\pi$.

A well-known example of hyperbolic geometry is **one-sheeted hyperboloid** (Fig. 9.5). It is defined by

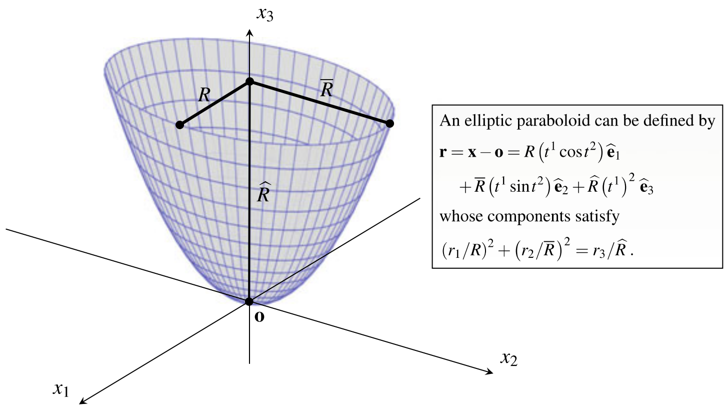


Fig. 9.4 Elliptic paraboloid

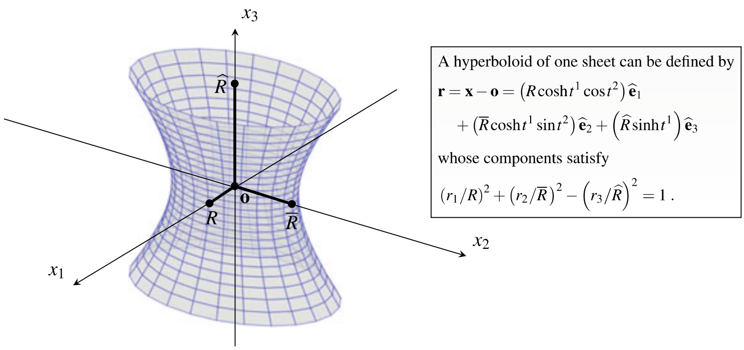


Fig. 9.5 Hyperboloid of one sheet

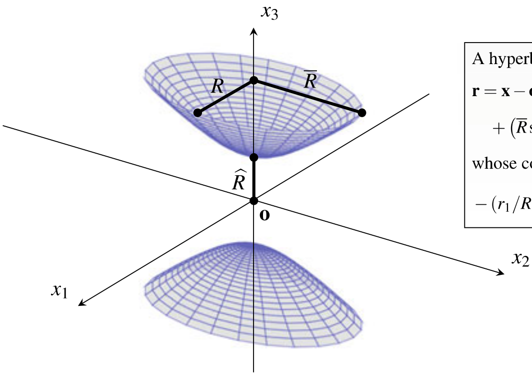
$$\underbrace{x_1 = R \cosh t^1 \cos t^2 \quad , \quad x_2 = \overline{R} \cosh t^1 \sin t^2 \quad , \quad x_3 = \widehat{R} \sinh t^1}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + \left(x_2/\overline{R}\right)^2 - \left(x_3/\widehat{R}\right)^2 = 1} , \tag{9.7}$$

where R, $\overline{R}$, $\widehat{R}$ are positive real numbers, $-\infty < t^1 < \infty$ and $0 \leq t^2 < 2\pi$.

Another example in hyperbolic geometry regards a **two-sheeted hyperboloid** (Fig. 9.6) defined by

$$\underbrace{x_1 = R \sinh t^1 \cos t^2 \quad , \quad x_2 = \overline{R} \sinh t^1 \sin t^2 \quad , \quad x_3 = \pm\widehat{R} \cosh t^1}_{\text{these coordinates satisfy the implicit relation } -(x_1/R)^2 - \left(x_2/\overline{R}\right)^2 + \left(x_3/\widehat{R}\right)^2 = 1} , \tag{9.8}$$

where R, $\overline{R}$, $\widehat{R}$ are positive real numbers, $0 \leq t^1 < \infty$ and $0 \leq t^2 < 2\pi$.



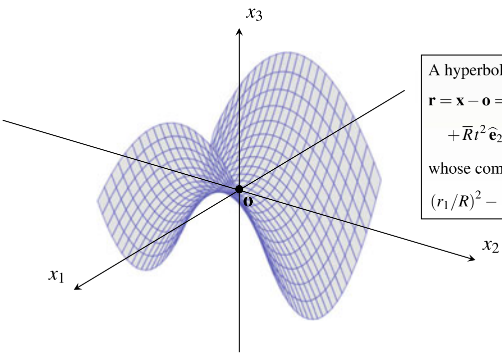
A hyperboloid of two sheets can be defined by

$$\mathbf{r} = \mathbf{x} - \mathbf{o} = \left(R \sinh t^1 \cos t^2\right) \widehat{\mathbf{e}}_1 + \left(\overline{R} \sinh t^1 \sin t^2\right) \widehat{\mathbf{e}}_2 + \left(\pm \widehat{R} \cosh t^1\right) \widehat{\mathbf{e}}_3$$

whose components satisfy

$$-\left(r_1/R\right)^2 - \left(r_2/\overline{R}\right)^2 + \left(r_3/\widehat{R}\right)^2 = 1 \,.$$

Fig. 9.6 Hyperboloid of two sheets



A hyperbolic paraboloid can be defined by

$$\mathbf{r} = \mathbf{x} - \mathbf{o} = R t^1 \widehat{\mathbf{e}}_1 + \overline{R} t^2 \widehat{\mathbf{e}}_2 + \widehat{R} \left[\left(t^1\right)^2 - \left(t^2\right)^2\right] \widehat{\mathbf{e}}_3$$

whose components satisfy

$$\left(r_1/R\right)^2 - \left(r_2/\overline{R}\right)^2 = r_3/\widehat{R} \,.$$

Fig. 9.7 Hyperbolic paraboloid

The last example here presents a **hyperbolic paraboloid** (Fig. 9.7) described by

$$\underbrace{x_1 = R t^1 \quad , \quad x_2 = \overline{R} t^2 \quad , \quad x_3 = \widehat{R} \left[\left(t^1\right)^2 - \left(t^2\right)^2\right]}_{\text{these coordinates satisfy the implicit relation } \left(x_1/R\right)^2 - \left(x_2/\overline{R}\right)^2 = x_3/\widehat{R}} \,, \tag{9.9}$$

where $R, \overline{R}, \widehat{R}$ are positive real numbers, $-\infty < t^1 < \infty$ and $-\infty < t^2 < \infty$. ✸

9.1.2 Surface Basis Vectors and Metric Coefficients

The *surface covariant basis vectors*, at an arbitrary point on the surface, are defined by

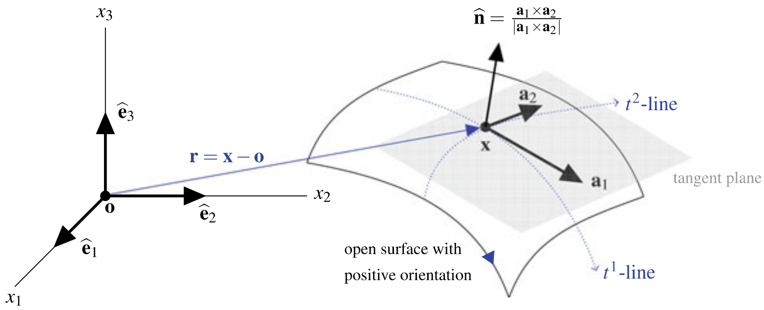


Fig. 9.8 Ambient basis at an arbitrary point of a surface encompassed by three-dimensional Euclidean space

$$\mathbf{a}_\alpha = \frac{\partial \mathbf{x}}{\partial t^\alpha}$$
$$= \lim_{h \to 0} \frac{\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1 + h\delta^1_\alpha, t^2 + h\delta^2_\alpha\right) - \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right)}{h} \; , \quad \longleftarrow \text{ see Exercise 9.1} \tag{9.10}$$

where δ^β_α denotes the two-dimensional Kronecker delta. These two **linearly independent** vectors are tangential to the coordinate lines corresponding to the Gaussian coordinates and, therefore, define a plane called *tangential plane*. These tangent vectors can be completed to provide a basis in the three-dimensional ambient space by means of the *surface normal vector*

$$\widehat{\mathbf{n}} = \pm \frac{\mathbf{a}_1 \times \mathbf{a}_2}{|\mathbf{a}_1 \times \mathbf{a}_2|} = \mathbf{a}_3 \; , \tag{9.11}$$

satisfying

$$\widehat{\mathbf{n}} \cdot \mathbf{a}_\alpha = 0 \quad , \quad \widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1 \; . \tag{9.12}$$

The appropriate sign in (9.11) must be chosen according to the previous discussions on orientation of surfaces given in Chap. 8. See Fig. 9.8 for a geometrical interpretation. Now, any arbitrary tensorial variable in the ambient space can be expressed with respect to the covariant basis

$$\{\mathbf{a}_i\} := \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} \quad \text{where} \quad \mathbf{a}_3 = \widehat{\mathbf{n}} \; . \tag{9.13}$$

It is assumed that the surface covariant basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly independent,[8] that is,

[8] This means that the Jacobian matrix

$$[\,J\,] = \begin{bmatrix} \frac{\partial x_1}{\partial t^1} & \frac{\partial x_1}{\partial t^2} \\ \frac{\partial x_2}{\partial t^1} & \frac{\partial x_2}{\partial t^2} \\ \frac{\partial x_3}{\partial t^1} & \frac{\partial x_3}{\partial t^2} \end{bmatrix} ,$$

$$\boxed{\mathbf{a}_1 \times \mathbf{a}_2 \neq 0 \,.} \tag{9.14}$$

This implies that the tangent vectors should not vanish and have various directions. Indeed, this condition is set to guarantee the existence of tangential planes and accordingly have well-defined normal vectors. With regard to this, a point at which this condition holds true is said to be a *regular point*; otherwise it is referred to as a *singular point*. And a surface $\mathcal{S}$ whose all normal vectors are well-defined is said to be a *regular surface*. Note that all surfaces in this text are assumed to be sufficiently smooth and regular.

Let

$$\left\{\mathbf{a}^i\right\} := \left\{\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3\right\} \,, \tag{9.15}$$

be the unique **dual** basis of $\{\mathbf{a}_i\}$. The procedure to calculate this ambient contravariant basis will be demonstrated in the following. ✳

To begin with, one can write the ambient covariant metric coefficients $(5.38)_3$ for the problem at hand as

$$\left[\, g_{ij} \,\right] = \begin{bmatrix} \left[a_{\alpha\beta}\right]_{2\times 2} & [\mathbf{0}]_{2\times 1} \\ [\mathbf{0}]^{\mathrm{T}}_{1\times 2} & 1 \end{bmatrix} \,, \tag{9.16}$$

where $a_{\alpha\beta}$ denote the *surface covariant metric coefficients*; given by,

$$\boxed{a_{\alpha\beta} = \mathbf{a}_\alpha \cdot \mathbf{a}_\beta = \mathbf{a}_\beta \cdot \mathbf{a}_\alpha = a_{\beta\alpha} \,,} \tag{9.17}$$

with

$$\boxed{\left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 & \mathbf{a}_1 \cdot \mathbf{a}_2 \\ \mathbf{a}_1 \cdot \mathbf{a}_2 & \mathbf{a}_2 \cdot \mathbf{a}_2 \end{bmatrix} \,.} \tag{9.18}$$

Recall that for a regular surface the condition (9.14) holds true. Having in mind the positive-definite property of the dot product, one can infer that

$$\boxed{a_{11} = \mathbf{a}_1 \cdot \mathbf{a}_1 > 0 \quad \text{and} \quad a_{22} = \mathbf{a}_2 \cdot \mathbf{a}_2 > 0 \,.} \tag{9.19}$$

Using (1.11), (1.78a) and $(9.17)_1$, one can also deduce that

---

has rank 2. In other words,

$$J_1^2 + J_2^2 + J_3^2 \neq 0 \,,$$

where

$$J_1 = \det\begin{bmatrix} \frac{\partial x_2}{\partial t^1} & \frac{\partial x_3}{\partial t^1} \\ \frac{\partial x_2}{\partial t^2} & \frac{\partial x_3}{\partial t^2} \end{bmatrix} \,, \quad J_2 = \det\begin{bmatrix} \frac{\partial x_3}{\partial t^1} & \frac{\partial x_1}{\partial t^1} \\ \frac{\partial x_3}{\partial t^2} & \frac{\partial x_1}{\partial t^2} \end{bmatrix} \,, \quad J_3 = \det\begin{bmatrix} \frac{\partial x_1}{\partial t^1} & \frac{\partial x_2}{\partial t^1} \\ \frac{\partial x_1}{\partial t^2} & \frac{\partial x_2}{\partial t^2} \end{bmatrix} \,.$$

$$
\begin{aligned}
|\mathbf{a}_1 \times \mathbf{a}_2|^2 &= (\mathbf{a}_1 \times \mathbf{a}_2) \cdot (\mathbf{a}_1 \times \mathbf{a}_2) \\
&= (\mathbf{a}_1 \cdot \mathbf{a}_1)(\mathbf{a}_2 \cdot \mathbf{a}_2) - (\mathbf{a}_1 \cdot \mathbf{a}_2)(\mathbf{a}_2 \cdot \mathbf{a}_1) \\
&= \boxed{a_{11} a_{22} - (a_{12})^2 > 0\ .}
\end{aligned} \tag{9.20}
$$

As a result, the symmetric matrix $\left[a_{\alpha\beta} \right]$ is **positive definite** since the determinants corresponding to its two upper-left submatrices are positive, see (2.125a)–(2.125b). Such a surface metric is then referred to as the *Riemannian metric*.

Recall from $(5.51)_4$ that $\left[g^{ij} \right] = \left[g_{ij} \right]^{-1}$. First, the inverse of $\left[g_{ij} \right]$ in (9.16) renders

$$
\left[g_{ij} \right]^{-1} = \begin{bmatrix} \left[a_{\alpha\beta} \right]^{-1}_{2\times 2} & [\mathbf{0}]_{2\times 1} \\ [\mathbf{0}]^{\mathrm{T}}_{1\times 2} & 1 \end{bmatrix} . \tag{9.21}
$$

Then, similarly to (9.16), the ambient contravariant metric coefficients $(5.46)_3$ can here be written as

$$
\left[g^{ij} \right] = \begin{bmatrix} \left[a^{\alpha\beta} \right]_{2\times 2} & [\mathbf{q}]_{2\times 1} \\ [\mathbf{q}]^{\mathrm{T}}_{1\times 2} & r \end{bmatrix} , \tag{9.22}
$$

where

$$
r = \mathbf{a}^3 \cdot \mathbf{a}^3 \quad , \quad \left[\mathbf{q} \right] = \left[\mathbf{a}^1 \cdot \mathbf{a}^3 \ \ \mathbf{a}^2 \cdot \mathbf{a}^3 \right]^{\mathrm{T}} , \tag{9.23}
$$

and $a^{\alpha\beta}$ present the *surface contravariant metric coefficients*; given by,

$$
\boxed{a^{\alpha\beta} = \mathbf{a}^\alpha \cdot \mathbf{a}^\beta = \mathbf{a}^\beta \cdot \mathbf{a}^\alpha = a^{\beta\alpha}\ ,} \tag{9.24}
$$

whose matrix form renders

$$
\boxed{\left[a^{\alpha\beta} \right] = \begin{bmatrix} \mathbf{a}^1 \cdot \mathbf{a}^1 & \mathbf{a}^1 \cdot \mathbf{a}^2 \\ \mathbf{a}^1 \cdot \mathbf{a}^2 & \mathbf{a}^2 \cdot \mathbf{a}^2 \end{bmatrix} .} \tag{9.25}
$$

Now, comparing (9.21) and (9.22) immediately gives

$$
\boxed{\underbrace{a_{\alpha\gamma} a^{\gamma\beta} = \delta^\beta_\alpha}_{\text{or } a^{\alpha\gamma} a_{\gamma\beta} = \delta^\alpha_\beta} \quad \text{or} \quad \underbrace{\left[a_{\alpha\beta} \right] = \left[a^{\alpha\beta} \right]^{-1}}_{\text{or } \left[a^{\alpha\beta} \right] = \left[a_{\alpha\beta} \right]^{-1}} .} \tag{9.26}
$$

One can also deduce that $\mathbf{a}^3$ is a vector of unit length in a direction perpendicular to a plane spanned by $\mathbf{a}^1$ and $\mathbf{a}^2$, that is,

$$
\mathbf{a}^3 \cdot \mathbf{a}^3 = 1 \quad , \quad \mathbf{a}^3 \cdot \mathbf{a}^\alpha = 0\ . \tag{9.27}
$$

Having obtained the required data, one can use (5.49) to finally arrive at the *surface contravariant basis vectors*

$$\boxed{\mathbf{a}^{\alpha} = a^{\alpha\beta}\mathbf{a}_{\beta}\ ,} \tag{9.28}$$

and, having in mind $\mathbf{a}_3 = \widehat{\mathbf{n}}$ together with $\widehat{\mathbf{n}} \cdot \mathbf{a}_{\alpha} = 0$,

$$\mathbf{a}^{3} = \mathbf{a}_3 = \widehat{\mathbf{n}} \quad \text{satisfying} \quad \widehat{\mathbf{n}} \cdot \mathbf{a}^{\alpha} = 0 \quad \text{and} \quad \widehat{\mathbf{n}} \cdot \mathbf{a}_{\alpha} = 0\ . \quad \circledast \tag{9.29}$$

The above results imply that the surface contravariant basis vectors $\mathbf{a}^1$ and $\mathbf{a}^2$ should also lie in the tangential plane defined by their corresponding dual vectors (see Fig. 9.10). The expression (9.11) thus admits the alternative representation

$$\widehat{\mathbf{n}} = \pm \frac{\mathbf{a}^1 \times \mathbf{a}^2}{\left|\mathbf{a}^1 \times \mathbf{a}^2\right|}\ . \tag{9.30}$$

From now on, all surfaces in this text are assumed to be **positively oriented** in such a way that

$$\boxed{\mathbf{a}_1 \times \mathbf{a}_2 = |\mathbf{a}_1 \times \mathbf{a}_2|\,\widehat{\mathbf{n}} \quad , \quad \mathbf{a}^1 \times \mathbf{a}^2 = \left|\mathbf{a}^1 \times \mathbf{a}^2\right|\widehat{\mathbf{n}}\ .} \tag{9.31}$$

From $(9.26)_1$ and (9.28), taking into account the replacement property of the Kronecker delta, one can obtain

$$\mathbf{a}_{\alpha} = a_{\alpha\beta}\mathbf{a}^{\beta}\ . \tag{9.32}$$

It is then easy to see that

$$\boxed{\mathbf{a}^{\alpha} \cdot \mathbf{a}_{\beta} = \delta^{\alpha}_{\beta} \quad , \quad \mathbf{a}_{\alpha} \cdot \mathbf{a}^{\beta} = \delta^{\beta}_{\alpha}\ .} \tag{9.33}$$

Regarding the established bases (9.13), (9.15) with $(9.29)_2$, the identity tensor in (5.78) now admits the following forms

$$\mathbf{I} = \overline{\mathbf{Z}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} = \overline{\mathbf{C}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} = \widetilde{\mathbf{Z}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} = \widetilde{\mathbf{C}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}\ , \tag{9.34}$$

where

$$\overline{\mathbf{Z}} = \mathbf{a}_{\alpha} \otimes \mathbf{a}^{\alpha}\ , \tag{9.35a}$$

$$\overline{\mathbf{C}} = \overline{\mathbf{Z}}^{\mathrm{T}}\overline{\mathbf{Z}} = a_{\alpha\beta}\mathbf{a}^{\alpha} \otimes \mathbf{a}^{\beta}\ , \tag{9.35b}$$

$$\widetilde{\mathbf{Z}} = \mathbf{a}^{\alpha} \otimes \mathbf{a}_{\alpha}\ , \tag{9.35c}$$

$$\widetilde{\mathbf{C}} = \widetilde{\mathbf{Z}}^{\mathrm{T}}\widetilde{\mathbf{Z}} = a^{\alpha\beta}\mathbf{a}_{\alpha} \otimes \mathbf{a}_{\beta}\ . \tag{9.35d}$$

The *surface permutation symbol* is defined by

$$\varepsilon_{\alpha\beta} = \beta - \alpha \quad \text{and, for consistency,} \quad \varepsilon^{\alpha\beta} = \beta - \alpha\ . \tag{9.36}$$

It is then easy to see that

$$\varepsilon_{\alpha\gamma}\varepsilon^{\gamma\beta} = -\delta_{\alpha}^{\beta} \quad , \quad \varepsilon^{\alpha\gamma}\varepsilon_{\gamma\beta} = -\delta_{\beta}^{\alpha} \quad , \quad \varepsilon^{\alpha\beta}\varepsilon_{\gamma\delta} = \delta_{\gamma}^{\alpha}\delta_{\delta}^{\beta} - \delta_{\delta}^{\alpha}\delta_{\gamma}^{\beta} \ . \tag{9.37}$$

From (9.37)$_3$, it follows that

$$\varepsilon^{\alpha\beta}\varepsilon_{\alpha\beta} = 2 \ . \tag{9.38}$$

One can finally establish

$$\det\left[\, a_{\alpha\beta} \,\right]\varepsilon_{\alpha\beta} = a_{\alpha\gamma}\varepsilon^{\gamma\delta}a_{\delta\beta} \quad , \quad \det\left[\, a^{\alpha\beta} \,\right]\varepsilon^{\alpha\beta} = a^{\alpha\gamma}\varepsilon_{\gamma\delta}a^{\delta\beta} \ . \tag{9.39}$$

9.1.3 Tensor Property of Surface Metric Coefficients

Consider a change of surface coordinates from $\left(t^1, t^2\right)$ to $\left(\bar{t}^1, \bar{t}^2\right)$. These **old** and **new** surface coordinates are linked by the following mutually inverse relations

$$\bar{t}^{\alpha} = \bar{t}^{\alpha}\left(t^1, t^2\right) \quad , \quad t^{\alpha} = t^{\alpha}\left(\bar{t}^1, \bar{t}^2\right) \ . \qquad \longleftarrow \text{see (5.54)} \tag{9.40}$$

As a result, the surface version of (5.55) can be written as

$$\mathbf{x} = \hat{\bar{\mathbf{x}}}^{s}\left(\bar{t}^1, \bar{t}^2\right) = \hat{\mathbf{x}}^{s}\left(t^1, t^2\right) \quad \text{with} \quad \left.\begin{array}{l} \bar{\mathbf{a}}_{\alpha} = \dfrac{\partial \mathbf{x}}{\partial \bar{t}^{\alpha}} \\ \mathbf{a}_{\alpha} = \dfrac{\partial \mathbf{x}}{\partial t^{\alpha}} \end{array}\right\} \ , \quad \left.\begin{array}{l} \bar{\mathbf{a}}^{\alpha} = \dfrac{\partial \bar{t}^{\alpha}}{\partial \mathbf{x}} \\ \mathbf{a}^{\alpha} = \dfrac{\partial t^{\alpha}}{\partial \mathbf{x}} \end{array}\right\} \ . \tag{9.41}$$

Now, the ambient relations (5.58a)–(5.60) translate to

$$\bar{\mathbf{a}}^{\alpha} \cdot \mathbf{a}_{\beta} = \frac{\partial \bar{t}^{\alpha}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial t^{\beta}} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\beta}} \ , \tag{9.42a}$$

$$\mathbf{a}^{\alpha} \cdot \bar{\mathbf{a}}_{\beta} = \frac{\partial t^{\alpha}}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \bar{t}^{\beta}} = \frac{\partial t^{\alpha}}{\partial \bar{t}^{\beta}} \ , \tag{9.42b}$$

$$\bar{\mathbf{a}}^{\alpha} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\beta}}\mathbf{a}^{\beta} \quad , \quad \mathbf{a}^{\alpha} = \frac{\partial t^{\alpha}}{\partial \bar{t}^{\beta}}\bar{\mathbf{a}}^{\beta} \ , \tag{9.42c}$$

$$\bar{\mathbf{a}}_{\alpha} = \frac{\partial t^{\beta}}{\partial \bar{t}^{\alpha}}\mathbf{a}_{\beta} \quad , \quad \mathbf{a}_{\alpha} = \frac{\partial \bar{t}^{\beta}}{\partial t^{\alpha}}\bar{\mathbf{a}}_{\beta} \ . \tag{9.42d}$$

Consequently, the variant $a_{\alpha\beta} = \mathbf{a}_{\alpha} \cdot \mathbf{a}_{\beta}$ may be called the *surface covariant metric tensor* because it obeys the following tensor transformation law

$$\boxed{\bar{a}_{\alpha\beta} = \frac{\partial t^{\theta}}{\partial \bar{t}^{\alpha}} a_{\theta\rho} \frac{\partial t^{\rho}}{\partial \bar{t}^{\beta}}} \quad \longleftarrow \text{see (5.61)} \tag{9.43}$$

With regard to this, the object $\delta^{\alpha}_{\beta} = \mathbf{a}^{\alpha} \cdot \mathbf{a}_{\beta}$ is also a tensor owing to

$$\bar{\delta}^{\alpha}_{\beta} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\theta}} \delta^{\theta}_{\rho} \frac{\partial t^{\rho}}{\partial \bar{t}^{\beta}} . \quad \longleftarrow \text{see (5.62)} \tag{9.44}$$

In a similar manner, the object $a^{\alpha\beta}$ - which may be called the *surface contravariant metric tensor* - transforms according to

$$\boxed{\bar{a}^{\alpha\beta} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\theta}} a^{\theta\rho} \frac{\partial \bar{t}^{\beta}}{\partial t^{\rho}}} \quad \longleftarrow \text{see (5.63)} \tag{9.45}$$

9.1.4 Shift Tensors and Related Identities

A *shift tensor* is a mathematical entity relating surface and ambient objects. Its important role in making the connection between the surface and ambient basis vectors, metric coefficients and tensorial variables will be demonstrated in the following.

The two surface covariant basis vectors may be expressed with respect to the three ambient covariant basis vectors as

$$\mathbf{a}_{\alpha} = \mathbf{I}\mathbf{a}_{\alpha} = \left(\mathbf{g}_i \otimes \mathbf{g}^i\right) \mathbf{a}_{\alpha} = \overline{Z}^{i}_{\alpha}\, \mathbf{g}_i \ , \tag{9.46}$$

where

$$\boxed{\overline{Z}^{i}_{\alpha} = \mathbf{g}^i \cdot \mathbf{a}_{\alpha} = \frac{\partial \Theta^i}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial t^{\alpha}} = \frac{\partial \Theta^i}{\partial t^{\alpha}}} . \tag{9.47}$$

Accordingly, one can rewrite (9.35a) as

$$\overline{\mathbf{Z}} = \mathbf{a}_{\alpha} \otimes \mathbf{a}^{\alpha} = \overline{Z}^{i}_{\alpha}\, \mathbf{g}_i \otimes \mathbf{a}^{\alpha} . \tag{9.48}$$

Notice that the surface tangent vectors remain unchanged under the action of the symmetric contra-covariant tensor $\overline{\mathbf{Z}}$, i.e. $\mathbf{a}_{\alpha} = \overline{\mathbf{Z}}\mathbf{a}_{\alpha}$. This **singular** tensor with the following matrix form

$$\left[\overline{\mathbf{Z}}\right] = \left[\overline{Z}^{i}_{\alpha}\right] = \begin{bmatrix} \mathbf{g}^1 \cdot \mathbf{a}_1 & \mathbf{g}^1 \cdot \mathbf{a}_2 \\ \mathbf{g}^2 \cdot \mathbf{a}_1 & \mathbf{g}^2 \cdot \mathbf{a}_2 \\ \mathbf{g}^3 \cdot \mathbf{a}_1 & \mathbf{g}^3 \cdot \mathbf{a}_2 \end{bmatrix} , \tag{9.49}$$

relative to the dyads $\mathbf{g}_i \otimes \mathbf{a}^\alpha$, is referred to as the *shift tensor*. It may be viewed as an object describing the tangent space. Notice that $\overline{Z}^i_\alpha$ can precisely be written as $\overline{Z}^{i}_{\,.\,\alpha}$ to declare that the first index in the shift tensor is an ambient index. The placeholder has been dropped here for notational simplicity.

Using (9.48)$_2$, the symmetric covariant tensor $\overline{\mathbf{C}}$, defined in (9.35b), can be rewritten as

$$\overline{\mathbf{C}} = \overline{\mathbf{Z}}^{\mathrm{T}}\overline{\mathbf{Z}} = \left(\overline{Z}^i_\alpha \, \mathbf{a}^\alpha \otimes \mathbf{g}_i\right)\left(\overline{Z}^j_\beta \, \mathbf{g}_j \otimes \mathbf{a}^\beta\right) = \overline{Z}^i_\alpha g_{ij} \overline{Z}^j_\beta \mathbf{a}^\alpha \otimes \mathbf{a}^\beta \; . \tag{9.50}$$

Comparing (9.35b)$_2$ and (9.50)$_3$ now leads to the following relationship between the surface and ambient covariant metric coefficients

$$\boxed{a_{\alpha\beta} = \overline{Z}^i_\alpha g_{ij} \overline{Z}^j_\beta \; .} \quad \longleftarrow \text{see (9.64)} \tag{9.51}$$

In matrix notation,

$$\left[\, a_{\alpha\beta} \,\right] = \left[\, \overline{\mathbf{C}} \,\right] = \left[\, \overline{\mathbf{Z}} \,\right]^{\mathrm{T}} \left[\, \mathbf{g} \,\right] \left[\, \overline{\mathbf{Z}} \,\right] = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} . \tag{9.52}$$

Note that in practice, the ambient space is usually referred to Cartesian coordinates with the metric coefficients being the Kronecker delta and all Christoffel elements being zero. In this case, the result (9.51) takes the form

$$\boxed{a_{\alpha\beta} = \overline{Z}_{i\alpha} \overline{Z}_{i\beta} \; .} \tag{9.53}$$

Making use of (9.53), the determinant of the symmetric matrix $\left[\, a_{\alpha\beta} \,\right]$ can be represented by

$$\begin{aligned} a := \det\left[\, a_{\alpha\beta} \,\right] &= a_{11}a_{22} - (a_{12})^2 \\ &= \overline{Z}_{i1}\, \overline{Z}_{i1}\, \overline{Z}_{j2}\, \overline{Z}_{j2} - \overline{Z}_{i1}\, \overline{Z}_{i2}\, \overline{Z}_{j1}\, \overline{Z}_{j2} \; . \end{aligned} \tag{9.54}$$

This result helps establish

$$\boxed{|\mathbf{a}_1 \times \mathbf{a}_2| = \sqrt{a} \; ,} \quad \longleftarrow \text{note that } a{>}0 \text{ according to (9.20)} \tag{9.55}$$

since, having in mind the bilinearity of the dot product and cross product,

$$\begin{aligned} |\mathbf{a}_1 \times \mathbf{a}_2|^2 &\xlongequal[\text{(1.11) and (9.46)}]{\text{from}} \left(\overline{Z}_{i1}\, \widehat{\mathbf{e}}_i \times \overline{Z}_{j2}\, \widehat{\mathbf{e}}_j\right) \cdot \left(\overline{Z}_{k1}\, \widehat{\mathbf{e}}_k \times \overline{Z}_{l2}\, \widehat{\mathbf{e}}_l\right) \\ &\xlongequal[\text{(1.64)}]{\text{from}} \left(\overline{Z}_{i1}\, \overline{Z}_{j2}\, \varepsilon_{ijm} \widehat{\mathbf{e}}_m\right) \cdot \left(\overline{Z}_{k1}\, \overline{Z}_{l2}\, \varepsilon_{kln} \widehat{\mathbf{e}}_n\right) \\ &\xlongequal[\text{(1.35)}]{\text{from}} \overline{Z}_{i1}\, \overline{Z}_{j2}\, \overline{Z}_{k1}\, \overline{Z}_{l2}\, \varepsilon_{ijm} \delta_{mn} \varepsilon_{kln} \end{aligned}$$

$$\xlongequal[\text{(1.36) and (1.58a)}]{\text{from}} \overline{Z}_{i1}\,\overline{Z}_{i1}\,\overline{Z}_{j2}\,\overline{Z}_{j2} - \overline{Z}_{i1}\,\overline{Z}_{i2}\,\overline{Z}_{j1}\,\overline{Z}_{j2}$$

$$\xlongequal[\text{(9.54)}]{\text{from}} a \ .$$

The interested reader may want to verify the result (9.55) in an alternative way:

$$\begin{aligned}
|\mathbf{a}_1 \times \mathbf{a}_2|^2 &\xlongequal[\text{(1.50)}]{\text{from}} |\mathbf{a}_1|^2\,|\mathbf{a}_2|^2 \sin^2\theta\,(\mathbf{a}_1,\mathbf{a}_2) \\
&\xlongequal[\text{(1.11)}]{\text{from}} (\mathbf{a}_1\cdot\mathbf{a}_1)\,(\mathbf{a}_2\cdot\mathbf{a}_2)\left[1-\cos^2\theta\,(\mathbf{a}_1,\mathbf{a}_2)\right] \\
&\xlongequal[\text{(1.12) and (9.17)}]{\text{from}} a_{11}a_{22}\left[1-\frac{(\mathbf{a}_1\cdot\mathbf{a}_2)^2}{|\mathbf{a}_1|^2\,|\mathbf{a}_2|^2}\right] \\
&\xlongequal[\text{(1.11) and (9.17)}]{\text{from}} a_{11}a_{22}-(a_{12})^2 \\
&\xlongequal[\text{(9.54)}]{\text{from}} a \ .
\end{aligned}$$

The result (9.55) helps compute the **area element** dA via the following expression (see Fig. 9.9)

$$\boxed{dA = \sqrt{a}\, dt^1 dt^2} \ , \tag{9.56}$$

noting that $d\mathbf{A} = \left(\mathbf{a}_1 dt^1\right) \times \left(\mathbf{a}_2 dt^2\right)$. From $(9.31)_1$, $(9.36)_1$ and (9.55), it follows that

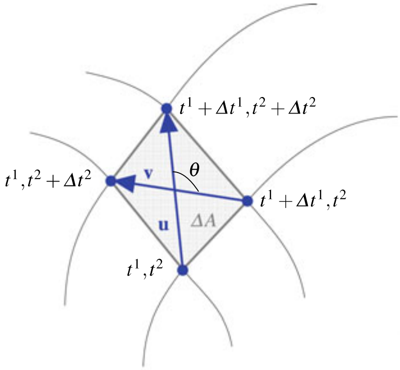
$$\boxed{\mathbf{a}_1 \times \mathbf{a}_2 = \sqrt{a}\,\widehat{\mathbf{n}} \quad \text{and, generally,} \quad \mathbf{a}_\alpha \times \mathbf{a}_\beta = \sqrt{a}\varepsilon_{\alpha\beta}\widehat{\mathbf{n}}\ .} \tag{9.57}$$

One can also establish

$$\boxed{\widehat{\mathbf{n}} \times \mathbf{a}_\alpha = \sqrt{a}\varepsilon_{\alpha\beta}\mathbf{a}^\beta} \ , \tag{9.58}$$

because

$$\begin{aligned}
\varepsilon^{\theta\rho}\sqrt{a}\varepsilon_{\theta\rho}\widehat{\mathbf{n}} \times \mathbf{a}_\alpha &\xlongequal[\text{hand from (9.38)}]{\text{on the one}} \left(2\sqrt{a}\right)\widehat{\mathbf{n}} \times \mathbf{a}_\alpha \\
&\xlongequal[\text{hand from (9.57)}]{\text{on the other}} \varepsilon^{\theta\rho}\left(\mathbf{a}_\theta \times \mathbf{a}_\rho\right) \times \mathbf{a}_\alpha \\
&\xlongequal[\text{(1.72)}]{\text{from}} \varepsilon^{\theta\rho}\left(\mathbf{a}_\theta\cdot\mathbf{a}_\alpha\right)\mathbf{a}_\rho - \varepsilon^{\theta\rho}\left(\mathbf{a}_\rho\cdot\mathbf{a}_\alpha\right)\mathbf{a}_\theta \\
&\xlongequal[\text{(9.17) and (9.32)}]{\text{from}} \left(a_{\alpha\theta}\varepsilon^{\theta\rho}a_{\rho\beta} - a_{\beta\theta}\varepsilon^{\theta\rho}a_{\rho\alpha}\right)\mathbf{a}^\beta \\
&\xlongequal[\text{(9.39) and (9.54)}]{\text{from}} a\left(\varepsilon_{\alpha\beta} - \varepsilon_{\beta\alpha}\right)\mathbf{a}^\beta \\
&\xlongequal[\text{(9.36)}]{\text{from}} \left(2\sqrt{a}\right)\sqrt{a}\varepsilon_{\alpha\beta}\mathbf{a}^\beta \ .
\end{aligned}$$



In this figure,

$\mathbf{u} = \Delta t^1 \mathbf{a}_1 + \Delta t^2 \mathbf{a}_2$,

$\mathbf{v} = -\Delta t^1 \mathbf{a}_1 + \Delta t^2 \mathbf{a}_2$,

$|\mathbf{u}|^2 = \Delta t^1 a_{11} \Delta t^1 + 2\Delta t^1 a_{12} \Delta t^2 + \Delta t^2 a_{22} \Delta t^2$,

$|\mathbf{v}|^2 = \Delta t^1 a_{11} \Delta t^1 - 2\Delta t^1 a_{12} \Delta t^2 + \Delta t^2 a_{22} \Delta t^2$,

$|\mathbf{u}|^2 |\mathbf{v}|^2 = \left[\Delta t^1 a_{11} \Delta t^1 + \Delta t^2 a_{22} \Delta t^2\right]^2 - \left[2\Delta t^1 a_{12} \Delta t^2\right]^2$,

$\mathbf{u} \cdot \mathbf{v} = -\Delta t^1 a_{11} \Delta t^1 + \Delta t^2 a_{22} \Delta t^2$.

Suppose that Δt^1 and Δt^2 are small enough such that $\mathbf{u}$ and $\mathbf{v}$ define a flat surface patch. The area, ΔA, of this patch is then given by

$$\Delta A = \frac{1}{2} |\mathbf{u}| |\mathbf{v}| \sin\theta = \frac{1}{2} |\mathbf{u}| |\mathbf{v}| \sqrt{1-\cos^2\theta} = \frac{1}{2} |\mathbf{u}| |\mathbf{v}| \sqrt{1 - \frac{(\mathbf{u}\cdot\mathbf{v})^2}{|\mathbf{u}|^2 |\mathbf{v}|^2}} = \frac{1}{2}\sqrt{|\mathbf{u}|^2 |\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2}$$

$$= \frac{1}{2}\sqrt{\cancel{a_{11}^2 (\Delta t^1)^4} + \cancel{a_{22}^2 (\Delta t^2)^4} + \left[2a_{11}a_{22} - 4a_{12}^2 + 2a_{11}a_{22}\right] (\Delta t^1)^2 (\Delta t^2)^2 - \cancel{a_{11}^2 (\Delta t^1)^4} - \cancel{a_{22}^2 (\Delta t^2)^4}}$$

$$= \sqrt{a_{11}a_{22} - a_{12}^2}\, \Delta t^1 \Delta t^2 = \sqrt{a}\, \Delta t^1 \Delta t^2 \ .$$

Fig. 9.9 Surface area element

Similarly to (9.46), the two surface contravariant basis vectors are now expressed with respect to the three ambient contravariant basis vectors as

$$\mathbf{a}^\alpha = \mathbf{I}\mathbf{a}^\alpha = \left(\mathbf{g}^i \otimes \mathbf{g}_i\right) \mathbf{a}^\alpha = \widetilde{Z}_i^\alpha \, \mathbf{g}^i \ , \tag{9.59}$$

where

$$\boxed{\widetilde{Z}_i^\alpha = \mathbf{g}_i \cdot \mathbf{a}^\alpha = \frac{\partial \mathbf{x}}{\partial \Theta^i} \cdot \frac{\partial t^\alpha}{\partial \mathbf{x}} = \frac{\partial t^\alpha}{\partial \Theta^i} \ .} \tag{9.60}$$

Consequently, the co-contravariant tensor $\widetilde{\mathbf{Z}}$ in (9.35c) can be rewritten as

$$\widetilde{\mathbf{Z}} = \mathbf{a}^\alpha \otimes \mathbf{a}_\alpha = \widetilde{Z}_i^\alpha \, \mathbf{g}^i \otimes \mathbf{a}_\alpha \ . \tag{9.61}$$

In this case, the surface contravariant basis vectors remain preserved under the action of this linear transformation, i.e. $\mathbf{a}^\alpha = \widetilde{\mathbf{Z}}\mathbf{a}^\alpha$. Its matrix form

$$\left[\widetilde{\mathbf{Z}}\right] = \left[\widetilde{Z}_i^\alpha\right] = \begin{bmatrix} \mathbf{g}_1 \cdot \mathbf{a}^1 & \mathbf{g}_1 \cdot \mathbf{a}^2 \\ \mathbf{g}_2 \cdot \mathbf{a}^1 & \mathbf{g}_2 \cdot \mathbf{a}^2 \\ \mathbf{g}_3 \cdot \mathbf{a}^1 & \mathbf{g}_3 \cdot \mathbf{a}^2 \end{bmatrix} , \tag{9.62}$$

relative to the dyads $\mathbf{g}^i \otimes \mathbf{a}_\alpha$, also acts as a **shift tensor**. Note that, similarly to $\overline{Z}^{\,i}_{\alpha}$, the placeholder has been dropped here for notational simplicity (indeed, $\widetilde{Z}_i^\alpha$ should be considered as $\widetilde{Z}_i{}^{\cdot\,\alpha}$). Now, substituting (9.61)$_2$ into (9.35d)$_1$ yields

$$\widetilde{\mathbf{C}} = \widetilde{\mathbf{Z}}^{\mathrm{T}}\widetilde{\mathbf{Z}} = \left(\widetilde{Z}_i^\alpha\, \mathbf{a}_\alpha \otimes \mathbf{g}^i\right)\left(\widetilde{Z}_j^\beta\, \mathbf{g}^j \otimes \mathbf{a}_\beta\right) = \widetilde{Z}_i^\alpha\, g^{ij}\, \widetilde{Z}_j^\beta \mathbf{a}_\alpha \otimes \mathbf{a}_\beta \; . \tag{9.63}$$

The uniqueness of the components of a tensor with respect to a given basis finally implies that

$$\boxed{a^{\alpha\beta} = \widetilde{Z}_i^\alpha\, g^{ij}\, \widetilde{Z}_j^\beta \; .} \quad \longleftarrow \text{see (9.117)} \tag{9.64}$$

In matrix representation,

$$\left[\, a^{\alpha\beta} \,\right] = \left[\, \widetilde{\mathbf{C}} \,\right] = \left[\, \widetilde{\mathbf{Z}} \,\right]^{\mathrm{T}} \left[\, \mathbf{g} \,\right]^{-1} \left[\, \widetilde{\mathbf{Z}} \,\right] = \frac{1}{a}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{bmatrix} . \tag{9.65}$$

In the case that the ambient space is coordinated with Cartesian coordinates, the result (9.64) reduces to

$$\boxed{a^{\alpha\beta} = \widetilde{Z}_i^\alpha\, \widetilde{Z}_i^\beta \; .} \tag{9.66}$$

From (9.26)$_1$, (9.54)$_1$ and (9.66), one will have

$$\begin{aligned} \frac{1}{a} &= \det\left[\, a^{\alpha\beta} \,\right] = a^{11}a^{22} - \left(a^{12}\right)^2 \\ &= \widetilde{Z}_i^1 \widetilde{Z}_i^1 \widetilde{Z}_j^2 \widetilde{Z}_j^2 - \widetilde{Z}_i^1 \widetilde{Z}_i^2 \widetilde{Z}_j^1 \widetilde{Z}_j^2 \; . \end{aligned} \tag{9.67}$$

Following similar procedures that led to (9.55) then reveals

$$\boxed{\left|\mathbf{a}^1 \times \mathbf{a}^2\right| = \frac{1}{\sqrt{a}} \; .} \tag{9.68}$$

From (9.31)$_2$, (9.36)$_2$ and (9.68), one immediately obtains

$$\boxed{\mathbf{a}^1 \times \mathbf{a}^2 = \frac{1}{\sqrt{a}}\widehat{\mathbf{n}} \quad \text{and, generally,} \quad \mathbf{a}^\alpha \times \mathbf{a}^\beta = \frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\widehat{\mathbf{n}} \; .} \tag{9.69}$$

One can also write

$$\boxed{\widehat{\mathbf{n}} \times \mathbf{a}^\alpha = \frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\mathbf{a}_\beta \; ,} \tag{9.70}$$

owing to

$$
\begin{aligned}
a^{\alpha\theta}\widehat{\mathbf{n}} \times \mathbf{a}_\theta
&\overset{\text{on the one hand}}{\underset{\text{from (9.28)}}{=\!=\!=\!=}} \widehat{\mathbf{n}} \times \mathbf{a}^\alpha \\
&\overset{\text{on the other hand}}{\underset{\text{from (9.28) and (9.58)}}{=\!=\!=\!=}} \sqrt{a}a^{\alpha\theta}\varepsilon_{\theta\rho}a^{\rho\beta}\mathbf{a}_\beta \\
&\overset{\text{from}}{\underset{\text{(9.39) and (9.67)}}{=\!=\!=\!=}} \frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\mathbf{a}_\beta \ .
\end{aligned}
$$

Let the ambient object

$$
\widehat{\mathbf{n}} = \underline{\widehat{n}}^{\,i}\,\mathbf{g}_i = \underline{\widehat{n}}_{\,i}\,\mathbf{g}^i \ , \tag{9.71}
$$

be the outward unit normal field acting along a positively oriented surface $\mathcal{S}$. The vector $\underline{\widehat{n}}^{\,i}$ and its corresponding covector $\underline{\widehat{n}}_{\,i}$ then render

$$
\underline{\widehat{n}}^{\,i} = \varepsilon^{ijk}\frac{\sqrt{a}\widetilde{Z}^\alpha_j\varepsilon_{\alpha\beta}\widetilde{Z}^\beta_k}{2J} \ , \tag{9.72a}
$$

$$
\underline{\widehat{n}}_{\,i} = \varepsilon_{ijk}\frac{J\overline{Z}^j_\alpha\varepsilon^{\alpha\beta}\overline{Z}^k_\beta}{2\sqrt{a}} \ . \tag{9.72b}
$$

Having in mind the bilinearity of the cross product, the relation (9.72b) follows from

$$
\begin{aligned}
\widehat{\mathbf{n}} &\overset{\text{from}}{\underset{\text{(9.11)}}{=\!=\!=}} \frac{\mathbf{a}_1 \times \mathbf{a}_2}{|\mathbf{a}_1 \times \mathbf{a}_2|} \\
&\overset{\text{from}}{\underset{\text{(9.46) and (9.55)}}{=\!=\!=\!=}} \frac{\overline{Z}^j_1\,\mathbf{g}_j \times \overline{Z}^k_2\,\mathbf{g}_k}{\sqrt{a}} \\
&\overset{\text{from}}{\underset{\text{(5.33)}}{=\!=\!=}} \frac{J\overline{Z}^j_1\,\overline{Z}^k_2}{\sqrt{a}}\varepsilon_{jki}\mathbf{g}^i \\
&\overset{\text{from the skew-symmetric}}{\underset{\text{property of the permutation symbol}}{=\!=\!=\!=\!=}} \frac{J\overline{Z}^j_\alpha\varepsilon^{\alpha\beta}\overline{Z}^k_\beta}{2\sqrt{a}}\varepsilon_{ijk}\mathbf{g}^i \ ,
\end{aligned}
$$

noting that

$$
\begin{aligned}
\overline{Z}^j_\alpha\varepsilon^{\alpha\beta}\overline{Z}^k_\beta\varepsilon_{ijk} &= \overline{Z}^j_1\overline{Z}^k_2\varepsilon_{jki} - \overline{Z}^j_2\overline{Z}^k_1\varepsilon_{jki} \\
&= \overline{Z}^j_1\overline{Z}^k_2\varepsilon_{jki} - \overline{Z}^k_2\overline{Z}^j_1\varepsilon_{kji} \\
&= \overline{Z}^j_1\overline{Z}^k_2\varepsilon_{jki} + \overline{Z}^k_2\overline{Z}^j_1\varepsilon_{jki} \\
&= 2\overline{Z}^j_1\overline{Z}^k_2\varepsilon_{jki} \ .
\end{aligned}
$$

And the result (9.72a) can be shown in an analogous manner.

The following identity holds true

$$\boxed{\overline{Z}^{i}_{\alpha}\widetilde{Z}^{\beta}_{i} = \delta^{\beta}_{\alpha}}\ , \tag{9.73}$$

since

$$\xrightarrow[\text{(9.33), (9.46) and (9.59)}]{\text{from}} \overline{Z}^{i}_{\alpha}\,\mathbf{g}_i \cdot \widetilde{Z}^{\beta}_{j}\,\mathbf{g}^j = \delta^{\beta}_{\alpha}$$
$$\xrightarrow[\text{dot product and (5.27)}]{\text{using the bilinearity of}} \overline{Z}^{i}_{\alpha}\delta^{j}_{i}\widetilde{Z}^{\beta}_{j} = \delta^{\beta}_{\alpha}$$
$$\xrightarrow[\text{(5.14)}]{\text{from}} \overline{Z}^{i}_{\alpha}\widetilde{Z}^{\beta}_{i} = \delta^{\beta}_{\alpha}\ .$$

The following identity also holds true

$$\boxed{\overline{Z}^{i}_{\alpha}\widetilde{Z}^{\alpha}_{j} + \underline{\widehat{n}}^{i}\,\underline{\widehat{n}}_{j} = \delta^{i}_{j}}\ , \tag{9.74}$$

since

$$\xrightarrow[\text{(9.34) and (9.35a)}]{\text{by using}} \mathbf{a}_\alpha \otimes \mathbf{a}^\alpha + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} = \mathbf{I}$$
$$\xrightarrow[\text{along with (9.46) and (9.59)}]{\text{by using the bilinearity of tensor product}} \overline{Z}^{i}_{\alpha}\widetilde{Z}^{\alpha}_{j}\,\mathbf{g}_i \otimes \mathbf{g}^j + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} = \mathbf{I}$$
$$\xrightarrow[\text{along with (5.78) and (9.71)}]{\text{by using the bilinearity of tensor product}} \overline{Z}^{i}_{\alpha}\widetilde{Z}^{\alpha}_{j}\,\mathbf{g}_i \otimes \mathbf{g}^j + \underline{\widehat{n}}^{i}\,\underline{\widehat{n}}_{j}\,\mathbf{g}_i \otimes \mathbf{g}^j = \mathbf{g}_i \otimes \mathbf{g}^i$$
$$\xrightarrow[\text{along with (2.43), (5.14) and (5.27)}]{\text{by premultiplying both sides with } \mathbf{g}^k} \overline{Z}^{k}_{\alpha}\widetilde{Z}^{\alpha}_{j}\,\mathbf{g}^j + \underline{\widehat{n}}^{k}\,\underline{\widehat{n}}_{j}\,\mathbf{g}^j = \mathbf{g}^k$$
$$\xrightarrow[\text{along with (5.14) and (5.27)}]{\text{by multiplying both sides with } \mathbf{g}_l} \overline{Z}^{k}_{\alpha}\widetilde{Z}^{\alpha}_{l} + \underline{\widehat{n}}^{k}\,\underline{\widehat{n}}_{l} = \delta^{k}_{l}$$
$$\xrightarrow[\text{and also } l \to j]{\text{by renaming } k \to i} \overline{Z}^{i}_{\alpha}\widetilde{Z}^{\alpha}_{j} + \underline{\widehat{n}}^{i}\,\underline{\widehat{n}}_{j} = \delta^{i}_{j}.$$

This identity, for instance, may be used to show that any vector lying in the tangent plane (called *surface vector*) satisfies

$$\hat{\mathbf{h}} \cdot \widehat{\mathbf{n}} = \underbrace{\underline{\hat{h}}^{\alpha}\overline{Z}^{i}_{\alpha}\,\underline{\widehat{n}}_{i}}_{= \underline{\hat{h}}^{j}\left(\widetilde{Z}^{\alpha}_{j}\overline{Z}^{i}_{\alpha}\right)\underline{\widehat{n}}_{i}} = \underbrace{\underline{\hat{h}}^{j}\left(\delta^{i}_{j} - \underline{\widehat{n}}^{i}\,\underline{\widehat{n}}_{j}\right)\underline{\widehat{n}}_{i}}_{= \underline{\hat{h}}^{i}\,\underline{\widehat{n}}_{i} - \underline{\hat{h}}^{j}\,\underline{\widehat{n}}_{j}} = 0\ . \quad \longleftarrow \text{see (9.84) and (9.85a)} \tag{9.75}$$

9.1.5 Tensor Property of Shift Tensors

The object $\overline{Z}^{i}_{\alpha}$ transforms tensorially as implied by its name. To show this, consider the **old** and **new** surface coordinates $\left(t^1, t^2\right)$ and $\left(\bar{t}^{\,1}, \bar{t}^{\,2}\right)$, respectively. Consider also the **old** and **new** ambient coordinates $\left(\Theta^1, \Theta^2, \Theta^3\right)$ and $\left(\bar{\Theta}^1, \bar{\Theta}^2, \bar{\Theta}^3\right)$, respectively. Then, these systems of coordinates are related through

$$\bar{\Theta}^i\left(\bar{t}^\alpha\right) = \bar{\Theta}^i\left\{\Theta^j\left[t^\beta\left(\bar{t}^\alpha\right)\right]\right\} . \tag{9.76}$$

Consequently, the shift tensor $\overline{Z}^{\,i}_{\alpha}$, according to (9.47)$_3$, obeys

$$\boxed{\overline{Z}^{\,i}_{\alpha} = \frac{\partial \bar{\Theta}^i}{\partial \bar{t}^\alpha} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial t^\beta}\frac{\partial t^\beta}{\partial \bar{t}^\alpha} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j} Z^{\,j}_{\beta}\frac{\partial t^\beta}{\partial \bar{t}^\alpha} .} \tag{9.77}$$

Consistent with this, the variant $\widetilde{Z}^{\alpha}_{i}$ in (9.60)$_3$ also preserves the tensor property (indeed, it is deserved to be called a tensor). Making use of

$$\bar{t}^\alpha\left(\bar{\Theta}^i\right) = \bar{t}^\alpha\left\{t^\beta\left[\Theta^j\left(\bar{\Theta}^i\right)\right]\right\} , \tag{9.78}$$

this is indicated by

$$\boxed{\bar{\widetilde{Z}}^{\alpha}_{i} = \frac{\partial \bar{t}^\alpha}{\partial \bar{\Theta}^i} = \frac{\partial \bar{t}^\alpha}{\partial t^\beta}\frac{\partial t^\beta}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial \bar{\Theta}^i} = \frac{\partial \bar{t}^\alpha}{\partial t^\beta}\widetilde{Z}^{\beta}_{j}\frac{\partial \Theta^j}{\partial \bar{\Theta}^i} .} \tag{9.79}$$

9.1.6 Surface Vectors and Tensors

Let **u** be a vector with the following representations

$$\mathbf{u} = \underline{u}^\alpha \mathbf{a}_\alpha = \underline{u}_\alpha \mathbf{a}^\alpha . \tag{9.80}$$

Such a vector, that lies in the tangential plane, is called the *surface vector*. It also admits the ambient forms (5.64a)–(5.64b). The contravariant and covariant components of a surface vector are related by

$$\underline{u}_\alpha = a_{\alpha\beta}\,\underline{u}^\beta \quad , \quad \underline{u}^\alpha = a^{\alpha\beta}\,\underline{u}_\beta . \quad \longleftarrow \text{see (5.66a)} \tag{9.81}$$

In a similar manner, a *surface tensor* **A** can **intrinsically** be expressed as

$$\mathbf{A} = \underline{A}^{\alpha\beta}\mathbf{a}_\alpha \otimes \mathbf{a}_\beta = \underline{A}^{\alpha}_{\,.\beta}\mathbf{a}_\alpha \otimes \mathbf{a}^\beta = \underline{A}_{\alpha}^{\,.\beta}\mathbf{a}^\alpha \otimes \mathbf{a}_\beta = \underline{A}_{\alpha\beta}\mathbf{a}^\alpha \otimes \mathbf{a}^\beta . \tag{9.82}$$

And it **extrinsically** admits the decompositions (5.73a)–(5.73d). By index juggling, one will have

$$\underline{A}_{\alpha\beta} = a_{\alpha\gamma}\,\underline{A}^{\gamma}_{\,.\beta} = \underline{A}_{\alpha}^{\,.\gamma}a_{\gamma\beta} = a_{\alpha\gamma}\,\underline{A}^{\gamma\delta}a_{\delta\beta} , \quad \longleftarrow \text{see (5.76a)} \tag{9.83a}$$

$$\underline{A}_{\alpha}^{\,.\beta} = a_{\alpha\gamma}\,\underline{A}^{\gamma\beta} = \underline{A}_{\alpha\gamma}a^{\gamma\beta} = a_{\alpha\gamma}\,\underline{A}^{\gamma}_{\,.\delta}a^{\delta\beta} , \tag{9.83b}$$

$$\underline{A}^{\alpha}_{\ .\beta} = a^{\alpha\gamma}\, \underline{A}_{\gamma\beta} = \underline{A}^{\alpha\delta} a_{\delta\beta} = a^{\alpha\gamma}\, \underline{A}_{\gamma}^{\ .\delta} a_{\delta\beta} \ , \tag{9.83c}$$

$$\underline{A}^{\alpha\beta} = a^{\alpha\gamma}\, \underline{A}_{\gamma}^{\ .\beta} = \underline{A}^{\alpha}_{\ .\gamma}\, a^{\gamma\beta} = a^{\alpha\gamma}\, \underline{A}_{\gamma\delta} a^{\delta\beta} \ . \tag{9.83d}$$

All surface vector and tensor variables in this text satisfy

$$\boxed{\mathbf{u} \cdot \widehat{\mathbf{n}} = \widehat{\mathbf{n}} \cdot \mathbf{u} = 0 \quad , \quad \mathbf{A}\widehat{\mathbf{n}} = \widehat{\mathbf{n}}\mathbf{A} = \mathbf{0} \ .} \tag{9.84}$$

Let $\mathbf{u} = \underline{u}^i \mathbf{g}_i = \underline{u}^{\alpha} \mathbf{a}_{\alpha}$ or $\mathbf{u} = \underline{u}_i \mathbf{g}^i = \underline{u}_{\alpha} \mathbf{a}^{\alpha}$ be a surface vector. Then, its ambient and surface components will be related through the shift tensors $\overline{Z}^i_{\alpha}$ and $\widetilde{Z}^{\alpha}_i$ as follows:

$$\underline{u}^i = \overline{Z}^i_{\alpha}\, \underline{u}^{\alpha} \quad , \quad \underline{u}^{\alpha} = \widetilde{Z}^{\alpha}_i\, \underline{u}^i \ , \tag{9.85a}$$

$$\underline{u}_i = \widetilde{Z}^{\alpha}_i\, \underline{u}_{\alpha} \quad , \quad \underline{u}_{\alpha} = \overline{Z}^i_{\alpha}\, \underline{u}_i \ . \tag{9.85b}$$

As can be seen, the object $\overline{Z}^i_{\alpha}$ helps shift the contravariant (covariant) components of $\mathbf{u}$ from surface (ambient) to ambient (surface). And the object $\widetilde{Z}^{\alpha}_i$ helps shift the contravariant (covariant) components of $\mathbf{u}$ from ambient (surface) to surface (ambient). It is not then difficult to see that

$$\underline{A}^{ij} = \overline{Z}^i_{\alpha}\, \underline{A}^{\alpha\beta} \overline{Z}^j_{\beta} \quad , \quad \underline{A}^{\alpha\beta} = \widetilde{Z}^{\alpha}_i\, \underline{A}^{ij} \widetilde{Z}^{\beta}_j \ , \tag{9.86a}$$

$$\underline{A}^{i}_{\ .j} = \overline{Z}^i_{\alpha}\, \underline{A}^{\alpha}_{\ .\beta} \widetilde{Z}^{\beta}_j \quad , \quad \underline{A}^{\alpha}_{\ .\beta} = \widetilde{Z}^{\alpha}_i\, \underline{A}^{i}_{\ .j} \overline{Z}^j_{\beta} \ , \tag{9.86b}$$

$$\underline{A}_i^{\ .j} = \widetilde{Z}^{\alpha}_i\, \underline{A}_{\alpha}^{\ .\beta} \overline{Z}^j_{\beta} \quad , \quad \underline{A}_{\alpha}^{\ .\beta} = \overline{Z}^i_{\alpha}\, \underline{A}_i^{\ .j} \widetilde{Z}^{\beta}_j \ , \tag{9.86c}$$

$$\underline{A}_{ij} = \widetilde{Z}^{\alpha}_i\, \underline{A}_{\alpha\beta} \widetilde{Z}^{\beta}_j \quad , \quad \underline{A}_{\alpha\beta} = \overline{Z}^i_{\alpha}\, \underline{A}_{ij} \overline{Z}^j_{\beta} \ . \tag{9.86d}$$

Tensor property of components of surface vectors and tensors. Recall from the relations in (5.104) that a vector was an invariant while its components, according to (5.105a)–(5.105b), did not remain invariant under a change of coordinates. In this regard, the variant $\underline{u}^i$ ($\underline{u}_i$) was called a contravariant (covariant) tensor because it obeyed the transformation rule (5.105a) ((5.105b)).

Now, let $\mathbf{u}$ be a surface vector and consider a change of surface coordinates from (t^1, t^2) to $(\bar{t}^1, \bar{t}^2)$. The invariance of this object is then written by

$$\begin{aligned} \mathbf{u} &= \underline{\bar{u}}^{\alpha} \bar{\mathbf{a}}_{\alpha} = \underline{\bar{u}}_{\alpha} \bar{\mathbf{a}}^{\alpha} \\ &= \underline{u}^{\alpha} \mathbf{a}_{\alpha} = \underline{u}_{\alpha} \mathbf{a}^{\alpha} \ , \end{aligned} \tag{9.87}$$

while the tensor property of its components is characterized by the following rules

$$\underline{\bar{u}}^{\alpha} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\beta}}\, \underline{u}^{\alpha} \ , \qquad \longleftarrow \text{see (9.42c)} \tag{9.88a}$$

$$\underline{\bar{u}}_{\alpha} = \frac{\partial t^{\beta}}{\partial \bar{t}^{\alpha}}\, \underline{u}_{\beta} \ . \tag{9.88b}$$

Accordingly, the variant $\underline{u}^{\alpha}$ ($\underline{u}_{\alpha}$) represents a first-order contravariant (covariant) surface tensor. In a similar manner,

$$\bar{\underline{A}}^{\alpha\beta} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}} \, \underline{A}^{\gamma\delta} \frac{\partial \bar{t}^{\beta}}{\partial t^{\delta}} \,, \tag{9.89a}$$

$$\bar{\underline{A}}^{\alpha}_{\cdot\beta} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}} \, \underline{A}^{\gamma}_{\cdot\delta} \frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} \,, \tag{9.89b}$$

$$\bar{\underline{A}}_{\alpha}^{\cdot\beta} = \frac{\partial t^{\gamma}}{\partial \bar{t}^{\alpha}} \, \underline{A}_{\gamma}^{\cdot\delta} \frac{\partial \bar{t}^{\beta}}{\partial t^{\delta}} \,, \tag{9.89c}$$

$$\bar{\underline{A}}_{\alpha\beta} = \frac{\partial t^{\gamma}}{\partial \bar{t}^{\alpha}} \, \underline{A}_{\gamma\delta} \frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} \,. \tag{9.89d}$$

Note that any vector (or tensor) in this chapter is a surface vector (or surface tensor), if not otherwise stated. And surface tensorial field variables naturally vary from point to point within the surface (note that they are undefined outside the surface). Thus, such tensor fields can only be differentiated along the surface.

Recall from Chap. 7 that the gradient of an ambient tensor field was established on the basis of the partial derivative of that tensor field with respect to the curvilinear coordinates. And the partial derivative of a tensorial field variable was expressed in terms of the covariant derivative of its components. Recall also that the covariant derivative crucially relied on the Christoffel symbols. The same procedure can be used here. But some important properties established for the ambient space do not properly work here due to curvature. In what follows, the goal is thus to characterize the Christoffel symbols and covariant derivative on the surface.

9.2 Gauss and Weingarten Formulas

9.2.1 Surface Christoffel Symbols of Second Kind

Similarly to (7.5), consider the following family of objects

$$\boldsymbol{\Gamma}_{\alpha\beta} := \frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} \,, \tag{9.90}$$

with the symmetry in the indices α and β, that is,

$$\boldsymbol{\Gamma}_{\alpha\beta} = \frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} = \frac{\partial^2 \mathbf{x}}{\partial t^{\beta} \partial t^{\alpha}} = \frac{\partial^2 \mathbf{x}}{\partial t^{\alpha} \partial t^{\beta}} = \frac{\partial \mathbf{a}_{\beta}}{\partial t^{\alpha}} = \boldsymbol{\Gamma}_{\beta\alpha} \,. \tag{9.91}$$

Then,

$$\boxed{\Gamma^{\gamma}_{\alpha\beta} = \frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} \cdot \mathbf{a}^{\gamma} \quad \text{satisfying} \quad \Gamma^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\beta\alpha} \,,} \tag{9.92}$$

are referred to as *surface Christoffel symbols of the second kind* or simply *surface Christoffel symbols* (see the pioneering work of Christoffel [10]). They are also known as the *connection coefficients*. For the two-dimensional space under consideration, they exhibit an array of 6 independent quantities.

Recall from *Hint* on Sect. 7.1.1 that the ambient Christoffel symbols were not the components of a third-order tensor. The object $\Gamma^{\gamma}_{\alpha\beta}$ is also not a tensor as implied by its name. To show its nontensorial transformation, consider an old basis $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}$ and a new basis $\{\bar{\mathbf{a}}_1, \bar{\mathbf{a}}_2, \widehat{\mathbf{n}}\}$ with the corresponding Christoffel symbols $\Gamma^{\gamma}_{\alpha\beta}$ and $\bar{\Gamma}^{\gamma}_{\alpha\beta}$. They are associated with the old and new surface coordinates $\left(t^1, t^2\right)$ and $\left(\bar{t}^1, \bar{t}^2\right)$ which are related through the relations given in (9.40). Following similar procedures that led to (7.23) then reveals

$$\boxed{\bar{\Gamma}^{\gamma}_{\alpha\beta} = \underbrace{\frac{\partial t^{\delta}}{\partial \bar{t}^{\alpha}} \frac{\partial t^{\theta}}{\partial \bar{t}^{\beta}} \frac{\partial \bar{t}^{\gamma}}{\partial t^{\rho}} \Gamma^{\rho}_{\delta\theta}}_{\text{tensorial part}} + \underbrace{\frac{\partial^2 t^{\rho}}{\partial \bar{t}^{\alpha} \partial \bar{t}^{\beta}} \frac{\partial \bar{t}^{\gamma}}{\partial t^{\rho}}}_{\text{nontensorial portion}} \ .} \tag{9.93}$$

9.2.2 Surface Second-Order Curvature Tensor

In contrast to (7.8), it is impossible to have $\partial \mathbf{a}_{\alpha} / \partial t^{\beta} = \Gamma^{\gamma}_{\alpha\beta} \mathbf{a}_{\gamma}$ for curved surfaces. The reason is that the sensitivity of the surface tangent vectors with respect to the Gaussian coordinates are not restricted to lie in the tangent plane. In general, they may have components in the normal direction.

With regard to this, the surface version of (7.8), known as the *Gauss formulas*, reads

$$\boxed{\frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} = \Gamma^{\gamma}_{\alpha\beta} \mathbf{a}_{\gamma} + \underline{b}_{\alpha\beta} \widehat{\mathbf{n}} \ ,} \tag{9.94}$$

where $\underline{b}_{\alpha\beta}$ is known as the *surface covariant curvature tensor*. Using $(9.12)_{1-2}$, (9.90), $(9.91)_5$ and (9.94), one will have

$$\underline{b}_{\alpha\beta} = \frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} \cdot \widehat{\mathbf{n}} = \frac{\partial \mathbf{a}_{\beta}}{\partial t^{\alpha}} \cdot \widehat{\mathbf{n}} = \underline{b}_{\beta\alpha} \ . \tag{9.95}$$

The reason for introducing the variant $\underline{b}_{\alpha\beta}$ as a tensor is that, under a change of surface coordinates from $\left(t^1, t^2\right)$ to $\left(\bar{t}^1, \bar{t}^2\right)$, it **tensorially** transforms according to

$$\begin{aligned}
\bar{\underline{b}}_{\alpha\beta} &= \frac{\partial}{\partial \bar{t}^{\beta}} \left[\bar{\mathbf{a}}_{\alpha} \right] \cdot \widehat{\mathbf{n}} \\
&= \frac{\partial}{\partial \bar{t}^{\beta}} \left[\frac{\partial t^{\theta}}{\partial \bar{t}^{\alpha}} \mathbf{a}_{\theta} \right] \cdot \widehat{\mathbf{n}} \\
&= \frac{\partial^2 t^{\theta}}{\partial \bar{t}^{\beta} \partial \bar{t}^{\alpha}} \mathbf{a}_{\theta} \cdot \widehat{\mathbf{n}} + \frac{\partial t^{\theta}}{\partial \bar{t}^{\alpha}} \frac{\partial \mathbf{a}_{\theta}}{\partial \bar{t}^{\beta}} \cdot \widehat{\mathbf{n}}
\end{aligned}$$

$$= \frac{\partial t^{\theta}}{\partial \bar{t}^{\alpha}} \frac{\partial \mathbf{a}_{\theta}}{\partial t^{\rho}} \frac{\partial t^{\rho}}{\partial \bar{t}^{\beta}} \cdot \widehat{\mathbf{n}}$$
$$= \frac{\partial t^{\theta}}{\partial \bar{t}^{\alpha}} \underline{b}_{\theta\rho} \frac{\partial t^{\rho}}{\partial \bar{t}^{\beta}} \; . \qquad \longleftarrow \text{see (9.43)} \tag{9.96}$$

The surface covariant curvature tensor can also be written in the following equivalent versions

$$\begin{aligned} \underline{b}_{\alpha\beta} &= \frac{\partial^2 \mathbf{x}}{\partial t^{\alpha} \partial t^{\beta}} \cdot \widehat{\mathbf{n}} \\ &= -\frac{\partial \mathbf{x}}{\partial t^{\alpha}} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^{\beta}} \\ &= -\frac{1}{2} \left[\frac{\partial \mathbf{x}}{\partial t^{\alpha}} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^{\beta}} + \frac{\partial \mathbf{x}}{\partial t^{\beta}} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^{\alpha}} \right] . \end{aligned} \tag{9.97}$$

This covariant symmetric second-order tensor in matrix representation renders

$$\left[\underline{b}_{\alpha\beta} \right] = \begin{bmatrix} \underline{b}_{11} & \underline{b}_{12} \\ \underline{b}_{12} & \underline{b}_{22} \end{bmatrix} . \tag{9.98}$$

As can be seen from the Gauss relations (9.94), the nontensorial object $\Gamma^{\gamma}_{\alpha\beta}$ and tensorial variable $\underline{b}_{\alpha\beta}$ help express the partial derivatives $\partial \mathbf{a}_{\alpha} / \partial t^{\beta}$ with respect to the ambient basis $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}$. See Fig. 9.10 for a geometrical interpretation. Notice that the surface covariant curvature tensor has an **extrinsic** nature. A more detailed discussion of curvature is given in Sect. 9.7.

Other useful relations regard the partial derivatives of the unit normal vector to the surface with respect to the Gaussian coordinates. They are referred to as the *Weingarten formulas*[9]:

$$\boxed{\frac{\partial \widehat{\mathbf{n}}}{\partial t^{\alpha}} = -\,\underline{b}_{\alpha}^{\cdot\beta} \mathbf{a}_{\beta} = -\,\underline{b}_{\alpha\beta} \mathbf{a}^{\beta} \; ,} \tag{9.99}$$

where

$$\underline{b}_{\alpha}^{\cdot\beta} = \underline{b}_{\alpha\gamma} a^{\gamma\beta} = a^{\beta\gamma} \underline{b}_{\gamma\alpha} = \underline{b}^{\beta}_{\cdot\alpha} \; , \tag{9.100}$$

or

$$\boxed{\underline{b}_{\alpha}^{\cdot\beta} = -\frac{\partial \widehat{\mathbf{n}}}{\partial t^{\alpha}} \cdot \mathbf{a}^{\beta} \; ,} \tag{9.101}$$

[9] The proof is not difficult. First, by means of the relations $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$ and $\widehat{\mathbf{n}} \cdot \mathbf{a}_{\beta} = 0$, one will have $\widehat{\mathbf{n}} \cdot \partial \widehat{\mathbf{n}} / \partial t^{\alpha} = 0$ which implies that $\partial \widehat{\mathbf{n}} / \partial t^{\alpha} = T_{\alpha}^{\cdot\beta} \mathbf{a}_{\beta}$ where $T_{\alpha}^{\cdot\beta}$ are unknown components to be determined. Then, by using the Gauss formulas, the partial derivatives of $\widehat{\mathbf{n}} \cdot \mathbf{a}_{\gamma} = 0$ lead to $\partial \widehat{\mathbf{n}} / \partial t^{\alpha} \cdot \mathbf{a}_{\gamma} = -\,\underline{b}_{\alpha\gamma}$ or $T_{\alpha}^{\cdot\beta} a_{\beta\gamma} = -\,\underline{b}_{\alpha\gamma}$. Finally, using the identity $a_{\beta\gamma} a^{\gamma\theta} = \delta^{\theta}_{\beta}$, one will finally arrive at $T_{\alpha}^{\cdot\theta} = -\,\underline{b}_{\alpha\gamma} a^{\gamma\theta} = -\,\underline{b}_{\alpha}^{\cdot\theta}$ or $T_{\alpha}^{\cdot\beta} = -\,\underline{b}_{\alpha}^{\cdot\beta}$.

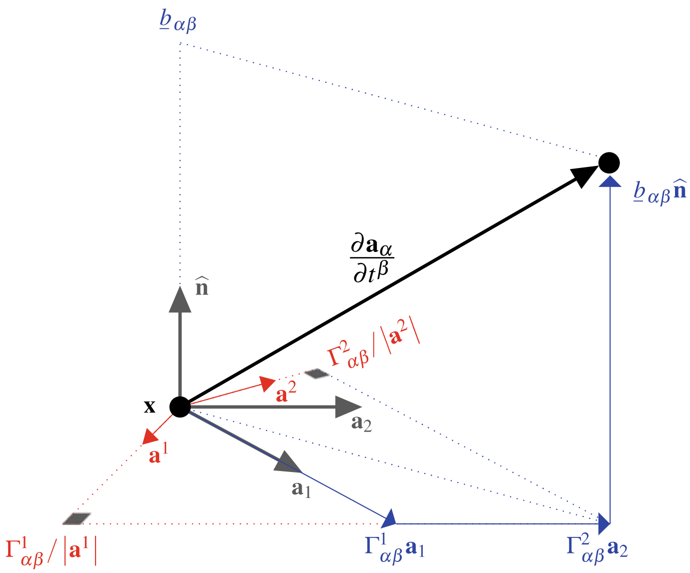


Fig. 9.10 Projection of the partial derivative $\partial\mathbf{a}_\alpha/\partial t^\beta$ onto the ambient basis $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}$

is called the *surface mixed curvature tensor*. The way that the unit normal vector to the surface tilts as one moves infinitesimally from one point to another is thus captured by this crucially important tensor. In matrix form, it renders

$$\left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \begin{bmatrix} \underline{b}_{1}^{\cdot 1} & \underline{b}_{1}^{\cdot 2} \\ \underline{b}_{2}^{\cdot 1} & \underline{b}_{2}^{\cdot 2} \end{bmatrix} = \begin{bmatrix} \underline{b}_{11}a^{11} + \underline{b}_{12}a^{12} & \underline{b}_{11}a^{12} + \underline{b}_{12}a^{22} \\ \underline{b}_{12}a^{11} + \underline{b}_{22}a^{12} & \underline{b}_{12}a^{12} + \underline{b}_{22}a^{22} \end{bmatrix} . \tag{9.102}$$

9.2.3 Mean and Gaussian Curvatures

The principal scalar invariants of the matrix $(9.102)_1$ are of crucial importance in differential geometry of surfaces. The half of its trace is called the *mean curvature*:

$$\boxed{\overline{\mathrm{H}} = \frac{1}{2}\mathrm{tr}\left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \frac{1}{2}\,\underline{b}_{\alpha}^{\cdot\alpha} = \frac{1}{2}\left(\underline{b}_{1}^{\cdot 1} + \underline{b}_{2}^{\cdot 2}\right) .} \tag{9.103}$$

And its determinant is referred to as the *Gaussian curvature*:

$$\boxed{\overline{\mathrm{K}} = \det\left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \underline{b}_{1}^{\cdot 1}\,\underline{b}_{2}^{\cdot 2} - \underline{b}_{1}^{\cdot 2}\,\underline{b}_{2}^{\cdot 1} .} \tag{9.104}$$

The mean curvature $\overline{\mathrm{H}}$ of a regular surface $\mathcal{S}$ at a generic point P is basically a measure of the rate of change of the area of a bounded domain of $\mathcal{S}$ in a neighborhood of P. With regard to this, a surface is said to be *minimal* if its mean curvature vanishes everywhere. For the characterization of such a surface of minimum area, see Exercise 9.11. A well-known example of a minimal surface regards **catenoid** (see Figs. 9.37 and 9.42).

It is important to note that the Gaussian curvature can solely be captured by measuring distances on the surfaces. And this reveals the fact this fundamental object has basically an **intrinsic** nature, see (9.483).

From (1.78a), $(9.69)_1$, (9.101) and $(9.104)_2$, the Gaussian curvature can be represented by

$$\boxed{\overline{\mathrm{K}} = \frac{1}{\sqrt{a}} \left(\frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \times \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} \right) \cdot \widehat{\mathbf{n}} \ .} \tag{9.105}$$

9.2.4 *Surface Christoffel Symbols of First Kind and Levi-Civita Connection*

Of interest here is to focus on the discussions regarding the Christoffel symbols. The objects in (9.92) may also be represented by

$$\boxed{\Gamma^{\alpha}_{\beta\gamma} = -\frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\gamma}} \cdot \mathbf{a}_{\beta} \ ,} \tag{9.106}$$

since

$$\text{from} \quad \mathbf{a}^{\alpha} \cdot \mathbf{a}_{\rho} = \delta^{\alpha}_{\rho} \quad \text{one obtains} \quad \frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\beta}} \cdot \mathbf{a}_{\rho} + \mathbf{a}^{\alpha} \cdot \left[\Gamma^{\gamma}_{\rho\beta} \, \mathbf{a}_{\gamma} + \underline{b}_{\rho\beta} \, \widehat{\mathbf{n}} \right] = 0 \ ,$$

and, therefore,

$$\frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\beta}} \cdot \mathbf{a}_{\rho} + \Gamma^{\gamma}_{\rho\beta} \delta^{\alpha}_{\gamma} = 0 \quad \text{gives} \quad \Gamma^{\alpha}_{\rho\beta} = -\frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\beta}} \cdot \mathbf{a}_{\rho} \quad \text{or} \quad \Gamma^{\alpha}_{\beta\gamma} = -\frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\gamma}} \cdot \mathbf{a}_{\beta} \ .$$

With the aid of $(9.12)_1$ and (9.94), the partial derivatives of the surface covariant metric coefficients $a_{\alpha\beta} = \mathbf{a}_{\alpha} \cdot \mathbf{a}_{\beta}$ render

$$\frac{\partial a_{\alpha\beta}}{\partial t^{\gamma}} = \Gamma^{\rho}_{\gamma\alpha} \, a_{\rho\beta} + \Gamma^{\rho}_{\gamma\beta} \, a_{\rho\alpha} \ . \tag{9.107}$$

It is now easy to arrive at

$$\frac{\partial a^{\alpha\beta}}{\partial t^{\gamma}} = -\Gamma^{\alpha}_{\gamma\rho} \, a^{\rho\beta} - \Gamma^{\beta}_{\gamma\rho} \, a^{\rho\alpha} \ , \tag{9.108}$$

and

$$\frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\beta}} = -\Gamma^{\alpha}_{\beta\gamma}\, \mathbf{a}^{\gamma} + \underline{b}^{\alpha}_{.\beta}\, \widehat{\mathbf{n}} \quad \text{where} \quad \underline{b}^{\alpha}_{.\beta} = a^{\alpha\theta}\, \underline{b}_{\theta\beta} = \underline{b}_{\beta}^{.\alpha} \ . \tag{9.109}$$

Making use of (9.28), i.e. $\mathbf{a}^{\theta} = a^{\theta\rho}\, \mathbf{a}_{\rho}$, and $(9.26)_1$, i.e. $a_{\gamma\theta} a^{\theta\rho} = \delta^{\rho}_{\gamma}$, the superscript index of $\Gamma^{\theta}_{\alpha\beta}$ in (9.92) can be lowered:

$$\Gamma_{\alpha\beta\gamma} = \frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} \cdot \mathbf{a}_{\gamma} \ . \tag{9.110}$$

The quantities $\Gamma_{\alpha\beta\gamma}$ that also possess symmetry in the first two indices are called *surface Christoffel symbols of the first kind*. These 6 independent quantities

$$\Gamma_{\alpha\beta\gamma} = \Gamma^{\theta}_{\alpha\beta}\, a_{\theta\gamma} \ , \tag{9.111}$$

help establish

$$\frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} = \Gamma_{\alpha\beta\gamma}\, \mathbf{a}^{\gamma} + \underline{b}_{\alpha\beta} \widehat{\mathbf{n}} \ . \tag{9.112}$$

By means of (9.107) and (9.111), the surface version of (7.14) becomes

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2} \left(\frac{\partial a_{\gamma\alpha}}{\partial t^{\beta}} + \frac{\partial a_{\gamma\beta}}{\partial t^{\alpha}} - \frac{\partial a_{\alpha\beta}}{\partial t^{\gamma}} \right) \ . \tag{9.113}$$

With the aid of $(9.26)_1$, (9.111) and (9.113), the surface counterpart of (7.15) will take the following form

$$\Gamma^{\gamma}_{\alpha\beta} = \frac{1}{2} \left(\frac{\partial a_{\rho\alpha}}{\partial t^{\beta}} + \frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} - \frac{\partial a_{\alpha\beta}}{\partial t^{\rho}} \right) a^{\rho\gamma} \ . \tag{9.114}$$

This is known as the *Levi-Civita* (or *Riemann* or *Christoffel*) *connection* (see, e.g., Levi-Civita [11]). Note that (9.92) presents the **extrinsic** definition of the Christoffel symbols while (9.114) renders their **intrinsic** definition. This is similar to the covariant metric coefficients which can be calculated extrinsically using the covariant basis vectors and intrinsically by measuring distances on the surfaces.

The surface analogue of $\Gamma^{i}_{ki} = \partial \ln J / \partial \Theta^{k}$, given in (7.17), renders

$$\Gamma^{\alpha}_{\gamma\alpha} = \frac{\partial \ln \sqrt{a}}{\partial t^{\gamma}} \ , \tag{9.115}$$

because

$$\frac{\partial \ln \sqrt{a}}{\partial t^\gamma} = \frac{1}{2a}\frac{\partial a}{\partial t^\gamma} = \frac{1}{2a}\frac{\partial a}{\partial a^{\alpha\beta}}\frac{\partial a^{\alpha\beta}}{\partial t^\gamma} = \frac{1}{2a}\left(-a\, a_{\alpha\beta}\right)\left(-\Gamma^\alpha_{\gamma\rho}\, a^{\rho\beta} - \Gamma^\beta_{\gamma\rho}\, a^{\rho\alpha}\right)$$
$$= \frac{1}{2}\Gamma^\alpha_{\gamma\rho}\delta^\rho_\alpha + \frac{1}{2}\Gamma^\beta_{\gamma\rho}\delta^\rho_\beta = \frac{1}{2}\Gamma^\rho_{\gamma\rho} + \frac{1}{2}\Gamma^\rho_{\gamma\rho} = \Gamma^\alpha_{\gamma\alpha} \; .$$

The ambient useful identity $\partial\left(J\mathbf{g}^i\right)/\partial\Theta^i = \mathbf{0}$, according to (7.18), does not hold true for curved surfaces due to curvature. Indeed,

$$\boxed{\frac{\partial\left(\sqrt{a}\mathbf{a}^\alpha\right)}{\partial t^\alpha} = \sqrt{a}\ \underline{b}^{\alpha}_{.\alpha}\,\widehat{\mathbf{n}}\ ,} \tag{9.116}$$

since

$$\frac{\partial\left(\sqrt{a}\,\mathbf{a}^\alpha\right)}{\partial t^\alpha} \xlongequal[\text{the product rule}]{\text{by using}} \frac{\partial\sqrt{a}}{\partial t^\alpha}\mathbf{a}^\alpha + \sqrt{a}\,\frac{\partial\mathbf{a}^\alpha}{\partial t^\alpha}$$
$$\xlongequal[\text{(9.109) and in light of (9.115)}]{\text{by using}} \sqrt{a}\,\Gamma^\gamma_{\alpha\gamma}\,\mathbf{a}^\alpha + \sqrt{a}\left[-\Gamma^\alpha_{\alpha\gamma}\,\mathbf{a}^\gamma + \underline{b}^{\alpha}_{.\alpha}\,\widehat{\mathbf{n}}\right]$$
$$\xlongequal[\text{and switching the names of dummy indices}]{\text{by considering the symmetry of Christoffel symbols}} \sqrt{a}\,\Gamma^\alpha_{\alpha\gamma}\,\left[\mathbf{a}^\gamma - \mathbf{a}^\gamma\right] + \sqrt{a}\ \underline{b}^{\alpha}_{.\alpha}\,\widehat{\mathbf{n}}$$
$$\xlongequal[\text{the properties of vector space}]{\text{by using}} \sqrt{a}\ \underline{b}^{\alpha}_{.\alpha}\,\widehat{\mathbf{n}}\ .$$

The surface and ambient Christoffel symbols are related through the following equation

$$\boxed{\Gamma^\gamma_{\alpha\beta} = \overline{Z}^i_\alpha\overline{Z}^j_\beta\widetilde{Z}^\gamma_k\,\mathit{\Gamma}^k_{ij} + \widetilde{Z}^\gamma_i\,\frac{\partial\overline{Z}^i_\alpha}{\partial t^\beta}\ ,} \quad \longleftarrow \text{see (9.218)} \tag{9.117}$$

since

$$\Gamma^\gamma_{\alpha\beta} \xlongequal[\text{(9.46), (9.59) and (9.92)}]{\text{by using}} \frac{\partial}{\partial t^\beta}\left[\overline{Z}^i_\alpha\,\mathbf{g}_i\right]\cdot\left(\widetilde{Z}^\gamma_k\,\mathbf{g}^k\right)$$
$$\xlongequal[\text{rule of differentiation}]{\text{by using the product}} \left[\frac{\partial\overline{Z}^i_\alpha}{\partial t^\beta}\mathbf{g}_i + \overline{Z}^i_\alpha\frac{\partial\mathbf{g}_i}{\partial t^\beta}\right]\cdot\left(\widetilde{Z}^\gamma_k\,\mathbf{g}^k\right)$$
$$\xlongequal[\text{rule of differentiation}]{\text{by using the chain}} \frac{\partial\overline{Z}^i_\alpha}{\partial t^\beta}\widetilde{Z}^\gamma_k\left(\mathbf{g}_i\cdot\mathbf{g}^k\right) + \overline{Z}^i_\alpha\widetilde{Z}^\gamma_k\frac{\partial\Theta^j}{\partial t^\beta}\frac{\partial\mathbf{g}_i}{\partial\Theta^j}\cdot\mathbf{g}^k$$
$$\xlongequal[\text{(5.14), (5.27), (7.7) and (9.47)}]{\text{by using}} \frac{\partial\overline{Z}^i_\alpha}{\partial t^\beta}\widetilde{Z}^\gamma_i + \overline{Z}^i_\alpha\widetilde{Z}^\gamma_k\overline{Z}^j_\beta\mathit{\Gamma}^k_{ij}\ .$$

Recall from the ambient space that the major role of the Christoffel symbols was to facilitate the covariant differentiation. This also holds true for curved surfaces. In what follows, the goal is thus to study the surface covariant derivative.

9.3 Surface Covariant Derivative

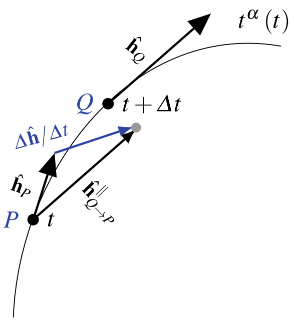
9.3.1 Surface First Covariant Differentiation

The goal here is to geometrically describe the *surface covariant derivative* of a vector and covector. The results will then be extended to tensors of higher ranks (note that this useful differential operator provides tensors out of tensors in accord with the ambient space). The procedure greatly relies on what is known as the *parallel transport* explained throughout the development. So far, it has intuitively been assumed that the partial derivative of a surface vector satisfies the product rule. In the following, it will be verified that both partial and covariant differential operators satisfy this fundamental rule of differentiation.

The surface covariant derivative of a vector field is illustrated in Fig. 9.11.

Let $\mathcal{S}$ be a surface embracing a parametrized curve $t^{\alpha}(t)$ with $\mathbf{a}_t = (dt^{\alpha}/dt)\,\mathbf{a}_{\alpha}$, according to (9.240)$_3$. Further, let $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}$ be a smooth vector field defined in the tangent planes of that surface. Now, consider a point P corresponding to t. Consider

Note that $\hat{\mathbf{h}}_P$ and $\hat{\mathbf{h}}_Q$ are not in the same tangent plane. Consequently, $\hat{\mathbf{h}}_Q - \hat{\mathbf{h}}_P$ is not well-defined. This is the reason for parallel transporting $\hat{\mathbf{h}}$ from Q to P.



Let $t^{\alpha}(t)$ be a parametrerized curve embedded in a surface $\mathcal{S}$. Further, let $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}$ be a smooth vector field defined in the tangent planes of $\mathcal{S}$. The rate of change in $\hat{\mathbf{h}}$ along the curve is then defined by

$$\frac{d\hat{\mathbf{h}}}{dt} := \lim_{\Delta t\to 0}\frac{\hat{\mathbf{h}}^{\parallel}_{Q\to P} - \hat{\mathbf{h}}_P}{\Delta t} \quad \text{or} \quad \frac{d\,\underline{\hat{h}}^{\alpha}}{dt} := \lim_{\Delta t\to 0}\frac{\underline{\hat{h}}^{\alpha\,\parallel}_{Q\to P} - \underline{\hat{h}}^{\alpha}_P}{\Delta t}\,,$$

where $\hat{\mathbf{h}}^{\parallel}_{Q\to P}$ denotes the *parallel transport* of $\hat{\mathbf{h}}$ from Q to P. Suppose that t is one of the coordinate lines. After some manipulations, the above definition leads to

$$\frac{\partial\hat{\mathbf{h}}}{\partial t^{\beta}} = \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}\mathbf{a}_{\alpha} + \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\mathbf{a}_{\alpha} + \underline{\hat{h}}^{\alpha}\,\underline{b}_{\alpha\beta}\,\hat{\mathbf{n}}\,.$$

The orthogonal projection of $\partial\hat{\mathbf{h}}/\partial t^{\beta}$ onto the plane tangent to $\mathcal{S}$ is known as the *surface covariant derivtive* of $\hat{\mathbf{h}}$:

$$\hat{\mathbf{h}}|_{\beta} = \left(\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}} + \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\right)\mathbf{a}_{\alpha}\,.$$

Fig. 9.11 Surface covariant derivative

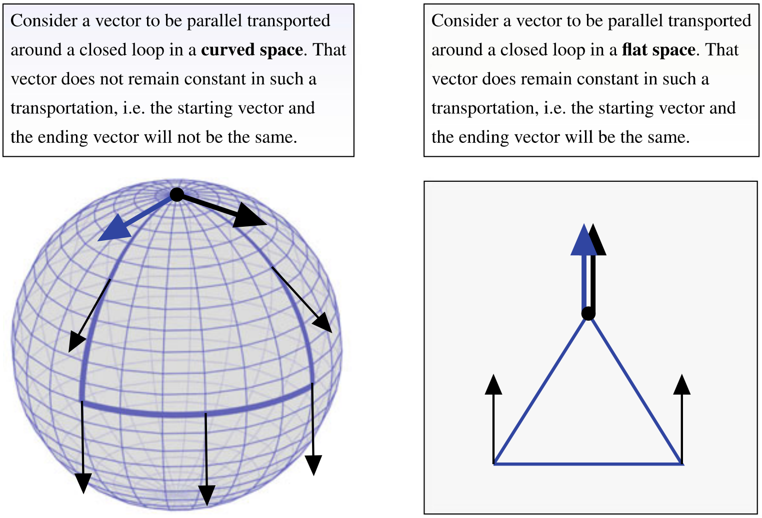


Fig. 9.12 Parallel transport

also a nearby point Q corresponding to $t + \Delta t$. The derivative of $\hat{\mathbf{h}}$ in the direction $\mathbf{a}_t$ is defined by

$$\frac{d\hat{\mathbf{h}}}{dt} := \lim_{\Delta t \to 0} \frac{\hat{\mathbf{h}}^{\parallel}_{Q\to P} - \hat{\mathbf{h}}_P}{\Delta t} \quad \text{or} \quad \frac{d\,\underline{\hat{h}}^{\alpha}}{dt} := \lim_{\Delta t \to 0} \frac{\underline{\hat{h}}^{\alpha\,\parallel}_{Q\to P} - \underline{\hat{h}}^{\alpha}_{P}}{\Delta t} \,, \tag{9.118}$$

where $\hat{\mathbf{h}}^{\parallel}_{Q\to P}$ represents the so-called *parallel transport* of $\hat{\mathbf{h}}$ from Q to P. The key point in this definition is that the difference $\hat{\mathbf{h}}^{\parallel}_{Q\to P} - \hat{\mathbf{h}}_P$ is well-defined. Note that $\hat{\mathbf{h}}_Q - \hat{\mathbf{h}}_P$ does not make any sense because the vectors lie in different tangent planes. This problem can be resolved by adopting a proper transportation. For instance, one can move a vector around by keeping its length and direction as constant as possible. This is known as the parallel transport of that vector. It basically allows one to connect the different tangent spaces in a curved space. For a geometrical interpretation, see Fig. 9.12.

The parallel transport of $\underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}$ from Q to P only affects the basis vectors, as implied by its name. Thus, in coordinate representation, the definition $(9.118)_1$ renders

$$\frac{d\hat{\mathbf{h}}}{dt} = \lim_{\Delta t \to 0} \frac{\underline{\hat{h}}^{\alpha}_{Q} \left(\mathbf{a}_{\alpha}\right)^{\parallel}_{Q\to P} - \underline{\hat{h}}^{\alpha}_{P} \left(\mathbf{a}_{\alpha}\right)_{P}}{\Delta t} \,, \tag{9.119}$$

where $\underline{\hat{h}}^{\alpha}_{Q} := \underline{\hat{h}}^{\alpha}\left(t + \Delta t\right)$ can be related to $\underline{\hat{h}}^{\alpha}_{P} := \underline{\hat{h}}^{\alpha}\left(t\right)$ via the following first-order Taylor series expansion

$$\underline{\hat{h}}^{\alpha}_{Q} = \underline{\hat{h}}^{\alpha}_{P} + \left.\frac{d\,\underline{\hat{h}}^{\alpha}}{dt}\right|_{P} \Delta t \,. \tag{9.120}$$

← see (6.24)–(6.25)

To proceed, one needs to realize that $(\mathbf{a}_\alpha)^{\parallel}_{Q\to P}$ is an **ambient** vector sitting at P. It can thus be decomposed with respect to the ambient basis $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}|_P$ as

$$(\mathbf{a}_\alpha)^{\parallel}_{Q\to P} = \xi_{\alpha\ P}^{\beta}\left(\mathbf{a}_\beta\right)_P + \varsigma_{\alpha\ P}\left(\widehat{\mathbf{n}}\right)_P \ . \tag{9.121}$$

For small enough Δt, the coefficients $\xi_{\alpha\ P}^{\beta}$ and $\varsigma_{\alpha\ P}$ are nearly δ_α^β and $(\mathbf{0})_\alpha$, respectively. Consequently, using (9.119)-(9.121), one can arrive at

$$\frac{d\hat{\mathbf{h}}}{dt} = \frac{d\,\underline{\hat{h}}^\alpha}{dt}\mathbf{a}_\alpha + \left(\lim_{\Delta t\to 0}\frac{\xi_\theta^\alpha - \delta_\theta^\alpha}{\Delta t}\right)\underline{\hat{h}}^\theta\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\left(\lim_{\Delta t\to 0}\frac{\varsigma_\alpha}{\Delta t}\right)\widehat{\mathbf{n}}\ , \tag{9.122}$$

where the subscript P has been dropped for convenience. Suppose that t is one of the coordinate curves. The above relation then takes the form

$$\frac{\partial\hat{\mathbf{h}}}{\partial t^\beta} = \frac{\partial\,\underline{\hat{h}}^\alpha}{\partial t^\beta}\mathbf{a}_\alpha + \left(\lim_{\Delta t^\beta\to 0}\frac{\xi_\theta^\alpha - \delta_\theta^\alpha}{\Delta t^\beta}\right)\underline{\hat{h}}^\theta\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\left(\lim_{\Delta t^\beta\to 0}\frac{\varsigma_\alpha}{\Delta t^\beta}\right)\widehat{\mathbf{n}}\ . \tag{9.123}$$

Let

$$\Gamma_{\beta\theta}^\alpha := \lim_{\Delta t^\beta\to 0}\frac{\xi_\theta^\alpha - \delta_\theta^\alpha}{\Delta t^\beta}\quad,\quad \underline{b}_{\alpha\beta} := \lim_{\Delta t^\beta\to 0}\frac{\varsigma_\alpha}{\Delta t^\beta}\ . \tag{9.124}$$

← for the characterization of $\Gamma_{\beta\theta}^\alpha$, see (9.144)–(9.151)

Then,

$$\frac{\partial\hat{\mathbf{h}}}{\partial t^\beta} = \left(\frac{\partial\,\underline{\hat{h}}^\alpha}{\partial t^\beta} + \Gamma_{\beta\theta}^\alpha\,\underline{\hat{h}}^\theta\right)\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\,\underline{b}_{\alpha\beta}\widehat{\mathbf{n}}\ . \tag{9.125}$$

One can also have

$$\frac{d\hat{\mathbf{h}}}{dt} = \left(\frac{d\,\underline{\hat{h}}^\alpha}{dt} + \Gamma_{\beta\theta}^\alpha\frac{dt^\beta}{dt}\,\underline{\hat{h}}^\theta\right)\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\,\underline{b}_{\alpha\beta}\frac{dt^\beta}{dt}\widehat{\mathbf{n}}\ . \tag{9.126}$$

The orthogonal projection of $\partial\hat{\mathbf{h}}/\partial t^\beta$ onto the tangent plane defines the *surface covariant derivative* of $\hat{\mathbf{h}}$. Here, it is denoted by

$$\hat{\mathbf{h}}\Big|_{\mathbf{a}_\beta} = \hat{\mathbf{h}}\Big|_\beta = \underline{\hat{h}}^\alpha\Big|_\beta\,\mathbf{a}_\alpha\ , \tag{9.127}$$

in the literature, this is sometimes denoted by $\nabla_{\mathbf{a}_\beta}\hat{\mathbf{h}} = \nabla_\beta\hat{\mathbf{h}} = \left(\nabla_\beta\,\underline{\hat{h}}^\alpha\right)\mathbf{a}_\alpha$

where

$$\boxed{\underline{\hat{h}}^\alpha\Big|_\beta = \frac{\partial\,\underline{\hat{h}}^\alpha}{\partial t^\beta} + \Gamma_{\beta\theta}^\alpha\,\underline{\hat{h}}^\theta\ .} \tag{9.128}$$

As can be seen, the covariant derivative operator relies on the connection coefficients. It is sometimes called *connection* because it basically helps connect the different tangent spaces. In general, there are many possible ways to define these coefficients (note that the most commonly used connection coefficients in general relativity are governed by the Levi-Civita connection). Indeed, each space can possibly have its own definition. And this means that the result of covariant derivative (and accordingly parallel transport) may vary from space to space.

The surface covariant differentiation is eventually the only rate of change that the two-dimensional inhabitants of the surface can measure. It should thus be regarded as an **intrinsic** operation. Indeed, the surface covariant derivative of a surface vector field is the partial derivative of such an invariant object with the normal component subtracted.

One can readily show that the surface covariant derivative coincides with the ordinary partial derivative for smooth scalar functions of the Gaussian coordinates:

$$\bar{h}\big|_{\beta} = \frac{\partial \bar{h}}{\partial t^{\beta}} \ . \tag{9.129}$$

This is eventually the directional derivative of $\bar{h}$ in the direction of $\mathbf{a}_\beta$:

$$\begin{aligned} D_{\mathbf{a}_\beta}\bar{h} &= \frac{\partial \bar{h}}{\partial \mathbf{x}} \cdot \mathbf{a}_\beta = \frac{\partial \bar{h}}{\partial t^{\alpha}} \frac{\partial t^{\alpha}}{\partial \mathbf{x}} \cdot \mathbf{a}_\beta \\ &= \frac{\partial \bar{h}}{\partial t^{\alpha}} \mathbf{a}^{\alpha} \cdot \mathbf{a}_\beta = \frac{\partial \bar{h}}{\partial t^{\alpha}} \delta^{\alpha}_{\beta} \\ &= \frac{\partial \bar{h}}{\partial t^{\beta}} \ . \quad \longleftarrow \text{see (7.4a)} \end{aligned} \tag{9.130}$$

Following similar procedures which led to (9.125), one can obtain

$$\boxed{\frac{\partial \mathbf{a}_\alpha}{\partial t^{\beta}} = \mathbf{a}_\alpha|_\beta + \underline{b}_{\alpha\beta}\widehat{\mathbf{n}} \quad \text{where} \quad \mathbf{a}_\alpha|_\beta = \Gamma^{\gamma}_{\alpha\beta}\mathbf{a}_\gamma \ .} \tag{9.131}$$

These relations basically represent the Gauss formulas demonstrated in (9.94). This helps verify that the partial derivative of $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_\alpha$ satisfies the product rule in the sense that

$$\frac{\partial \hat{\mathbf{h}}}{\partial t^{\beta}} = \frac{\partial \left(\underline{\hat{h}}^{\alpha}\mathbf{a}_\alpha\right)}{\partial t^{\beta}} = \frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}\mathbf{a}_\alpha + \underline{\hat{h}}^{\alpha}\frac{\partial \mathbf{a}_\alpha}{\partial t^{\beta}} = \underline{\hat{h}}^{\alpha}\Big|_{\beta}\mathbf{a}_\alpha + \underline{\hat{h}}^{\alpha}\,\underline{b}_{\alpha\beta}\,\widehat{\mathbf{n}} \ . \tag{9.132}$$

The relation $(9.109)_1$ can now be rephrased as

$$\boxed{\frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\beta}} = \mathbf{a}^{\alpha}|_\beta + \underline{b}^{\alpha}_{\cdot\beta}\widehat{\mathbf{n}} \quad \text{where} \quad \mathbf{a}^{\alpha}|_\beta = -\Gamma^{\alpha}_{\beta\gamma}\,\mathbf{a}^{\gamma} \ .} \tag{9.133}$$

As can be seen from (9.131)$_2$ and (9.133)$_2$, the ambient metrinilic property $\mathbf{g}_i|_j = \mathbf{0}$ and $\mathbf{g}^i\big|_j = \mathbf{0}$, according to (7.36a)–(7.36b), does not hold true for curved surfaces.

Let $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha} \mathbf{a}_{\alpha}$, $\mathbf{u} = \underline{u}^{\beta} \mathbf{a}_{\beta}$ be two smooth vector fields. Similar procedures which led to (9.125) can then be followed to show that the partial derivative of $\hat{\mathbf{h}} \cdot \mathbf{u}$ satisfies the product rule, that is,

$$\boxed{\frac{\partial \left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)}{\partial t^{\beta}} = \frac{\partial \hat{\mathbf{h}}}{\partial t^{\beta}} \cdot \mathbf{u} + \hat{\mathbf{h}} \cdot \frac{\partial \mathbf{u}}{\partial t^{\beta}} .} \quad \leftarrow \text{the proof is given in Exercise 9.12} \tag{9.134}$$

Indeed,

$$\frac{\partial \left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)}{\partial t^{\beta}} = \underline{\hat{h}}^{\alpha}\Big|_{\beta} \left(a_{\alpha\gamma}\, \underline{u}^{\gamma}\right) + \left(\underline{\hat{h}}^{\alpha} a_{\alpha\gamma}\right) \underline{u}^{\gamma}\big|_{\beta} , \tag{9.135}$$

which is in alignment with

$$\begin{aligned} \frac{\partial \left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)}{\partial t^{\beta}} &= \frac{\partial \left(\underline{\hat{h}}^{\alpha} a_{\alpha\gamma}\, \underline{u}^{\gamma}\right)}{\partial t^{\beta}} \qquad \swarrow \text{ note that } \frac{\partial a_{\alpha\gamma}}{\partial t^{\beta}} = \Gamma^{\rho}_{\beta\alpha}\, a_{\rho\gamma} + \Gamma^{\rho}_{\beta\gamma}\, a_{\rho\alpha} \\ &= \frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\beta}} a_{\alpha\gamma}\, \underline{u}^{\gamma} + \underline{\hat{h}}^{\alpha} \frac{\partial a_{\alpha\gamma}}{\partial t^{\beta}}\, \underline{u}^{\gamma} + \underline{\hat{h}}^{\alpha} a_{\alpha\gamma} \frac{\partial\, \underline{u}^{\gamma}}{\partial t^{\beta}} . \end{aligned}$$

Consider the fact that $\hat{\mathbf{h}} \cdot \mathbf{u}$ is a scalar quantity whose partial derivative and covariant differentiation coincide. Thus, one can establish the useful product rule

$$\boxed{\left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)\Big|_{\beta} = \left(\hat{\mathbf{h}}\Big|_{\beta}\right) \cdot \mathbf{u} + \hat{\mathbf{h}} \cdot \left(\mathbf{u}|_{\beta}\right) ,} \tag{9.136}$$

which is known as the *metric compatibility property*. This helps characterize some important properties addressed throughout the development.

Let α , $\beta \in \mathrm{R}$ be two arbitrary constants and $\bar{h} \in \mathrm{R}$ be a smooth scalar function of the surface coordinates. Further, let $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathscr{E}^{o3}_{\mathrm{r}}$ be three smooth vector fields. It should not be difficult now to verify that the surface covariant derivative satisfies the following properties (Steinmann [12])

$$(\alpha\mathbf{u} + \beta\mathbf{v})|_{\mathbf{w}} = \alpha\left(\mathbf{u}|_{\mathbf{w}}\right) + \beta\left(\mathbf{v}|_{\mathbf{w}}\right) , \quad \leftarrow \text{linearity property in the field} \tag{9.137a}$$

$$\mathbf{u}|_{(\alpha\mathbf{v}+\beta\mathbf{w})} = \alpha\left(\mathbf{u}|_{\mathbf{v}}\right) + \beta\left(\mathbf{u}|_{\mathbf{w}}\right) , \quad \leftarrow \text{linearity property in the direction vector} \tag{9.137b}$$

$$\left(\bar{h}\mathbf{u}\right)\big|_{\mathbf{v}} = \left(D_{\mathbf{v}}\bar{h}\right)\mathbf{u} + \bar{h}\left(\mathbf{u}|_{\mathbf{v}}\right) , \quad \leftarrow \text{Leibniz rule of differentiation} \tag{9.137c}$$

$$\mathbf{u}|_{(\bar{h}\mathbf{v})} = \bar{h}\left(\mathbf{u}|_{\mathbf{v}}\right) . \tag{9.137d}$$

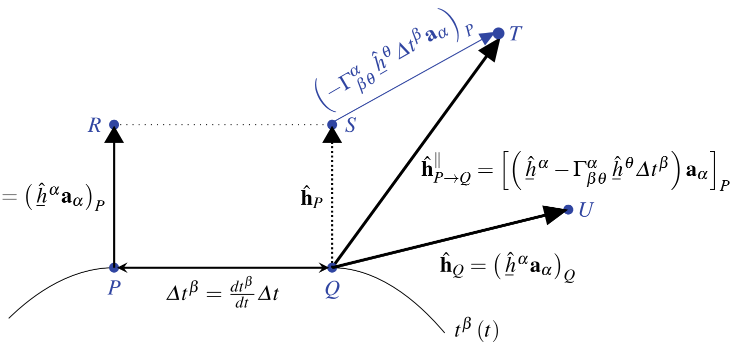


Fig. 9.13 Parallel transport of vector field

Let $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}$ be a smooth vector field. And consider its components (basis vectors) as scalar functions (vectors). Using (9.130)$_5$, (9.131)$_2$ and (9.137c), one can then verify that the surface covariant derivative of $\hat{\mathbf{h}}$ with respect to $\mathbf{a}_{\beta}$ satisfies the product rule in the sense that

$$\hat{\mathbf{h}}\Big|_{\mathbf{a}_{\beta}} = \left(\underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}\right)\Big|_{\mathbf{a}_{\beta}} = \left(D_{\mathbf{a}_{\beta}}\,\underline{\hat{h}}^{\alpha}\right)\mathbf{a}_{\alpha} + \underline{\hat{h}}^{\alpha}\left(\mathbf{a}_{\alpha}|_{\mathbf{a}_{\beta}}\right) = \underline{\hat{h}}^{\alpha}\Big|_{\beta}\,\mathbf{a}_{\alpha}\ . \tag{9.138}$$

Let $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}$, $\mathbf{u} = \underline{u}^{\beta}\mathbf{a}_{\beta}$ be two smooth vector fields. Making use of (9.137b) and (9.138)$_3$, the surface covariant derivative of $\hat{\mathbf{h}}$ along $\mathbf{u}$ can then be established according to

$$\hat{\mathbf{h}}\Big|_{\mathbf{u}} = \hat{\mathbf{h}}\Big|_{\left(\underline{u}^{\beta}\mathbf{a}_{\beta}\right)} = \hat{\mathbf{h}}\Big|_{\mathbf{a}_{\beta}}\,\underline{u}^{\beta} = \underline{\hat{h}}^{\alpha}\Big|_{\beta}\,\underline{u}^{\beta}\mathbf{a}_{\alpha}\ . \tag{9.139}$$

The condition for the **parallel transport** of $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}$ along $\mathbf{u} = \underline{u}^{\beta}\mathbf{a}_{\beta}$ reads

$$\boxed{\hat{\mathbf{h}}\Big|_{\mathbf{u}} = \mathbf{0} \quad \text{or} \quad \underline{\hat{h}}^{\alpha}\Big|_{\beta} = 0 \quad \text{or} \quad \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}} = -\Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\ .} \tag{9.140}$$

Accordingly, the change in a vector from a point t to an infinitesimally close point $t + dt$ when it is parallel transported along a curve $t^{\beta}(t)$ represents (see Fig. 9.13)

$$\boxed{d\,\underline{\hat{h}}^{\alpha} = \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}dt^{\beta} = -\Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}dt^{\beta} = -\Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\frac{dt^{\beta}}{dt}dt\ .} \tag{9.141}$$

Consider a vector $\hat{\mathbf{h}}$ which has been parallel transported along $\mathbf{a}_{\beta}$ (this basically defines a vector field along the corresponding coordinate curve). Using (9.130)$_5$, (9.136) and (9.140), one can write

$$\left(\hat{\mathbf{h}} \cdot \hat{\mathbf{h}}\right)\Big|_\beta = \left(\hat{\mathbf{h}}\Big|_\beta\right) \cdot \hat{\mathbf{h}} + \hat{\mathbf{h}} \cdot \left(\hat{\mathbf{h}}\Big|_\beta\right) = 0 \quad \text{or} \quad \frac{\partial \left(\hat{\mathbf{h}} \cdot \hat{\mathbf{h}}\right)}{\partial t^\beta} = 0 \,. \tag{9.142}$$

This result states that:

> The length of a vector remains constant during its parallel transport along a curve.

Now, consider the two vectors $\hat{\mathbf{h}}$ and $\mathbf{u}$ (with the same origin) which have been parallel transported along $\mathbf{a}_\beta$ (to form two vector fields along the corresponding coordinate line). From (9.130)$_5$, (9.136) and (9.140), it follows that

$$\left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)\Big|_\beta = \left(\hat{\mathbf{h}}\Big|_\beta\right) \cdot \mathbf{u} + \hat{\mathbf{h}} \cdot \left(\mathbf{u}|_\beta\right) = 0 \quad \text{or} \quad \frac{\partial \left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)}{\partial t^\beta} = 0 \,. \tag{9.143}$$

One then concludes that:

> The angle between two vectors remains constant when they are parallel transported along a curve.

Notice that the results (9.142) and (9.143) are two important consequences of the metric compatibility property. Moreover, one can again use this property to characterize the coefficients $\Gamma^\alpha_{\beta\theta}$ in (9.124). This is demonstrated in the following. ✭

Let $\hat{\mathbf{h}} = \mathbf{a}_\alpha$ and $\mathbf{u} = \mathbf{a}_\rho$. The metric compatibility property then renders

$$\left(\mathbf{a}_\alpha \cdot \mathbf{a}_\rho\right)\Big|_\beta = \left(\mathbf{a}_\alpha|_\beta\right) \cdot \mathbf{a}_\rho + \mathbf{a}_\alpha \cdot \left(\mathbf{a}_\rho\Big|_\beta\right) \,, \tag{9.144}$$

or, using (9.17)$_1$, (9.130)$_5$ and (9.131)$_2$,

$$\frac{\partial a_{\alpha\rho}}{\partial t^\beta} = \Gamma^\gamma_{\alpha\beta} \mathbf{a}_\gamma \cdot \mathbf{a}_\rho + \mathbf{a}_\alpha \cdot \Gamma^\gamma_{\rho\beta} \mathbf{a}_\gamma = \Gamma^\gamma_{\alpha\beta} a_{\gamma\rho} + \Gamma^\gamma_{\rho\beta} a_{\alpha\gamma} \,. \tag{9.145}$$

Here, the coefficients $\Gamma^\gamma_{\alpha\beta}$ are generally regarded as 8 unknowns. Considering the symmetry of the metric tensor components, notice that one only has 6 independent equations. A natural treatment is discussed below.

Let $\Gamma^\gamma_{\alpha\beta}$ be additively decomposed according to

$$\underbrace{\Gamma^\gamma_{\alpha\beta}}_{\text{8 unknowns}} = \underbrace{\frac{1}{2}\left(\Gamma^\gamma_{\alpha\beta} + \Gamma^\gamma_{\beta\alpha}\right)}_{\text{6 unknowns}} + \underbrace{\frac{1}{2}\left(\Gamma^\gamma_{\alpha\beta} - \Gamma^\gamma_{\beta\alpha}\right)}_{\text{2 unknowns}} \,, \tag{9.146}$$

where (twice) its skew-symmetric part is referred to as the *torsion tensor* (see the pioneering work of Cartan [13]):

$$\boxed{T^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\alpha\beta} - \Gamma^{\gamma}_{\beta\alpha}\ .} \tag{9.147}$$

The problem can be resolved by assuming that the torsion tensor vanishes:

$$\boxed{T^{\gamma}_{\alpha\beta} = 0\ .} \tag{9.148}$$

This is a popular (but debatable) assumption frequently used, for instance, in general relativity. One is thus left with a system of 6 linear equations in six variables $\left(\Gamma^1_{11}, \Gamma^1_{22}, \Gamma^1_{12}, \Gamma^2_{11}, \Gamma^2_{22}, \Gamma^2_{12}\right)$.

The result $(9.145)_2$, by renaming the free indices $\alpha \to \rho$, $\rho \to \beta$ and $\beta \to \alpha$, can be written as

$$-\frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} = -\,\Gamma^{\gamma}_{\rho\alpha} a_{\gamma\beta} - \Gamma^{\gamma}_{\beta\alpha} a_{\rho\gamma}\ . \tag{9.149}$$

In a similar manner,

$$\frac{\partial a_{\beta\alpha}}{\partial t^{\rho}} = \Gamma^{\gamma}_{\beta\rho} a_{\gamma\alpha} + \Gamma^{\gamma}_{\alpha\rho} a_{\beta\gamma}\ . \tag{9.150}$$

From $(9.145)_2$, (9.149) and (9.150), one can finally arrive at

$$2\Gamma^{\gamma}_{\rho\beta} a_{\alpha\gamma} = \frac{\partial a_{\alpha\rho}}{\partial t^{\beta}} + \frac{\partial a_{\beta\alpha}}{\partial t^{\rho}} - \frac{\partial a_{\rho\beta}}{\partial t^{\alpha}}\ , \tag{9.151}$$

or $\Gamma^{\theta}_{\alpha\beta} a_{\rho\theta} = \frac{1}{2}\left(\frac{\partial a_{\rho\alpha}}{\partial t^{\beta}} + \frac{\partial a_{\beta\rho}}{\partial t^{\alpha}} - \frac{\partial a_{\alpha\beta}}{\partial t^{\rho}}\right)$ or $\Gamma^{\gamma}_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial a_{\rho\alpha}}{\partial t^{\beta}} + \frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} - \frac{\partial a_{\alpha\beta}}{\partial t^{\rho}}\right) a^{\rho\gamma}$

which is the Levi-Civita connection given in (9.114). ✭

Let $\mathbf{u}$ and $\hat{\mathbf{h}}$ be two smooth vector fields. Space is said to be *torsion-free* if no gap appears when these vectors are parallel transported along each other. If the so-called *commutator* of $\mathbf{u}$ and $\hat{\mathbf{h}}$ vanishes,[10] this condition is written by

---

[10] The torsion tensor $T^{\alpha}_{\beta\gamma}$, introduced in (9.147), is basically defined by its operation on two smooth vector fields as follows:

$$\begin{aligned}
\mathbf{t}\left(\mathbf{u}, \hat{\mathbf{h}}\right) &= \hat{\mathbf{h}}\Big|_{\mathbf{u}} - \mathbf{u}|_{\hat{\mathbf{h}}} - \left[\mathbf{u}, \hat{\mathbf{h}}\right] \\
&= \underline{\hat{h}}^{\alpha}\Big|_{\beta}\, \underline{u}^{\beta} \mathbf{a}_{\alpha} - \underline{u}^{\alpha}\Big|_{\beta}\, \underline{\hat{h}}^{\beta} \mathbf{a}_{\alpha} - \left(\underline{u}^{\theta} \frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta} \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}}\right) \mathbf{a}_{\alpha} \\
&= \left(\frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}\, \underline{u}^{\beta} + \Gamma^{\alpha}_{\beta\gamma}\, \underline{\hat{h}}^{\gamma}\, \underline{u}^{\beta} - \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\beta}}\, \underline{\hat{h}}^{\beta} - \Gamma^{\alpha}_{\beta\gamma}\, \underline{u}^{\gamma}\, \underline{\hat{h}}^{\beta} - \frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}\, \underline{u}^{\beta} + \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\beta}}\, \underline{\hat{h}}^{\beta}\right) \mathbf{a}_{\alpha} \\
&= \left(\Gamma^{\alpha}_{\beta\gamma}\, \underline{\hat{h}}^{\gamma}\, \underline{u}^{\beta} - \Gamma^{\alpha}_{\beta\gamma}\, \underline{u}^{\gamma}\, \underline{\hat{h}}^{\beta}\right) \mathbf{a}_{\alpha} \\
&= \left(\Gamma^{\alpha}_{\beta\gamma} - \Gamma^{\alpha}_{\gamma\beta}\right) \underline{\hat{h}}^{\gamma}\, \underline{u}^{\beta} \mathbf{a}_{\alpha}
\end{aligned}$$

$$\boxed{\hat{\mathbf{h}}\Big|_{\mathbf{u}} = \mathbf{u}|_{\hat{\mathbf{h}}} \ .} \quad \longleftarrow \text{ see the expression (9.619) and also Fig. 9.34} \tag{9.152}$$

One can then conclude that the connection coefficients have symmetry in their lower two indices. To show this, suppose that $\hat{\mathbf{h}} = \mathbf{a}_\alpha$, $\mathbf{u} = \mathbf{a}_\beta$. Then,

$$\boxed{\mathbf{a}_\alpha|_\beta - \mathbf{a}_\beta\Big|_\alpha = \left(\Gamma^\gamma_{\alpha\beta} - \Gamma^\gamma_{\beta\alpha}\right)\mathbf{a}_\gamma = \mathbf{0} \text{ implies that } \Gamma^\gamma_{\alpha\beta} = \Gamma^\gamma_{\beta\alpha} \ .} \tag{9.153}$$

The goal is now to derive the surface covariant derivative of a covector. This will be demonstrated in the following. ✈

Let $\bar{h} = \underline{u}^\alpha\, \underline{\hat{h}}_\alpha$ be a smooth scalar function whose covariant derivative is

$$\left(\underline{u}^\alpha\, \underline{\hat{h}}_\alpha\right)\Big|_\beta = \frac{\partial\left(\underline{u}^\alpha\, \underline{\hat{h}}_\alpha\right)}{\partial t^\beta} = \frac{\partial\, \underline{u}^\alpha}{\partial t^\beta}\, \underline{\hat{h}}_\alpha + \underline{u}^\alpha \frac{\partial\, \underline{\hat{h}}_\alpha}{\partial t^\beta} \ . \tag{9.154}$$

The demand for satisfying the product rule then helps represent

$$\begin{aligned}
\left(\underline{u}^\alpha\, \underline{\hat{h}}_\alpha\right)\Big|_\beta &= \left(\underline{u}^\alpha\big|_\beta\right)\underline{\hat{h}}_\alpha + \underline{u}^\alpha\left(\underline{\hat{h}}_\alpha\Big|_\beta\right) \\
&= \frac{\partial\, \underline{u}^\alpha}{\partial t^\beta}\, \underline{\hat{h}}_\alpha + \underline{\hat{h}}_\alpha \Gamma^\alpha_{\beta\theta}\, \underline{u}^\theta + \underline{\hat{h}}_\alpha\Big|_\beta\, \underline{u}^\alpha \\
&= \frac{\partial\, \underline{u}^\alpha}{\partial t^\beta}\, \underline{\hat{h}}_\alpha + \underline{\hat{h}}_\theta \Gamma^\theta_{\beta\alpha}\, \underline{u}^\alpha + \underline{\hat{h}}_\alpha\Big|_\beta\, \underline{u}^\alpha \ .
\end{aligned} \tag{9.155}$$

From (9.154) and (9.155), it follows that

$$\underline{\hat{h}}_\alpha\Big|_\beta\, \underline{u}^\alpha = \frac{\partial\, \underline{\hat{h}}_\alpha}{\partial t^\beta}\, \underline{u}^\alpha - \underline{\hat{h}}_\theta \Gamma^\theta_{\beta\alpha}\, \underline{u}^\alpha \ .$$

Considering the fact that $\underline{u}^\alpha$ is entirely arbitrary finally helps establish

$$\boxed{\underline{\hat{h}}_\alpha\Big|_\beta = \frac{\partial\, \underline{\hat{h}}_\alpha}{\partial t^\beta} - \Gamma^\theta_{\alpha\beta}\, \underline{\hat{h}}_\theta \ .} \quad ✈ \tag{9.156}$$

Consequently, in accord with (9.132), one will have

$$= T^\alpha_{\beta\gamma}\, \underline{u}^\beta\, \underline{\hat{h}}^\gamma \mathbf{a}_\alpha \ .$$

Here, $\left[\mathbf{u}, \hat{\mathbf{h}}\right]$ denotes the *commutator* (or *Lie bracket*) of **u** and $\hat{\mathbf{h}}$. This sophisticated tool eventually presents the *Lie derivative* of $\hat{\mathbf{h}}$ with respect to **u**, see (9.550) and (9.617). Guided by (9.563), note that the commutator of the basis vectors vanishes.

$$\frac{\partial \hat{\mathbf{h}}}{\partial t^{\beta}} = \frac{\partial \left(\underline{\hat{h}}_{\alpha} \mathbf{a}^{\alpha} \right)}{\partial t^{\beta}} = \frac{\partial \underline{\hat{h}}_{\alpha}}{\partial t^{\beta}} \mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha} \frac{\partial \mathbf{a}^{\alpha}}{\partial t^{\beta}} = \underline{\hat{h}}_{\alpha} \Big|_{\beta} \mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha} \underline{b}^{\alpha}_{.\beta} \widehat{\mathbf{n}} \,. \tag{9.157}$$

The relations (9.140)$_2$ and (9.141) now translate to

$$\boxed{\underline{\hat{h}}_{\alpha} \Big|_{\beta} = 0 \quad \text{or} \quad \frac{\partial \underline{\hat{h}}_{\alpha}}{\partial t^{\beta}} = \Gamma^{\theta}_{\alpha\beta} \underline{\hat{h}}_{\theta} \,,} \tag{9.158}$$

and

$$\boxed{d \underline{\hat{h}}_{\alpha} = \frac{\partial \underline{\hat{h}}_{\alpha}}{\partial t^{\beta}} dt^{\beta} = \Gamma^{\theta}_{\alpha\beta} \underline{\hat{h}}_{\theta} dt^{\beta} = \Gamma^{\theta}_{\alpha\beta} \underline{\hat{h}}_{\theta} \frac{dt^{\beta}}{dt} dt \,.} \tag{9.159}$$

In summary, some fundamental properties of the surface covariant derivative are highlighted here. This mathematical object

- ➪ coincides with the ordinary partial derivative for scalar functions varying on the surface,
- ➪ satisfies the sum and product rules, and
- ➪ violates the commutative law when the surface under consideration possesses a nonvanishing Riemann-Christoffel curvature tensor (see (9.201a)-(9.201b)).

Analogously to the ambient expression (7.26), the coordinate index on the surface covariant derivative can be raised to define the *surface contravariant derivative* as follows:

$$\boxed{\underline{\hat{h}}^{\alpha} \Big|^{\beta} = \underline{\hat{h}}^{\alpha} \Big|_{\theta} a^{\theta\beta} \quad , \quad \underline{\hat{h}}_{\alpha} \Big|^{\beta} = \underline{\hat{h}}_{\alpha} \Big|_{\theta} a^{\theta\beta} \,.} \tag{9.160}$$

Let $\tilde{\mathbf{H}} = \tilde{\mathbf{H}} \left(t^1, t^2 \right)$ be a smooth surface tensor field which can be decomposed according to (9.82)$_{1-4}$. In the following, the aim is to characterize the surface covariant derivative of such a tensor field. ➡

Recall from (9.132) and (9.138) that both partial and covariant derivatives of an object satisfied the product rule. Moreover, recall that the surface covariant differentiation of an invariant was equal to its partial derivative with all contributions possessing extrinsic objects subtracted. Now, using (9.131)$_{1-2}$ and (9.133)$_{1-2}$, the rate of change in $\tilde{\mathbf{H}}$ can be decomposed as

$$\frac{\partial \tilde{\mathbf{H}}}{\partial t^{\gamma}} = \underline{\tilde{H}}^{\alpha\beta} \Big|_{\gamma} \mathbf{a}_{\alpha} \otimes \mathbf{a}_{\beta} + \underline{\tilde{H}}^{\alpha\beta} \underline{b}_{\beta\gamma} \mathbf{a}_{\alpha} \otimes \widehat{\mathbf{n}} + \underline{\tilde{H}}^{\alpha\beta} \underline{b}_{\alpha\gamma} \widehat{\mathbf{n}} \otimes \mathbf{a}_{\beta} \,, \tag{9.161a}$$

$$\frac{\partial \tilde{\mathbf{H}}}{\partial t^{\gamma}} = \underline{\tilde{H}}^{\alpha}_{.\beta} \Big|_{\gamma} \mathbf{a}_{\alpha} \otimes \mathbf{a}^{\beta} + \underline{\tilde{H}}^{\alpha}_{.\beta} \underline{b}^{\beta}_{.\gamma} \mathbf{a}_{\alpha} \otimes \widehat{\mathbf{n}} + \underline{\tilde{H}}^{\alpha}_{.\beta} \underline{b}_{\alpha\gamma} \widehat{\mathbf{n}} \otimes \mathbf{a}^{\beta} \,, \tag{9.161b}$$

$$\frac{\partial \tilde{\mathbf{H}}}{\partial t^{\gamma}} = \underline{\tilde{H}}_{\alpha}^{.\beta} \Big|_{\gamma} \mathbf{a}^{\alpha} \otimes \mathbf{a}_{\beta} + \underline{\tilde{H}}_{\alpha}^{.\beta} \underline{b}_{\beta\gamma} \mathbf{a}^{\alpha} \otimes \widehat{\mathbf{n}} + \underline{\tilde{H}}_{\alpha}^{.\beta} \underline{b}^{\alpha}_{.\gamma} \widehat{\mathbf{n}} \otimes \mathbf{a}_{\beta} \,, \tag{9.161c}$$

$$\frac{\partial \tilde{\mathbf{H}}}{\partial t^\gamma} = \underline{\tilde{H}}_{\alpha\beta}\Big|_\gamma \, \mathbf{a}_\alpha \otimes \mathbf{a}_\beta + \underline{\tilde{H}}_{\alpha\beta}\, \underline{b}^{\beta}_{.\gamma} \mathbf{a}^\alpha \otimes \widehat{\mathbf{n}} + \underline{\tilde{H}}_{\alpha\beta}\, \underline{b}^{\alpha}_{.\gamma} \widehat{\mathbf{n}} \otimes \mathbf{a}^\beta \ , \tag{9.161d}$$

where

$$\underline{\tilde{H}}^{\alpha\beta}\Big|_\gamma = \frac{\partial\, \underline{\tilde{H}}^{\alpha\beta}}{\partial t^\gamma} + \Gamma^\alpha_{\gamma\theta}\, \underline{\tilde{H}}^{\theta\beta} + \Gamma^\beta_{\gamma\theta}\, \underline{\tilde{H}}^{\alpha\theta} \ , \tag{9.162a}$$

$$\underline{\tilde{H}}^{\alpha}_{\ .\beta}\Big|_\gamma = \frac{\partial\, \underline{\tilde{H}}^{\alpha}_{\ .\beta}}{\partial t^\gamma} + \Gamma^\alpha_{\gamma\theta}\, \underline{\tilde{H}}^{\theta}_{\ .\beta} - \Gamma^\theta_{\gamma\beta}\, \underline{\tilde{H}}^{\alpha}_{\ .\theta} \ , \tag{9.162b}$$

$$\underline{\tilde{H}}_{\alpha}^{\ .\beta}\Big|_\gamma = \frac{\partial\, \underline{\tilde{H}}_{\alpha}^{\ .\beta}}{\partial t^\gamma} - \Gamma^\theta_{\gamma\alpha}\, \underline{\tilde{H}}_{\theta}^{\ .\beta} + \Gamma^\beta_{\gamma\theta}\, \underline{\tilde{H}}_{\alpha}^{\ .\theta} \ , \tag{9.162c}$$

$$\underline{\tilde{H}}_{\alpha\beta}\Big|_\gamma = \frac{\partial\, \underline{\tilde{H}}_{\alpha\beta}}{\partial t^\gamma} - \Gamma^\theta_{\gamma\alpha}\, \underline{\tilde{H}}_{\theta\beta} - \Gamma^\theta_{\gamma\beta}\, \underline{\tilde{H}}_{\alpha\theta} \ . \tag{9.162d}$$

In alignment with (7.28), the surface contravariant derivative of these components are given by

$$\left.\begin{array}{l} \underline{\tilde{H}}^{\alpha\beta}\Big|^\gamma = \underline{\tilde{H}}^{\alpha\beta}\Big|_\theta \, a^{\theta\gamma} \\ \underline{\tilde{H}}^{\alpha}_{\ .\beta}\Big|^\gamma = \underline{\tilde{H}}^{\alpha}_{\ .\beta}\Big|_\theta \, a^{\theta\gamma} \end{array}\right\} \ , \quad \left.\begin{array}{l} \underline{\tilde{H}}_{\alpha}^{\ .\beta}\Big|^\gamma = \underline{\tilde{H}}_{\alpha}^{\ .\beta}\Big|_\theta \, a^{\theta\gamma} \\ \underline{\tilde{H}}_{\alpha\beta}\Big|^\gamma = \underline{\tilde{H}}_{\alpha\beta}\Big|_\theta \, a^{\theta\gamma} \end{array}\right\} \ . \tag{9.163}$$

From $(9.145)_2$ and (9.162d), one can immediately conclude that the surface covariant differentiation of the covariant metric coefficients vanishes. And this exhibits a **metrically connected space**. Indeed, for such a geometry,

$$\boxed{a^{\alpha\beta}\big|_\gamma = 0 \quad , \quad \delta^\alpha_\beta\big|_\gamma = 0 \quad , \quad a_{\alpha\beta}\big|_\gamma = 0 \ .} \tag{9.164}$$

In alignment with the ambient relations (7.34) and (7.35), the metrinilic property of covariant differentiation applies to all surface metrics. But, the covariant and contravariant basis vectors do not remain covariantly constant, i.e. $\mathbf{a}_\alpha|_\beta \neq \mathbf{0}$ and $\mathbf{a}^\alpha|_\beta \neq \mathbf{0}$. It is then easy to see that

$$a_{\beta\rho}\left(\underline{\hat{h}}^\rho\Big|_\gamma\right) = \underline{\hat{h}}_\beta\Big|_\gamma \quad , \quad a^{\beta\rho}\left(\underline{\hat{h}}_\rho\Big|_\gamma\right) = \underline{\hat{h}}^\beta\Big|_\gamma \ . \quad \longleftarrow \text{see (7.37)} \tag{9.165}$$

Suppose one is given $a^{-1/2}\varepsilon^{\alpha\beta}\mathbf{a}_\alpha \otimes \mathbf{a}_\beta$ and $a^{+1/2}\varepsilon_{\alpha\beta}\mathbf{a}^\alpha \otimes \mathbf{a}^\beta$. The interested reader may then want to verify the following metrinilic property of the covariant derivative

$$\left(\frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\right)\Bigg|_\gamma = 0 \quad , \quad \left(\sqrt{a}\varepsilon_{\alpha\beta}\right)\Big|_\gamma = 0 \ . \quad \longleftarrow \text{see (9.670)} \tag{9.166}$$

To characterize the surface Riemann-Christoffel curvature tensor, one needs to have the second-order covariant derivatives of a surface vector. This motivates to represent the second-order partial and covariant derivatives of a surface vector in the following.

9.3.2 Surface Second Covariant Differentiation

By means of the relations (9.99)$_{1-2}$, (9.128), (9.131)$_{1-2}$, (9.132)$_3$, (9.133)$_{1-2}$, (9.156) and (9.157)$_3$ along with the product rule of differentiation, the surface version of (7.40a)–(7.40b) can be expressed as

$$\frac{\partial^2 \hat{\mathbf{h}}}{\partial t^\gamma \partial t^\delta} = \left[\underline{\hat{h}}^\alpha{}_{|\delta\gamma} - \underline{\hat{h}}^\theta \, \underline{b}_{\theta\delta} \, \underline{b}_\gamma^{\cdot\alpha} \right] \mathbf{a}_\alpha + \left[\underline{\hat{h}}^\alpha \Big|_\delta \, \underline{b}_{\alpha\gamma} + \frac{\partial}{\partial t^\gamma} \left(\underline{\hat{h}}^\alpha \, \underline{b}_{\alpha\delta} \right) \right] \widehat{\mathbf{n}} \,, \tag{9.167a}$$

$$\frac{\partial^2 \hat{\mathbf{h}}}{\partial t^\gamma \partial t^\delta} = \left[\underline{\hat{h}}_{\alpha|\delta\gamma} - \underline{\hat{h}}_\theta \, \underline{b}^\theta{}_{\cdot\delta} \, \underline{b}_{\gamma\alpha} \right] \mathbf{a}^\alpha + \left[\underline{\hat{h}}_\alpha \Big|_\delta \, \underline{b}^\alpha{}_{\cdot\gamma} + \frac{\partial}{\partial t^\gamma} \left(\underline{\hat{h}}_\alpha \, \underline{b}^\alpha{}_{\cdot\delta} \right) \right] \widehat{\mathbf{n}} \,, \tag{9.167b}$$

where

$$\underline{\hat{h}}^\alpha{}_{|\delta\gamma} = \frac{\partial^2 \underline{\hat{h}}^\alpha}{\partial t^\gamma \partial t^\delta} + \frac{\partial \Gamma^\alpha_{\delta\theta}}{\partial t^\gamma} \underline{\hat{h}}^\theta + \Gamma^\alpha_{\gamma\rho} \Gamma^\rho_{\delta\theta} \underline{\hat{h}}^\theta + \Gamma^\alpha_{\delta\theta} \frac{\partial \underline{\hat{h}}^\theta}{\partial t^\gamma} + \Gamma^\alpha_{\gamma\theta} \frac{\partial \underline{\hat{h}}^\theta}{\partial t^\delta} \,, \tag{9.168a}$$

$$\underline{\hat{h}}_{\alpha|\delta\gamma} = \frac{\partial^2 \underline{\hat{h}}_\alpha}{\partial t^\gamma \partial t^\delta} + \Gamma^\rho_{\gamma\alpha} \Gamma^\theta_{\rho\delta} \underline{\hat{h}}_\theta - \frac{\partial \Gamma^\theta_{\alpha\delta}}{\partial t^\gamma} \underline{\hat{h}}_\theta - \Gamma^\theta_{\alpha\delta} \frac{\partial \underline{\hat{h}}_\theta}{\partial t^\gamma} - \Gamma^\theta_{\gamma\alpha} \frac{\partial \underline{\hat{h}}_\theta}{\partial t^\delta} \,. \tag{9.168b}$$

The true *surface second-order covariant derivatives* are then given by

$$\begin{aligned} \underline{\hat{h}}^\alpha \Big|_{\delta\gamma} &= \left(\underline{\hat{h}}^\alpha \Big|_\delta \right) \Big|_\gamma \\ &= \frac{\partial}{\partial t^\gamma} \left(\underline{\hat{h}}^\alpha \Big|_\delta \right) + \Gamma^\alpha_{\gamma\theta} \left(\underline{\hat{h}}^\theta \Big|_\delta \right) - \Gamma^\theta_{\gamma\delta} \left(\underline{\hat{h}}^\alpha \Big|_\theta \right) \\ &= \boxed{ \underline{\hat{h}}^\alpha{}_{|\delta\gamma} - \Gamma^\rho_{\gamma\delta} \left(\frac{\partial \underline{\hat{h}}^\alpha}{\partial t^\rho} + \Gamma^\alpha_{\rho\theta} \underline{\hat{h}}^\theta \right) } \,, \end{aligned} \tag{9.169a}$$

$$\begin{aligned} \underline{\hat{h}}_\alpha \Big|_{\delta\gamma} &= \left(\underline{\hat{h}}_\alpha \Big|_\delta \right) \Big|_\gamma \\ &= \frac{\partial}{\partial t^\gamma} \left(\underline{\hat{h}}_\alpha \Big|_\delta \right) - \Gamma^\theta_{\gamma\alpha} \left(\underline{\hat{h}}_\theta \Big|_\delta \right) - \Gamma^\theta_{\gamma\delta} \left(\underline{\hat{h}}_\alpha \Big|_\theta \right) \\ &= \boxed{ \underline{\hat{h}}_{\alpha|\delta\gamma} - \Gamma^\rho_{\gamma\delta} \left(\frac{\partial \underline{\hat{h}}_\alpha}{\partial t^\rho} - \Gamma^\theta_{\alpha\rho} \underline{\hat{h}}_\theta \right) } \,. \end{aligned} \tag{9.169b}$$

It follows that

$$\underline{\hat{h}}^{\alpha}\Big|_{\delta\gamma} - \underline{\hat{h}}^{\alpha}\Big|_{\gamma\delta} = \underline{\hat{h}}^{\alpha}{}_{|\delta\gamma} - \underline{\hat{h}}^{\alpha}{}_{|\gamma\delta} \quad , \quad \underline{\hat{h}}_{\alpha}\Big|_{\delta\gamma} - \underline{\hat{h}}_{\alpha}\Big|_{\gamma\delta} = \underline{\hat{h}}_{\alpha|\delta\gamma} - \underline{\hat{h}}_{\alpha|\gamma\delta} \,. \tag{9.170}$$

The above relations can be extended to surface tensors of higher orders in a lengthy but straightforward manner. This remains to be done by the ambitious reader.

9.3.3 Invariance of Surface Covariant Differentiation

Let $\left(t^1, t^2\right)$ and $\left(\bar{t}^1, \bar{t}^2\right)$ be an **old** and a **new** surface coordinates, respectively, which are related according to the relationships (9.40). Consider now the surface vector $\hat{\mathbf{h}} = \underline{\bar{\hat{h}}}^{\alpha}\bar{\mathbf{a}}_{\alpha} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha} = \underline{\bar{\hat{h}}}_{\alpha}\bar{\mathbf{a}}^{\alpha} = \underline{\hat{h}}_{\alpha}\mathbf{a}^{\alpha}$ whose components and basis vectors tensorially transform according to

$$\underline{\bar{\hat{h}}}^{\alpha} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\beta}}\,\underline{\hat{h}}^{\beta} \quad , \quad \bar{\mathbf{a}}_{\alpha} = \frac{\partial t^{\gamma}}{\partial \bar{t}^{\alpha}}\mathbf{a}_{\gamma} \,, \tag{9.171a}$$

$$\underline{\bar{\hat{h}}}_{\alpha} = \frac{\partial t^{\beta}}{\partial \bar{t}^{\alpha}}\,\underline{\hat{h}}_{\beta} \quad , \quad \bar{\mathbf{a}}^{\alpha} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}}\mathbf{a}^{\gamma} \,. \tag{9.171b}$$

← see (7.29b)

The surface partial derivative $\partial/\partial t^{\alpha}$ preserves the tensor property when its operands are invariant field variables. But, this operator does not provide tensors when applied to variants. For instance,

$$\frac{\partial\,\underline{\bar{\hat{h}}}^{\alpha}}{\partial \bar{t}^{\beta}} = \frac{\partial}{\partial \bar{t}^{\beta}}\left[\frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}}\,\underline{\hat{h}}^{\gamma}\right] = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}}\frac{\partial\,\underline{\hat{h}}^{\gamma}}{\partial t^{\delta}}\frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} + \frac{\partial^2 \bar{t}^{\alpha}}{\partial t^{\delta}\partial t^{\gamma}}\frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}}\,\underline{\hat{h}}^{\gamma} \,, \tag{9.172a}$$

$$\frac{\partial\,\underline{\bar{\hat{h}}}_{\alpha}}{\partial \bar{t}^{\beta}} = \frac{\partial}{\partial \bar{t}^{\beta}}\left[\frac{\partial t^{\gamma}}{\partial \bar{t}^{\alpha}}\,\underline{\hat{h}}_{\gamma}\right] = \frac{\partial t^{\gamma}}{\partial \bar{t}^{\alpha}}\frac{\partial\,\underline{\hat{h}}_{\gamma}}{\partial t^{\delta}}\frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} + \frac{\partial^2 t^{\gamma}}{\partial \bar{t}^{\beta}\partial \bar{t}^{\alpha}}\,\underline{\hat{h}}_{\gamma} \,. \tag{9.172b}$$

← see (7.30b)

Note that in these transformations, the deficiency stems from the development of the last nontensorial terms. This problem is resolved by using the technique of surface covariant derivative:

$$\boxed{\underline{\bar{\hat{h}}}^{\alpha}\Big|_{\beta} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}}\left(\underline{\hat{h}}^{\gamma}\Big|_{\delta}\right)\frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} \,,} \tag{9.173a}$$

$$\boxed{\underline{\bar{\hat{h}}}_{\alpha}\Big|_{\beta} = \frac{\partial t^{\gamma}}{\partial \bar{t}^{\alpha}}\left(\underline{\hat{h}}_{\gamma}\Big|_{\delta}\right)\frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} \,.} \tag{9.173b}$$

← see (7.31b)

In a similar manner,

$$\underline{\bar{\tilde{H}}}^{\alpha\beta}\Big|_{\gamma} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\delta}} \frac{\partial \bar{t}^{\beta}}{\partial t^{\theta}} \left(\underline{\tilde{H}}^{\delta\theta}\Big|_{\rho} \right) \frac{\partial t^{\rho}}{\partial \bar{t}^{\gamma}} , \tag{9.174a}$$

$$\underline{\bar{\tilde{H}}}^{\alpha}{}_{.\beta}\Big|_{\gamma} = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\delta}} \frac{\partial t^{\theta}}{\partial \bar{t}^{\beta}} \left(\underline{\bar{\tilde{H}}}^{\delta}{}_{.\theta}\Big|_{\rho} \right) \frac{\partial t^{\rho}}{\partial \bar{t}^{\gamma}} , \tag{9.174b}$$

$$\underline{\bar{\tilde{H}}}_{\alpha}{}^{.\beta}\Big|_{\gamma} = \frac{\partial t^{\delta}}{\partial \bar{t}^{\alpha}} \frac{\partial \bar{t}^{\beta}}{\partial t^{\theta}} \left(\underline{\tilde{H}}_{\delta}{}^{.\theta}\Big|_{\rho} \right) \frac{\partial t^{\rho}}{\partial \bar{t}^{\gamma}} , \tag{9.174c}$$

$$\underline{\bar{\tilde{H}}}_{\alpha\beta}\Big|_{\gamma} = \frac{\partial t^{\delta}}{\partial \bar{t}^{\alpha}} \frac{\partial t^{\theta}}{\partial \bar{t}^{\beta}} \left(\underline{\tilde{H}}_{\delta\theta}\Big|_{\rho} \right) \frac{\partial t^{\rho}}{\partial \bar{t}^{\gamma}} . \quad \leftarrow \text{see (7.32d)} \tag{9.174d}$$

9.3.4 Surface Covariant Differentiation of Invariant Objects with Surface and Ambient Indices

In the literature, the surface covariant derivative of invariant tensor fields is sometimes introduced in a slightly different way, see Naghdi [14] and Grinfeld [15]. This strategy will be followed in the next section for the study of calculus of moving surfaces. Recall from (9.129) that the surface covariant derivative coincides with the surface partial differentiation when applied to an invariant object of order 0. The goal here is to extend this result to invariant tensor fields of higher ranks. For instance,

$$\boxed{\hat{\mathbf{h}}\Big|_{\beta} := \frac{\partial \hat{\mathbf{h}}}{\partial t^{\beta}} ,} \quad \leftarrow \text{see (7.39)} \tag{9.175}$$

where $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha} = \underline{\hat{h}}_{\alpha}\mathbf{a}^{\alpha}$ denotes an invariant surface object of order 1. Referred to the surface covariant and contravariant basis vectors, the partial derivative $\partial\hat{\mathbf{h}}/\partial t^{\beta}$ has been represented in $(9.132)_3$ and $(9.157)_3$, respectively. Consistent with (7.38a)–(7.38b), of interest here is to demand that the surface covariant derivative satisfies the sum and product rules in the sense that

$$\hat{\mathbf{h}}\Big|_{\beta} = \left(\underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}\right)\Big|_{\beta} = \left(\underline{\hat{h}}^{\alpha}\Big|_{\beta}\right)\mathbf{a}_{\alpha} + \underline{\hat{h}}^{\alpha}\left(\mathbf{a}_{\alpha}|_{\beta}\right) , \tag{9.176a}$$

$$\hat{\mathbf{h}}\Big|_{\beta} = \left(\underline{\hat{h}}_{\alpha}\mathbf{a}^{\alpha}\right)\Big|_{\beta} = \left(\underline{\hat{h}}_{\alpha}\Big|_{\beta}\right)\mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha}\left(\mathbf{a}^{\alpha}|_{\beta}\right) . \tag{9.176b}$$

This immediately implies that

$$\boxed{\mathbf{a}_{\alpha}|_{\beta} = \underline{b}_{\alpha\beta}\widehat{\mathbf{n}} ,} \tag{9.177a}$$

$$\boxed{\mathbf{a}^{\alpha}|_{\beta} = \underline{b}^{\alpha}{}_{.\beta}\,\widehat{\mathbf{n}} .} \tag{9.177b}$$

The demand for keeping the above structure for any arbitrary ambient object will lead to some useful properties. This will be characterized in the following.

To begin with, let $\hat{\mathbf{h}} = \underline{\hat{h}}^i \mathbf{g}_i$ be a smooth first-order tensor field of the ambient coordinates $\left(\Theta^1, \Theta^2, \Theta^3\right)$ which themselves are functions of the Gaussian coordinates $\left(t^1, t^2\right)$. The demand for

$$\begin{aligned}
\hat{\mathbf{h}}\Big|_\alpha &= \frac{\partial \hat{\mathbf{h}}}{\partial t^\alpha} \\
&\xlongequal[\text{product rules of differentiation}]{\text{by using the chain and}} \frac{\partial\, \underline{\hat{h}}^i}{\partial t^\alpha}\mathbf{g}_i + \underline{\hat{h}}^i \frac{\partial \mathbf{g}_i}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial t^\alpha} \\
&\xlongequal[\text{(7.7), (7.8) and (9.47)}]{\text{from}} \frac{\partial\, \underline{\hat{h}}^i}{\partial t^\alpha}\mathbf{g}_i + \underline{\hat{h}}^i \Gamma^m_{ji} \overline{Z}^j_\alpha \mathbf{g}_m \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} \left(\frac{\partial\, \underline{\hat{h}}^i}{\partial t^\alpha} + \overline{Z}^j_\alpha \Gamma^i_{jm}\, \underline{\hat{h}}^m \right) \mathbf{g}_i \ ,
\end{aligned} \tag{9.178}$$

along with insisting on satisfying the sum and product rules

$$\hat{\mathbf{h}}\Big|_\alpha = \left(\underline{\hat{h}}^i\Big|_\alpha \right)\mathbf{g}_i + \underline{\hat{h}}^i \left(\mathbf{g}_i|_\alpha \right) \ , \tag{9.179}$$

then helps establish

$$\underline{\hat{h}}^i\Big|_\alpha = \frac{\partial\, \underline{\hat{h}}^i}{\partial t^\alpha} + \overline{Z}^j_\alpha \Gamma^i_{jm}\, \underline{\hat{h}}^m = \left(\frac{\partial\, \underline{\hat{h}}^i}{\partial \Theta^j} + \Gamma^i_{jm}\, \underline{\hat{h}}^m \right)\overline{Z}^j_\alpha = \underline{\hat{h}}^i\Big|_j \overline{Z}^j_\alpha \ , \tag{9.180}$$

and

$$\boxed{\mathbf{g}_i|_\alpha = \mathbf{0}\ .} \tag{9.181}$$

In a similar manner, for an arbitrary ambient first-order tensor field $\hat{\mathbf{h}} = \underline{\hat{h}}_i \mathbf{g}^i$, one will have

$$\underline{\hat{h}}_i\Big|_\alpha = \frac{\partial\, \underline{\hat{h}}_i}{\partial t^\alpha} - \overline{Z}^j_\alpha \Gamma^m_{ji}\, \underline{\hat{h}}_m = \left(\frac{\partial\, \underline{\hat{h}}_i}{\partial \Theta^j} - \Gamma^m_{ij}\, \underline{\hat{h}}_m \right)\overline{Z}^j_\alpha = \underline{\hat{h}}_i\Big|_j \overline{Z}^j_\alpha \ , \tag{9.182}$$

and

$$\boxed{\mathbf{g}^i\Big|_\alpha = \mathbf{0}\ .} \tag{9.183}$$

The expressions $(9.180)_{1-2}$ and $(9.182)_{1-2}$ show that how the ambient vectors and covectors should covariantly be differentiated along a surface while the results (9.181) and (9.183) demonstrate that the surface covariant derivative is metrinilic with respect to the ambient basis vectors, see Grinfeld [15] for more details. Notice that the properties $\mathbf{g}_i|_\alpha = \mathbf{0}$ and $\mathbf{g}^i\Big|_\alpha = \mathbf{0}$ can also be obtained by formally applying the surface covariant derivative to these ambient basis vectors, that is,

$$\mathbf{g}_i\big|_\alpha \xlongequal[\text{(9.182)}]{\text{from}} \mathbf{g}_i\big|_j \, \overline{Z}^{\,j}_\alpha \xlongequal[\text{(7.36a)}]{\text{from}} \mathbf{0} \quad , \quad \mathbf{g}^i\big|_\alpha \xlongequal[\text{(9.180)}]{\text{from}} \mathbf{g}^i\big|_j \, \overline{Z}^{\,j}_\alpha \xlongequal[\text{(7.36b)}]{\text{from}} \mathbf{0} \ .$$

Consequently, the surface covariant derivative should also be metrinilic with respect to the ambient metric coefficients:

$$\boxed{g^{ij}\big|_\alpha = 0 \quad , \quad \delta^i_j\big|_\alpha = 0 \quad , \quad g_{ij}\big|_\alpha = 0 \ .} \quad \longleftarrow \text{see (7.34)–(7.35)} \tag{9.184}$$

Let the ambient object $\widehat{\mathbf{n}}$ be the unit normal vector to the surface. With the aid of (9.99)$_{1-2}$ and (9.175), one then writes

$$\boxed{\widehat{\mathbf{n}}\big|_\alpha = \frac{\partial \widehat{\mathbf{n}}}{\partial t^\alpha} = -\, \underline{b}_{\alpha}^{\cdot\beta} \mathbf{a}_\beta = -\, \underline{b}_{\alpha\beta} \mathbf{a}^\beta \ .} \tag{9.185}$$

Consider now the decomposition $\widehat{\mathbf{n}} = \underline{\widehat{n}}^i \mathbf{g}_i$. Making use of (9.46)$_3$, the surface covariant derivative in (9.185)$_2$ can be rephrased as $\widehat{\mathbf{n}}\big|_\alpha = -\, \underline{b}_{\alpha}^{\cdot\beta} \overline{Z}^{\,i}_\beta \mathbf{g}_i$. The metrinilic property of the surface covariant differentiation with respect to the ambient basis vectors then implies that

$$\boxed{\underline{\widehat{n}}^i\big|_\alpha = -\, \underline{b}_{\alpha}^{\cdot\beta} \overline{Z}^{\,i}_\beta \ .} \tag{9.186}$$

In a similar manner,

$$\boxed{\underline{\widehat{n}}_i\big|_\alpha = -\, \underline{b}_{\alpha\beta} \widetilde{Z}^{\beta}_i \ .} \tag{9.187}$$

Next, let $\tilde{\mathbf{H}} = \underline{\tilde{H}}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j = \underline{\tilde{H}}^{i}_{\ \cdot j} \mathbf{g}_i \otimes \mathbf{g}^j = \underline{\tilde{H}}_{i}^{\ \cdot j} \mathbf{g}^i \otimes \mathbf{g}_j = \underline{\tilde{H}}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j$ be a smooth tensor field of $\left(\Theta^1\left(t^1, t^2\right), \Theta^2\left(t^1, t^2\right), \Theta^3\left(t^1, t^2\right)\right)$. It is then easy to see that

$$\underline{\tilde{H}}^{ij}\Big|_\alpha = \underline{\tilde{H}}^{ij}\Big|_k \overline{Z}^{\,k}_\alpha \ , \quad \longleftarrow \text{see (7.27a)} \tag{9.188a}$$

$$\underline{\tilde{H}}^{i}_{\ \cdot j}\Big|_\alpha = \underline{\tilde{H}}^{i}_{\ \cdot j}\Big|_k \overline{Z}^{\,k}_\alpha \ , \tag{9.188b}$$

$$\underline{\tilde{H}}_{i}^{\ \cdot j}\Big|_\alpha = \underline{\tilde{H}}_{i}^{\ \cdot j}\Big|_k \overline{Z}^{\,k}_\alpha \ , \tag{9.188c}$$

$$\underline{\tilde{H}}_{ij}\Big|_\alpha = \underline{\tilde{H}}_{ij}\Big|_k \overline{Z}^{\,k}_\alpha \ . \tag{9.188d}$$

Regarding the ambient covariant metric coefficients, the relation (9.188d) with the aid of the results (7.27d)$_2$ and (9.184)$_3$ delivers

$$\frac{\partial g_{ij}}{\partial t^\alpha} = \overline{Z}^{\,k}_\alpha \Gamma^m_{ki} g_{mj} + \overline{Z}^{\,k}_\alpha \Gamma^m_{kj} g_{im} \ . \tag{9.189}$$

Consider now two smooth ambient vector fields $\hat{\mathbf{h}}_1 = \underline{\hat{h}}^i_1 \mathbf{g}_i$ and $\hat{\mathbf{h}}_2 = \underline{\hat{h}}^j_2 \mathbf{g}_j$ whose scalar product represents $\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2 = \underline{\hat{h}}^i_1 g_{ij} \, \underline{\hat{h}}^j_2$. Then, by using (9.129) and (9.189) along with the product rule of differentiation,

$$\left(\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2 \right) \Big|_\alpha = \frac{\partial}{\partial t^\alpha} \left(\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2 \right) = \frac{\partial \, \underline{\hat{h}}^i_1}{\partial t^\alpha} g_{ij} \, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 \frac{\partial g_{ij}}{\partial t^\alpha} \, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 g_{ij} \frac{\partial \, \underline{\hat{h}}^j_2}{\partial t^\alpha}$$
$$= \frac{\partial \, \underline{\hat{h}}^i_1}{\partial t^\alpha} g_{ij} \, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 \overline{Z}^k_\alpha \Gamma^m_{ki} g_{mj} \, \underline{\hat{h}}^j_2$$
$$+ \underline{\hat{h}}^i_1 \overline{Z}^k_\alpha \Gamma^m_{kj} g_{im} \, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 g_{ij} \frac{\partial \, \underline{\hat{h}}^j_2}{\partial t^\alpha} . \tag{9.190}$$

It should not be difficult now to verify that

$$\boxed{\left(\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2 \right) \Big|_\alpha = \left(\hat{\mathbf{h}}_1 \Big|_\alpha \right) \cdot \hat{\mathbf{h}}_2 + \hat{\mathbf{h}}_1 \cdot \left(\hat{\mathbf{h}}_2 \Big|_\alpha \right) .} \quad \longleftarrow \text{see (9.136)} \tag{9.191}$$

At the end, the surface covariant derivative of the shift tensors $\overline{Z}^i_\alpha = \mathbf{g}^i \cdot \mathbf{a}_\alpha$ and $\widetilde{Z}^\alpha_i = \mathbf{g}_i \cdot \mathbf{a}^\alpha$, respectively given in $(9.47)_1$ and $(9.60)_1$, will be characterized for completeness. Having in mind (9.177a)-(9.177b), the metrinilic property of the surface covariant derivative relative to the ambient basis vectors then helps represent

$$\boxed{\overline{Z}^i_\alpha \Big|_\beta = \mathbf{g}^i \cdot \mathbf{a}_\alpha|_\beta = \underline{b}_{\alpha\beta} \mathbf{g}^i \cdot \widehat{\mathbf{n}} = \underline{b}_{\alpha\beta} \, \underline{\widehat{n}}^i \ ,} \tag{9.192}$$

or $\underline{b}_{\alpha\beta} = \overline{Z}^i_\alpha \Big|_\beta \, \underline{\widehat{n}}_i$ considering the identity $\underline{\widehat{n}}^i \, \underline{\widehat{n}}_i = 1$, and

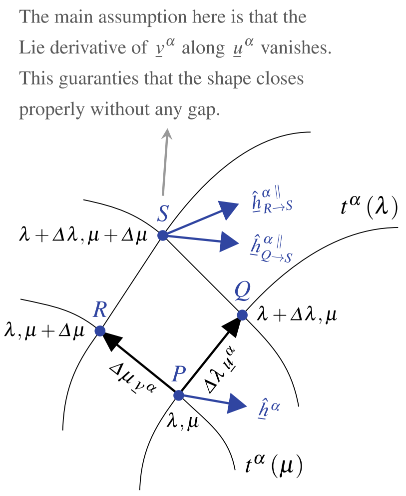
$$\boxed{\widetilde{Z}^\alpha_i \big|_\beta = \mathbf{g}_i \cdot \mathbf{a}^\alpha|_\beta = \underline{b}^\alpha_{.\beta} \mathbf{g}_i \cdot \widehat{\mathbf{n}} = \underline{b}^\alpha_{.\beta} \, \underline{\widehat{n}}_i \ ,} \tag{9.193}$$

or $\underline{b}^\alpha_{.\beta} = \widetilde{Z}^\alpha_i \big|_\beta \, \underline{\widehat{n}}^i$ considering the identity $\underline{\widehat{n}}_i \, \underline{\widehat{n}}^i = 1$.

9.4 Surface Riemann-Christoffel Curvature Tensor

The aim here is to geometrically describe the *surface Riemann-Christoffel curvature* using the powerful tool of parallel transport. It will be shown that how such an important object in differential geometry helps establish noncommutativity of the covariant derivative when applied to tensors. The fundamental properties of the surface Riemann-Christoffel curvature tensor will also be addressed.

Let $t^\alpha(\lambda)$ and $t^\alpha(\mu)$ be two parametrized curves embedded in a surface $\mathcal{S}$ as illustrated in Fig. 9.14.



Let $t^\alpha(\lambda)$ and $t^\alpha(\mu)$ be two parametrized curves embedded in a surface $\mathcal{S}$. The tangent vector to $t^\alpha(\lambda)$ $(t^\alpha(\mu))$ at P is $\underline{u}^\alpha$ $(\underline{v}^\alpha)$. Let $\hat{\underline{h}}^\alpha$ be a surface vector at P. Moreover, let $\hat{\underline{h}}^{\alpha\,\|}_{Q\to S}$ $(\hat{\underline{h}}^{\alpha\,\|}_{R\to S})$ be the parallel transport of $\hat{\underline{h}}^\alpha$ from P to Q (R) and subsequently from Q (R) to S. After these transportation processes, the change in $\hat{\underline{h}}^\alpha$ is governed by

$$\delta\,\hat{\underline{h}}^\alpha = \lim_{\Delta\lambda,\Delta\mu\to 0} \frac{\hat{\underline{h}}^{\alpha\,\|}_{Q\to S} - \hat{\underline{h}}^{\alpha\,\|}_{R\to S}}{\Delta\lambda\Delta\mu} .$$

This finally leads to

$$\delta\,\hat{\underline{h}}^\alpha = \underline{\mathbb{R}}^{\alpha}{}_{.\beta\gamma\delta}\,\hat{\underline{h}}^\beta\,\underline{v}^\gamma\,\underline{u}^\delta ,$$

where the tensorial object

$$\underline{\mathbb{R}}^{\alpha}{}_{.\beta\gamma\delta} = \frac{\partial\Gamma^\alpha_{\beta\delta}}{\partial t^\gamma} - \frac{\partial\Gamma^\alpha_{\beta\gamma}}{\partial t^\delta} + \Gamma^\alpha_{\gamma\theta}\Gamma^\theta_{\delta\beta} - \Gamma^\alpha_{\delta\theta}\Gamma^\theta_{\gamma\beta} ,$$

is called the surface Riemann-Christoffel curvature.

Fig. 9.14 Surface Riemann-Christoffel curvature

Consider an infinitesimal closed loop $PQSR$. Let $\hat{\underline{h}}^\alpha$ be a smooth surface vector sitting at point P. According to (9.240)$_3$, the tangent vector to $t^\alpha(\lambda)$ $(t^\alpha(\mu))$ at P is $\mathbf{a}_\lambda = \underline{u}^\alpha\mathbf{a}_\alpha$ $(\mathbf{a}_\mu = \underline{v}^\alpha\mathbf{a}_\alpha)$ where $\underline{u}^\alpha = dt^\alpha/d\lambda$ $(\underline{v}^\alpha = dt^\alpha/d\mu)$. The goal is to compare the vectors $\hat{\underline{h}}^{\alpha\,\|}_{Q\to S}$ and $\hat{\underline{h}}^{\alpha\,\|}_{R\to S}$ which help characterize the Riemann-Christoffel curvature tensor. Here, $\hat{\underline{h}}^{\alpha\,\|}_{Q\to S}$ $(\hat{\underline{h}}^{\alpha\,\|}_{R\to S})$ denotes the parallel transport of $\hat{\underline{h}}^\alpha$ from P to Q (R) and subsequently from Q (R) to S. Note that one can alternatively compare the original vector $\hat{\underline{h}}^\alpha$ with one obtained by its parallel transport around the whole loop. Although the final result is the same, the computations described below are easier.

Guided by (9.141)$_3$, the change in $\hat{\underline{h}}^\alpha$ when it is parallel transported from P to Q is

$$\Delta\,\hat{\underline{h}}^\theta = -\left(\Gamma^\theta_{\sigma\tau}\,\hat{\underline{h}}^\tau\,\frac{dt^\sigma}{d\lambda}\right)_P \Delta\lambda \quad \text{or} \quad \hat{\underline{h}}^{\theta\,\|}_{P\to Q} = \hat{\underline{h}}^\theta - \Gamma^\theta_{\sigma\tau}\,\hat{\underline{h}}^\tau\,\underline{u}^\sigma\,\Delta\lambda . \tag{9.194}$$

In a similar manner,

$$\underline{v}^{\beta\,\|}_{P\to Q} = \underline{v}^\beta - \Gamma^\beta_{\phi\varphi}\,\underline{v}^\varphi\,\underline{u}^\phi\,\Delta\lambda . \tag{9.195}$$

The connection coefficients $\Gamma^\alpha_{\beta\theta\ Q} := \Gamma^\alpha_{\beta\theta}\left(t^1_Q, t^2_Q\right)$ can be expressed in terms of $\Gamma^\alpha_{\beta\theta}$ and their partial derivatives via the following first-order Taylor series expansion

$$\begin{aligned} \Gamma^\alpha_{\beta\theta\ Q} &= \Gamma^\alpha_{\beta\theta\ P} + \left(\frac{\partial\Gamma^\alpha_{\beta\theta}}{\partial t^\gamma}\frac{dt^\gamma}{d\lambda}\right)\Bigg|_P \Delta\lambda \\ &= \Gamma^\alpha_{\beta\theta} + \frac{\partial\Gamma^\alpha_{\beta\theta}}{\partial t^\gamma}\,\underline{u}^\gamma\,\Delta\lambda . \quad \longleftarrow \text{see (6.24)–(6.25)} \end{aligned} \tag{9.196}$$

Now, the change in $\underline{\hat{h}}^{\alpha}$ as it is parallel transported from Q to S renders

$$\begin{aligned}
\underline{\hat{h}}^{\alpha\,\|}_{Q\to S} &= \underline{\hat{h}}^{\alpha\,\|}_{P\to Q} - \Gamma^{\alpha}_{\beta\theta\;Q}\,\underline{\hat{h}}^{\theta\,\|}_{P\to Q}\,\underline{v}^{\beta\,\|}_{P\to Q}\Delta\mu \\
&= \underline{\hat{h}}^{\alpha} - \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\,\underline{u}^{\beta}\Delta\lambda \\
&\quad - \underbrace{\left(\Gamma^{\alpha}_{\beta\theta} + \frac{\partial\Gamma^{\alpha}_{\beta\theta}}{\partial t^{\gamma}}\,\underline{u}^{\gamma}\Delta\lambda\right)\left(\underline{\hat{h}}^{\theta} - \Gamma^{\theta}_{\sigma\tau}\,\underline{\hat{h}}^{\tau}\,\underline{u}^{\sigma}\Delta\lambda\right)}_{=\Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta} - \Gamma^{\alpha}_{\beta\theta}\Gamma^{\theta}_{\sigma\tau}\,\underline{\hat{h}}^{\tau}\,\underline{u}^{\sigma}\Delta\lambda + \frac{\partial\Gamma^{\alpha}_{\beta\theta}}{\partial t^{\gamma}}\,\underline{u}^{\gamma}\,\underline{\hat{h}}^{\theta}\Delta\lambda}\left(\underline{v}^{\beta} - \Gamma^{\beta}_{\phi\varphi}\,\underline{v}^{\varphi}\,\underline{u}^{\phi}\Delta\lambda\right)\Delta\mu \\
&= \underline{\hat{h}}^{\alpha} - \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\,\underline{u}^{\beta}\Delta\lambda - \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\,\underline{v}^{\beta}\Delta\mu \\
&\quad + \left(-\frac{\partial\Gamma^{\alpha}_{\beta\theta}}{\partial t^{\gamma}}\,\underline{\hat{h}}^{\theta}\,\underline{v}^{\beta}\,\underline{u}^{\gamma} + \Gamma^{\alpha}_{\beta\theta}\Gamma^{\theta}_{\sigma\tau}\,\underline{\hat{h}}^{\tau}\,\underline{v}^{\beta}\,\underline{u}^{\sigma} + \Gamma^{\alpha}_{\beta\theta}\Gamma^{\beta}_{\phi\varphi}\,\underline{\hat{h}}^{\theta}\,\underline{v}^{\varphi}\,\underline{u}^{\phi}\right)\Delta\lambda\Delta\mu\,,
\end{aligned}$$

where the higher-order terms have not been written here (since, in the limit when $\Delta\lambda$ and $\Delta\mu$ go to zero, they vanish). In a similar manner,

$$\begin{aligned}
\underline{\hat{h}}^{\alpha\,\|}_{R\to S} &= \underline{\hat{h}}^{\alpha} - \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\,\underline{v}^{\beta}\Delta\mu - \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\,\underline{u}^{\beta}\Delta\lambda \\
&\quad + \left(-\frac{\partial\Gamma^{\alpha}_{\beta\theta}}{\partial t^{\gamma}}\,\underline{\hat{h}}^{\theta}\,\underline{v}^{\gamma}\,\underline{u}^{\beta} + \Gamma^{\alpha}_{\beta\theta}\Gamma^{\theta}_{\sigma\tau}\,\underline{\hat{h}}^{\tau}\,\underline{v}^{\sigma}\,\underline{u}^{\beta} + \Gamma^{\alpha}_{\beta\theta}\Gamma^{\beta}_{\phi\varphi}\,\underline{\hat{h}}^{\theta}\,\underline{v}^{\phi}\,\underline{u}^{\varphi}\right)\Delta\mu\Delta\lambda\,.
\end{aligned}$$

The difference $\underline{\hat{h}}^{\alpha\,\|}_{Q\to S} - \underline{\hat{h}}^{\alpha\,\|}_{R\to S}$ per unit (coordinate) area is a measure of curvature (note that curvature is a **local** concept defined at each point of manifold). Let

$$\delta\,\underline{\hat{h}}^{\alpha} = \lim_{\Delta\lambda,\Delta\mu\to 0}\frac{\underline{\hat{h}}^{\alpha\,\|}_{Q\to S} - \underline{\hat{h}}^{\alpha\,\|}_{R\to S}}{\Delta\lambda\Delta\mu}\,. \tag{9.197}$$

This deviation of $\underline{\hat{h}}^{\alpha}$ is then given by

$$\boxed{\delta\,\underline{\hat{h}}^{\alpha} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}\,\underline{\hat{h}}^{\beta}\,\underline{v}^{\gamma}\,\underline{u}^{\delta}\,,} \tag{9.198}$$

where $\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}$ is known as the *surface Riemann-Christoffel curvature tensor* (see Carroll [16] and Dalarsson and Dalarsson [17]):

$$\boxed{\underbrace{\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta} = \frac{\partial\Gamma^{\alpha}_{\beta\delta}}{\partial t^{\gamma}} - \frac{\partial\Gamma^{\alpha}_{\beta\gamma}}{\partial t^{\delta}} + \Gamma^{\alpha}_{\gamma\theta}\Gamma^{\theta}_{\delta\beta} - \Gamma^{\alpha}_{\delta\theta}\Gamma^{\theta}_{\gamma\beta}}_{\text{note that although } \Gamma^{\alpha}_{\beta\delta} \text{ is not a tensor, this object truly represents a tensor}}\,.} \tag{9.199}$$

The fourth-order tensor $\mathbb{R}$ in (9.198) may be thought of as an operator which takes the objects **u** and **v**, characterizing the orientation of an infinitesimal parallelogram, and acts on a vector $\hat{\mathbf{h}}$ to measure how much this vector changes when it is parallel transported around such a parallelogram. Notice that all elements used to define the

Riemann-Christoffel curvature were intrinsic objects. And this means that the object $\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot}$ has an **intrinsic** nature. See the pioneering work of Riemann [18].

Let $\underline{\hat{h}}_{\alpha}$ be a generic covector. Following similar procedures which led to (9.198) then reveals

$$\boxed{\delta\,\underline{\hat{h}}_{\beta} = -\,\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot}\;\underline{\hat{h}}_{\alpha}\,\underline{v}^{\gamma}\,\underline{u}^{\delta}\;.} \tag{9.200}$$

Hint: The procedure outlined above relied on the assumption that the Lie derivative of the tangent vectors with respect to each other vanishes. Otherwise, the parallelogram may not be closed properly. See Fig. 9.32 for a geometrical interpretation.

In what follows, the object $\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot}$ is alternatively derived using the second covariant differentiation of a vector or covector and its basic properties are introduced.

Noncommutativity of the surface covariant derivative is governed by

$$\boxed{\underline{\hat{h}}^{\alpha}\Big|_{\delta\gamma} - \underline{\hat{h}}^{\alpha}\Big|_{\gamma\delta} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot}\;\underline{\hat{h}}^{\beta}\;,} \tag{9.201a}$$

$$\boxed{\underline{\hat{h}}_{\beta}\Big|_{\gamma\delta} - \underline{\hat{h}}_{\beta}\Big|_{\delta\gamma} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot}\;\underline{\hat{h}}_{\alpha}\;,} \tag{9.201b}$$

where

$$\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot} = \frac{\partial \Gamma^{\alpha}_{\beta\delta}}{\partial t^{\gamma}} - \frac{\partial \Gamma^{\alpha}_{\beta\gamma}}{\partial t^{\delta}} + \Gamma^{\alpha}_{\gamma\theta}\Gamma^{\theta}_{\delta\beta} - \Gamma^{\alpha}_{\delta\theta}\Gamma^{\theta}_{\gamma\beta} \;\; \text{satisfying} \;\; \underline{\mathbb{R}}^{\alpha}{}_{.\,\alpha\gamma\delta}^{\;\cdot\;\cdot\;\cdot} = 0\;, \tag{9.202}$$

present the mixed components of the fourth-order Riemann-Christoffel curvature tensor

$$\mathbb{R} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot}\;\mathbf{a}_{\alpha}\otimes\mathbf{a}^{\beta}\otimes\mathbf{a}^{\gamma}\otimes\mathbf{a}^{\delta}\;. \tag{9.203}$$

The fully covariant form of this tensor renders

$$\mathbb{R} = \underline{\mathbb{R}}_{\alpha\beta\gamma\delta}\,\mathbf{a}^{\alpha}\otimes\mathbf{a}^{\beta}\otimes\mathbf{a}^{\gamma}\otimes\mathbf{a}^{\delta}\;, \tag{9.204}$$

where the components may be written as

$$\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = a_{\alpha\rho}\,\underline{\mathbb{R}}^{\rho}{}_{.\,\beta\gamma\delta}^{\;\cdot\;\cdot\;\cdot} = \frac{\partial \Gamma_{\beta\delta\alpha}}{\partial t^{\gamma}} - \frac{\partial \Gamma_{\beta\gamma\alpha}}{\partial t^{\delta}} + \Gamma_{\alpha\delta\rho}\Gamma^{\rho}_{\beta\gamma} - \Gamma_{\alpha\gamma\rho}\Gamma^{\rho}_{\beta\delta}\;, \tag{9.205}$$

or

$$\begin{aligned}\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = {} & \frac{1}{2}\left(\frac{\partial^2 a_{\alpha\delta}}{\partial t^{\gamma}\partial t^{\beta}} + \frac{\partial^2 a_{\beta\gamma}}{\partial t^{\delta}\partial t^{\alpha}} - \frac{\partial^2 a_{\beta\delta}}{\partial t^{\gamma}\partial t^{\alpha}} - \frac{\partial^2 a_{\alpha\gamma}}{\partial t^{\delta}\partial t^{\beta}}\right) \\ & + a^{\theta\rho}\left(\Gamma_{\alpha\delta\theta}\Gamma_{\beta\gamma\rho} - \Gamma_{\alpha\gamma\theta}\Gamma_{\beta\delta\rho}\right)\;.\end{aligned} \tag{9.206}$$

These covariant components possess the following major symmetries and minor antisymmetries

$$\boxed{\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \underline{\mathbb{R}}_{\gamma\delta\alpha\beta} \quad , \quad \underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = -\,\underline{\mathbb{R}}_{\alpha\beta\delta\gamma}\ ,} \tag{9.207}$$

which imply that

$$\boxed{\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = -\,\underline{\mathbb{R}}_{\beta\alpha\gamma\delta}\ .} \tag{9.208}$$

These properties also hold true for the mixed components (9.199); for instance,

$$\boxed{\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot} = -\,\underline{\mathbb{R}}_{\beta}{}^{\cdot\,\alpha\,\cdot\,\cdot}_{\ \,.\,\gamma\delta}\ ,} \tag{9.209}$$

can be used to derive (9.201b) from (9.201a), see the footnote on Sect. 7.1.6.

An extension of (9.201a) is given by

$$\underline{\tilde{H}}^{\alpha\beta}\Big|_{\delta\gamma} - \underline{\tilde{H}}^{\alpha\beta}\Big|_{\gamma\delta} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\theta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}^{\theta\beta} + \underline{\mathbb{R}}^{\beta}{}_{.\,\theta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}^{\alpha\theta}\ . \tag{9.210}$$

In a similar manner, (9.201b) can be extended to

$$\underline{\tilde{H}}_{\alpha\beta}\Big|_{\gamma\delta} - \underline{\tilde{H}}_{\alpha\beta}\Big|_{\delta\gamma} = \underline{\mathbb{R}}^{\theta}{}_{.\,\alpha\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}_{\theta\beta} + \underline{\mathbb{R}}^{\theta}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}_{\alpha\theta}\ . \tag{9.211}$$

In terms of the mixed components, the relations (9.210) and (9.211) render

$$\underline{\tilde{H}}^{\alpha}_{.\beta}\Big|_{\delta\gamma} - \underline{\tilde{H}}^{\alpha}_{.\beta}\Big|_{\gamma\delta} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\theta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}^{\theta}_{.\beta} - \underline{\mathbb{R}}^{\theta}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}^{\alpha}_{.\theta}\ , \tag{9.212a}$$

$$\underline{\tilde{H}}_{\alpha}^{.\beta}\Big|_{\gamma\delta} - \underline{\tilde{H}}_{\alpha}^{.\beta}\Big|_{\delta\gamma} = \underline{\mathbb{R}}^{\theta}{}_{.\,\alpha\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}_{\theta}^{.\beta} - \underline{\mathbb{R}}^{\beta}{}_{.\,\theta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\ \underline{\tilde{H}}_{\alpha}^{.\theta}\ . \tag{9.212b}$$

The Riemann-Christoffel curvature tensor satisfies the Bianchi identities (see Bianchi [19])

$$\boxed{\left.\begin{aligned} \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot} + \underline{\mathbb{R}}^{\alpha}{}_{.\,\gamma\delta\beta}^{\ \cdot\,\cdot\,\cdot} + \underline{\mathbb{R}}^{\alpha}{}_{.\,\delta\beta\gamma}^{\ \cdot\,\cdot\,\cdot} &= 0 \\ \underline{\mathbb{R}}_{\alpha\beta\gamma\delta} + \underline{\mathbb{R}}_{\alpha\gamma\delta\beta} + \underline{\mathbb{R}}_{\alpha\delta\beta\gamma} &= 0 \end{aligned}\right\}\ ,} \tag{9.213}$$

and

$$\boxed{\left.\begin{aligned} \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\Big|_{\theta} + \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\delta\theta}^{\ \cdot\,\cdot\,\cdot}\Big|_{\gamma} + \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\theta\gamma}^{\ \cdot\,\cdot\,\cdot}\Big|_{\delta} &= 0 \\ \underline{\mathbb{R}}_{\alpha\beta\gamma\delta}\Big|_{\theta} + \underline{\mathbb{R}}_{\alpha\beta\delta\theta}\Big|_{\gamma} + \underline{\mathbb{R}}_{\alpha\beta\theta\gamma}\Big|_{\delta} &= 0 \end{aligned}\right\}\ ,} \tag{9.214}$$

where

$$\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}\Big|_{\theta} = \frac{\partial\,\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}^{\ \cdot\,\cdot\,\cdot}}{\partial t^{\theta}}$$

$$+ \Gamma^{\alpha}_{\theta\rho} \underline{\mathbb{R}}^{\rho}{}_{.\,\beta\gamma\delta} - \Gamma^{\rho}_{\beta\theta} \underline{\mathbb{R}}^{\alpha}{}_{.\,\rho\gamma\delta} - \Gamma^{\rho}_{\gamma\theta} \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\rho\delta} - \Gamma^{\rho}_{\delta\theta} \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\rho} \;, \tag{9.215a}$$

$$\underline{\mathbb{R}}_{\alpha\beta\gamma\delta}\Big|_{\theta} = \frac{\partial\, \underline{\mathbb{R}}_{\alpha\beta\gamma\delta}}{\partial t^{\theta}} - \Gamma^{\rho}_{\alpha\theta} \underline{\mathbb{R}}_{\rho\beta\gamma\delta} - \Gamma^{\rho}_{\beta\theta} \underline{\mathbb{R}}_{\alpha\rho\gamma\delta} - \Gamma^{\rho}_{\gamma\theta} \underline{\mathbb{R}}_{\alpha\beta\rho\delta} - \Gamma^{\rho}_{\delta\theta} \underline{\mathbb{R}}_{\alpha\beta\gamma\rho} \;. \tag{9.215b}$$

Notice that the relations demonstrated in (9.201a)-(9.215b) have the same structure as the ambient equations (7.49a)–(7.64b). They are adopted here for convenience.

The major symmetric and minor antisymmetric properties (9.207)-(9.208) leave only four nonzero components with one degree of freedom:

$$\underline{\mathbb{R}}_{1212} = \underline{\mathbb{R}}_{2121} = - \underline{\mathbb{R}}_{1221} = - \underline{\mathbb{R}}_{2112} \;. \tag{9.216}$$

Using $(9.36)_1$, these relations can be unified according to

$$\boxed{\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \underline{\mathbb{R}}_{1212}\, \varepsilon_{\alpha\beta}\, \varepsilon_{\gamma\delta} \;.} \tag{9.217}$$

The surface and ambient Riemann-Christoffel curvature tensors are related through the following expression

$$\boxed{\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta} = \widetilde{Z}^{\alpha}_{i} \overline{Z}^{j}_{\beta} \overline{Z}^{k}_{\gamma} \overline{Z}^{l}_{\delta}\, \mathbb{R}^{i}{}_{.\,jkl} + \underline{b}^{\alpha}{}_{.\,\gamma}\, \underline{b}_{\beta\delta} - \underline{b}^{\alpha}{}_{.\,\delta}\, \underline{b}_{\beta\gamma} \;,} \quad \longleftarrow \text{see (9.51) and (9.479b)} \tag{9.218}$$

since

$$\underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}\, \underline{\hat{h}}^{\beta} \overset{\text{by using}}{\underset{\text{(9.85a) and (9.201a)}}{=\!=\!=}} \left[\left(\widetilde{Z}^{\alpha}_{i}\, \underline{\hat{h}}^{i} \right)\Big|_{\delta} \right]\Big|_{\gamma} - \left[\left(\widetilde{Z}^{\alpha}_{i}\, \underline{\hat{h}}^{i} \right)\Big|_{\gamma} \right]\Big|_{\delta}$$

$$\overset{\text{by using (9.180) and (9.193)}}{\underset{\text{along with the product rule}}{=\!=\!=}} \underbrace{\left[\underline{b}^{\alpha}{}_{.\,\delta}\, \underline{\widehat{n}}_{i}\, \underline{\hat{h}}^{i} + \underline{\hat{h}}^{i}\Big|_{l} \overline{Z}^{l}_{\delta} \widetilde{Z}^{\alpha}_{i} \right]}_{\text{note that } \underline{\widehat{n}}_{i}\, \underline{\hat{h}}^{i} = 0, \text{ see (9.75)}}\Big|_{\gamma} - \left[\underline{\hat{h}}^{i}\Big|_{k} \overline{Z}^{k}_{\gamma} \widetilde{Z}^{\alpha}_{i} \right]\Big|_{\delta}$$

$$\overset{\text{by using (9.180),(9.192) and (9.193)}}{\underset{\text{along with the product rule}}{=\!=\!=}} \left[\underline{\hat{h}}^{i}\Big|_{lk} - \underline{\hat{h}}^{i}\Big|_{kl} \right] \overline{Z}^{k}_{\gamma} \overline{Z}^{l}_{\delta} \widetilde{Z}^{\alpha}_{i}$$

$$+ \underbrace{\left[\underline{\hat{h}}^{i}\Big|_{l}\, \underline{b}_{\delta\gamma}\, \underline{\widehat{n}}^{l} - \underline{\hat{h}}^{i}\Big|_{k}\, \underline{b}_{\gamma\delta}\, \underline{\widehat{n}}^{k} \right] \widetilde{Z}^{\alpha}_{i}}_{= 0, \text{ by renaming the dummy indices and using (9.95)}}$$

$$+ \underbrace{\left[\underline{\hat{h}}^{i}\Big|_{l} \overline{Z}^{l}_{\delta}\, \underline{b}^{\alpha}{}_{.\,\gamma} - \underline{\hat{h}}^{i}\Big|_{k} \overline{Z}^{k}_{\gamma}\, \underline{b}^{\alpha}{}_{.\,\delta} \right] \underline{\widehat{n}}_{i}}_{= \left[\underline{\widehat{n}}_{i}|_{k} \overline{Z}^{k}_{\gamma}\, \underline{b}^{\alpha}{}_{.\,\delta} - \underline{\widehat{n}}_{i}|_{l} \overline{Z}^{l}_{\delta}\, \underline{b}^{\alpha}{}_{.\,\gamma} \right] \underline{\hat{h}}^{i}, \text{ since from } \underline{\hat{h}}^{i}\, \underline{\widehat{n}}_{i} = 0 \text{ one obtains } \underline{\hat{h}}^{i}\big|_{l}\, \underline{\widehat{n}}_{i} + \underline{\widehat{n}}_{i}|_{l}\, \underline{\hat{h}}^{i} = 0}$$

$$\overset{\text{by using}}{\underset{\text{(7.49a) and (9.182)}}{=\!=\!=}} \mathbb{R}^{i}{}_{.\,jkl}\, \underline{\hat{h}}^{j} \overline{Z}^{k}_{\gamma} \overline{Z}^{l}_{\delta} \widetilde{Z}^{\alpha}_{i} + \left[\underline{\widehat{n}}_{i}\Big|_{\gamma} \underline{b}^{\alpha}{}_{.\,\delta} - \underline{\widehat{n}}_{i}\Big|_{\delta} \underline{b}^{\alpha}{}_{.\,\gamma} \right] \underline{\hat{h}}^{i}$$

$$\overset{\text{by using}}{\underset{\text{(9.187)}}{=\!=\!=}} \mathbb{R}^{i}{}_{.\,jkl}\, \underline{\hat{h}}^{j} \overline{Z}^{k}_{\gamma} \overline{Z}^{l}_{\delta} \widetilde{Z}^{\alpha}_{i} + \left[\underline{b}^{\alpha}{}_{.\,\gamma}\, \underline{b}_{\delta\beta} - \underline{b}^{\alpha}{}_{.\,\delta}\, \underline{b}_{\gamma\beta} \right] \widetilde{Z}^{\beta}_{i}\, \underline{\hat{h}}^{i}$$

$$\xlongequal[\text{(9.85a) and (9.95)}]{\text{by using}} \left[\widetilde{Z}_i^{\alpha} \overline{Z}_{\beta}^{j} \overline{Z}_{\gamma}^{k} \overline{Z}_{\delta}^{l} \, \underline{\mathbb{R}}^{i}_{\cdot jkl} + \underline{b}^{\alpha}_{\cdot\gamma} \, \underline{b}_{\beta\delta} - \underline{b}^{\alpha}_{\cdot\delta} \, \underline{b}_{\beta\gamma} \right] \underline{\hat{h}}^{\beta} \ ,$$

noting that $\underline{\hat{h}}^{\beta}$ is an arbitrary object. By index juggling, one can further establish

$$\boxed{\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \overline{Z}_{\alpha}^{i} \overline{Z}_{\beta}^{j} \overline{Z}_{\gamma}^{k} \overline{Z}_{\delta}^{l} \, \underline{\mathbb{R}}_{ijkl} + \underline{b}_{\alpha\gamma} \, \underline{b}_{\beta\delta} - \underline{b}_{\alpha\delta} \, \underline{b}_{\beta\gamma} \ .} \tag{9.219}$$

A two-dimensional surface with vanishing Riemann-Christoffel tensor represents an **intrinsically** flat space. For instance, cylinders are intrinsically flat although they are curved when viewed externally from the enveloping space. And this means that their second-order curvature tensor is nonzero. Another example regards planes which are intrinsically and extrinsically flat (their second-order curvature tensor also vanishes identically). If the Riemann-Christoffel tensor is not zero, then the surface under consideration is intrinsically as well as extrinsically curved. Some well-known examples of such a curved space include spheres, one-sheeted hyperboloids and two-sheeted hyperboloids.

The Riemann-Christoffel tensor characterizes curvature of space from an intrinsic point of view. This is because the metric coefficients can be obtained by measuring distances on the surfaces which means that the metric tensor can also be regarded as an intrinsic object bypassing the position vector and surface basis vectors. As a result, the Christoffel symbols and subsequently Riemann-Christoffel tensor turn out to be intrinsic objects. This reveals the intrinsic nature of the Riemann-Christoffel curvature which relies on a priori given metric tensor.

In what follows, the goal is to represent two important differential operators for surface scalar and vector fields. This relies on a slight modification in the conventional definition of differential operators introduced in Chap. 7.

To begin with, consider the gradient operator

$$\boxed{\operatorname{grad}(\bullet) := \frac{\partial(\bullet)}{\partial \mathbf{x}} = \frac{\partial(\bullet)}{\partial t^{\alpha}} \mathbf{a}^{\alpha} \ .} \quad \longleftarrow \text{see (7.65)} \tag{9.220}$$

In a similar manner, the Nabla operator of vector calculus in this context takes the following form

$$\boxed{\boldsymbol{\nabla} := \frac{\partial}{\partial \mathbf{x}} = \frac{\partial}{\partial t^{\alpha}} \mathbf{a}^{\alpha} \ .} \quad \longleftarrow \text{see (7.66)} \tag{9.221}$$

Surface gradient. Let $\bar{h}\left(t^1, t^2\right)$ and $\hat{\mathbf{h}}\left(t^1, t^2\right)$ be a given surface scalar and vector field, respectively. The surface gradient of a surface field variable, $\operatorname{grad}^{\mathrm{s}}(\bullet)$, is then defined as its ordinary gradient with all normal components to the surface subtracted:

$$\boxed{\operatorname{grad}^{\mathrm{s}}\bar{h} = \operatorname{grad}\bar{h} - [\widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}] \operatorname{grad}\bar{h} \quad \text{where} \quad \operatorname{grad}\bar{h} = \frac{\partial \bar{h}}{\partial \mathbf{x}} = \frac{\partial \bar{h}}{\partial t^{\alpha}} \mathbf{a}^{\alpha} \ ,} \tag{9.222a}$$

$$\boxed{\operatorname{grad}^{s}\hat{\mathbf{h}} = \operatorname{grad}\hat{\mathbf{h}} - [\hat{\mathbf{n}} \otimes \hat{\mathbf{n}}]\operatorname{grad}\hat{\mathbf{h}} \quad \text{where} \quad \operatorname{grad}\hat{\mathbf{h}} = \frac{\partial \hat{\mathbf{h}}}{\partial \mathbf{x}} = \frac{\partial \hat{\mathbf{h}}}{\partial t^{\beta}} \otimes \mathbf{a}^{\beta} \ .} \tag{9.222b}$$

Note that $\partial \bar{h}/\partial t^{\alpha} = \bar{h}\big|_{\alpha}$ and $\partial \hat{\mathbf{h}}/\partial t^{\beta}$ is given in (9.132) and (9.157). Now, it is easy to express the surface gradients with respect to the surface bases as

$$\operatorname{grad}^{s}\bar{h} = \bar{h}\big|_{\alpha}\,\mathbf{a}^{\alpha} = \bar{h}\big|^{\alpha}\,\mathbf{a}_{\alpha} \quad \text{where} \quad \bar{h}\big|^{\alpha} = a^{\alpha\beta}\,\bar{h}\big|_{\beta} \ , \tag{9.223a}$$

$$\begin{aligned} \operatorname{grad}^{s}\hat{\mathbf{h}} &= \underline{\hat{h}}^{\alpha}\Big|_{\beta}\,\mathbf{a}_{\alpha} \otimes \mathbf{a}^{\beta} = \underline{\hat{h}}_{\alpha}\Big|_{\beta}\,\mathbf{a}^{\alpha} \otimes \mathbf{a}^{\beta} \\ &= \underline{\hat{h}}^{\alpha}\Big|^{\beta}\,\mathbf{a}_{\alpha} \otimes \mathbf{a}_{\beta} = \underline{\hat{h}}_{\alpha}\Big|^{\beta}\,\mathbf{a}^{\alpha} \otimes \mathbf{a}_{\beta} \ . \end{aligned} \tag{9.223b}$$

Surface divergence. The surface divergence of a surface vector field $\hat{\mathbf{h}}\left(t^{1}, t^{1}\right)$ is a scalar field defined by

$$\boxed{\operatorname{div}^{s}\hat{\mathbf{h}} = \operatorname{grad}^{s}\hat{\mathbf{h}} : \mathbf{I} = \operatorname{tr}\left(\operatorname{grad}^{s}\hat{\mathbf{h}}\right) \ ,} \quad \longleftarrow \text{see (7.77)} \tag{9.224}$$

where $\operatorname{tr}\mathbf{A} = \mathbf{A} : \mathbf{I} = \mathbf{I} : \mathbf{A}$, according to (2.83), and the identity tensor $\mathbf{I}$ is given in (9.34). This scalar field variable can be expressed in terms of the components as

$$\operatorname{div}^{s}\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha}\Big|_{\alpha} = a^{\alpha\beta}\,\underline{\hat{h}}_{\alpha}\Big|_{\beta} = a_{\alpha\beta}\,\underline{\hat{h}}^{\alpha}\Big|^{\beta} = \underline{\hat{h}}_{\alpha}\Big|^{\alpha} \ . \tag{9.225}$$

The Voss-Weyl formula (7.81) holds true here, that is,

$$\boxed{\operatorname{div}^{s}\hat{\mathbf{h}} = \frac{1}{\sqrt{a}}\frac{\partial}{\partial t^{\alpha}}\left[\sqrt{a}\,\underline{\hat{h}}^{\alpha}\right] \ ,} \tag{9.226}$$

owing to

$$\begin{aligned} \frac{1}{\sqrt{a}}\frac{\partial}{\partial t^{\alpha}}\left[\sqrt{a}\,\underline{\hat{h}}^{\alpha}\right] &\xlongequal[\text{rule of differentiation}]{\text{by using the product}} \frac{1}{\sqrt{a}}\frac{\partial\sqrt{a}}{\partial t^{\alpha}}\,\underline{\hat{h}}^{\alpha} + \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\alpha}} \\ &\xlongequal[(9.115)]{\text{by using}} \Gamma^{\gamma}_{\alpha\gamma}\,\underline{\hat{h}}^{\alpha} + \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\alpha}} \\ &\xlongequal[(9.92)]{\text{by using}} \Gamma^{\gamma}_{\gamma\alpha}\,\underline{\hat{h}}^{\alpha} + \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\alpha}} \\ &\xlongequal[(9.128)]{\text{by using}} \underline{\hat{h}}^{\alpha}\Big|_{\alpha} \\ &\xlongequal[(9.225)]{\text{by using}} \operatorname{div}^{s}\hat{\mathbf{h}} \ . \end{aligned}$$

9.5 Fundamental Forms of Surfaces

The so-called *fundamental forms* of a surface can be classified into three types. In the literature, much attention is devoted to the first and second fundamental forms since the last one can be expressed in terms of them. The first fundamental form is an **intrinsic** object while the second one demonstrates an **extrinsic** feature of a surface. They provide the most important pieces of data associated with any curved surface. The both fundamental forms are utilized to not only measure the length of a curve embedded in a surface and the area of a patch of that surface but also help determine some important surface characteristics such as the Gaussian curvature.

9.5.1 First Fundamental Form

The *first fundamental form* (or *first groundform*) characterizes the metric properties of a surface that depend on the length of a curve and the angle between two curves. Indeed, it contains the **intrinsic** data regarding measurements on the surface and represents a **quadratic** expression of the form

$$\begin{aligned} \mathrm{I}_{\mathbf{r}} = ds^2 = d\mathbf{r}\cdot d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial t^{\alpha}}\cdot\frac{\partial \mathbf{r}}{\partial t^{\beta}}dt^{\alpha}dt^{\beta} = \mathbf{a}_{\alpha}\cdot\mathbf{a}_{\beta}dt^{\alpha}dt^{\beta} = a_{\alpha\beta}dt^{\alpha}dt^{\beta} \\ &= \boxed{\mathrm{E}_{\mathbf{r}}\left(dt^1\right)^2 + 2\mathrm{F}_{\mathbf{r}}dt^1dt^2 + \mathrm{G}_{\mathbf{r}}\left(dt^2\right)^2\ ,} \end{aligned} \tag{9.227}$$

where the coefficients

$$\mathrm{E}_{\mathbf{r}} = a_{11}\quad,\quad \mathrm{G}_{\mathbf{r}} = a_{22}\quad,\quad \mathrm{F}_{\mathbf{r}} = a_{12} = a_{21}\ , \tag{9.228}$$

present smooth functions of the Gaussian coordinates. Note that the expressions written in (9.227) for the **arc length element** of a surface generalize the classical equation for the arc length element in the two-dimensional flat Euclidean space, that is,

$$ds^2 = dx^2 + dy^2\ \left(\text{considering } t^1 = x\ ,\ t^2 = y \text{ and } \mathrm{E}_{\mathbf{r}} = \mathrm{G}_{\mathbf{r}} = 1\ ,\ \mathrm{F}_{\mathbf{r}} = 0\right)\ . \tag{9.229}$$

Moreover, the partial derivatives of the position vector $\mathbf{r}$ were written up to order 1. Indeed,

$$\underbrace{\mathbf{r}\left(t^1 + dt^1, t^2 + dt^2\right) \approx \mathbf{r}\left(t^1, t^2\right) + \frac{\partial \mathbf{r}}{\partial t^{\alpha}}dt^{\alpha}}_{\text{or } d\mathbf{r}\approx\mathbf{a}_{\alpha}dt^{\alpha}}\ . \tag{9.230}$$

The first fundamental form in matrix representation renders

$$\mathrm{I}_{\mathbf{r}} = [\,d\mathbf{t}\,]^{\mathrm{T}}[\,\mathbf{I}_{\mathbf{r}}\,][\,d\mathbf{t}\,]\quad \text{where}\quad [\,d\mathbf{t}\,] = \begin{bmatrix} dt^1 \\ dt^2 \end{bmatrix}\ ,\ [\,\mathbf{I}_{\mathbf{r}}\,] = \begin{bmatrix} \mathrm{E}_{\mathbf{r}} & \mathrm{F}_{\mathbf{r}} \\ \mathrm{F}_{\mathbf{r}} & \mathrm{G}_{\mathbf{r}} \end{bmatrix}\ . \tag{9.231}$$

Consequently, the surface covariant metric coefficients (9.17) take the form

$$\boxed{\left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} \mathrm{E}_{\mathbf{r}} & \mathrm{F}_{\mathbf{r}} \\ \mathrm{F}_{\mathbf{r}} & \mathrm{G}_{\mathbf{r}} \end{bmatrix} \quad \text{with} \quad a = \det\left[\, a_{\alpha\beta} \,\right] = \mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^2 \,.} \tag{9.232}$$

Then, the surface contravariant metric coefficients $(9.26)_4$ become

$$\boxed{\left[\, a^{\alpha\beta} \,\right] = \frac{1}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^2} \begin{bmatrix} \mathrm{G}_{\mathbf{r}} & -\mathrm{F}_{\mathbf{r}} \\ -\mathrm{F}_{\mathbf{r}} & \mathrm{E}_{\mathbf{r}} \end{bmatrix} \text{ with } \frac{1}{a} = \det\left[\, a^{\alpha\beta} \,\right] = \frac{1}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^2} \,.} \tag{9.233}$$

The surface Christoffel symbols (9.114) can be computed in terms of the coefficients of the first groundform and their partial derivatives. To show this, consider the expressions $\widehat{\mathbf{n}} \cdot \mathbf{a}_\alpha = 0$, $a_{\alpha\beta} = \mathbf{a}_\alpha \cdot \mathbf{a}_\beta$, $\partial \mathbf{a}_\alpha / \partial t^\beta = \Gamma^\gamma_{\alpha\beta} \mathbf{a}_\gamma + \underline{b}_{\alpha\beta} \widehat{\mathbf{n}}$ and establish

$$\begin{aligned} \frac{1}{2}\frac{\partial a_{11}}{\partial t^1} &= \left(\frac{\partial \mathbf{a}_1}{\partial t^1}\right) \cdot \mathbf{a}_1 = \left(\Gamma^1_{11}\mathbf{a}_1 + \Gamma^2_{11}\mathbf{a}_2 + \underline{b}_{11}\widehat{\mathbf{n}}\right) \cdot \mathbf{a}_1 \\ &= \Gamma^1_{11}\, a_{11} + \Gamma^2_{11}\, a_{12} \,, \end{aligned} \tag{9.234a}$$

$$\begin{aligned} \frac{1}{2}\frac{\partial a_{11}}{\partial t^2} &= \left(\frac{\partial \mathbf{a}_1}{\partial t^2}\right) \cdot \mathbf{a}_1 = \left(\Gamma^1_{12}\mathbf{a}_1 + \Gamma^2_{12}\mathbf{a}_2 + \underline{b}_{12}\widehat{\mathbf{n}}\right) \cdot \mathbf{a}_1 \\ &= \Gamma^1_{12}\, a_{11} + \Gamma^2_{12}\, a_{12} \,, \end{aligned} \tag{9.234b}$$

$$\begin{aligned} \frac{1}{2}\frac{\partial a_{22}}{\partial t^1} &= \left(\frac{\partial \mathbf{a}_2}{\partial t^1}\right) \cdot \mathbf{a}_2 = \left(\Gamma^1_{12}\mathbf{a}_1 + \Gamma^2_{12}\mathbf{a}_2 + \underline{b}_{12}\widehat{\mathbf{n}}\right) \cdot \mathbf{a}_2 \\ &= \Gamma^1_{12}\, a_{12} + \Gamma^2_{12}\, a_{22} \,, \end{aligned} \tag{9.234c}$$

$$\begin{aligned} \frac{1}{2}\frac{\partial a_{22}}{\partial t^2} &= \left(\frac{\partial \mathbf{a}_2}{\partial t^2}\right) \cdot \mathbf{a}_2 = \left(\Gamma^1_{22}\mathbf{a}_1 + \Gamma^2_{22}\mathbf{a}_2 + \underline{b}_{22}\widehat{\mathbf{n}}\right) \cdot \mathbf{a}_2 \\ &= \Gamma^1_{22}\, a_{12} + \Gamma^2_{22}\, a_{22} \,, \end{aligned} \tag{9.234d}$$

$$\begin{aligned} \frac{\partial a_{12}}{\partial t^1} - \frac{1}{2}\frac{\partial a_{11}}{\partial t^2} &= \left(\frac{\partial \mathbf{a}_1}{\partial t^1}\right) \cdot \mathbf{a}_2 = \left(\Gamma^1_{11}\mathbf{a}_1 + \Gamma^2_{11}\mathbf{a}_2 + \underline{b}_{11}\widehat{\mathbf{n}}\right) \cdot \mathbf{a}_2 \\ &= \Gamma^1_{11}\, a_{12} + \Gamma^2_{11}\, a_{22} \,, \end{aligned} \tag{9.234e}$$

$$\frac{\partial a_{12}}{\partial t^2} - \frac{1}{2}\frac{\partial a_{22}}{\partial t^1} = \left(\frac{\partial \mathbf{a}_2}{\partial t^2}\right) \cdot \mathbf{a}_1 = \left(\Gamma^1_{22}\mathbf{a}_1 + \Gamma^2_{22}\mathbf{a}_2 + \underline{b}_{22}\widehat{\mathbf{n}}\right) \cdot \mathbf{a}_1$$
$$= \Gamma^1_{22}\, a_{11} + \Gamma^2_{22}\, a_{12} \ . \tag{9.234f}$$

It is then easy to see that

$$\boxed{\begin{bmatrix} \Gamma^1_{11} \\ \Gamma^2_{11} \end{bmatrix} = \frac{1}{2\left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right)} \begin{bmatrix} \mathrm{G_r} & -\mathrm{F_r} \\ -\mathrm{F_r} & \mathrm{E_r} \end{bmatrix} \begin{bmatrix} \partial \mathrm{E_r}/\partial t^1 \\ 2\partial \mathrm{F_r}/\partial t^1 - \partial \mathrm{E_r}/\partial t^2 \end{bmatrix} ,} \tag{9.235a}$$

$$\boxed{\begin{bmatrix} \Gamma^1_{12} \\ \Gamma^2_{12} \end{bmatrix} = \begin{bmatrix} \Gamma^1_{21} \\ \Gamma^2_{21} \end{bmatrix} = \frac{1}{2\left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right)} \begin{bmatrix} \mathrm{G_r} & -\mathrm{F_r} \\ -\mathrm{F_r} & \mathrm{E_r} \end{bmatrix} \begin{bmatrix} \partial \mathrm{E_r}/\partial t^2 \\ \partial \mathrm{G_r}/\partial t^1 \end{bmatrix} ,} \tag{9.235b}$$

$$\boxed{\begin{bmatrix} \Gamma^1_{22} \\ \Gamma^2_{22} \end{bmatrix} = \frac{1}{2\left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right)} \begin{bmatrix} \mathrm{G_r} & -\mathrm{F_r} \\ -\mathrm{F_r} & \mathrm{E_r} \end{bmatrix} \begin{bmatrix} 2\partial \mathrm{F_r}/\partial t^2 - \partial \mathrm{G_r}/\partial t^1 \\ \partial \mathrm{G_r}/\partial t^2 \end{bmatrix} .} \tag{9.235c}$$

Accordingly, one finds that

$$\Gamma^1_{11} + \Gamma^2_{12} = \frac{\partial}{\partial t^1}\left(\log\sqrt{\mathrm{E_r G_r} - \mathrm{F_r^2}}\right) , \tag{9.236a}$$
$$\Gamma^1_{12} + \Gamma^2_{22} = \frac{\partial}{\partial t^2}\left(\log\sqrt{\mathrm{E_r G_r} - \mathrm{F_r^2}}\right) . \tag{9.236b}$$

Using $(9.19)_{1-2}$, $(9.20)_3$ and $(9.228)_{1-3}$, one will have

$$\boxed{\mathrm{E_r} > 0 \quad , \quad \mathrm{G_r} > 0 \quad , \quad \mathrm{E_r G_r} - \mathrm{F_r^2} > 0 \ .} \tag{9.237}$$

And this implies that the first fundamental form of a surface is always **positive**. This is verified by reformulating $(9.227)_6$ according to

$$\boxed{\mathrm{I_r} = \mathrm{E_r^{-1}}\left(\mathrm{E_r}dt^1 + \mathrm{F_r}dt^2\right)^2 + \mathrm{E_r^{-1}}\left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right)\left(dt^2\right)^2 > 0 \ .} \tag{9.238}$$

Suppose that the **parametric equations** $t \to \left(t^1(t), t^2(t)\right)$ describe a smooth curve $\boldsymbol{\gamma}$ as

$$\boxed{\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1(t), t^2(t)\right) = \mathbf{x}^{\mathrm{c}}(t) \ ,} \tag{9.239}$$

for which

$$\boxed{\mathbf{a}_t = \frac{d\mathbf{x}^{\mathrm{c}}}{dt} = \frac{\partial \hat{\mathbf{x}}^{\mathrm{s}}}{\partial t^\alpha}\frac{dt^\alpha}{dt} = \frac{dt^\alpha}{dt}\mathbf{a}_\alpha \ ,} \tag{9.240}$$

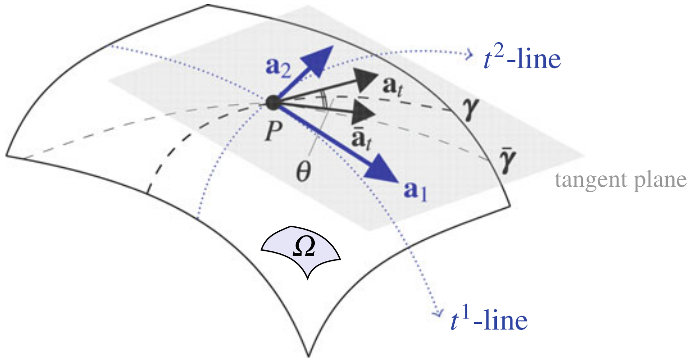


Fig. 9.15 Surface curves $\boldsymbol{\gamma}$, $\bar{\boldsymbol{\gamma}}$ with corresponding tangent vectors $\mathbf{a}_t$, $\bar{\mathbf{a}}_t$ and surface patch Ω

and

$$\boxed{\mathbf{a}_t \cdot \mathbf{a}_t = \frac{dt^\alpha}{dt} a_{\alpha\beta} \frac{dt^\beta}{dt} \ .} \tag{9.241}$$

See Fig. 9.15 for a geometrical interpretation. The length of this curve from a point P at $\mathbf{x} = \mathbf{x}^{\mathrm{c}}(0)$ to another point Q at $\mathbf{x} = \mathbf{x}^{\mathrm{c}}(t)$ is computed via

$$\begin{aligned} L &= \int_{\mathbf{x}^{\mathrm{c}}(0)}^{\mathbf{x}^{\mathrm{c}}(t)} \sqrt{d\mathbf{x} \cdot d\mathbf{x}} = \int_0^t \sqrt{\mathbf{a}_t \cdot \mathbf{a}_t}\, d\hat{t} = \int_0^t \sqrt{\frac{dt^\alpha}{d\hat{t}} a_{\alpha\beta} \frac{dt^\beta}{d\hat{t}}}\, d\hat{t} \\ &= \int_0^t \sqrt{\mathrm{E}_{\mathbf{r}} \left(\frac{dt^1}{d\hat{t}}\right)^2 + 2\mathrm{F}_{\mathbf{r}} \frac{dt^1}{d\hat{t}} \frac{dt^2}{d\hat{t}} + \mathrm{G}_{\mathbf{r}} \left(\frac{dt^2}{d\hat{t}}\right)^2}\, d\hat{t} \ . \end{aligned} \tag{9.242}$$

Suppose one is given the two parametric surface curves $\boldsymbol{\gamma} : t \to \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1(t), t^2(t)\right)$ and $\bar{\boldsymbol{\gamma}} : t \to \hat{\mathbf{x}}^{\mathrm{s}}\left(\bar{t}^1(t), \bar{t}^2(t)\right)$ that intersect at the point P as shown in Fig. 9.15. The tangent vectors at this point will be

$$\mathbf{a}_t = \frac{d\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1(t), t^2(t)\right)}{dt} = \frac{dt^\alpha}{dt}\mathbf{a}_\alpha \quad , \quad \bar{\mathbf{a}}_t = \frac{d\hat{\mathbf{x}}^{\mathrm{s}}\left(\bar{t}^1(t), \bar{t}^2(t)\right)}{dt} = \frac{d\bar{t}^\beta}{dt}\mathbf{a}_\beta \ . \tag{9.243}$$

Using (1.11)–(1.12), $(9.17)_1$, $(9.228)_{1-4}$, $(9.243)_2$ and $(9.243)_4$, the angle between these surface curves then renders

$$\theta = \arccos\left(\frac{\mathbf{a}_t \cdot \bar{\mathbf{a}}_t}{|\mathbf{a}_t|\,|\bar{\mathbf{a}}_t|}\right) , \tag{9.244}$$

where

$$\mathbf{a}_t \cdot \bar{\mathbf{a}}_t = \mathrm{E}_{\mathbf{r}} \frac{dt^1}{dt}\frac{d\bar{t}^1}{dt} + \mathrm{F}_{\mathbf{r}}\left(\frac{dt^1}{dt}\frac{d\bar{t}^2}{dt} + \frac{dt^2}{dt}\frac{d\bar{t}^1}{dt}\right) + \mathrm{G}_{\mathbf{r}} \frac{dt^2}{dt}\frac{d\bar{t}^2}{dt} \ , \tag{9.245a}$$

$$|\mathbf{a}_t| = \sqrt{\mathrm{E}_\mathbf{r}\left(\frac{dt^1}{dt}\right)^2 + 2\mathrm{F}_\mathbf{r}\frac{dt^1}{dt}\frac{dt^2}{dt} + \mathrm{G}_\mathbf{r}\left(\frac{dt^2}{dt}\right)^2}\ , \tag{9.245b}$$

$$|\bar{\mathbf{a}}_t| = \sqrt{\mathrm{E}_\mathbf{r}\left(\frac{d\bar{t}^1}{dt}\right)^2 + 2\mathrm{F}_\mathbf{r}\frac{d\bar{t}^1}{dt}\frac{d\bar{t}^2}{dt} + \mathrm{G}_\mathbf{r}\left(\frac{d\bar{t}^2}{dt}\right)^2}\ . \tag{9.245c}$$

If $\mathbf{a}_t$ is orthogonal to $\bar{\mathbf{a}}_t$, then (9.245a) takes the form

$$\underbrace{\mathrm{E}_\mathbf{r}\, dt^1 d\bar{t}^1 + \mathrm{F}_\mathbf{r}\left(dt^1 d\bar{t}^2 + dt^2 d\bar{t}^1\right) + \mathrm{G}_\mathbf{r}\, dt^2 d\bar{t}^2 = 0}_{\text{or } \mathrm{E}_\mathbf{r} + \mathrm{F}_\mathbf{r}\left(\frac{dt^2}{dt^1} + \frac{d\bar{t}^2}{d\bar{t}^1}\right) + \mathrm{G}_\mathbf{r}\left(\frac{dt^2}{dt^1}\frac{d\bar{t}^2}{d\bar{t}^1}\right) = 0}\ . \tag{9.246}$$

Guided by $(9.54)_2$, (9.55) and $(9.228)_{1-3}$, the area A of a surface patch over a region Ω is given by

$$A = \int_\Omega dA = \int_\Omega \sqrt{a}\, dt^1 dt^2 = \int_\Omega \sqrt{a_{11}a_{22} - (a_{12})^2}\, dt^1 dt^2$$
$$= \int_\Omega \sqrt{\mathrm{E}_\mathbf{r}\mathrm{G}_\mathbf{r} - (\mathrm{F}_\mathbf{r})^2}\, dt^1 dt^2\ . \tag{9.247}$$

Note that $(9.242)_4$, (9.244)-(9.245c) and $(9.247)_4$ clearly show the key role of the first fundamental form in determining the **geometrical characteristics** of a surface. This reveals the fact that the coefficients $\mathrm{E}_\mathbf{r}$, $\mathrm{F}_\mathbf{r}$ and $\mathrm{G}_\mathbf{r}$ with intrinsic nature should be obtained, in principle, from some measurements carried out on the surface. This is described below.

Consider a generic point $\left(t_0^1, t_0^2\right)$. The goal is to calculate the coefficients $\mathrm{E}_\mathbf{r}$, $\mathrm{F}_\mathbf{r}$ and $\mathrm{G}_\mathbf{r}$ at this point. Note that one is supposed to solve the problem by only measuring the lengths of curve segments.

To begin with, consider the parametric curve $\left(t_0^1 + t, t_0^2\right)$ passing through $\left(t_0^1, t_0^2\right)$ at $t = 0$. Notice that this curve basically represents the coordinate line associated with t^1. Using $(9.242)_4$, the length of such curve, $L_{\mathrm{E}_\mathrm{r}}(t)$, becomes

$$L_{\mathrm{E}_\mathrm{r}}(t) = \int_0^t \sqrt{\mathrm{E}_\mathbf{r}\left(\hat{t}\right)}\, d\hat{t} \quad \text{which provides} \quad \mathrm{E}_\mathbf{r} = \left(\left.\frac{dL_{\mathrm{E}_\mathrm{r}}}{dt}\right|_{t=0}\right)^2\ . \tag{9.248}$$

Denoting by $L_{\mathrm{G}_\mathrm{r}}(t)$ the length of the parametric curve $\left(t_0^1, t_0^2 + t\right)$ from $\left(t_0^1, t_0^2\right)$, one then similarly arrives at

$$L_{\mathrm{G}_\mathrm{r}}(t) = \int_0^t \sqrt{\mathrm{G}_\mathbf{r}\left(\hat{t}\right)}\, d\hat{t} \quad \text{and, immediately,} \quad \mathrm{G}_\mathbf{r} = \left(\left.\frac{dL_{\mathrm{G}_\mathrm{r}}}{dt}\right|_{t=0}\right)^2\ . \tag{9.249}$$

Consider this time the parametric curve $\left(t_0^1 + t, t_0^2 + t\right)$ passing through $\left(t_0^1, t_0^2\right)$ at $t = 0$. By means of $(9.242)_4$, $(9.248)_2$ and $(9.249)_2$, the length of this curve, $L_{\mathrm{F}_\mathrm{r}}(t)$, finally helps obtain

$$\mathrm{F}_{\mathbf{r}} = \left(\frac{dL_{\mathrm{F}_{\mathbf{r}}}}{\sqrt{2}dt} \bigg|_{t=0} \right)^2 - \left(\frac{dL_{\mathrm{E}_{\mathbf{r}}}}{\sqrt{2}dt} \bigg|_{t=0} \right)^2 - \left(\frac{dL_{\mathrm{G}_{\mathbf{r}}}}{\sqrt{2}dt} \bigg|_{t=0} \right)^2 . \tag{9.250}$$

9.5.2 Second Fundamental Form

The *second fundamental form* (or *second groundform*) measures how a surface deviates from its tangent plane in a neighborhood of a given point. The second groundform contains the **extrinsic** data regarding the curvature of a surface since it keeps track of the twisting of the unit normal field to that surface.

The second fundamental form is defined in some alternative forms within **second-order** terms. One well-known definition is

$$\begin{aligned} \mathrm{II}_{\mathbf{r}} &= -d\mathbf{r} \cdot d\,\widehat{\mathbf{n}} = -\frac{\partial \mathbf{r}}{\partial t^\alpha} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^\beta} dt^\alpha dt^\beta = -\mathbf{a}_\alpha \cdot \left(-\underline{b}_{\beta\gamma} \mathbf{a}^\gamma \right) dt^\alpha dt^\beta = \underline{b}_{\beta\alpha} dt^\alpha dt^\beta \\ &= \boxed{\mathrm{e}_{\mathbf{r}} \left(dt^1 \right)^2 + 2\mathrm{f}_{\mathbf{r}} dt^1 dt^2 + \mathrm{g}_{\mathbf{r}} \left(dt^2 \right)^2 \ ,} \end{aligned} \tag{9.251}$$

← see (9.260) and (9.262)

where

$$\mathrm{e}_{\mathbf{r}} = \underline{b}_{11} = \frac{\partial^2 \mathbf{r}}{\partial t^1 \partial t^1} \cdot \widehat{\mathbf{n}} = -\frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \ , \tag{9.252a}$$

$$\mathrm{g}_{\mathbf{r}} = \underline{b}_{22} = \frac{\partial^2 \mathbf{r}}{\partial t^2 \partial t^2} \cdot \widehat{\mathbf{n}} = -\frac{\partial \mathbf{r}}{\partial t^2} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} \ , \tag{9.252b}$$

$$\mathrm{f}_{\mathbf{r}} = \underline{b}_{12} = \underline{b}_{21} = \frac{\partial^2 \mathbf{r}}{\partial t^1 \partial t^2} \cdot \widehat{\mathbf{n}} = -\frac{1}{2} \left[\frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} + \frac{\partial \mathbf{r}}{\partial t^2} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \right] . \tag{9.252c}$$

Note that $\mathrm{II}_{\mathbf{r}}$ is a symmetric bilinear form similarly to $\mathrm{I}_{\mathbf{r}}$ but it is not, in general, positive. Moreover, the differential position vector $d\mathbf{r}$ and the normal vector increment $d\,\widehat{\mathbf{n}}$ have linearly been expanded, see (9.230).

The second fundamental form in matrix representation renders

$$\mathrm{II}_{\mathbf{r}} = [\,d\mathbf{t}\,]^{\mathrm{T}} [\,\mathbf{II}_{\mathbf{r}}\,] [\,d\mathbf{t}\,] \quad \text{where} \quad [\,d\mathbf{t}\,] = \begin{bmatrix} dt^1 \\ dt^2 \end{bmatrix} , \ [\,\mathbf{II}_{\mathbf{r}}\,] = \begin{bmatrix} \mathrm{e}_{\mathbf{r}} & \mathrm{f}_{\mathbf{r}} \\ \mathrm{f}_{\mathbf{r}} & \mathrm{g}_{\mathbf{r}} \end{bmatrix} . \tag{9.253}$$

Subsequently, the matrix form of the surface covariant curvature tensor $\underline{b}_{\alpha\beta}$ renders

$$\boxed{[\,\underline{b}_{\alpha\beta}\,] = \begin{bmatrix} \mathrm{e}_{\mathbf{r}} & \mathrm{f}_{\mathbf{r}} \\ \mathrm{f}_{\mathbf{r}} & \mathrm{g}_{\mathbf{r}} \end{bmatrix} ,} \tag{9.254}$$

with

$$\boxed{b = \det [\,\underline{b}_{\alpha\beta}\,] = \underline{b}_{11}\,\underline{b}_{22} - \underline{b}_{12}^2 = \mathrm{e}_{\mathbf{r}} \mathrm{g}_{\mathbf{r}} - \mathrm{f}_{\mathbf{r}}^2 \ .} \tag{9.255}$$

Using (2.94)$_2$, (9.100)$_1$, (9.233)$_1$ and (9.254), the surface mixed curvature tensor $\underline{b}_{\alpha}^{\cdot\beta}$ then admits the following matrix form

$$\boxed{\left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \left[\, \underline{b}_{\alpha\gamma} \,\right]\left[\, a^{\gamma\beta} \,\right] = \frac{1}{\mathrm{E_r G_r - F_r^2}} \begin{bmatrix} \mathrm{e_r G_r - f_r F_r} & \mathrm{-e_r F_r + f_r E_r} \\ \mathrm{f_r G_r - g_r F_r} & \mathrm{-f_r F_r + g_r E_r} \end{bmatrix} .} \tag{9.256}$$

The mean curvature (9.103)$_1$ can now be written as

$$\boxed{\overline{\mathrm{H}} = \frac{1}{2} \mathrm{tr} \left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \frac{\mathrm{e_r G_r - 2 f_r F_r + g_r E_r}}{2\left(\mathrm{E_r G_r - F_r^2}\right)} .} \tag{9.257}$$

And the Gaussian curvature (9.104)$_1$ takes the following form

$$\boxed{\overline{\mathrm{K}} = \det \left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \det \left[\, \underline{b}_{\alpha\gamma} \,\right] \det \left[\, a^{\gamma\beta} \,\right] = \frac{b}{a} = \frac{\mathrm{e_r g_r - f_r^2}}{\mathrm{E_r G_r - F_r^2}} .} \tag{9.258}$$

Hint: It is important to note that changing the orientation of surface according to $\overline{\widehat{\mathbf{n}}} = -\widehat{\mathbf{n}}$ will lead to $\overline{\mathrm{II}}_{\mathbf{r}} = -d\mathbf{r} \cdot d\overline{\widehat{\mathbf{n}}} = +d\mathbf{r} \cdot d\widehat{\mathbf{n}} = -\mathrm{II}_{\mathbf{r}}$ and thus $\overline{\underline{b}}_{\alpha\beta} = -\underline{b}_{\alpha\beta}$. As a result, the mean curvature will be affected while the Gaussian curvature remains unaffected:

$$\boxed{\overline{\overline{\mathrm{H}}} = -\overline{\mathrm{H}} \quad , \quad \overline{\overline{\mathrm{K}}} = +\overline{\mathrm{K}} \,.} \tag{9.259}$$

And the sensitivity (insensitivity) of $\overline{\mathrm{H}}$ ($\overline{\mathrm{K}}$) with respect to the choice of $\widehat{\mathbf{n}}$ basically means that the mean (Gaussian) curvature has an extrinsic (intrinsic) attribute.

The second groundform may also be defined as

$$\boxed{\mathrm{II}_{\mathbf{r}} = 2d\mathbf{r} \cdot \widehat{\mathbf{n}} \,,} \quad \leftarrow \text{see (9.292)} \tag{9.260}$$

where here the quadratic approximation of the position vector at $\left(t^1 + dt^1, t^2 + dt^2\right)$ according to

$$\mathbf{r}\left(t^1 + dt^1, t^2 + dt^2\right) = \mathbf{r}\left(t^1, t^2\right) + \frac{\partial \mathbf{r}}{\partial t^{\alpha}} dt^{\alpha} + \frac{1}{2} \frac{\partial^2 \mathbf{r}}{\partial t^{\beta} \partial t^{\alpha}} dt^{\alpha} dt^{\beta} + R \,, \tag{9.261}$$

or $d\mathbf{r} = \mathbf{a}_{\alpha} dt^{\alpha} + \frac{1}{2} \frac{\partial \mathbf{a}_{\alpha}}{\partial t^{\beta}} dt^{\alpha} dt^{\beta} + R$ where $R = o\left(\left(dt^1\right)^2, dt^1 dt^2, \left(dt^2\right)^2\right)$

helps obtain the identical result

$$\mathrm{II}_{\mathbf{r}} = 2d\mathbf{r} \cdot \widehat{\mathbf{n}} = 2\mathbf{a}_{\alpha} \cdot \widehat{\mathbf{n}} dt^{\alpha} + \left[\Gamma_{\alpha\beta}^{\gamma} \mathbf{a}_{\gamma} + \underline{b}_{\alpha\beta} \widehat{\mathbf{n}}\right] \cdot \widehat{\mathbf{n}} dt^{\alpha} dt^{\beta} = \underline{b}_{\alpha\beta} dt^{\alpha} dt^{\beta} \,,$$

noting that $\widehat{\mathbf{n}} \cdot \mathbf{a}_{\alpha} = 0$ and $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$, according to (9.12).

To provide another definition for the second groundform, consider the following one-parameter family of regular surfaces

$$\mathbf{R}\left(t^1, t^2, h\right) = \mathbf{r}\left(t^1, t^2\right) - h\widehat{\mathbf{n}}\left(t^1, t^2\right) \quad \text{with} \quad h \in (-\varepsilon, \varepsilon) \ , \tag{9.262}$$

whose first fundamental form is given by

$$\mathrm{I}_{\mathbf{R}} = \mathrm{E}_{\mathbf{R}}\left(dt^1\right)^2 + 2\mathrm{F}_{\mathbf{R}}dt^1dt^2 + \mathrm{G}_{\mathbf{R}}\left(dt^2\right)^2 \ , \tag{9.263}$$

where

$$\mathrm{E}_{\mathbf{R}} = \frac{\partial \mathbf{R}}{\partial t^1} \cdot \frac{\partial \mathbf{R}}{\partial t^1} \quad , \quad \mathrm{F}_{\mathbf{R}} = \frac{\partial \mathbf{R}}{\partial t^1} \cdot \frac{\partial \mathbf{R}}{\partial t^2} \quad , \quad \mathrm{G}_{\mathbf{R}} = \frac{\partial \mathbf{R}}{\partial t^2} \cdot \frac{\partial \mathbf{R}}{\partial t^2} \ . \tag{9.264}$$

Then, the second groundform of $\mathbf{r}\left(t^1, t^2\right)$ is obtained from the first groundform of $\mathbf{R}\left(t^1, t^2, h\right)$ via the following expression

$$\boxed{\mathrm{II}_{\mathbf{r}} = \frac{1}{2} \left.\frac{\partial \mathrm{I}_{\mathbf{R}}}{\partial h}\right|_{h=0} \ ,} \tag{9.265}$$

since

$$\begin{aligned}
\mathrm{e}_{\mathbf{r}} &= \frac{1}{2} \left.\frac{\partial \mathrm{E}_{\mathbf{R}}}{\partial h}\right|_{h=0} = \frac{1}{2} \frac{\partial}{\partial h} \left[\frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \mathbf{r}}{\partial t^1} - 2h \frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} + h^2 \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^1}\right]\Bigg|_{h=0} \\
&= -\frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \ , \\
\mathrm{g}_{\mathbf{r}} &= \frac{1}{2} \left.\frac{\partial \mathrm{G}_{\mathbf{R}}}{\partial h}\right|_{h=0} = \frac{1}{2} \frac{\partial}{\partial h} \left[\frac{\partial \mathbf{r}}{\partial t^2} \cdot \frac{\partial \mathbf{r}}{\partial t^2} - 2h \frac{\partial \mathbf{r}}{\partial t^2} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} + h^2 \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2}\right]\Bigg|_{h=0} \\
&= -\frac{\partial \mathbf{r}}{\partial t^2} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} \ , \\
\mathrm{f}_{\mathbf{r}} &= \frac{1}{2} \left.\frac{\partial \mathrm{F}_{\mathbf{R}}}{\partial h}\right|_{h=0} = \frac{1}{2} \frac{\partial}{\partial h} \left[\frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \mathbf{r}}{\partial t^2} - h \frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} - h \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \cdot \frac{\partial \mathbf{r}}{\partial t^2} + h^2 \frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2}\right]\Bigg|_{h=0} \\
&= -\frac{1}{2} \left[\frac{\partial \mathbf{r}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} + \frac{\partial \mathbf{r}}{\partial t^2} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^1}\right] \ .
\end{aligned}$$

Consider two infinitesimally close points P and Q which are respectively located at $\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right)$ and $\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1 + dt^1, t^2 + dt^2\right)$ as shown in Fig. 9.16. This figure basically illustrates how a surface curves in the embedding ambient space. It is precisely the second groundform that helps determine the distance from Q to the tangent plane at P. Namely,

$$\boxed{d = \left|d\mathbf{x} \cdot \widehat{\mathbf{n}}\left(t^1, t^2\right)\right| = \left|d\mathbf{r} \cdot \widehat{\mathbf{n}}\right| = \frac{1}{2} \left|\mathrm{II}_{\mathbf{r}}\right| \ .} \tag{9.266}$$

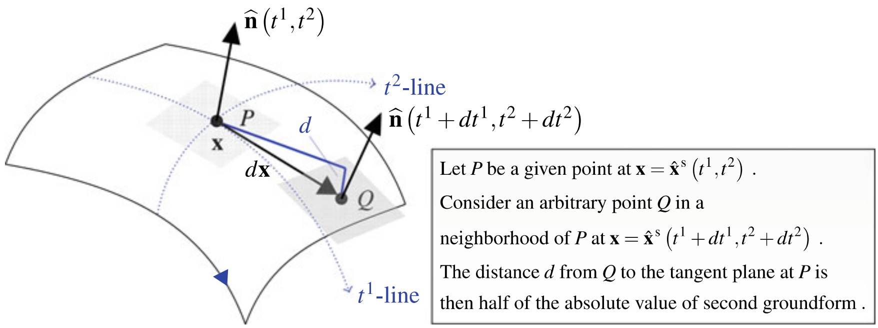


Fig. 9.16 Local structure of surface

The position of Q with respect to the tangent plane at P basically depends on the way that the surface is embedded in the surrounding ambient space. In general, four different types of point can be identified with respect to that embedding. This is described in the following.

9.5.2.1 Asymptotic Direction

An *asymptotic direction* of a regular surface $\mathcal{S}$ at a point P is a direction for which the so-called *normal curvature*, according to (9.270), becomes zero. And an *asymptotic curve* is one whose tangent line at each point over its whole domain of definition coincides with the asymptotic direction of the surface in which it embeds (note that an asymptotic curve can simply be a straight line).

Guided by $(9.270)_1$, $\kappa^{\mathrm{n}} = 0$ implies that $\mathrm{II}_{\mathbf{r}} = 0$. Denoting this direction by $\lambda = dt^2/dt^1$, one then has

$$\mathrm{g}_{\mathbf{r}}\lambda^2 + 2\mathrm{f}_{\mathbf{r}}\lambda + \mathrm{e}_{\mathbf{r}} = 0 \,. \tag{9.267}$$

Assuming $\mathrm{g}_{\mathbf{r}} \neq 0$, the roots of this quadratic equation render

$$\boxed{\lambda = \frac{-\mathrm{f}_{\mathbf{r}} \pm \sqrt{\mathrm{f}_{\mathbf{r}}^2 - \mathrm{g}_{\mathbf{r}}\mathrm{e}_{\mathbf{r}}}}{\mathrm{g}_{\mathbf{r}}} \,.} \tag{9.268}$$

Then, a point where

- $\mathrm{f}_{\mathbf{r}}^2 - \mathrm{g}_{\mathbf{r}}\mathrm{e}_{\mathbf{r}} < 0$, is called *elliptic point*. There is no (real) root and, therefore, there will be no asymptotic direction. At elliptic points, the surface is dome shaped and (locally) lies on one side of its tangential plane. As an example, any point on the surface of a sphere is an elliptic point.
- $\mathrm{f}_{\mathbf{r}}^2 - \mathrm{g}_{\mathbf{r}}\mathrm{e}_{\mathbf{r}} = 0$, is referred to as *parabolic point*. There is one repeated root and, therefore, there will be one asymptotic direction given by $\lambda = -\mathrm{g}_{\mathbf{r}}^{-1}\mathrm{f}_{\mathbf{r}}$. For instance, any point on the surface of a cylinder is a parabolic point and the

only one vertical line passing through that point represents the corresponding asymptotic line.

➪ $\mathrm{f}_{\mathbf{r}}^2 - \mathrm{g}_{\mathbf{r}}\mathrm{e}_{\mathbf{r}} > 0$, is said to be *hyperbolic* (or *saddle*) *point*. In this case, there are two roots. Accordingly, there will be two asymptotic directions determined by $\lambda = \mathrm{g}_{\mathbf{r}}^{-1}\left(-\mathrm{f}_{\mathbf{r}} \pm \sqrt{\mathrm{f}_{\mathbf{r}}^2 - \mathrm{g}_{\mathbf{r}}\mathrm{e}_{\mathbf{r}}}\right)$. At hyperbolic points, the surface is saddle shaped. A well-known example regards the one-sheeted hyperboloid which entirely consists of such points.

A point where $\mathrm{g}_{\mathbf{r}} = \mathrm{f}_{\mathbf{r}} = \mathrm{e}_{\mathbf{r}} = 0$, is termed *flat* (or *planar*) *point*. It is clear that the surface covariant curvature tensor identically vanishes for this special case. As a simple example, any point on a flat surface (or plane) is a flat point. And any curve immersed in a plane, called *plane curve*, is an asymptotic curve.

Next, suppose that $\mathrm{g}_{\mathbf{r}} = 0$ and $\mathrm{e}_{\mathbf{r}} \neq 0$. One then needs to find the roots of the quadratic equation $\mathrm{e}_{\mathbf{r}}\bar{\lambda}^2 + 2\mathrm{f}_{\mathbf{r}}\bar{\lambda} = 0$ where $\bar{\lambda} = dt^1/dt^2$.

At the end, consider the case for which $\mathrm{g}_{\mathbf{r}} = \mathrm{e}_{\mathbf{r}} = 0$ and $\mathrm{f}_{\mathbf{r}} \neq 0$. Then, $\mathrm{f}_{\mathbf{r}}\, dt^1 dt^2 = 0$ implies that the coordinate curves $t^1 = c$ and $t^2 = d$ will be the two asymptotic curves. Here, c and d denote real constants.

9.5.3 Third Fundamental Form

The *third fundamental form* (or *third groundform*) of a surface identifies the principal linear part of growth of the angle between unit normal vector fields at two infinitesimally close points. This is indicated by

$$\begin{aligned}
\mathrm{III}_{\mathbf{r}} &= d\widehat{\mathbf{n}} \cdot d\widehat{\mathbf{n}} \\
&= \frac{\partial \widehat{\mathbf{n}}}{\partial t^\alpha} \cdot \frac{\partial \widehat{\mathbf{n}}}{\partial t^\beta} dt^\alpha dt^\beta \\
&= \underline{b}_{\alpha}^{\cdot\gamma}\, a_{\gamma\theta}\, \underline{b}_{\beta}^{\cdot\theta} dt^\alpha dt^\beta \\
&= \underline{b}_{\alpha\gamma} a^{\gamma\theta}\, \underline{b}_{\theta\beta} dt^\alpha dt^\beta \; .
\end{aligned} \tag{9.269}$$

The third groundform of a surface contains no new information since it can be expressed in terms of the first and second fundamental forms of that surface. This is described below. ➥

9.5.4 Relation Between First, Second and Third Fundamental Forms

The first and second groundforms are related through the so-called *normal curvature*. This object, denoted here by κ^{n}, is defined as the ratio of the second fundamental form to the first one:

$$\kappa^{\mathrm{n}} = \frac{\mathrm{II}_{\mathbf{r}}}{\mathrm{I}_{\mathbf{r}}} = \frac{\underline{b}_{\alpha\beta} dt^{\alpha} dt^{\beta}}{a_{\gamma\delta} dt^{\gamma} dt^{\delta}} = \frac{\mathrm{e}_{\mathbf{r}} \left(dt^1\right)^2 + 2\mathrm{f}_{\mathbf{r}} dt^1 dt^2 + \mathrm{g}_{\mathbf{r}} \left(dt^2\right)^2}{\mathrm{E}_{\mathbf{r}} \left(dt^1\right)^2 + 2\mathrm{F}_{\mathbf{r}} dt^1 dt^2 + \mathrm{G}_{\mathbf{r}} \left(dt^2\right)^2} \ . \quad \longleftarrow \text{see (9.348)} \tag{9.270}$$

To obtain the relationship between the fundamental forms, one first needs to rewrite the Cayley-Hamilton equation (4.21) for two-dimensional spaces. In this case, it reads

$$\mathbf{A}^2 - I_1(\mathbf{A})\,\mathbf{A} + I_2(\mathbf{A})\,\mathbf{I} = \mathbf{O} \quad \text{where} \quad I_1(\mathbf{A}) = \mathrm{tr}\,[\,\mathbf{A}\,]\ ,\ I_2(\mathbf{A}) = \det\,[\,\mathbf{A}\,]\ . \tag{9.271}$$

Having in mind $(9.103)_1$ and $(9.104)_1$, this equation for the surface mixed curvature tensor will take the following form

$$\boxed{\underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta}^{\cdot\rho} - 2\overline{\mathrm{H}}\, \underline{b}_{\alpha}^{\cdot\rho} + \overline{\mathrm{K}}\delta_{\alpha}^{\rho} = 0\ ,} \tag{9.272}$$

in indicial notation. Multiplying both sides of (9.272) with $a_{\rho\beta}$, taking into account the relation $\underline{b}_{\theta}^{\cdot\rho} a_{\rho\beta} = \underline{b}_{\theta\beta}$ and the replacement property of the Kronecker delta, then yields

$$\underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta\beta} - 2\overline{\mathrm{H}}\, \underline{b}_{\alpha\beta} + \overline{\mathrm{K}} a_{\alpha\beta} = 0\ , \tag{9.273}$$

or

$$\underline{b}_{\alpha\gamma}\, a^{\gamma\theta}\, \underline{b}_{\theta\beta} - 2\overline{\mathrm{H}}\, \underline{b}_{\alpha\beta} + \overline{\mathrm{K}} a_{\alpha\beta} = 0\ . \tag{9.274}$$

It is now easy to arrive at the desired result

$$\boxed{\mathrm{III}_{\mathbf{r}} - 2\overline{\mathrm{H}}\,\mathrm{II}_{\mathbf{r}} + \overline{\mathrm{K}}\,\mathrm{I}_{\mathbf{r}} = 0\ .} \tag{9.275}$$

In the following, some useful identities are formulated for the subsequent developments.

Contracting the free indices in (9.272) gives

$$\boxed{\underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta}^{\cdot\alpha} - 2\overline{\mathrm{H}}\, \underline{b}_{\alpha}^{\cdot\alpha} + 2\overline{\mathrm{K}} = 0 \quad \text{or} \quad \mathrm{tr}\left[\,\underline{b}_{\alpha}^{\cdot\beta}\right]^2 = \left(2\overline{\mathrm{H}}\right)^2 - 2\overline{\mathrm{K}}\ .} \tag{9.276}$$

And this can be extended to

$$\boxed{\underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta}^{\cdot\rho}\, \underline{b}_{\rho}^{\cdot\alpha} - 2\overline{\mathrm{H}}\, \underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta}^{\cdot\alpha} + \overline{\mathrm{K}}\, \underline{b}_{\alpha}^{\cdot\alpha} = 0 \quad \text{or} \quad \mathrm{tr}\left[\,\underline{b}_{\alpha}^{\cdot\beta}\right]^3 = \left(2\overline{\mathrm{H}}\right)^3 - 6\overline{\mathrm{K}}\,\overline{\mathrm{H}}\ .} \tag{9.277}$$

9.6 Embedded Curves in Three-Dimensional Euclidean Space and Ruled Surfaces

Recall that curves or one-dimensional manifolds were **intrinsically flat** for which the Riemann-Christoffel curvature tensor identically vanished. Their Euclidean attribute thus admits a Cartesian coordinate frame. Embedded surfaces in the three-dimensional Euclidean space were partially analyzed and curves embedded in surfaces (or *surface curves*) were briefly studied within the developments achieved so far. In this section, mathematical description of curves immersed in the ambient Euclidean space (or *space curves*) are first represented. Then, their local properties such as **curvature** and **torsion** are introduced. This part provides important material required for the further study of curved surfaces. This section will finally end up with what are known as *ruled surfaces* which are generated from the space curves. The so-called *developable surface*, which represents a special ruled surface, and its three different types (namely; *generalized cylinder*, *generalized cone* and *tangent developable*) are also studied.

9.6.1 Mathematical Description of Space Curves

Consider a **space curve**[11] as a set of connected points in the surrounding external space such that any of its arbitrary connected subset can be pulled into a straight line segment without any distortion. It then makes sense to naturally define such subsets via smooth functions. In this text, a parametrized curve is defined as a **smooth mapping** from an open interval of the real line R into $\mathscr{E}_{\mathrm{p}}^3$ (i.e. the three-dimensional Euclidean point space). This is indicated by $\boldsymbol{\gamma}(t) : \mathrm{I} \to \mathscr{E}_{\mathrm{p}}^3$ where the function $\boldsymbol{\gamma}$ is defined on the interval $\mathrm{I} = (a, b) \subset \mathrm{R}$ and the real variable $t \in \mathrm{I}$ is called the *parameter* of $\boldsymbol{\gamma}$. Smooth mapping basically means that the point function

$$\underbrace{\mathbf{x} = \mathbf{x}^{\mathrm{c}}(t)}_{\text{one could also introduce the vector function } \mathbf{r} = \mathbf{x} - \mathbf{o} = \mathbf{x}^{\mathrm{c}}(t) - \mathbf{o} = \mathbf{r}^{\mathrm{c}}(t)} \quad , \tag{9.278}$$

has derivatives of all orders at all points over its domain of definition. The set $\mathbf{x}^{\mathrm{c}}(I)$ represents a subset of $\mathscr{E}_{\mathrm{p}}^3$ which is known as the *trace* of $\mathbf{x}^{\mathrm{c}}$. It is important to note that distinct parametrizations do not necessarily lead to different traces.[12]

---

[11] A *plane curve* is a curve lying in a single plane. In contrast, a *space curve* is one that may be defined over any region of its embedding three-dimensional space.

[12] As an example, consider the distinct parametrized curves

$$\boldsymbol{\alpha}(t) = (a\cos t, b\sin t) \quad , \quad \boldsymbol{\beta}(t) = (a\cos 3t, b\sin 3t) \quad \text{where} \quad t \in (0 - \varepsilon, 2\pi + \varepsilon) \,, \; \varepsilon > 0 \,.$$

Notice that their tangent vectors will not be the same:

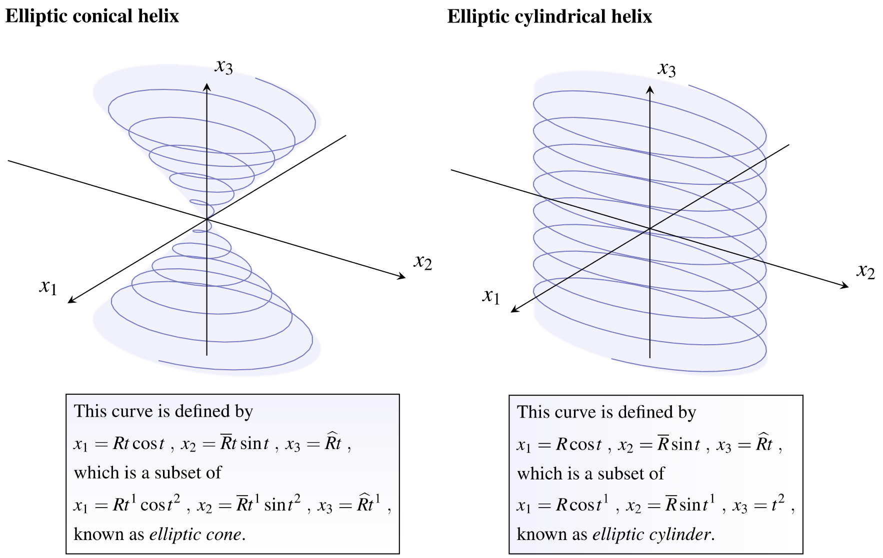


Fig. 9.17 Helix

Suppose one is given a parametrized curve $\mathbf{x}^{\mathrm{c}}(t) : \mathrm{I} \to \mathscr{E}^3_{\mathrm{p}}$ where $\mathrm{I} \subset \mathrm{R}$ and $t \in \mathrm{I}$. Such a curve is said to be *regular* if $d\mathbf{x}^{\mathrm{c}}(t)/dt \neq \mathbf{0}$ for all parameters $t \in \mathrm{I}$. For a regular curve, there exists a well-defined straight line, called the *tangent line*, at every interior point within the domain of definition. A space curve in general may contain the so-called *singular point* satisfying $d\mathbf{x}^{\mathrm{c}}(t)/dt = \mathbf{0}$ at any point corresponding to t. Of interest in this text is to only consider the regular curves which are essential for subsequent developments.

In alignment with $(9.2)_{1-3}$, let the ambient space be coordinated with a Cartesian coordinate frame for simplicity. Then,

$$\underbrace{x_1 = x_1^{\mathrm{c}}(t) \quad , \quad x_2 = x_2^{\mathrm{c}}(t) \quad , \quad x_3 = x_3^{\mathrm{c}}(t)}_{\text{or } \mathbf{r} = \mathbf{r}^{\mathrm{c}}(t) = r_1^{\mathrm{c}}(t)\,\widehat{\mathbf{e}}_1 + r_2^{\mathrm{c}}(t)\,\widehat{\mathbf{e}}_2 + r_3^{\mathrm{c}}(t)\,\widehat{\mathbf{e}}_3} \ . \tag{9.279}$$

As an example, consider an **elliptic conical helix** as shown in Fig. 9.17. It is defined by

$$\boxed{\underbrace{x_1 = Rt\cos t \quad , \quad x_2 = \overline{R}t\sin t \quad , \quad x_3 = \widehat{R}t}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + \left(x_2/\overline{R}\right)^2 = \left(x_3/\widehat{R}\right)^2} \quad ,} \tag{9.280}$$

$$\frac{d\boldsymbol{\beta}}{dt} = 3\frac{d\boldsymbol{\alpha}}{dt} \ .$$

But, they have the same trace which is nothing but the ellipse $x^2/a^2 + y^2/b^2 = 1$.

where $R, \overline{R}, \widehat{R}$ are positive real numbers and $-\infty < t < \infty$. This curve is basically a subset of a surface called *elliptic cone*.

In practice, one often deals with an **elliptic cylindrical helix**. It is defined by

$$\boxed{\underbrace{x_1 = R\cos t \quad , \quad x_2 = \overline{R}\sin t \quad , \quad x_3 = \widehat{R}t}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + \left(x_2/\overline{R}\right)^2 = 1} \quad ,} \tag{9.281}$$

where $R, \overline{R}, \widehat{R}$ are positive real numbers and $-\infty < t < \infty$. This curve lies on a surface called *elliptic cylinder* as illustrated in Fig. 9.17. Note that both elliptic conical and cylindrical helices shown in this figure are right-handed curves.

Analogously to $(9.10)_{1-2}$, the tangent vector (or *velocity vector*) to the space curve (9.278) is defined as

$$\mathbf{a}_t = \frac{d\mathbf{x}}{dt} = \lim_{h\to 0} \frac{\mathbf{x}^{\mathrm{c}}(t+h) - \mathbf{x}^{\mathrm{c}}(t)}{h} \;. \quad \longleftarrow \text{ see Exercise 9.9} \tag{9.282}$$

A straight line passing through the generic point $\mathbf{x}^{\mathrm{c}}(t_0)$ is called the *tangent line* if it has the same direction as the tangent vector $\mathbf{a}_t(t_0)$. The covariant metric tensor defined in $(9.17)_1$ now takes the form

$$a_{tt} = \mathbf{a}_t \cdot \mathbf{a}_t \quad \text{with the trivial matrix representation} \quad [a_{tt}] = \left[\; \mathbf{a}_t \cdot \mathbf{a}_t \;\right] . \tag{9.283}$$

Curve (or surface) parametrization is not unique. The curve parameter t can be any arbitrary defined quantity such as time or the so-called *arc length*. Let's for clarity distinguish between curve parametrization by the arc length parameter s (s-parametrized curve) and any other quantity t (t-parametrized curve). The curve parametrization by the arc length is known as *natural parametrization* and accordingly s may be referred to as the *natural parameter*. The natural parametrization of curves is very popular although it is sometimes difficult to come up with such useful parametrization. It basically helps develop the upcoming mathematical formulation in a more convenient and efficient way. The key point is that the tangent vector to a space curve parametrized by the natural parameter has **unit length**.[13] This is described below.

---

[13] Consider a t-parametrized space curve $\mathbf{x} = \mathbf{x}^{\mathrm{c}}(t)$ which by means of $t = t(s)$ can naturally be parametrized according to $\mathbf{x} = \mathbf{x}^{\mathrm{c}}(t) = \hat{\mathbf{x}}^{\mathrm{c}}(s)$. Then, the tangent vector $\widehat{\mathbf{a}}_1^{\mathrm{c}} = d\mathbf{x}/ds$ is of unit length owing to

$$\widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} = \frac{d\mathbf{x}}{ds} \cdot \frac{d\mathbf{x}}{ds} = \frac{ds^2}{ds^2} = 1 \; .$$

It is interesting to point out that the condition $\left|\widehat{\mathbf{a}}_1^{\mathrm{c}}\right| = 1$ is often introduced in advance as the definition of the arc length, that is,

$$ds^2 = d\mathbf{x} \cdot d\mathbf{x} \quad \Rightarrow \quad s(t) = \int_{\mathbf{x}^{\mathrm{c}}(t_1)}^{\mathbf{x}^{\mathrm{c}}(t)} \sqrt{d\mathbf{x} \cdot d\mathbf{x}} \; .$$

.

Using (9.242)$_2$, (9.279)$_{1-3}$ and (9.282)$_1$, the length of a curve segment between two points corresponding to t_1 and t renders

$$s(t) = \int_{t_1}^{t} \sqrt{\mathbf{a}_t(\hat{t}) \cdot \mathbf{a}_t(\hat{t})}\, d\hat{t}$$
$$= \int_{t_1}^{t} \sqrt{\left(\frac{dx_1^{\mathrm{c}}(\hat{t})}{dt}\right)^2 + \left(\frac{dx_2^{\mathrm{c}}(\hat{t})}{dt}\right)^2 + \left(\frac{dx_3^{\mathrm{c}}(\hat{t})}{dt}\right)^2}\, d\hat{t}\,. \tag{9.284}$$

Consequently,

$$\boxed{\frac{ds}{dt} = |\mathbf{a}_t(t)| \quad \text{or} \quad \frac{dt}{ds} = \frac{1}{|\mathbf{a}_t(s)|}\,,} \tag{9.285}$$

noting that the space curve under consideration does not contain any singular point over its whole domain of definition. It is then easy to represent

$$t(s) = \int_{s_1}^{s} \frac{1}{|\mathbf{a}_t(\hat{s})|}\, d\hat{s} \quad \text{where} \quad s_1 = s(t_1)\ . \tag{9.286}$$

This result reveals the fact that any regular space curve, in principal, admits the natural parametrization

$$\mathbf{x} = \mathbf{x}^{\mathrm{c}}(t(s)) = \hat{\mathbf{x}}^{\mathrm{c}}(s)\ , \tag{9.287}$$

with

$$\widehat{\mathbf{a}}_1^{\mathrm{c}} = \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} = \frac{d\mathbf{x}^{\mathrm{c}}}{dt}\frac{dt}{ds} = \frac{\mathbf{a}_t}{|\mathbf{a}_t|}\ , \quad \longleftarrow \text{see (9.814a)} \tag{9.288}$$

whose length apparently equals 1. With regard to this, natural parametrization may be viewed as traversing the space curve with the unit speed which may be called *unit-speed parametrization*.

9.6.2 Curvature and Torsion of Space Curves

The goal here is to study the local properties of space curves using the arc length approach. Such properties aim at describing the behavior of a space curve in a neighborhood of a generic point. To this end, let $\hat{\mathbf{x}}^{\mathrm{c}}(s) : \mathrm{I} \to \mathscr{E}_{\mathrm{p}}^3$ be a regular space curve parametrized by the natural parameter $s \in \mathrm{I}$. The condition $\left|\widehat{\mathbf{a}}_1^{\mathrm{c}}\right| = \left|d\hat{\mathbf{x}}^{\mathrm{c}}/ds\right| = 1$ immediately implies that

$$\widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = 0\ . \tag{9.289}$$

And this reveals the fact that the rate of change of the tangent vector $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ with respect to the arc length parameter s is perpendicular to itself. The new established vector $d\widehat{\mathbf{a}}_1^{\mathrm{c}}/ds = d^2\hat{\mathbf{x}}^{\mathrm{c}}/ds^2$ is called the *acceleration vector*. This object plays a major role in differential geometry of curves. Its magnitude

$$\boxed{\kappa^{\mathrm{c}} = \left| \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \right| = \left| \frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \right| ,} \quad \leftarrow \text{see (9.808)} \tag{9.290}$$

is called the *curvature* of $\hat{\mathbf{x}}^{\mathrm{c}}$ at a generic point corresponding to $s = s_0$. And its unit vector

$$\widehat{\mathbf{a}}_2^{\mathrm{c}} = \left| \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \right|^{-1} \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = \frac{1}{\kappa^{\mathrm{c}}} \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = \left| \frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \right|^{-1} \frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} , \quad \leftarrow \text{see (9.814b)} \tag{9.291}$$

is referred to as the *principal normal vector* at s_0. The unit vectors $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ at a point define a plane called *osculating plane* at that point (see Fig. 9.18).

The fact that $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ cannot be detected by the inhabitants living on the curve implies that κ^{c} should be an **extrinsic** entity. With regard to this, a curve is similar to the surface of a cylinder which is intrinsically flat but extrinsically curved. Note that, by considering (1.9d), (1.11), and $(9.290)_1$, the curvature of a space curve renders a nonnegative scalar. At points where curvature vanishes, the principal normal vector is undefined. Technically, a *singular point of order 0* refers to the point at which $\widehat{\mathbf{a}}_1^{\mathrm{c}} = \mathbf{0}$ (this issue was treated by assuming regularity). And one deals with a *singular point of order 1* when $d\widehat{\mathbf{a}}_1^{\mathrm{c}}/ds = \mathbf{0}$. Notice that $d\widehat{\mathbf{a}}_1^{\mathrm{c}}/ds = d^2\,\mathbf{x}^{\mathrm{c}}/ds^2 = \mathbf{0}$ results in the straight line $\hat{\mathbf{x}}^{\mathrm{c}}(s) = \hat{\mathbf{c}}^{\mathrm{c}} s + \hat{\mathbf{d}}^{\mathrm{c}}$ where $\hat{\mathbf{c}}^{\mathrm{c}}$, with unit length, and $\hat{\mathbf{d}}^{\mathrm{c}}$ denote constant vectors.

Of interest here is to consider the regular space curves which do not contain any singular point of order 1 over their whole domain of definition. As a result, the curvature of any space curve in this text is positive.

Recall from $(9.251)_4$ and (9.260) that $2d\mathbf{r} \cdot \widehat{\mathbf{n}} = \underline{b}_{\alpha\beta} dt^{\alpha} dt^{\beta}$. The curve analogue of this relation is given by

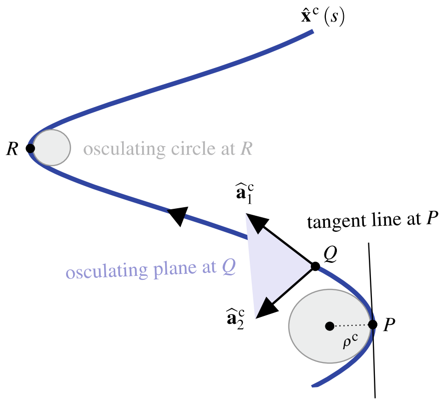
$$\boxed{\underbrace{2d\mathbf{r} \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}} = \kappa^{\mathrm{c}} ds^2}_{\text{note that } d\mathbf{r} = \frac{d\mathbf{r}}{ds} ds + \frac{1}{2} \frac{d^2\mathbf{r}}{ds^2} ds^2 + o\left(ds^2\right) = \widehat{\mathbf{a}}_1^{\mathrm{c}} ds + \frac{1}{2} \kappa^{\mathrm{c}} \widehat{\mathbf{a}}_2^{\mathrm{c}} ds^2 + o\left(ds^2\right)} .} \tag{9.292}$$

And this means that κ^{c} in curve theory plays the role of $\underline{b}_{\alpha\beta}$ in (local) theorem of surfaces.

It is important to point out that the curvature of a space curve remains **invariant** under a change of orientation.[14] This geometric object - with a long history which

[14] To show this, consider a space curve $\boldsymbol{\gamma}(s) : \mathrm{I} \to \mathscr{E}_{\mathrm{p}}^3$ parametrized by the natural parameter $s \in \mathrm{I}$. Suppose that this curve does not contain any singular point of order 0 or 1 in I. Let $\overline{\boldsymbol{\gamma}}(\overline{s}) = \boldsymbol{\gamma}(s)$ where $\overline{s} = -s$. Then,

Note that the osculating circle has exactly the same curvature as the curve at a given point. And such a circle lies on the concave side of the curve.



Let $\hat{\mathbf{x}}^{c}(s)$ be a smooth curve without any singular point of order 0 or 1. The (unit) tangent vector to this curve then renders

$\widehat{\mathbf{a}}_1^{c} = \dfrac{d\hat{\mathbf{x}}^{c}}{ds}$.

It follows that

$\dfrac{d\widehat{\mathbf{a}}_1^{c}}{ds} = \kappa^{c}\widehat{\mathbf{a}}_2^{c}$,

where κ^{c} and $\widehat{\mathbf{a}}_2^{c}$ respectively denote the *curvature* and *principal normal vector* of $\hat{\mathbf{x}}^{c}$. The quantity κ^{c} at a given point P controls the motion of $\hat{\mathbf{x}}^{c}$ along $\widehat{\mathbf{a}}_2^{c}$ in the *osculating plane* at P. At the end, $\rho^{c} = 1/\kappa^{c}$ represents the *radius of curvature* of $\hat{\mathbf{x}}^{c}$ at P. It is basically the radius of a circle known as the *osculating circle* at that point.

Fig. 9.18 Tangent vector, principal normal vector, osculating plane and osculating circle of a smooth curve with natural parametrization

dates back to the ancient times - measures how fast a curve deviates from the tangent line in a neighborhood of a generic point. In other words, the curvature of a space curve is a measure of its failure to be a straight line. This is equivalent to saying that the non-linearity of a curve at a point is captured by its curvature at that point. Physically, such a quantity illustrates the bending of a straight line segment at a certain plane. Its reciprocal according to

$$\boxed{\rho^{c} = \frac{1}{\kappa^{c}}\ ,} \tag{9.293}$$

is referred to as the *radius of curvature* of $\hat{\mathbf{x}}^{c}$ at a point corresponding to $s = s_0$. This positive scalar at a point on the curve equals the radius of a circle that best fits the curve at that point. Such a circle is called *osculating circle* (or *kissing circle*), see Fig. 9.18. This geometric object has exactly the same curvature (and tangent line) as the curve at a given point and lies in the concave (or inner) side of the curve. As a simple example, consider a circle which is basically its own osculating circle. Then, the radius of curvature will be its radius. Notice that applying more bending provides bigger curvature and accordingly the (radius of) osculating circle becomes smaller.

$$\frac{d\boldsymbol{\gamma}}{ds} = -\frac{d\overline{\boldsymbol{\gamma}}}{d\overline{s}}\ , \quad \frac{d^2\boldsymbol{\gamma}}{ds^2} = +\frac{d^2\overline{\boldsymbol{\gamma}}}{d\overline{s}^2}\ .$$

This reveals the fact that by changing the positive sense of $\boldsymbol{\gamma}(s)$, the tangent vector $d\boldsymbol{\gamma}/ds$ will be affected whereas the acceleration vector $d^2\boldsymbol{\gamma}/ds^2$ and accordingly its length, $\kappa^{c} = |d^2\boldsymbol{\gamma}/ds^2|$, remain unaffected.

The identities $\widehat{\mathbf{a}}_1^{\rm c} \cdot \widehat{\mathbf{a}}_1^{\rm c} = 1$, $\widehat{\mathbf{a}}_2^{\rm c} \cdot \widehat{\mathbf{a}}_2^{\rm c} = 1$ and $\widehat{\mathbf{a}}_2^{\rm c} \cdot \widehat{\mathbf{a}}_1^{\rm c} = 0$ with the aid of (9.291)$_2$ help obtain

$$\widehat{\mathbf{a}}_2^{\rm c} \cdot \frac{d\,\widehat{\mathbf{a}}_2^{\rm c}}{ds} = 0 \quad \text{and} \quad \frac{d\,\widehat{\mathbf{a}}_2^{\rm c}}{ds} \cdot \widehat{\mathbf{a}}_1^{\rm c} + \underbrace{\widehat{\mathbf{a}}_2^{\rm c} \cdot \frac{d\,\widehat{\mathbf{a}}_1^{\rm c}}{ds}}_{= \kappa^{\rm c}\,\widehat{\mathbf{a}}_2^{\rm c} \cdot \widehat{\mathbf{a}}_2^{\rm c} = \kappa^{\rm c}\,\widehat{\mathbf{a}}_1^{\rm c} \cdot \widehat{\mathbf{a}}_1^{\rm c}} = 0 \,, \tag{9.294}$$

or

$$\left(\frac{d\,\widehat{\mathbf{a}}_2^{\rm c}}{ds} + \kappa^{\rm c}\,\widehat{\mathbf{a}}_1^{\rm c} \right) \cdot \widehat{\mathbf{a}}_1^{\rm c} = 0 \,. \tag{9.295}$$

It is not then difficult to see that

$$\left(\frac{d\,\widehat{\mathbf{a}}_2^{\rm c}}{ds} + \kappa^{\rm c}\,\widehat{\mathbf{a}}_1^{\rm c} \right) \cdot \widehat{\mathbf{a}}_2^{\rm c} = 0 \,. \tag{9.296}$$

Thus, the new established vector $d\,\widehat{\mathbf{a}}_2^{\rm c}/ds + \kappa^{\rm c}\,\widehat{\mathbf{a}}_1^{\rm c}$ is orthogonal to both $\widehat{\mathbf{a}}_1^{\rm c}$ and $\widehat{\mathbf{a}}_2^{\rm c}$. Denoting by $\tau^{\rm c}$ and $\widehat{\mathbf{a}}_3^{\rm c}$ its length and unit vector, respectively, one can write

$$\frac{d\,\widehat{\mathbf{a}}_2^{\rm c}}{ds} + \kappa^{\rm c}\,\widehat{\mathbf{a}}_1^{\rm c} = \tau^{\rm c}\,\widehat{\mathbf{a}}_3^{\rm c} \quad \text{or} \quad \frac{d\,\widehat{\mathbf{a}}_2^{\rm c}}{ds} = \tau^{\rm c}\,\widehat{\mathbf{a}}_3^{\rm c} - \kappa^{\rm c}\,\widehat{\mathbf{a}}_1^{\rm c} \,, \tag{9.297}$$

where $\tau^{\rm c}$ is called the *torsion* of $\hat{\mathbf{x}}^{\rm c}$ at a point corresponding to $s = s_0$ and $\widehat{\mathbf{a}}_3^{\rm c}$ is referred to as the *binormal vector* at s_0. The object is chosen such that the following triad of vectors

$$\left\{ \widehat{\mathbf{a}}_i^{\rm c} \right\} := \left\{ \widehat{\mathbf{a}}_1^{\rm c}, \widehat{\mathbf{a}}_2^{\rm c}, \widehat{\mathbf{a}}_3^{\rm c} \right\} \,, \tag{9.298}$$

forms a right-handed basis for the embedding ambient space, that is,

$$\underbrace{\widehat{\mathbf{a}}_3^{\rm c} = \widehat{\mathbf{a}}_1^{\rm c} \times \widehat{\mathbf{a}}_2^{\rm c}}_{\text{note that } \widehat{\mathbf{a}}_1^{\rm c} = \widehat{\mathbf{a}}_2^{\rm c} \times \widehat{\mathbf{a}}_3^{\rm c} \text{ and } \widehat{\mathbf{a}}_2^{\rm c} = \widehat{\mathbf{a}}_3^{\rm c} \times \widehat{\mathbf{a}}_1^{\rm c}} \,, \tag{9.299}$$

or

$$\left(\widehat{\mathbf{a}}_1^{\rm c} \times \widehat{\mathbf{a}}_2^{\rm c} \right) \cdot \widehat{\mathbf{a}}_3^{\rm c} = 1 \,. \tag{9.300}$$

Accordingly, one finds that

$$\mathbf{a}_3^{\rm c} = \left| \frac{d^2\hat{\mathbf{x}}^{\rm c}}{ds^2} \right|^{-1} \frac{d\hat{\mathbf{x}}^{\rm c}}{ds} \times \frac{d^2\hat{\mathbf{x}}^{\rm c}}{ds^2} \,. \quad \longleftarrow \text{see (9.814c)} \tag{9.301}$$

Hint: In the literature, the orthonormal basis $\left\{ \widehat{\mathbf{a}}_1^{\rm c}, \widehat{\mathbf{a}}_2^{\rm c}, \widehat{\mathbf{a}}_3^{\rm c} \right\}$ is usually denoted by $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$ or $\{\mathbf{t}, \mathbf{n}, \mathbf{b}\}$.

Following the above considerations, the introduced tangent vector $\widehat{\mathbf{a}}_1^{\rm c}$ and the principal normal vector $\widehat{\mathbf{a}}_2^{\rm c}$ have already been completed by means of the binormal vector $\widehat{\mathbf{a}}_3^{\rm c}$ to provide an orthonormal basis for the three-dimensional real vector space.

This basis, referred to as the *Frenet trihedron* (or *moving trihedron*), varies as it moves along the curve and forms what is called the *Frenet frame*. Such a frame can be served as local coordinate axes of the curve under consideration which helps describe its local properties such as curvature and torsion. See Fig. 9.19 for a geometrical interpretation.

Three mutually perpendicular planes can be constructed by means of the Frenet frame: (i) the plane spanned by $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ is called the *osculating plane*, (ii) the plane defined by $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{a}}_3^{\mathrm{c}}$ is known as the *rectifying plane* and (iii) the so-called *normal plane* is one generated by $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ and $\widehat{\mathbf{a}}_3^{\mathrm{c}}$ (see Fig. 9.19). They are known as the *Frenet planes*.

The torsion of a curve measures how sharply that curve deviates from the osculating plane in a neighborhood of a generic point. In other words, the torsion of a curve indicates a measure of its failure to be contained in a plane. This is equivalent to saying that the non-planarity of a curve at a generic point is captured by its torsion at that point. Recall that a space curve is intrinsically flat. This physically means that a curve in space can be constructed from a straight line by twisting and bending of that line.

The rate of change of the binormal vector $\widehat{\mathbf{a}}_3^{\mathrm{c}}$ is related to the principal normal vector $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ through the following relationship

$$\boxed{\frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds} = -\tau^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}}\ ,} \tag{9.302}$$

because

$$\begin{aligned}
\frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds} &\xlongequal[\text{(9.299)}]{\text{from}} \frac{d}{ds}\left[\widehat{\mathbf{a}}_1^{\mathrm{c}} \times \widehat{\mathbf{a}}_2^{\mathrm{c}}\right] \\
&\xlongequal[\text{(6.4i)}]{\text{from}} \left[\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds}\right] \times \widehat{\mathbf{a}}_2^{\mathrm{c}} + \widehat{\mathbf{a}}_1^{\mathrm{c}} \times \left[\frac{d\,\widehat{\mathbf{a}}_2^{\mathrm{c}}}{ds}\right] \\
&\xlongequal[\text{(9.291) and (9.297)}]{\text{from}} \left[\kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}}\right] \times \widehat{\mathbf{a}}_2^{\mathrm{c}} + \widehat{\mathbf{a}}_1^{\mathrm{c}} \times \left[\tau^{\mathrm{c}}\,\widehat{\mathbf{a}}_3^{\mathrm{c}} - \kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_1^{\mathrm{c}}\right] \\
&\xlongequal[\text{(1.49a), (1.51) and (9.299)}]{\text{from}} -\tau^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}}\ .
\end{aligned}$$

Consequently,

$$\boxed{\tau^{\mathrm{c}} = -\widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds}\ .} \tag{9.303}$$

One can further represent

$$\boxed{\tau^{\mathrm{c}} = \left|\frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2}\right|^{-2} \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} \cdot \left(\frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \times \frac{d^3\hat{\mathbf{x}}^{\mathrm{c}}}{ds^3}\right)\ ,} \quad \longleftarrow \text{see (9.809)} \tag{9.304}$$

since

$$\tau^{\mathrm{c}} \xlongequal[\text{(9.299) and (9.303)}]{\text{from (6.4i),}} -\widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \left(\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \times \widehat{\mathbf{a}}_2^{\mathrm{c}} \right) - \widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_1^{\mathrm{c}} \times \frac{d\,\widehat{\mathbf{a}}_2^{\mathrm{c}}}{ds} \right)$$

$$\xlongequal[\text{(1.73) and (9.291)}]{\text{from}} -\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \cdot \left(\widehat{\mathbf{a}}_2^{\mathrm{c}} \times \widehat{\mathbf{a}}_2^{\mathrm{c}} \right) - \widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_1^{\mathrm{c}} \times \frac{d}{ds} \left[\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{\kappa^{\mathrm{c}} ds} \right] \right)$$

$$\xlongequal[\text{(6.4a) and (9.291)}]{\text{from (1.40), (1.51),}} \frac{1}{(\kappa^{\mathrm{c}})^2} \frac{d\kappa^{\mathrm{c}}}{ds} \widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_1^{\mathrm{c}} \times \left[\kappa^{\mathrm{c}} \widehat{\mathbf{a}}_2^{\mathrm{c}} \right] \right) - \frac{1}{\kappa^{\mathrm{c}}} \widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_1^{\mathrm{c}} \times \frac{d^2 \widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds^2} \right)$$

$$\xlongequal[\text{(1.49a) and (1.73)}]{\text{from}} \frac{1}{\kappa^{\mathrm{c}}} \frac{d\kappa^{\mathrm{c}}}{ds} \widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_2^{\mathrm{c}} \times \widehat{\mathbf{a}}_2^{\mathrm{c}} \right) + \frac{1}{\kappa^{\mathrm{c}}} \widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_2^{\mathrm{c}} \times \frac{d^2 \widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds^2} \right)$$

$$\xlongequal[\text{(9.288), (9.290) and (9.291)}]{\text{from (1.40), (1.51),}} \left| \frac{d^2 \hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \right|^{-2} \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} \cdot \left(\frac{d^2 \hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \times \frac{d^3 \hat{\mathbf{x}}^{\mathrm{c}}}{ds^3} \right) .$$

The torsion of any plane curve (with possibly nonvanishing curvature) is identically zero. Unlike the curvature, the torsion of a space curve at a point can be negative. The sign of this entity is significant because it affects the local shape of a space curve. Indeed, one needs to distinguish between a right-handed curve with positive torsion and left-handed curve possessing negative torsion, see Fig. 9.19. It is worthwhile to point out that the torsion of a space curve remains invariant under a change of orientation. This is described below.

Consider a s-parametrized space curve $\boldsymbol{\gamma}(s) : \mathrm{I} \to \mathscr{E}_{\mathrm{p}}^3$ without any singular point of order 0 or 1 in I. Let $\overline{\boldsymbol{\gamma}}(\overline{s}) = \boldsymbol{\gamma}(s)$ where $\overline{s} = -s$. The goal is now to verify that $\overline{\tau}^{\mathrm{c}}(\overline{s}) = \tau^{\mathrm{c}}(s)$ where $\overline{\tau}^{\mathrm{c}}(\overline{s})$ and $\tau^{\mathrm{c}}(s)$ respectively present the torsion of $\overline{\boldsymbol{\gamma}}(\overline{s})$ and $\boldsymbol{\gamma}(s)$. Denoting by $\left\{ \overline{\widehat{\mathbf{a}}}_i^{\mathrm{c}} \right\}$ and $\left\{ \widehat{\mathbf{a}}_i^{\mathrm{c}} \right\}$ the natural triad of $\overline{\boldsymbol{\gamma}}(\overline{s})$ and $\boldsymbol{\gamma}(s)$, respectively, one will have

$$\overline{\widehat{\mathbf{a}}}_1^{\mathrm{c}} = -\widehat{\mathbf{a}}_1^{\mathrm{c}} \quad , \quad \overline{\widehat{\mathbf{a}}}_2^{\mathrm{c}} = +\widehat{\mathbf{a}}_2^{\mathrm{c}} \qquad \text{and} \qquad \overline{\widehat{\mathbf{a}}}_3^{\mathrm{c}} = -\widehat{\mathbf{a}}_3^{\mathrm{c}} \quad , \quad \frac{d\,\overline{\widehat{\mathbf{a}}}_3^{\mathrm{c}}}{d\,\overline{s}} = +\frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{d\,s} \, .$$

As a result,

$$\overline{\tau}^{\mathrm{c}} = -\frac{d\,\overline{\widehat{\mathbf{a}}}_3^{\mathrm{c}}}{d\,\overline{s}} \cdot \overline{\widehat{\mathbf{a}}}_2^{\mathrm{c}} = -\frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds} \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}} = \tau^{\mathrm{c}} \, .$$

As can be seen, by changing the positive sense of a space curve, its binormal vector will be affected whereas its torsion remains unaffected.

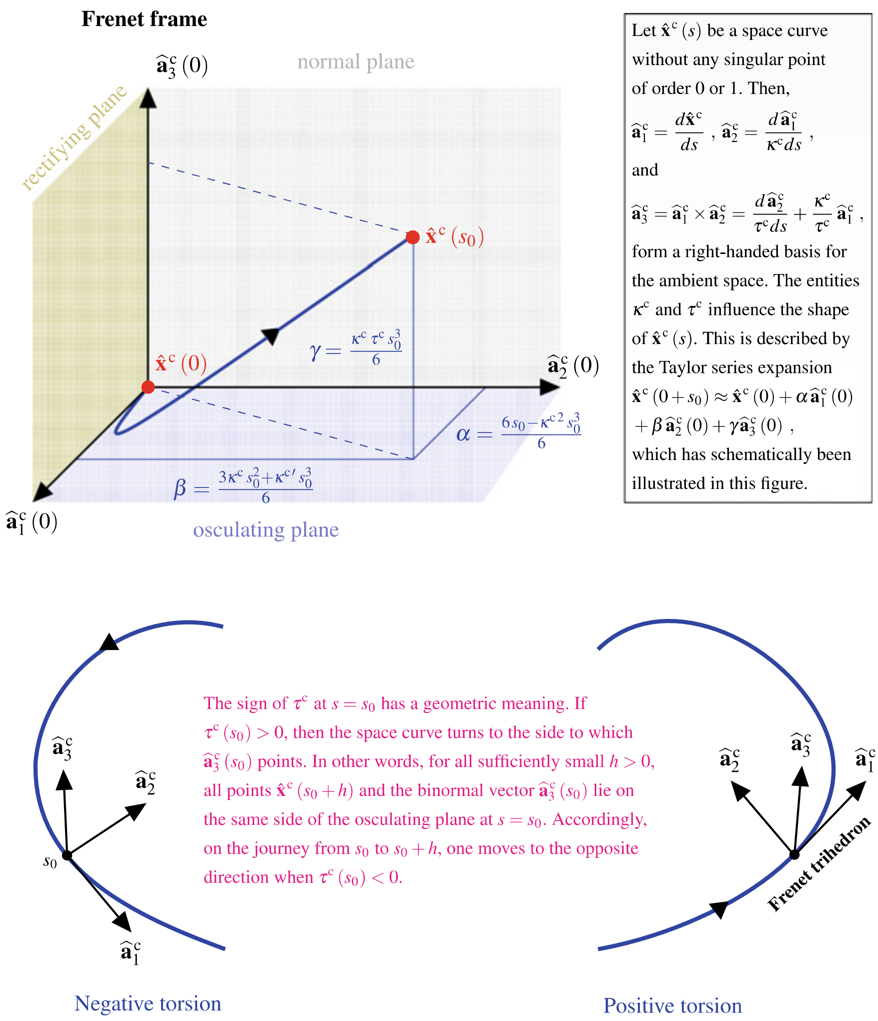


Fig. 9.19 Space curves with their local properties

9.6.3 *Frenet Formulas for Space Curves*

The so-called *Frenet* (or *Frenet-Serret*) *formulas* represent a linear system of ordinary differential equations governing the moving trihedron along a space curve (see the pioneering works of Frenet [20] and Serret [21]). They establish the rate of change of the tangent, principal normal and binormal vectors in terms of each other using the curvature and torsion. The Frenet formulas are the curve analogue of the formulas established by Gauss and Weingarten for curved surfaces. They are already derived in $(9.291)_2$, $(9.297)_2$ and (9.302). For convenience, they are listed in the following:

$$\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = \kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}} \ , \qquad \longleftarrow \text{see (9.820a) and (9.873a)} \tag{9.305a}$$

$$\frac{d\,\widehat{\mathbf{a}}_2^{\mathrm{c}}}{ds} = -\kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_1^{\mathrm{c}} + \tau^{\mathrm{c}}\,\widehat{\mathbf{a}}_3^{\mathrm{c}} \ , \tag{9.305b}$$

$$\frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds} = -\tau^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}} \ . \tag{9.305c}$$

The matrix form of the Frenet formulas then renders

$$\begin{bmatrix} \vdots & \vdots & \vdots \\ \frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} & \frac{d\,\widehat{\mathbf{a}}_2^{\mathrm{c}}}{ds} & \frac{d\,\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds} \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \widehat{\mathbf{a}}_1^{\mathrm{c}} & \widehat{\mathbf{a}}_2^{\mathrm{c}} & \widehat{\mathbf{a}}_3^{\mathrm{c}} \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & -\kappa^{\mathrm{c}} & 0 \\ \kappa^{\mathrm{c}} & 0 & -\tau^{\mathrm{c}} \\ 0 & \tau^{\mathrm{c}} & 0 \end{bmatrix} . \tag{9.306}$$

In these expressions, the functions $\kappa^{\mathrm{c}} = \kappa^{\mathrm{c}}(s)$ and $\tau^{\mathrm{c}} = \tau^{\mathrm{c}}(s)$ are called *natural* (or *intrinsic*) *equations*. Once they are known, the three unit vectors of the Frenet frame are known at each point (at least in principle). Technically, according to the so-called *fundamental theorem of space curves*, there exists a regular s-parametrized curve $\hat{\mathbf{x}}^{\mathrm{c}}(s) : \mathrm{I} \to \mathscr{E}_{\mathrm{p}}^3$ for the given sufficiently smooth functions $\kappa^{\mathrm{c}}(s) > 0$ and $\tau^{\mathrm{c}}(s)$. And any other curve, $\bar{\hat{\mathbf{x}}}^{\mathrm{c}}$, with identical conditions will differ from $\hat{\mathbf{x}}^{\mathrm{c}}(s)$ by a **rigid body motion** according to $\bar{\hat{\mathbf{x}}}^{\mathrm{c}} - \mathbf{o} = \mathbf{R}\left(\hat{\mathbf{x}}^{\mathrm{c}} - \mathbf{o}\right) + \mathbf{c}$. Here, $\mathbf{R}$ is a proper orthogonal tensor and $\mathbf{c}$ presents a real vector (they are independent of the arc length parameter). This relation basically states that any point of $\hat{\mathbf{x}}^{\mathrm{c}}(s)$ has been displaced in an identical manner. The proof of the theorem can be found in any decent book on differential geometry (see, for example, do Carmo [8]).

The Frenet formulas can be represented in an elegant way by introducing the useful vector

$$\boxed{\mathbf{d}^{\mathrm{c}} = \tau^{\mathrm{c}}\,\widehat{\mathbf{a}}_1^{\mathrm{c}} + \kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_3^{\mathrm{c}} \ ,} \qquad \longleftarrow \text{see (9.875)} \tag{9.307}$$

called *Darboux vector* (see the early pioneering work of Darboux [22]). One can now write

$$\boxed{\frac{d\,\widehat{\mathbf{a}}_i^{\mathrm{c}}}{ds} = \mathbf{d}^{\mathrm{c}} \times \widehat{\mathbf{a}}_i^{\mathrm{c}} \ , \quad i = 1, 2, 3 \ .} \tag{9.308}$$

The Frenet formulas may also be written as

$$\boxed{\frac{d\,\widehat{\mathbf{a}}_i^{\mathrm{c}}}{ds} = \mathbf{W}_s^{\mathrm{c}}\,\widehat{\mathbf{a}}_i^{\mathrm{c}} \ , \quad i = 1, 2, 3 \ ,} \tag{9.309}$$

where the skew-symmetric tensor $\mathbf{W}_s^{\mathrm{c}}$ presents

$$\boxed{\mathbf{W}_s^{\mathrm{c}} = \tau^{\mathrm{c}}\left(\widehat{\mathbf{a}}_3^{\mathrm{c}} \otimes \widehat{\mathbf{a}}_2^{\mathrm{c}} - \widehat{\mathbf{a}}_2^{\mathrm{c}} \otimes \widehat{\mathbf{a}}_3^{\mathrm{c}}\right) + \kappa^{\mathrm{c}}\left(\widehat{\mathbf{a}}_2^{\mathrm{c}} \otimes \widehat{\mathbf{a}}_1^{\mathrm{c}} - \widehat{\mathbf{a}}_1^{\mathrm{c}} \otimes \widehat{\mathbf{a}}_2^{\mathrm{c}}\right) \ .} \tag{9.310}$$

9.6.4 Ruled Surfaces

The goal here is to introduce an important class of surfaces called *ruled surfaces*. Let $\mathcal{S}$ be a surface parametrized by

$$\boxed{\mathbf{x}\left(t^1, t^2\right) = \boldsymbol{\alpha}\left(t^1\right) + t^2 \mathbf{w}\left(t^1\right) \quad , \quad a < t^1 < b \, , \; -\infty < t^2 < \infty \, ,} \tag{9.311}$$

where $\boldsymbol{\alpha}\left(t^1\right)$ denotes a point function whose trace presents a curve in space and $\mathbf{w}\left(t^1\right)$ is a vector function specifying a direction. This represents a *ruled surface*. It is basically swept out by a continuous (translational and/or rotational) motion of a straight line in space. This straight line, denoted by L_{t^2}, is known as a *ruling* (or *generator*) of $\mathcal{S}$ (see Fig. 9.20). And it has a representation of the form $t^2 \to \boldsymbol{\alpha}\left(t^1_*\right) + t^2 \mathbf{w}\left(t^1_*\right)$. The ruling L_{t^2} moves along the object $\boldsymbol{\alpha}\left(t^1\right)$ which is referred to as the *directrix* (or *base curve*) of $\mathcal{S}$. Note that in this context, one no longer insists on regularity of $\mathcal{S}$. Thus, a ruled surface is not necessarily a regular surface. It is assumed that the tangent vector to the curve of reference, $\boldsymbol{\alpha}\left(t^1\right)$, is never zero and $\mathbf{w}\left(t^1\right) \neq 0$. For convenience, it is assumed that $\mathbf{w}\left(t^1\right)$ is of unit length, i.e.

$$\boxed{\mathbf{w} \cdot \mathbf{w} = 1 \quad \text{which immediately implies that} \quad \mathbf{w} \cdot \mathbf{w}' = 0 \quad \text{where} \quad \bullet' := \frac{d\bullet}{dt^1} \, .} \tag{9.312}$$

Another assumption, generally made for developing the theory, is that

$$\boxed{\mathbf{w}'\left(t^1\right) \neq 0 \, .} \tag{9.313}$$

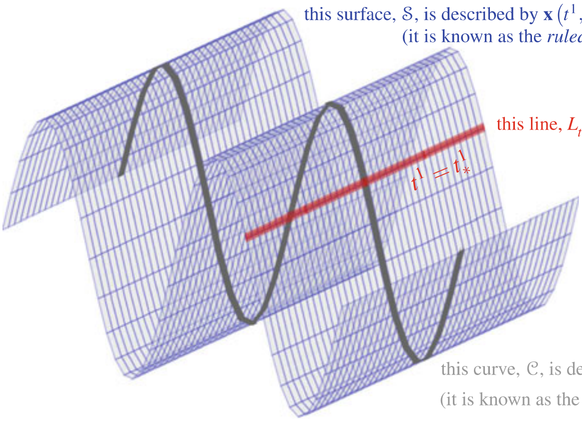


Fig. 9.20 Ruled surface

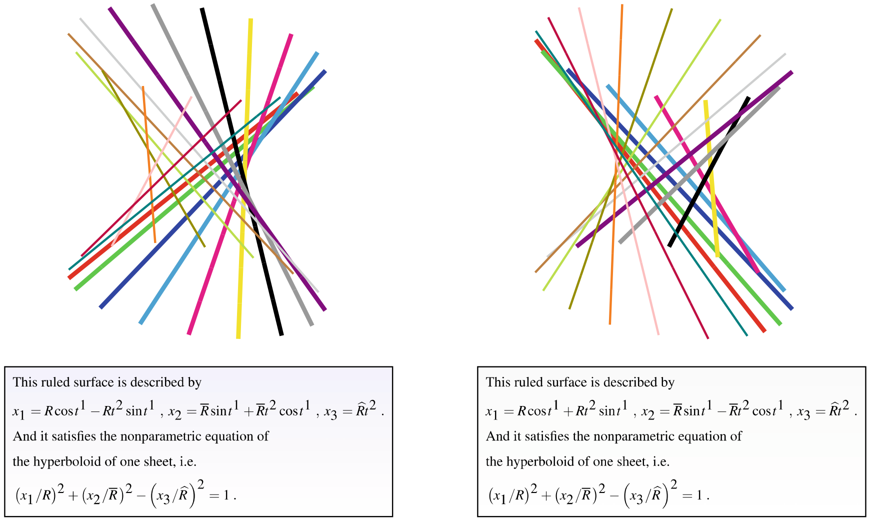


Fig. 9.21 Two families of rulings on one-sheeted hyperboloid

Accordingly, the ruled surface is said to be *noncylindrical*. This relies on the fact that $\mathbf{w}' = \mathbf{0}$ or $\mathbf{w} = \mathbf{constant} = \mathbf{w}_0$ leads to the special case $\mathbf{x}\left(t^1, t^2\right) = \boldsymbol{\alpha}\left(t^1\right) + t^2\mathbf{w}_0$ which is nothing but the parametric equation of a (generalized) cylinder, see (9.323).

Two well-known examples of ruled surfaces are cylinders and cones. It is interesting to note that the hyperboloid of one sheet is also a ruled surface. Indeed, it is a **doubly ruled** surface because the two parametric equations

$$\underbrace{x_1 = R\cos t^1 \mp Rt^2 \sin t^1 \quad , \quad x_2 = \overline{R}\sin t^1 \pm \overline{R}t^2 \cos t^1 \quad , \quad x_3 = \widehat{R}t^2}_{\text{note that } \boldsymbol{\alpha}\left(t^1\right) = \left(R\cos t^1, \overline{R}\sin t^1, 0\right) \text{ and } \mathbf{w}\left(t^1\right) = \mp R\sin t^1 \widehat{\mathbf{e}}_1 \pm \overline{R}t^2 \cos t^1 \widehat{\mathbf{e}}_2 + \widehat{R}\widehat{\mathbf{e}}_3} , \tag{9.314}$$

satisfy the following nonparametric equation of the single-sheeted hyperboloid

$$\left(\frac{x_1}{R}\right)^2 + \left(\frac{x_2}{\overline{R}}\right)^2 - \left(\frac{x_3}{\widehat{R}}\right)^2 = 1 . \quad \leftarrow \text{see (9.7)} \tag{9.315}$$

These two ruled surfaces are illustrated in Fig. 9.21.

Let $\hat{\mathbf{x}}^{\mathrm{c}}(s)$ be a unit-speed curve with positive curvature. Further, let $\widehat{\mathbf{a}}_1^{\mathrm{c}}$, $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ and $\widehat{\mathbf{a}}_3^{\mathrm{c}}$ be its unit tangent vector, unit principal normal vector and unit binormal vector, respectively. Three different ruled surfaces can then be associated with that curve. Accordingly, one can have

* the *tangent surface* (or *tangent developable*) of $\hat{\mathbf{x}}^{c}(s)$; described by,

$$\boxed{\hat{\mathbf{x}}^{s}\left(s, t^{2}\right) = \hat{\mathbf{x}}^{c}(s) + t^{2}\,\widehat{\mathbf{a}}_{1}^{c}(s)\ ,} \tag{9.316}$$

* the *principal normal surface* of $\hat{\mathbf{x}}^{c}(s)$; parametrized by,

$$\boxed{\hat{\mathbf{x}}^{s}\left(s, t^{2}\right) = \hat{\mathbf{x}}^{c}(s) + t^{2}\,\widehat{\mathbf{a}}_{2}^{c}(s)\ ,} \tag{9.317}$$

* and the *binormal surface* of $\hat{\mathbf{x}}^{c}(s)$; defined by,

$$\boxed{\hat{\mathbf{x}}^{s}\left(s, t^{2}\right) = \hat{\mathbf{x}}^{c}(s) + t^{2}\,\widehat{\mathbf{a}}_{3}^{c}(s)\ .} \tag{9.318}$$

Consider a general ruled surface of the form (9.311). For this surface, the ambient basis (9.13) can be written as

$$\mathbf{a}_{1} = \boldsymbol{\alpha}' + t^{2}\mathbf{w}' \quad , \quad \mathbf{a}_{2} = \mathbf{w} \quad , \quad \widehat{\mathbf{n}} = \frac{\boldsymbol{\alpha}' \times \mathbf{w} + t^{2}\mathbf{w}' \times \mathbf{w}}{\left|\boldsymbol{\alpha}' \times \mathbf{w} + t^{2}\mathbf{w}' \times \mathbf{w}\right|}\ . \tag{9.319}$$

Now, by using (1.49c), $(1.73)_{1}$, $(9.252\text{b})_{2}$, $(9.252\text{c})_{2}$ and $(9.319)_{1-3}$,

$$\begin{aligned} \mathrm{g}_{\mathbf{r}} &= \frac{\partial \mathbf{a}_{2}}{\partial t^{2}} \cdot \widehat{\mathbf{n}} = \mathbf{0} \cdot \widehat{\mathbf{n}} \\ &= \mathbf{0}\ , \end{aligned} \tag{9.320a}$$

$$\begin{aligned} \mathrm{f}_{\mathbf{r}} &= \frac{\partial \mathbf{a}_{1}}{\partial t^{2}} \cdot \widehat{\mathbf{n}} = \mathbf{w}' \cdot \frac{\boldsymbol{\alpha}' \times \mathbf{w} + t^{2}\mathbf{w}' \times \mathbf{w}}{\left|\boldsymbol{\alpha}' \times \mathbf{w} + t^{2}\mathbf{w}' \times \mathbf{w}\right|} \\ &= \frac{\boldsymbol{\alpha}' \cdot \left(\mathbf{w} \times \mathbf{w}'\right)}{\left|\boldsymbol{\alpha}' \times \mathbf{w} + t^{2}\mathbf{w}' \times \mathbf{w}\right|}\ . \end{aligned} \tag{9.320b}$$

Consequently, the Gaussian curvature $(9.258)_{4}$ becomes

$$\boxed{\overline{\mathrm{K}} = -\frac{\left[\boldsymbol{\alpha}' \cdot \left(\mathbf{w} \times \mathbf{w}'\right)\right]^{2}}{\left(\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^{2}\right)\left|\boldsymbol{\alpha}' \times \mathbf{w} + t^{2}\mathbf{w}' \times \mathbf{w}\right|^{2}}\ .} \tag{9.321}$$

Thus, the ruled surface is a surface with everywhere nonpositive Gaussian curvature.

9.6.4.1 Developable Surfaces

Let $\mathcal{S}$ be a (not necessarily noncylindrical) ruled surface of the form (9.311). This is called a *developable surface* (or simply *developable*) if, at regular points, its Gaussian curvature vanishes. The tangent planes are the same at all points on a ruling of a developable surface. This special surface is called developable because it can be developed into a plane without any stretching, folding or tearing.

Guided by (9.321), the ruled surface is developable if and only if

$$\boxed{\boldsymbol{\alpha}' \cdot \left(\mathbf{w} \times \mathbf{w}'\right) = 0 \,.} \tag{9.322}$$

This condition helps identify three different developable surfaces which are described in the following.

1. Consider the trivial case $\mathbf{w} \times \mathbf{w}' = \mathbf{0}$. This implies that $\mathbf{w}'$ is a scalar multiple of $\mathbf{w}$, i.e. $\mathbf{w}'\left(t^1\right) = \lambda\left(t^1\right)\mathbf{w}\left(t^1\right)$. From the relations $(9.312)_{1-2}$, it then follows that $\lambda\left(t^1\right) = 0$ which results in $\mathbf{w}\left(t^1\right) = \mathbf{w}_0$. Consequently, one can arrive at the following parametric equation of *generalized cylinder* (Fig. 9.22)

$$\mathbf{x}\left(t^1, t^2\right) = \boldsymbol{\alpha}\left(t^1\right) + t^2\mathbf{w}_0 \,. \tag{9.323}$$

2. Consider the case in which $\mathbf{w} \times \mathbf{w}' \neq \mathbf{0}$. The expression (9.322) then implies that the triplet of vectors $\boldsymbol{\alpha}'$, $\mathbf{w}$ and $\mathbf{w}'$ are linearly dependent. Thus, one can write $\boldsymbol{\alpha}'\left(t^1\right) = \lambda\left(t^1\right)\mathbf{w}\left(t^1\right) + \mu\left(t^1\right)\mathbf{w}'\left(t^1\right)$ where λ and μ are some smooth functions. At this stage, the curve $\boldsymbol{\alpha}\left(t^1\right)$ is reparametrized as

$$\tilde{\boldsymbol{\alpha}}\left(t^1\right) = \boldsymbol{\alpha}\left(t^1\right) - \mu\left(t^1\right)\mathbf{w}\left(t^1\right) \quad \text{with} \quad \tilde{\boldsymbol{\alpha}}'\left(t^1\right) = \left[\lambda\left(t^1\right) - \mu'\left(t^1\right)\right]\mathbf{w}\left(t^1\right) \,.$$

   Two different cases can now be considered which are listed in the following.

   a. Suppose that $\tilde{\boldsymbol{\alpha}}'\left(t^1\right) = \mathbf{0}$. One then has $\tilde{\boldsymbol{\alpha}}\left(t^1\right) = \tilde{\boldsymbol{\alpha}}_0$ which immediately yields $\boldsymbol{\alpha}\left(t^1\right) = \tilde{\boldsymbol{\alpha}}_0 + \mu\left(t^1\right)\mathbf{w}\left(t^1\right)$. Consequently, one can arrive at the following parametric equation of *generalized cone* (Fig. 9.22)

$$\mathbf{x}\left(\tilde{t}^1, \tilde{t}^2\right) = \tilde{\boldsymbol{\alpha}}_0 + \tilde{t}^2\mathbf{w}\left(\tilde{t}^1\right) \quad \text{where} \quad \tilde{t}^1 = t^1 \;,\; \tilde{t}^2 = t^2 + \mu\left(t^1\right) \,. \tag{9.324}$$

   b. Suppose that $\tilde{\boldsymbol{\alpha}}'\left(t^1\right) \neq \mathbf{0}$. One then finally arrives at the following parametric equation of *tangent developable* (Fig. 9.23)

$$\mathbf{x}\left(\tilde{t}^1, \tilde{t}^2\right) = \tilde{\boldsymbol{\alpha}}\left(\tilde{t}^1\right) + \tilde{t}^2\tilde{\boldsymbol{\alpha}}'\left(\tilde{t}^1\right) \quad \text{where} \quad \tilde{t}^2 = \frac{t^2 + \mu\left(t^1\right)}{\lambda\left(t^1\right) - \mu'\left(t^1\right)} \,. \tag{9.325}$$

Guided by these considerations, a developable surface is basically a union of some pieces of generalized cylinders, generalized cones and tangent developables.

Suppose one is given a regular space curve $\boldsymbol{\alpha}\left(t^1\right)$ with nowhere-vanishing curvature $\kappa^{\mathrm{c}} = \left|\boldsymbol{\alpha}'\right|^{-3}\left|\boldsymbol{\alpha}' \times \boldsymbol{\alpha}''\right|$, according to (9.808). Then,

(i) the corresponding generalized cylinder $\mathbf{x}\left(t^1, t^2\right) = \boldsymbol{\alpha}\left(t^1\right) + t^2\mathbf{w}_0$ is regular wherever $\mathbf{a}_1 \times \mathbf{a}_2 = \boldsymbol{\alpha}' \times \mathbf{w}_0 \neq \mathbf{0}$.

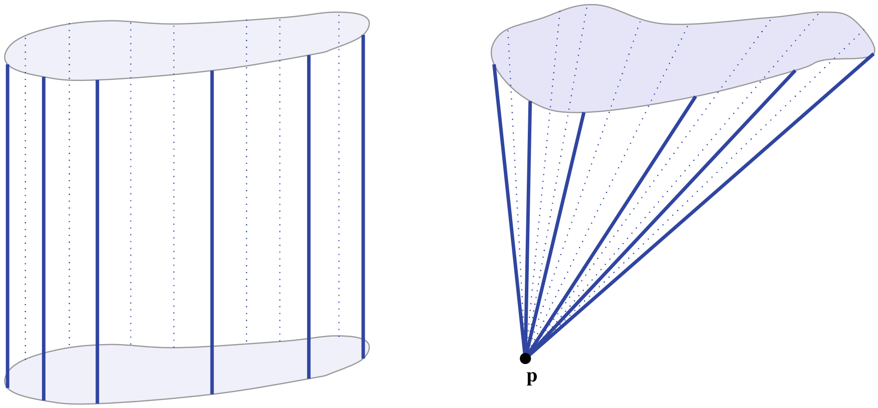


Fig. 9.22 Generalized cylinder and generalized cone

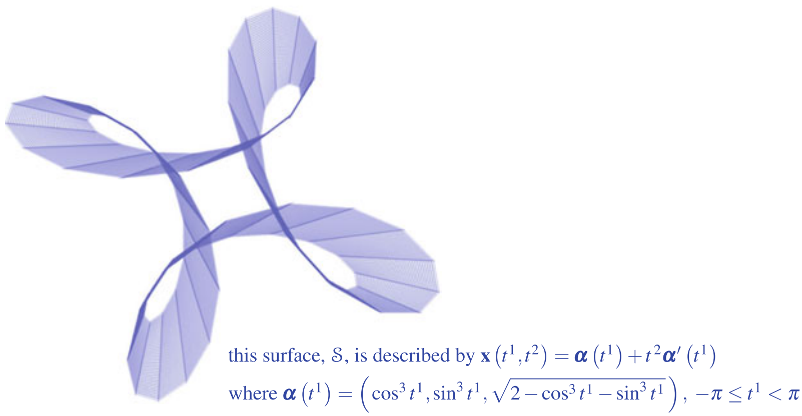


Fig. 9.23 Tangent developable

(ii) the corresponding generalized cone $\mathbf{x}\left(t^1, t^2\right) = \mathbf{p} + t^2\boldsymbol{\alpha}\left(t^1\right)$ is regular wherever $\mathbf{a}_1 \times \mathbf{a}_2 = t^2\boldsymbol{\alpha}' \times \boldsymbol{\alpha} \neq \mathbf{0}$. Such a surface is thus regular if there is no ruling tangent to the base curve. Notice that it can never be regular at its vertex $\mathbf{p}$ (at which $t^2 = 0$).

(iii) the corresponding tangent developable surface $\mathbf{x}\left(t^1, t^2\right) = \boldsymbol{\alpha}\left(t^1\right) + t^2\boldsymbol{\alpha}'\left(t^1\right)$ is regular wherever $\mathbf{a}_1 \times \mathbf{a}_2 = t^2\boldsymbol{\alpha}'' \times \boldsymbol{\alpha}' \neq \mathbf{0}$. Guided by $\kappa^{\mathrm{c}} \neq 0$, such a surface is regular at all points except those lying on the base curve (at which $t^2 = 0$).

9.6.4.2 Line of Striction and Distribution Parameter

Let $\mathcal{S}$ be a noncylindrical ruled surface of the form $\check{\mathbf{x}}\left(t^1, \check{t}^2\right) = \boldsymbol{\alpha}\left(t^1\right) + \check{t}^2 \mathbf{w}\left(t^1\right)$, according to (9.311). The goal here is to reparameterize $\mathcal{S}$ as

$$\mathbf{x}\left(t^1, t^2\right) = \boldsymbol{\beta}\left(t^1\right) + t^2 \mathbf{w}\left(t^1\right) \ , \tag{9.326}$$

where the parametrized curve $\boldsymbol{\beta}\left(t^1\right)$ is such that

$$\boldsymbol{\beta}'\left(t^1\right) \cdot \mathbf{w}'\left(t^1\right) = 0 \ . \tag{9.327}$$

This parametrized curve, $\boldsymbol{\beta}\left(t^1\right)$, is known as the *line of striction* of $\mathcal{S}$. And its points are referred to as the *central points* of $\mathcal{S}$. Let $u\left(t^1\right)$ be a smooth function measuring the distance from a point on the directrix $\boldsymbol{\alpha}\left(t^1\right)$ to a central point on the striction curve $\boldsymbol{\beta}\left(t^1\right)$. One then writes

$$\underbrace{\boldsymbol{\beta}\left(t^1\right) = \boldsymbol{\alpha}\left(t^1\right) + u\left(t^1\right)\mathbf{w}\left(t^1\right)}_{\text{with } \boldsymbol{\beta}'\left(t^1\right) = \boldsymbol{\alpha}'\left(t^1\right) + u'\left(t^1\right)\mathbf{w}\left(t^1\right) + u\left(t^1\right)\mathbf{w}'\left(t^1\right)} \ . \tag{9.328}$$

Bearing in mind that $\mathbf{w} \cdot \mathbf{w}' = 0$, the expression $\boldsymbol{\beta}' \cdot \mathbf{w}' = 0$ now helps determine $u\left(t^1\right)$ as follows:

$$\boxed{u\left(t^1\right) = -\frac{\boldsymbol{\alpha}'\left(t^1\right) \cdot \mathbf{w}'\left(t^1\right)}{\mathbf{w}'\left(t^1\right) \cdot \mathbf{w}'\left(t^1\right)} \ .} \tag{9.329}$$

As a result,

$$\begin{aligned}
\check{\mathbf{x}}\left(t^1, \check{t}^2\right) = \boldsymbol{\alpha}\left(t^1\right) + \check{t}^2 \mathbf{w}\left(t^1\right) &= \boldsymbol{\beta}\left(t^1\right) + \left[\check{t}^2 + \frac{\boldsymbol{\alpha}'\left(t^1\right) \cdot \mathbf{w}'\left(t^1\right)}{\mathbf{w}'\left(t^1\right) \cdot \mathbf{w}'\left(t^1\right)} := t^2\right] \mathbf{w}\left(t^1\right) \\
&= \boldsymbol{\beta}\left(t^1\right) + t^2 \mathbf{w}\left(t^1\right) \\
&:= \mathbf{x}\left(t^1, t^2\right) \ .
\end{aligned}$$

Notice that the striction curve of $\mathcal{S}$ becomes its directrix when $\boldsymbol{\alpha}'\left(t^1\right)$ is perpendicular to $\mathbf{w}'\left(t^1\right)$.

Let $\mathcal{S}$ be a noncylindrical ruled surface of the form (9.326). The *distribution parameter* of $\mathcal{S}$ is then defined by

$$\boxed{p\left(t^1\right) = \frac{\mathbf{w}'\left(t^1\right) \cdot \left[\boldsymbol{\beta}'\left(t^1\right) \times \mathbf{w}\left(t^1\right)\right]}{\mathbf{w}'\left(t^1\right) \cdot \mathbf{w}'\left(t^1\right)} \ .} \tag{9.330}$$

The Gaussian curvature of $\mathcal{S}$ can now be expressed in terms of $p\left(t^1\right)$. The goal is thus to simplify the Gaussian curvature (9.321) for the problem at hand. To do so, consider $\boldsymbol{\beta}' \cdot \mathbf{w}' = 0$ and $\mathbf{w} \cdot \mathbf{w}' = 0$. This implies that $\boldsymbol{\beta}' \times \mathbf{w}$ is a scalar multiple of $\mathbf{w}'$. Consequently, guided by (9.330), one will have $\boldsymbol{\beta}' \times \mathbf{w} = p\mathbf{w}'$. As a result,

$$\begin{aligned} \left|\boldsymbol{\beta}' \times \mathbf{w} + t^2\mathbf{w}' \times \mathbf{w}\right|^2 &= \left|p\mathbf{w}' + t^2\mathbf{w}' \times \mathbf{w}\right|^2 \\ &= p^2 \left|\mathbf{w}'\right|^2 + \left(t^2\right)^2 \left|\mathbf{w}' \times \mathbf{w}\right|^2 \\ &= \left[p^2 + \left(t^2\right)^2\right] \left|\mathbf{w}'\right|^2 \ . \end{aligned}$$

Consider now the covariant basis vectors

$$\left.\begin{aligned} \mathbf{a}_1 &= \boldsymbol{\beta}' + t^2\mathbf{w}' \\ \mathbf{a}_2 &= \mathbf{w} \end{aligned}\right\} \quad \text{with} \quad \left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} \boldsymbol{\beta}' \cdot \boldsymbol{\beta}' + \left(t^2\right)^2 \mathbf{w}' \cdot \mathbf{w}' & \boldsymbol{\beta}' \cdot \mathbf{w} \\ \boldsymbol{\beta}' \cdot \mathbf{w} & 1 \end{bmatrix} ,$$

and

$$\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^2 = \det\left[\, a_{\alpha\beta} \,\right] = \left|\mathbf{a}_1 \times \mathbf{a}_2\right|^2 = \left|\boldsymbol{\beta}' \times \mathbf{w} + t^2\mathbf{w}' \times \mathbf{w}\right|^2 = \left[p^2 + \left(t^2\right)^2\right] \left|\mathbf{w}'\right|^2 \ .$$

These results help finally obtain

$$\boxed{\overline{\mathrm{K}} = -\frac{p^2}{\left[p^2 + \left(t^2\right)^2\right]^2} \ .} \tag{9.331}$$

It should not be difficult now to verify that

$$\boxed{\overline{\mathrm{H}} = \frac{\left(\boldsymbol{\beta}'' + t^2\mathbf{w}''\right) \cdot \left[\left(\boldsymbol{\beta}' + t^2\mathbf{w}'\right) \times \mathbf{w}\right] - 2p\left|\mathbf{w}'\right|^2 \left(\boldsymbol{\beta}' \cdot \mathbf{w}\right)}{2\left[p^2 + \left(t^2\right)^2\right]^{3/2} \left|\mathbf{w}'\right|^3} \ .} \tag{9.332}$$

Let $\boldsymbol{\beta}\left(s\right)$ be a unit-speed curve, i.e. $\left|\boldsymbol{\beta}'\right| = 1$, with the positive curvature $\kappa^{\mathrm{c}}\left(s\right)$. Then, it can be shown that the mean curvature of the principal normal surface

$$\hat{\mathbf{x}}^{\mathrm{s}}\left(s, t^2\right) = \boldsymbol{\beta}\left(s\right) + \frac{t^2}{\kappa^{\mathrm{c}}\left(s\right)}\boldsymbol{\beta}''\left(s\right) \ , \tag{9.333}$$

takes the form

$$\overline{\mathrm{H}} = \frac{t^2\left[\tau^{\mathrm{c}\,\prime} + t^2\left(\kappa^{\mathrm{c}\,\prime}\tau^{\mathrm{c}} - \tau^{\mathrm{c}\,\prime}\kappa^{\mathrm{c}}\right)\right]}{2\left[p^2 + \left(t^2\right)^2\right]^{3/2}\left(\tau^{\mathrm{c}\,2} + \kappa^{\mathrm{c}\,2}\right)^{3/2}} \ , \quad \longleftarrow \text{the proof is given in Exercise 9.19} \tag{9.334}$$

where

$$p = \frac{\tau^{\mathrm{c}}}{\tau^{\mathrm{c}\,2} + \kappa^{\mathrm{c}\,2}} \ . \tag{9.335}$$

9.7 Curvature

Curvature represents one of the most richest topics in differential geometry which is fundamental in theory of manifolds. It quantifies the amount by which a manifold deviates from being straight. In general, there are two types of curvatures; namely, *intrinsic curvature* and *extrinsic curvature*. Intrinsic curvature is a property characterizing a manifold internally as detected by an inhabitant living on that manifold. Whereas extrinsic curvature characterizes a manifold externally as viewed by an outsider. A well-known example of intrinsic curvature regards the Gaussian curvature whereas the mean curvature has an extrinsic attribute. Various notions of curvature have been introduced so far. The main goal here is to study some other concepts of this geometric object with a long history. See Casey [23] for further discussions.

9.7.1 Curvature of Surface Curve

The curvature of a curve embedded in a regular surface can be studied from extrinsic and intrinsic points of view. And it is mainly addressed by the so-called *normal curvature* and *geodesic curvature*. The normal curvature has an extrinsic attribute whereas the geodesic curvature is an intrinsic object. The reason is that the normal curvature depends on the way in which the surface is embedded in the surrounding ambient space whereas the geodesic curvature is a measure of curving of a curve relative to the surface in which it is embedded. These important quantities are mathematically described in the following.

9.7.1.1 Normal and Geodesic Curvatures

To this end, consider a s-parametrized curve without any singularity that is defined by

$$\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1(s), t^2(s)\right) = \hat{\mathbf{x}}^{\mathrm{c}}(s) \ . \quad \leftarrow \text{see (9.239) and (9.287)} \tag{9.336}$$

One then has

$$\widehat{\mathbf{a}}_1^{\mathrm{c}} = \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} = \frac{\partial \hat{\mathbf{x}}^{\mathrm{s}}}{\partial t^\alpha}\frac{dt^\alpha}{ds} = \frac{dt^\alpha}{ds}\mathbf{a}_\alpha \ . \tag{9.337}$$

Note that the tangent vector $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ to the curve $\hat{\mathbf{x}}^{\mathrm{c}}$ at a point lies in the tangent plane of the surface in which it is embedded. Differentiation of (9.337) with respect to the arc length parameter s further yields

$$\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = \frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} = \frac{d^2t^\alpha}{ds^2}\mathbf{a}_\alpha + \frac{dt^\alpha}{ds}\frac{\partial \mathbf{a}_\alpha}{\partial t^\beta}\frac{dt^\beta}{ds} \ . \tag{9.338}$$

By means of the expressions (9.94) and (9.305a), the relation $(9.338)_2$ can be rephrased as

$$\underbrace{\kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}} = \left(\frac{d^2t^\gamma}{ds^2} + \frac{dt^\alpha}{ds}\Gamma^\gamma_{\alpha\beta}\frac{dt^\beta}{ds}\right)\mathbf{a}_\gamma + \frac{dt^\alpha}{ds}\,\underline{b}_{\alpha\beta}\frac{dt^\beta}{ds}\widehat{\mathbf{n}}}_{\text{with } d^2t^\gamma + \Gamma^\gamma_{\alpha\beta}dt^\alpha dt^\beta = \left(\widehat{\mathbf{a}}_2^{\mathrm{c}}\cdot\mathbf{a}^\gamma\right)\kappa^{\mathrm{c}}ds^2 \text{ and } \underline{b}_{\alpha\beta}dt^\alpha dt^\beta = \left(\widehat{\mathbf{a}}_2^{\mathrm{c}}\cdot\widehat{\mathbf{n}}\right)\kappa^{\mathrm{c}}ds^2} \quad . \tag{9.339}$$

Note that the vector on the left hand side of $(9.339)_1$ is basically the rate of change of the tangent vector. It is referred to as the *curvature vector*:

$$\mathbf{k}^{\mathrm{c}} = \kappa^{\mathrm{c}}\,\widehat{\mathbf{a}}_2^{\mathrm{c}} \; . \quad \longleftarrow \text{note that } \mathbf{k}^{\mathrm{c}} = d^2\hat{\mathbf{x}}^{\mathrm{c}}/ds^2 \tag{9.340}$$

This vector can be decomposed into the tangential component

$$\boxed{\mathbf{k}^{\mathrm{g}} = \left(\frac{d^2t^\gamma}{ds^2} + \frac{dt^\alpha}{ds}\Gamma^\gamma_{\alpha\beta}\frac{dt^\beta}{ds}\right)\mathbf{a}_\gamma \;,} \tag{9.341}$$

called the *geodesic curvature vector*, and the normal component

$$\boxed{\mathbf{k}^{\mathrm{n}} = \frac{dt^\alpha}{ds}\,\underline{b}_{\alpha\beta}\frac{dt^\beta}{ds}\widehat{\mathbf{n}} \;,} \tag{9.342}$$

termed the *normal curvature vector*. Up to a sign, the *geodesic curvature*, κ^{g}, is defined to be the length of the geodesic curvature vector:

$$\boxed{\kappa^{\mathrm{g}} = \pm\left|\mathbf{k}^{\mathrm{g}}\right| \; .} \tag{9.343}$$

In a similar manner, the so-called *normal curvature*, κ^{n}, is defined as

$$\boxed{\kappa^{\mathrm{n}} = \pm\left|\mathbf{k}^{\mathrm{n}}\right| \; .} \tag{9.344}$$

Now, one can trivially write

$$\mathbf{k}^{\mathrm{c}} = \mathbf{k}^{\mathrm{g}} + \mathbf{k}^{\mathrm{n}} \quad \text{with} \quad \mathbf{k}^{\mathrm{g}}\cdot\mathbf{k}^{\mathrm{n}} = 0 \quad \text{and} \quad \left(\kappa^{\mathrm{c}}\right)^2 = \left(\kappa^{\mathrm{n}}\right)^2 + \left(\kappa^{\mathrm{g}}\right)^2 \; . \tag{9.345}$$

Let ϕ be the angle between $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ and $\widehat{\mathbf{n}}$ at a given point as illustrated in Fig. 9.25. One then has

$$\boxed{\underbrace{\kappa^{\mathrm{g}} = \kappa^{\mathrm{c}}\sin\phi \quad , \quad \kappa^{\mathrm{n}} = \kappa^{\mathrm{c}}\cos\phi}_{\text{or } \kappa^{\mathrm{c}} = \kappa^{\mathrm{g}}\sin\phi + \kappa^{\mathrm{n}}\cos\phi} \;,} \tag{9.346}$$

and

$$\boxed{\cos\phi = \widehat{\mathbf{a}}_2^{\mathrm{c}}\cdot\widehat{\mathbf{n}} \quad , \quad \sin\phi = \widehat{\mathbf{a}}_3^{\mathrm{c}}\cdot\widehat{\mathbf{n}} \; .} \tag{9.347}$$

It is easy to express the normal curvature in terms of the first and second fundamental forms of the surface:

$$\boxed{\kappa^{\mathrm{n}} \xlongequal[\text{(9.342) and (9.344)}]{\text{from}} \frac{\underline{b}_{\alpha\beta} dt^\alpha dt^\beta}{ds ds} \xlongequal[\text{(9.251)}]{\text{from}} \frac{\mathrm{II}_{\mathbf{r}}}{ds^2} \xlongequal[\text{(9.227)}]{\text{from}} \frac{\mathrm{II}_{\mathbf{r}}}{\mathrm{I}_{\mathbf{r}}}\ ,} \tag{9.348}$$

which has already been given in (9.270). Considering the fact that $\widehat{\mathbf{n}}$ and $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ are orthogonal unit vectors leads to

$$\boxed{\kappa^{\mathrm{n}} = -\frac{d\,\widehat{\mathbf{n}}}{ds} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}\ .} \tag{9.349}$$

The interested reader may use (9.349) to arrive at $(9.348)_3$. Having in mind that the inner product is a symmetric bilinear form and using the chain rule of differentiation, this can be shown as follows:

$$\begin{aligned}
\kappa^{\mathrm{n}} = -\left(\frac{d\,\widehat{\mathbf{n}}}{ds}\right) \cdot \left[\widehat{\mathbf{a}}_1^{\mathrm{c}}\right] &\xlongequal[\text{(9.288)}]{\text{from}} -\left(\frac{d\,\widehat{\mathbf{n}}}{dt}\frac{dt}{ds}\right) \cdot \left[\frac{\mathbf{a}_t}{|\mathbf{a}_t|}\right] \\
&\xlongequal[\text{(9.285)}]{\text{from}} -\frac{1}{|\mathbf{a}_t|^2}\left(\frac{\partial\widehat{\mathbf{n}}}{\partial t^\alpha}\frac{dt^\alpha}{dt}\right) \cdot [\mathbf{a}_t] \\
&\xlongequal[\text{(9.99) and (9.240)}]{\text{from}} -\frac{1}{|\mathbf{a}_t|^2}\left(-\underline{b}_{\alpha\gamma}\mathbf{a}^\gamma\right) \cdot \left[\mathbf{a}_\beta\right] \frac{dt^\alpha}{dt}\frac{dt^\beta}{dt} \\
&\xlongequal[\text{(1.11) and (9.33)}]{\text{from}} \frac{\underline{b}_{\alpha\gamma}\delta_\beta^\gamma dt^\alpha dt^\beta}{(\mathbf{a}_t \cdot \mathbf{a}_t)\,dtdt} \\
&\xlongequal[\text{(5.14) and (9.241)}]{\text{from}} \frac{\underline{b}_{\alpha\beta} dt^\alpha dt^\beta}{a_{\gamma\delta} dt^\gamma dt^\delta} \\
&\xlongequal[\text{(9.227) and (9.251)}]{\text{from}} \frac{\mathrm{II}_{\mathbf{r}}}{\mathrm{I}_{\mathbf{r}}}\ .
\end{aligned}$$

One can also establish

$$\boxed{\kappa^{\mathrm{n}} = -\,\widehat{\mathbf{n}}|_{\widehat{\mathbf{a}}_1^{\mathrm{c}}} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}\ ,} \quad \longleftarrow \text{see (9.454)} \tag{9.350}$$

since

$$\begin{aligned}
-\,\widehat{\mathbf{n}}|_{\left(\frac{dt^\beta}{ds}\mathbf{a}_\beta\right)} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} = -\,\widehat{\mathbf{n}}|_\beta \frac{dt^\beta}{ds} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} &= -\frac{\partial\widehat{\mathbf{n}}}{\partial t^\beta}\frac{dt^\beta}{ds} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} \\
&= \underline{b}_{\beta\alpha}\frac{dt^\beta}{ds}\mathbf{a}^\alpha \cdot \frac{dt^\gamma}{ds}\mathbf{a}_\gamma = \underline{b}_{\alpha\beta}\frac{dt^\beta}{ds}\delta_\gamma^\alpha\frac{dt^\gamma}{ds} \\
&= \underline{b}_{\alpha\beta}\frac{dt^\alpha}{ds}\frac{dt^\beta}{ds}\ .
\end{aligned}$$

Consider a curve $\mathcal{C}$ with nowhere vanishing unit tangent vector $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and curvature κ^{c} which lies on the intersection of two regular surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$. Let $\widehat{\mathbf{n}}_1$ ($\widehat{\mathbf{n}}_2$) be the unit normal field to $\mathcal{S}_1$ ($\mathcal{S}_2$). Further, let κ_1^{n} (κ_2^{n}) be the normal curvature of $\mathcal{S}_1$ ($\mathcal{S}_2$) in the direction of $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and α be the angel between the unit normal fields, i.e. $\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2 = \cos\alpha$. Then, the curvature of $\mathcal{C}$ and the normal curvatures of $\mathcal{S}_1$ and $\mathcal{S}_2$ are related through the following equation

$$\boxed{\kappa^{\mathrm{c}\,2} \sin^2\alpha = \kappa_1^{\mathrm{n}\,2} - 2\kappa_1^{\mathrm{n}}\kappa_2^{\mathrm{n}} \cos\alpha + \kappa_2^{\mathrm{n}\,2}\ ,} \tag{9.351}$$

because

$$\begin{aligned}
\left|\kappa_1^{\mathrm{n}}\widehat{\mathbf{n}}_2 - \kappa_2^{\mathrm{n}}\widehat{\mathbf{n}}_1\right|^2 &\xlongequal[\text{from (9.346) and (9.347)}]{\text{on the one hand}} \left|\kappa^{\mathrm{c}}\left(\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}}\right)\widehat{\mathbf{n}}_2 - \kappa^{\mathrm{c}}\left(\widehat{\mathbf{n}}_2 \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}}\right)\widehat{\mathbf{n}}_1\right|^2 \\
&\xlongequal[\text{(1.72)}]{\text{from}} \left|\kappa^{\mathrm{c}}\left(\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2\right) \times \widehat{\mathbf{a}}_2^{\mathrm{c}}\right|^2 \\
&\xlongequal[\text{and (1.78a)}]{\text{from (1.11)}} \kappa^{\mathrm{c}\,2}\left\{\left[\left(\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2\right) \cdot \left(\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2\right)\right]\left[\widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}}\right] - \left[\left(\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2\right) \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}}\right]^2\right\} \\
&\xlongequal[\widehat{\mathbf{a}}_1^{\mathrm{c}} \perp \widehat{\mathbf{a}}_2^{\mathrm{c}} \text{ and } (\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2) \parallel \widehat{\mathbf{a}}_1^{\mathrm{c}}]{\text{from } \widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \widehat{\mathbf{a}}_2^{\mathrm{c}} = 1,} \kappa^{\mathrm{c}\,2}\left\{\left(\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2\right) \cdot \left(\widehat{\mathbf{n}}_1 \times \widehat{\mathbf{n}}_2\right)\right\} \\
&\xlongequal[\text{(1.78a)}]{\text{from}} \kappa^{\mathrm{c}\,2}\left\{\left(\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_1\right)\left(\widehat{\mathbf{n}}_2 \cdot \widehat{\mathbf{n}}_2\right) - \left(\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2\right)^2\right\} \\
&\xlongequal[\text{assumption}]{\text{by}} \kappa^{\mathrm{c}\,2}\left\{1 - \cos^2\alpha\right\} \\
&\xlongequal[\text{hand from (1.11)}]{\text{on the other}} \kappa_1^{\mathrm{n}\,2}\left(\widehat{\mathbf{n}}_2 \cdot \widehat{\mathbf{n}}_2\right)^2 - 2\kappa_1^{\mathrm{n}}\kappa_2^{\mathrm{n}}\left(\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2\right) + \kappa_2^{\mathrm{n}\,2}\left(\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_1\right)^2 \\
&\xlongequal[\text{assumption}]{\text{by}} \kappa_1^{\mathrm{n}\,2} - 2\kappa_1^{\mathrm{n}}\kappa_2^{\mathrm{n}} \cos\alpha + \kappa_2^{\mathrm{n}\,2}\ .
\end{aligned}$$

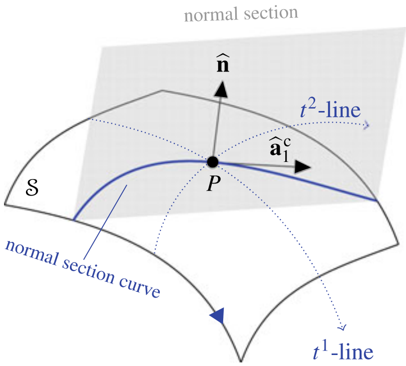
9.7.1.2 Normal Curvature of Coordinate Curves

Consider the fact that $dt^1 \neq 0$ and $dt^2 = 0$ on the t^1-curve. As a result, the normal curvature in the direction of the t^1-curve takes the form

$$\boxed{\kappa_{t^1}^{\mathrm{n}} = \frac{b_{11}}{a_{11}} = \frac{\mathrm{e}_{\mathbf{r}}}{\mathrm{E}_{\mathbf{r}}}\ .} \tag{9.352}$$

In a similar manner,

$$\boxed{\kappa_{t^2}^{\mathrm{n}} = \frac{b_{22}}{a_{22}} = \frac{\mathrm{g}_{\mathbf{r}}}{\mathrm{G}_{\mathbf{r}}}\ .} \tag{9.353}$$



Let $\widehat{\mathbf{n}}$ be the unit normal vector to a regular surface $\mathcal{S}$ at P. Further, let $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ be a given unit vector in the tangent space of $\mathcal{S}$ at P. The plane containing these vectors is called the *normal section*. And the curve cut out by this plane is referred to as the *normal section curve*. According to **Meusnier theorem**, all surface curves passing through a point on a surface and possessing the identical tangent line will have the same normal curvature at that point.

Fig. 9.24 Normal section

9.7.1.3 Meusnier Theorem

Let $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ be a unit vector in the tangent plane of a regular surface $\mathcal{S}$ at a given point P. Further, let $\widehat{\mathbf{n}}$ be the unit normal to $\mathcal{S}$ at P. Then, the plane spanned by $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{n}}$ is said to be the *normal section* of $\mathcal{S}$ at P in the direction of $\widehat{\mathbf{a}}_1^{\mathrm{c}}$. And the curve cut out by this plane is called the *normal section curve* (or simply *normal section*) of $\mathcal{S}$ at P along $\widehat{\mathbf{a}}_1^{\mathrm{c}}$, see Fig. 9.24.

Let $\hat{\lambda} = dt^2/dt^1$ be the direction of the tangent line to the normal section curve at a point P. Then, κ^{n} only depends on $\hat{\lambda}$ via the following relation

$$\boxed{\kappa^{\mathrm{n}} = \frac{\mathrm{e_r} + 2\mathrm{f_r}\hat{\lambda} + \mathrm{g_r}\hat{\lambda}^2}{\mathrm{E_r} + 2\mathrm{F_r}\hat{\lambda} + \mathrm{G_r}\hat{\lambda}^2}\ .} \quad \longleftarrow \text{see (9.270)} \tag{9.354}$$

This leads to the following theorem (see the pioneering work of Meusnier [24]):

Theorem A (*Meusnier*)

All surface curves passing through a point on a surface and possessing the identical tangent line will have the same normal curvature at that point.

The Meusnier theorem basically states that the quantity $\kappa^{\mathrm{n}} = \kappa^{\mathrm{c}} \cos\phi$ (Fig. 9.25) is an **invariant** object. This can be written as

$$\underbrace{\kappa^{\mathrm{n}} = \left(\widehat{\mathbf{a}}_1^{\mathrm{c}} \otimes \widehat{\mathbf{a}}_1^{\mathrm{c}}\right) : \mathbf{b} = \widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \mathbf{b}\widehat{\mathbf{a}}_1^{\mathrm{c}}}_{\text{notice that } \kappa^{\mathrm{n}} = \frac{dt^\alpha}{ds}\mathbf{a}_\alpha \cdot \left[\left(b_{\gamma\delta}\mathbf{a}^\gamma \otimes \mathbf{a}^\delta\right)\left(\frac{dt^\beta}{ds}\mathbf{a}_\beta\right)\right] = \frac{dt^\alpha}{ds}\, b_{\alpha\beta}\, \frac{dt^\beta}{ds}} \ . \tag{9.355}$$

As can be seen from Fig. 9.24, the geodesic curvature identically vanishes for a normal section curve at a given point P in the direction of $\widehat{\mathbf{a}}_1^{\mathrm{c}}$. This helps realize that

$\kappa^{\text{c}} = |\kappa^{\text{n}}|$ and in general $\widehat{\mathbf{a}}_2^{\text{c}} = \pm\widehat{\mathbf{n}}$ for such a curve. Regarding the normal section curve of a surface $\mathcal{S}$, one will have $\widehat{\mathbf{a}}_2^{\text{c}} = +\widehat{\mathbf{n}}$ ($\widehat{\mathbf{a}}_2^{\text{c}} = -\widehat{\mathbf{n}}$) when $\widehat{\mathbf{n}}$ points in the direction of concavity (convexity) of $\mathcal{S}$.

In the following, the goal is to provide a formula for computing the geodesic curvature. This helps introduce a new orthonormal basis for the surface curve under consideration. ❍

With the aid of $(9.339)_1$ and (9.341), considering the identities $\widehat{\mathbf{a}}_1^{\text{c}} \cdot \widehat{\mathbf{a}}_2^{\text{c}} = 0$ and $\widehat{\mathbf{a}}_1^{\text{c}} \cdot \widehat{\mathbf{n}} = 0$, one can deduce that $\mathbf{k}^{\text{g}} \cdot \widehat{\mathbf{a}}_1^{\text{c}} = 0$. From (9.342) and $(9.345)_2$, one can also deduce that $\mathbf{k}^{\text{g}} \cdot \widehat{\mathbf{n}} = 0$. Thus, $\mathbf{k}^{\text{g}}$ is perpendicular to both $\widehat{\mathbf{a}}_1^{\text{c}}$ and $\widehat{\mathbf{n}}$. This implies that $\mathbf{k}^{\text{g}}$ should be *collinear* with the unit vector $\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\text{c}}$, i.e. $\mathbf{k}^{\text{g}}$ is a scalar multiple of $\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\text{c}}$. The geodesic curvature is then exactly this proportionality factor:

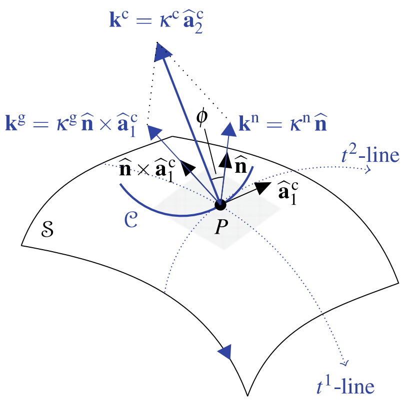
$$\mathbf{k}^{\text{g}} = \kappa^{\text{g}}\, \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\text{c}} \,. \tag{9.356}$$

This decomposition finally helps obtain the desired result

$$\boxed{\kappa^{\text{g}} = \left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\text{c}}\right) \cdot \mathbf{k}^{\text{g}} \,.} \quad ❍ \tag{9.357}$$

In addition to the (orthonormal) curve-based Frenet frame $\left\{\widehat{\mathbf{a}}_1^{\text{c}}, \widehat{\mathbf{a}}_2^{\text{c}}, \widehat{\mathbf{a}}_3^{\text{c}}\right\}$ and the (generally non-orthonormal) surface-based frame $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}$, one can now establish another (orthonormal) frame; namely $\left\{\widehat{\mathbf{a}}_1^{\text{c}}, \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\text{c}}, \widehat{\mathbf{n}}\right\}$. See Fig. 9.25 for a geometrical interpretation.

Note that the Frenet frame is a very useful tool for studying the geometry of space curves. In the case that a curve is embedded in a surface, its principal normal and binormal vectors are generally neither tangent nor perpendicular to that surface. And this means that the Frenet trihedron $\left\{\widehat{\mathbf{a}}_1^{\text{c}}, \widehat{\mathbf{a}}_2^{\text{c}}, \widehat{\mathbf{a}}_3^{\text{c}}\right\}$ will not be suitable for describing



Consider a curve $\mathcal{C}$ embedded in a surface $\mathcal{S}$ and parametrized by its arc length s. It is defined by $\mathbf{x} = \widehat{\mathbf{x}}^{\text{s}}\left(t^1(s), t^2(s)\right) = \widehat{\mathbf{x}}^{\text{c}}(s)$ whose tangent vector is $\widehat{\mathbf{a}}_1^{\text{c}} = d\widehat{\mathbf{x}}^{\text{c}}/ds$ and whose principal normal vector renders $\widehat{\mathbf{a}}_2^{\text{c}} = \left(\kappa^{\text{c}}\right)^{-1} d\widehat{\mathbf{a}}_1^{\text{c}}/ds$ where κ^{c} presents the curvature of $\mathcal{C}$ at a point P. One then has $\mathbf{k}^{\text{c}} = \mathbf{k}^{\text{n}} + \mathbf{k}^{\text{g}}$ or $\kappa^{\text{c}}\,\widehat{\mathbf{a}}_2^{\text{c}} = \kappa^{\text{n}}\,\widehat{\mathbf{n}} + \kappa^{\text{g}}\,\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\text{c}}$ where
$\mathbf{k}^{\text{c}}$ is the curvature vector,
$\mathbf{k}^{\text{n}}$ denotes the normal curvature vector,
$\mathbf{k}^{\text{g}}$ presents the geodesic curvature vector,
κ^{n} stands for the normal curvature,
κ^{g} represents the geodesic curvature, and
$\widehat{\mathbf{n}}$ is reserved for the unit normal vector to $\mathcal{S}$.

Fig. 9.25 Normal and geodesic curvature vectors the geometry of surface curves. This is the reason for introducing the new trihedron $\left\{\widehat{\mathbf{a}}_1^c, \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^c, \widehat{\mathbf{n}}\right\}$. Recall that the derivative of Frenet trihedron with respect to the arc length parameter relied on the (ordinary) curvature and torsion at a given point. But, the sensitivity of the new established moving trihedron relative to the arc length parameter relies on the normal curvature, geodesic curvature and another geometric object called *geodesic torsion*. Some alternative forms of this important quantity has been given in (9.468)-(9.473). And the resulting system of ordinary differential equations has been represented in (9.873a)-(9.873c).

9.7.1.4 Some Forms of Geodesic Curvature

The relation (9.357) can also be written as

$$\kappa^{\mathrm{g}} = \left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^c\right) \cdot \mathbf{k}^c \quad \text{where} \quad \widehat{\mathbf{a}}_1^c = \frac{d\hat{\mathbf{x}}^c}{ds} \quad \text{and} \quad \mathbf{k}^c = \frac{d\,\widehat{\mathbf{a}}_1^c}{ds} = \frac{d^2\hat{\mathbf{x}}^c}{ds^2} \,. \tag{9.358}$$

Having in mind the identity $(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{u} = (\mathbf{w} \times \mathbf{u}) \cdot \mathbf{v}$, this takes the useful form

$$\kappa^{\mathrm{g}} = \left(\frac{d\hat{\mathbf{x}}^c}{ds} \times \frac{d^2\hat{\mathbf{x}}^c}{ds^2}\right) \cdot \widehat{\mathbf{n}} \,. \tag{9.359}$$

The curvature κ^c and the normal curvature κ^n have extrinsic attribute while the geodesic curvature κ^g is an **intrinsic** object owing to

$$\begin{aligned}\kappa^{\mathrm{g}} = \sqrt{a}\Bigg[&+\Gamma_{11}^2\left(\frac{dt^1}{ds}\right)^3 - \Gamma_{22}^1\left(\frac{dt^2}{ds}\right)^3 \\ &+ \left(2\Gamma_{12}^2 - \Gamma_{11}^1\right)\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds} - \left(2\Gamma_{12}^1 - \Gamma_{22}^2\right)\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2 \\ &+ \frac{dt^1}{ds}\frac{d^2t^2}{ds^2} - \frac{d^2t^1}{ds^2}\frac{dt^2}{ds}\Bigg] \,. \quad \longleftarrow \text{the proof is given in Exercise 9.14}\end{aligned} \tag{9.360}$$

This relation can be represented in a more elegant way as

$$\kappa^{\mathrm{g}} = \sqrt{a}\,\det\begin{bmatrix} \frac{dt^1}{ds} & \frac{d^2t^1}{ds^2} + \Gamma_{11}^1\left[\frac{dt^1}{ds}\right]^2 + 2\Gamma_{12}^1\frac{dt^1}{ds}\frac{dt^2}{ds} + \Gamma_{22}^1\left[\frac{dt^2}{ds}\right]^2 \\ \frac{dt^2}{ds} & \frac{d^2t^2}{ds^2} + \Gamma_{11}^2\left[\frac{dt^1}{ds}\right]^2 + 2\Gamma_{12}^2\frac{dt^1}{ds}\frac{dt^2}{ds} + \Gamma_{22}^2\left[\frac{dt^2}{ds}\right]^2 \end{bmatrix} \,. \tag{9.361}$$

Another representation is

$$\kappa^{\mathrm{g}} = \frac{1}{\sqrt{a}}\left[\frac{\partial}{\partial t^1}\left(\mathbf{a}_2 \cdot \widehat{\mathbf{a}}_1^c\right) - \frac{\partial}{\partial t^2}\left(\mathbf{a}_1 \cdot \widehat{\mathbf{a}}_1^c\right)\right] \,. \quad \longleftarrow \text{the proof is given in Exercise 9.14} \tag{9.362}$$

Consider a surface curve described by $\varphi\left(t^1, t^2\right) = \text{constant}$. Its geodesic curvature is then given by

$$\kappa^{g} = \pm\frac{1}{\sqrt{a}}\left[\frac{\partial}{\partial t^1}\frac{a_{12}\partial\varphi/\partial t^2 - a_{22}\partial\varphi/\partial t^1}{\tilde{\Omega}}\right. \quad \longleftarrow \text{this is known as the } \textit{Bonnet formula} \text{ for the geodesic curvature, see (9.472)}$$

$$\left. + \frac{\partial}{\partial t^2}\frac{a_{12}\partial\varphi/\partial t^1 - a_{11}\partial\varphi/\partial t^2}{\tilde{\Omega}}\right], \quad \longleftarrow \text{the proof is given in Exercise 9.14} \tag{9.363}$$

where

$$\tilde{\Omega} := \sqrt{a_{11}\left(\frac{\partial\varphi}{\partial t^2}\right)^2 - 2a_{12}\frac{\partial\varphi}{\partial t^1}\frac{\partial\varphi}{\partial t^2} + a_{22}\left(\frac{\partial\varphi}{\partial t^1}\right)^2} . \tag{9.364}$$

Let $\mathcal{C}$ be a naturally represented regular curve, $\mathbf{x} = (x(s), y(s), z(s))$, embedded in an **implicit surface** $\mathcal{S}$ of the form

$$f(x, y, z) = 0 , \tag{9.365}$$

with

$$\mathbf{x}' = x'\widehat{\mathbf{e}}_x + y'\widehat{\mathbf{e}}_y + z'\widehat{\mathbf{e}}_z \quad \text{where} \quad \bullet' = \frac{d\bullet}{ds} , \tag{9.366a}$$

$$\mathbf{x}'' = x''\widehat{\mathbf{e}}_x + y''\widehat{\mathbf{e}}_y + z''\widehat{\mathbf{e}}_z \quad \text{where} \quad \bullet'' = \frac{d^2\bullet}{ds^2} , \tag{9.366b}$$

$$\widehat{\mathbf{n}} = \frac{\text{grad} f}{|\text{grad} f|} = \frac{f_x\widehat{\mathbf{e}}_x + f_y\widehat{\mathbf{e}}_y + f_z\widehat{\mathbf{e}}_z}{\sqrt{f_x^2 + f_y^2 + f_z^2}} \quad \text{where} \quad f_\bullet = \frac{\partial f}{\partial \bullet} . \tag{9.366c}$$

Here, $\left\{\widehat{\mathbf{e}}_x, \widehat{\mathbf{e}}_y, \widehat{\mathbf{e}}_z\right\}$ denotes the standard basis. Guided by (9.359), the geodesic curvature of $\mathcal{C}$ then renders

$$\kappa^{g} = \frac{\left(y'z'' - z'y''\right) f_x + \left(z'x'' - x'z''\right) f_y + \left(x'y'' - y'x''\right) f_z}{\sqrt{f_x^2 + f_y^2 + f_z^2}} . \tag{9.367}$$

9.7.1.5 Geodesic Curvature of Coordinate Curves

On a t^1-curve, one will have $dt^1 \neq 0$ and $dt^2 = 0$. Consequently, $\widehat{\mathbf{a}}_1^{c} = \left(dt^1/ds\right)\mathbf{a}_1$. And this implies that

$$\frac{dt^1}{ds} = \frac{1}{\sqrt{\mathrm{E}_{\mathbf{r}}}} . \tag{9.368}$$

Denoting by $\kappa^{g}_{t^1}$ the geodesic curvature of t^1-curve, one may use the the expression (9.360) to arrive at

$$\boxed{\kappa_{t^1}^{g} = \frac{1}{2E_r^{3/2}\sqrt{E_rG_r - F_r^2}}\left[-F_r\frac{\partial E_r}{\partial t^1} + 2E_r\frac{\partial F_r}{\partial t^1} - E_r\frac{\partial E_r}{\partial t^2}\right] .} \tag{9.369}$$

In a similar manner, consider a t^2-curve for which $dt^1 = 0$ and $dt^2 \neq 0$. As a result,

$$\frac{dt^2}{ds} = \frac{1}{\sqrt{G_r}} . \tag{9.370}$$

The geodesic curvature of $t^1 =$ constant, denoted by $\kappa_{t^2}^{g}$, then takes the form

$$\boxed{\kappa_{t^2}^{g} = \frac{-1}{2G_r^{3/2}\sqrt{E_rG_r - F_r^2}}\left[2G_r\frac{\partial F_r}{\partial t^2} - G_r\frac{\partial G_r}{\partial t^1} - F_r\frac{\partial G_r}{\partial t^2}\right] .} \tag{9.371}$$

Let $\mathcal{S}$ be a regular surface described by an *orthogonal parametrization* $\mathbf{x} = \hat{\mathbf{x}}^{s}\left(t^1, t^2\right)$. This means that the coordinate curves of that surface intersect orthogonally. It can be shown that such a parametrization does exist at all points of a regular surface. And this holds true if and only if the matrix $\left[\, a_{\alpha\beta} \,\right]$ is diagonal, i.e. $a_{12} = \mathbf{a}_1 \cdot \mathbf{a}_2 = F_r = 0$. In this case, the geodesic curvatures (9.369) and (9.371) reduce to

$$\hat{\kappa}_{t^1}^{g} = \frac{-1}{2E_r\sqrt{G_r}}\frac{\partial E_r}{\partial t^2} = \frac{-1}{\sqrt{G_r}}\frac{\partial}{\partial t^2}\left(\log\sqrt{E_r}\right) , \tag{9.372a}$$

$$\hat{\kappa}_{t^2}^{g} = \frac{1}{2G_r\sqrt{E_r}}\frac{\partial G_r}{\partial t^1} = \frac{1}{\sqrt{E_r}}\frac{\partial}{\partial t^1}\left(\log\sqrt{G_r}\right) . \tag{9.372b}$$

In the following, the aim is to characterize the geodesic curvature of a curve for the case in which the coordinate curves of its two-dimensional embedding space are orthogonal.

Let $\mathcal{S}$ be a regular surface with an orthogonal parametrization embedding a regular curve $\mathcal{C}$. Further, let the everywhere nonzero unit tangent vector to that curve, $\widehat{\mathbf{a}}_1^{c}$, makes an angle $\theta\left(t^1(s), t^2(s)\right)$ with $\mathbf{a}_1$. One then has

$$\underbrace{\widehat{\mathbf{a}}_1^{c} = \frac{dt^1}{ds}\mathbf{a}_1 + \frac{dt^2}{ds}\mathbf{a}_2 = \cos\theta\frac{\mathbf{a}_1}{|\mathbf{a}_1|} + \sin\theta\frac{\mathbf{a}_2}{|\mathbf{a}_2|}}_{\text{note that } |\mathbf{a}_1| = \sqrt{E_r}\,,\ |\mathbf{a}_2| = \sqrt{G_r} \text{ and } \mathbf{a}_1\cdot\widehat{\mathbf{a}}_1^{c} = \sqrt{E_r}\cos\theta\,,\ \mathbf{a}_2\cdot\widehat{\mathbf{a}}_1^{c} = \sqrt{G_r}\sin\theta} . \tag{9.373}$$

Consequently,

$$\frac{d\theta}{ds} = \frac{dt^1}{ds}\frac{\partial\theta}{\partial t^1} + \frac{dt^2}{ds}\frac{\partial\theta}{\partial t^2} = \frac{\cos\theta}{\sqrt{E_r}}\frac{\partial\theta}{\partial t^1} + \frac{\sin\theta}{\sqrt{G_r}}\frac{\partial\theta}{\partial t^2} . \tag{9.374}$$

These considerations help establish the well-known relation

$$\boxed{\kappa^{g} = \hat{\kappa}^{g}_{t^1} \cos\theta + \hat{\kappa}^{g}_{t^2} \sin\theta + \frac{d\theta}{ds}\,,} \quad \longleftarrow \text{this is known as the } \textit{Liouville formula} \tag{9.375}$$

because

$$\begin{aligned}
\kappa^{g} &\xlongequal[\text{(9.362) and (9.373)}]{\text{by using}} \frac{1}{\sqrt{\mathrm{E}_\mathbf{r} \mathrm{G}_\mathbf{r}}} \left[\frac{\partial}{\partial t^1} \left(\sqrt{\mathrm{G}_\mathbf{r}} \sin\theta \right) - \frac{\partial}{\partial t^2} \left(\sqrt{\mathrm{E}_\mathbf{r}} \cos\theta \right) \right] \\
&\xlongequal[\text{the product rule}]{\text{by using}} \frac{\sin\theta}{2\mathrm{G}_\mathbf{r}\sqrt{\mathrm{E}_\mathbf{r}}} \frac{\partial \mathrm{G}_\mathbf{r}}{\partial t^1} + \frac{\cos\theta}{\sqrt{\mathrm{E}_\mathbf{r}}} \frac{\partial\theta}{\partial t^1} - \frac{\cos\theta}{2\mathrm{E}_\mathbf{r}\sqrt{\mathrm{G}_\mathbf{r}}} \frac{\partial \mathrm{E}_\mathbf{r}}{\partial t^2} + \frac{\sin\theta}{\sqrt{\mathrm{G}_\mathbf{r}}} \frac{\partial\theta}{\partial t^2} \\
&\xlongequal[\text{(9.372a), (9.372b) and (9.374)}]{\text{by using}} \hat{\kappa}^{g}_{t^2} \sin\theta + \frac{d\theta}{ds} + \hat{\kappa}^{g}_{t^1} \cos\theta\ .
\end{aligned}$$

9.7.2 Geodesics

A curve with identically vanishing geodesic curvature at all points over its whole domain of definition is called a *geodesic*. As a result, every normal section curve is a geodesic. The great circle (or the largest circle) on the surface of a sphere is also a geodesic curve.

9.7.2.1 Geodesic Equations by Geodesic Curvature Vector

To provide differential equations governing a geodesic curve, consider a regular surface curve defined by (9.336), i.e. $\mathbf{x} = \hat{\mathbf{x}}^{s}\left(t^1(s), t^2(s)\right) = \hat{\mathbf{x}}^{c}(s)$. Recall from $(9.337)_3$ that its *velocity vector*, $\widehat{\mathbf{a}}^{c}_{1} = d\mathbf{x}/ds$, at a given point completely lies in the tangent plane of its two-dimensional embedding space. But, guided by (9.340)-(9.342), its *acceleration vector*, $\mathbf{k}^{c} = d^2\mathbf{x}/ds^2$, has both tangential and normal parts. Such a curve is called a geodesic if its acceleration vector is completely in the normal direction, i.e. $\mathbf{k}^{g} = \mathbf{0}$ (technically, this is referred to as a *pregeodesic* and a geodesic curve is one whose parameter is a constant times arc length). This results in the following second-order nonlinear ordinary differential equations

$$\boxed{\frac{d^2 t^\gamma}{ds^2} + \frac{dt^\alpha}{ds} \Gamma^\gamma_{\alpha\beta} \frac{dt^\beta}{ds} = 0\,,} \tag{9.376}$$

called the *geodesic equations*. The geodesic equations can also be derived by using the powerful tool of covariant differentiation. This is described in the following. ∎

9.7.2.2 Geodesic Equations Using Covariant Derivative

Let $\mathcal{C}$ be a naturally represented regular surface curve with everywhere nonzero unit tangent vector $\widehat{\mathbf{a}}_1^c$. This curve is called a geodesic if

$$\boxed{\widehat{\mathbf{a}}_1^c\Big|_{\widehat{\mathbf{a}}_1^c} = \mathbf{0}} \; , \tag{9.377}$$

because

$$\begin{aligned}
\mathbf{0} &\xlongequal[\text{(9.337) and (9.377)}]{\text{by using}} \left(\frac{dt^\alpha}{ds}\mathbf{a}_\alpha\right)\Bigg|_{\left(\frac{dt^\beta}{ds}\mathbf{a}_\beta\right)} \\
&\xlongequal[\text{(9.137b)}]{\text{by using}} \left(\frac{dt^\alpha}{ds}\mathbf{a}_\alpha\right)\Bigg|_{\mathbf{a}_\beta} \frac{dt^\beta}{ds} \\
&\xlongequal[\text{(9.137c)}]{\text{by using}} \left(D_{\mathbf{a}_\beta}\frac{dt^\alpha}{ds}\right)\frac{dt^\beta}{ds}\mathbf{a}_\alpha + \frac{dt^\alpha}{ds}\frac{dt^\beta}{ds}\left(\mathbf{a}_\alpha\big|_{\mathbf{a}_\beta}\right) \\
&\xlongequal[\text{(9.130) and (9.131)}]{\text{by using}} \left(\frac{\partial}{\partial t^\beta}\frac{dt^\alpha}{ds}\right)\frac{dt^\beta}{ds}\mathbf{a}_\alpha + \frac{dt^\alpha}{ds}\frac{dt^\beta}{ds}\Gamma^\gamma_{\alpha\beta}\mathbf{a}_\gamma \\
&\xlongequal[\text{and renaming the dummy indices}]{\text{by applying the chain rule}} \left(\frac{d^2t^\gamma}{ds^2} + \frac{dt^\alpha}{ds}\Gamma^\gamma_{\alpha\beta}\frac{dt^\beta}{ds}\right)\mathbf{a}_\gamma \; ,
\end{aligned}$$

noting that the surface basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly independent. A curve constructed by parallel transporting a unit vector along itself on a curved surface is thus considered a geodesic. ∎

9.7.2.3 Geodesic Equations by Calculus of Variations

Finding the shortest distance between two points is a simple problem within the context of *calculus of variations* (see Washizu [25], Oden and Reddy [26] and Jost and Li-Jost [27]). In what follows, the goal is to obtain the geodesic equations using variational principles for completeness.

Let $\mathcal{C}$ be a regular surface curve parametrized by $t^\alpha(t)$ where t ia any general parameter (not necessarily the arc length). Further, let P and Q be two points on that curve corresponding to $t = a$ and $t = b$, respectively. Consider now a curve $\tilde{\mathcal{C}}$ in a neighborhood of $\mathcal{C}$ joining P and Q. It is given by

$$\tilde{t}^\alpha(t, h) = t^\alpha(t) + h\eta^\alpha(t) \quad \text{such that} \quad \eta^\alpha(a) = \eta^\alpha(b) = 0 \; , \tag{9.378}$$

where η^α denotes an arbitrary differentiable function and $h \in (-\varepsilon, \varepsilon)$ where $\varepsilon > 0$ is sufficiently small. This defines a set of varied curves and $h\eta^\alpha$ presents the *(first) variation* of t^α which is usually denoted by δt^α. Indeed, it is known as an *admissible* *variation* because it satisfies the end conditions $\delta t^\alpha(a) = \delta t^\alpha(b) = 0$.[15] Note that, trivially,

$$\delta t^\alpha = h \lim_{h \to 0} \frac{\tilde{t}^\alpha - t^\alpha}{h} \ . \tag{9.379}$$

The goal is now to minimize

$$s = \int_a^b \Lambda\left(t^\alpha, \dot{t}^\alpha\right) dt \ , \tag{9.380}$$

subject to the conditions

$$t^\alpha(a) = t_a^\alpha \quad , \quad t^\alpha(b) = t_b^\alpha \ , \tag{9.381}$$

where

[15] The new differential operator δ is a linear operator which obeys the standard rules of differentiation such as the product and quotient rules. There is an analogy between the variational operator δ of variational calculus and the differential operator d of differential calculus. To demonstrate this, consider a function F of an independent variable x and two dependent variables $y(x)$ and $\dot{y}(x) := dy(x)/dx$, i.e.

$$F = \hat{F}(x, y, \dot{y}) \quad \text{whose total differential is} \quad dF = \frac{\partial \hat{F}}{\partial x} dx + \frac{\partial \hat{F}}{\partial y} dy + \frac{\partial \hat{F}}{\partial \dot{y}} d\dot{y} \ .$$

As can be seen, d operates on both dependent and independent variables. But, δ only acts on the dependent variables y and $\dot{y}$ because the independent variable x remains constant when y is varied to $y + \delta y$. Thus,

$$\delta F = \frac{\partial \hat{F}}{\partial y} \delta y + \frac{\partial \hat{F}}{\partial \dot{y}} \delta \dot{y} \ .$$

It is important to note that δ and d/dx commute:

$$\delta\left(\frac{dy}{dx}\right) = \frac{d}{dx}(\delta y) \ .$$

Moreover, one can interchange δ with $\int$ as follows:

$$\delta \int_a^b F \, dx = \int_a^b \delta F \, dx \ .$$

Hence,

$$\delta \int_a^b F \, dx = \int_a^b \left[\frac{\partial \hat{F}}{\partial y} \delta y + \frac{\partial \hat{F}}{\partial \dot{y}} \delta \dot{y}\right] dx = \int_a^b \left[\frac{\partial \hat{F}}{\partial y} \delta y + \frac{\partial \hat{F}}{\partial \dot{y}} \frac{d}{dx}(\delta y)\right] dx$$
$$= \int_a^b \left[\frac{\partial \hat{F}}{\partial y} - \frac{d}{dx} \frac{\partial \hat{F}}{\partial \dot{y}}\right] \delta y \, dx + \left[\frac{\partial \hat{F}}{\partial \dot{y}} \delta y\right]_a^b \ .$$

Note that the boundary term vanishes if δy is an admissible function.

$$\Lambda\left(t^{\alpha},\dot{t}^{\alpha}\right)=\sqrt{\dot{t}^{\alpha}a_{\alpha\beta}\dot{t}^{\beta}}=\dot{s} \quad , \quad \bullet:=\frac{d\bullet}{dt} \; , \quad \longleftarrow \text{see (9.242) and (9.285)} \tag{9.382}$$

and t_a^{α}, t_b^{α} are given constants. Notice that the relation (9.380) is an integral of a function of functions. This is called a *functional* which here represents the length of t^{α} from a to b. Accordingly, the length of $\tilde{t}^{\alpha}$ is given by

$$\tilde{s}\left(h\right)=\int_a^b \Lambda\left(t^{\alpha}+h\eta^{\alpha},\dot{t}^{\alpha}+h\dot{\eta}^{\alpha}\right)\,dt \; . \tag{9.383}$$

By Taylor expansion,

$$\tilde{s}\left(h\right)=\int_a^b \Lambda\left(t^{\alpha},\dot{t}^{\alpha}\right)\,dt+h\int_a^b\left(\frac{\partial\Lambda}{\partial t^{\alpha}}\eta^{\alpha}+\frac{\partial\Lambda}{\partial\dot{t}^{\alpha}}\dot{\eta}^{\alpha}\right)\,dt+o\left(h\right) \; . \tag{9.384}$$

By carrying out the integration by parts,

$$\tilde{s}\left(h\right)=s+h\int_a^b\left(\frac{\partial\Lambda}{\partial t^{\alpha}}-\frac{d}{dt}\frac{\partial\Lambda}{\partial\dot{t}^{\alpha}}\right)\eta^{\alpha}\,dt+\underbrace{\left[\frac{\partial\Lambda}{\partial\dot{t}^{\alpha}}\left(h\eta^{\alpha}\right)\right]_a^b}_{=\,0,\ \ \text{according to (9.378)}}+o\left(h\right) \; . \tag{9.385}$$

Consequently, the first variation of s takes the form

$$\delta s=h\lim_{h\to 0}\frac{\tilde{s}-s}{h}=\int_a^b\left(\frac{\partial\Lambda}{\partial t^{\alpha}}-\frac{d}{dt}\frac{\partial\Lambda}{\partial\dot{t}^{\alpha}}\right)\delta t^{\alpha}\,dt \; . \tag{9.386}$$

The **necessary condition** for s to have an extremum is that $\delta s=0$. By the useful *fundamental lemma of variational calculus*,[16] one can arrive at the following set of equations,

$$\frac{\partial\Lambda}{\partial t^{\rho}}-\frac{d}{dt}\frac{\partial\Lambda}{\partial\dot{t}^{\rho}}=0 \; , \tag{9.387}$$

[16] Consider the integral statement

$$\int_a^b G\left(x,y,\dot{y}\right)\delta y\,dx=0 \; ,$$

where $\delta y\left(x\right)=h\eta\left(x\right)$. Here, the arbitrary differentiable function $\eta\left(x\right)$ satisfies $\eta\left(a\right)=\eta\left(b\right)=0$ and $h\in\left(-\varepsilon,\varepsilon\right)$ where the positive number ε is infinitesimal. The Euler equation of this integral expression is $G=0$. This result relies on simple arguments discussed below.

Consider the fact that η is an arbitrary function. This allows one to choose $\eta=G$. Then,

$$\int_a^b G^2\,dx=0 \; .$$

Notice that the integral of a positive function is always positive. And this implies the required result $G=0$. This is known as the fundamental lemma of calculus of variations.

called *Euler equations*. To proceed, one needs to have

$$\frac{\partial \Lambda}{\partial t^{\rho}} = \frac{1}{2\Lambda}\frac{\partial a_{\alpha\beta}}{\partial t^{\rho}}\dot{t}^{\alpha}\dot{t}^{\beta} , \tag{9.388a}$$

$$\frac{\partial \Lambda}{\partial \dot{t}^{\rho}} = \frac{a_{\alpha\beta}}{2\Lambda}\left[\frac{\partial \dot{t}^{\alpha}}{\partial \dot{t}^{\rho}}\dot{t}^{\beta} + \dot{t}^{\alpha}\frac{\partial \dot{t}^{\beta}}{\partial \dot{t}^{\rho}}\right] = \frac{a_{\rho\beta}\dot{t}^{\beta} + a_{\alpha\rho}\dot{t}^{\alpha}}{2\Lambda} = \frac{a_{\rho\beta}\dot{t}^{\beta}}{\Lambda} , \tag{9.388b}$$

and, consequently,

$$\frac{d}{dt}\frac{\partial \Lambda}{\partial \dot{t}^{\rho}} = -\frac{1}{\Lambda^2}\frac{d\Lambda}{dt}a_{\rho\beta}\dot{t}^{\beta} + \frac{1}{\Lambda}\left[\frac{\partial a_{\rho\beta}}{\partial t^{\alpha}}\dot{t}^{\alpha}\dot{t}^{\beta} + a_{\rho\beta}\ddot{t}^{\beta}\right] . \tag{9.389}$$

It follows that

$$\frac{\partial \Lambda}{\partial t^{\rho}} - \frac{d}{dt}\frac{\partial \Lambda}{\partial \dot{t}^{\rho}} = \underbrace{-\frac{1}{2\Lambda}\left[\frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} + \frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} - \frac{\partial a_{\alpha\beta}}{\partial t^{\rho}}\right]\dot{t}^{\alpha}\dot{t}^{\beta}}_{\text{recall from (9.113) that } \frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} - \frac{\partial a_{\alpha\beta}}{\partial t^{\rho}} = 2\Gamma_{\alpha\beta\rho} - \frac{\partial a_{\rho\alpha}}{\partial t^{\beta}}} + \frac{1}{\Lambda^2}\frac{d\Lambda}{dt}a_{\rho\beta}\dot{t}^{\beta} - \frac{1}{\Lambda}a_{\rho\beta}\ddot{t}^{\beta}$$

$$= \underbrace{-\frac{1}{2\Lambda}\left[\frac{\partial a_{\rho\beta}}{\partial t^{\alpha}} + 2\Gamma_{\alpha\beta\rho} - \frac{\partial a_{\rho\alpha}}{\partial t^{\beta}}\right]\dot{t}^{\alpha}\dot{t}^{\beta}}_{= -\frac{1}{\Lambda}\Gamma_{\alpha\beta\rho}\dot{t}^{\alpha}\dot{t}^{\beta},\ \text{owing to } \frac{\partial a_{\rho\beta}}{\partial t^{\alpha}}\dot{t}^{\alpha}\dot{t}^{\beta} = \frac{\partial a_{\rho\alpha}}{\partial t^{\beta}}\dot{t}^{\beta}\dot{t}^{\alpha}} + \frac{1}{\Lambda^2}\frac{d\Lambda}{dt}a_{\rho\beta}\dot{t}^{\beta} - \frac{1}{\Lambda}a_{\rho\beta}\ddot{t}^{\beta} .$$

Thus, one can arrive at

$$\underbrace{a_{\rho\beta}\ddot{t}^{\beta} + \Gamma_{\alpha\beta\rho}\dot{t}^{\alpha}\dot{t}^{\beta} = \frac{1}{\Lambda}\frac{d\Lambda}{dt}a_{\rho\beta}\dot{t}^{\beta}}_{\text{or } a^{\gamma\rho}a_{\rho\beta}\ddot{t}^{\beta} + \Gamma_{\alpha\beta\rho}a^{\rho\gamma}\dot{t}^{\alpha}\dot{t}^{\beta} = \frac{1}{\Lambda}\frac{d\Lambda}{dt}a^{\gamma\rho}a_{\rho\beta}\dot{t}^{\beta}} ,$$

or, using $(9.26)_2$ and (9.111),

$$\boxed{\frac{d^2t^{\gamma}}{dt^2} + \frac{dt^{\alpha}}{dt}\Gamma^{\gamma}_{\alpha\beta}\frac{dt^{\beta}}{dt} = \frac{1}{\Lambda}\frac{d\Lambda}{dt}\frac{dt^{\gamma}}{dt} .} \tag{9.390}$$

At the end, suppose that $t = s$. Then, $\Lambda = 1$ and consequently $d\Lambda/dt = 0$. The geodesic equations (9.376) can thus be achieved. And this means that finding the straightest possible path on a curved space is eventually a variational problem.

As a simple example, consider a **plane** defined by

$$\boxed{\hat{\mathbf{x}}^{s}\left(t^1, t^2\right) = \mathbf{p} + t^1\mathbf{u} + t^2\mathbf{v} ,} \tag{9.391}$$

where $\mathbf{p}$ presents a fixed point on this flat surface and $\mathbf{u}$, $\mathbf{v}$ are constant vectors. The tangent vectors are simply $\mathbf{a}_1 = \mathbf{u}$, $\mathbf{a}_2 = \mathbf{v}$ and consequently all Christoffel symbols identically vanish. The general solution of the geodesic equations thus renders $t^1 = a_0 + a\,s$, $t^2 = b_0 + b\,s$ (where a_0, b_0 (a, b) basically specify the initial position (velocity) of $t^\alpha(s)$). Finally, one ends up with the geodesic curve

$$\hat{\mathbf{x}}^{\mathrm{c}}(s) = (\mathbf{p} + a_0\mathbf{u} + b_0\mathbf{v}) + s\,(a\mathbf{u} + b\mathbf{v}) \ , \tag{9.392}$$

which is nothing but the equation of a **straight line**. As a result, one can always say that the shortest path between two points on a flat surface is a geodesic. But, this does not hold true for a curved surface. As an example, consider two not diametrically opposed points on a great circle of a sphere which provides two curves on that circle. It is well-known that both these curves are geodesics (even the longer one). In this regard, a geodesic may best be described as the *straightest possible path* between two points on a curved space.

9.7.2.4 Geodesic Equations for Implicit Surfaces

Let $\mathcal{S}$ be an **implicit surface** of the form (9.365), i.e. $f(x, y, z) = 0$. Further, let $\mathcal{C}$ be a regular curve described by $\mathbf{x} = (x(s), y(s), z(s))$ on that surface whose geodesic curvature is given in (9.367). In this case, there are **three** differential equations. By setting $\kappa^{\mathrm{g}} = 0$, the first equation renders

$$\left(y'z'' - z'y''\right) f_x + \left(z'x'' - x'z''\right) f_y + \left(x'y'' - y'x''\right) f_z = 0 \ . \tag{9.393}$$

Consider the fact that the tangent vector $\mathbf{x}'$ and the curvature vector $\mathbf{x}''$ of a geodesic curve, according to (9.366a)-(9.366b), should be perpendicular to each other. The second equation thus takes the form

$$x'x'' + y'y'' + z'z'' = 0 \ . \tag{9.394}$$

The third equation is $d^2 f/ds^2 = 0$, i.e.

$$\begin{aligned} f_{xx}x'x' + f_{yy}y'y' + f_{zz}z'z' + 2f_{xy}x'y' + 2f_{xz}x'z' + 2f_{yz}y'z' \\ + f_x x'' + f_y y'' + f_z z'' = 0 \quad \text{where} \quad f_{\bullet\circ} = \frac{\partial^2 f}{\partial \bullet \partial \circ} \ . \end{aligned} \tag{9.395}$$

One can solve the above coupled system of three equations in $\left(x'', y'', z''\right)$. Let

$$\bar{\Omega}_1 = f_{xx}x'x' + f_{yy}y'y' + f_{zz}z'z' + 2f_{xy}x'y' + 2f_{xz}x'z' + 2f_{yz}y'z' \ , \tag{9.396a}$$

$$\bar{\Omega}_2 = \left(x'f_y - y'f_x\right)^2 + \left(x'f_z - z'f_x\right)^2 + \left(z'f_y - y'f_z\right)^2 \neq 0 \ . \tag{9.396b}$$

Then,

$$x'' = \frac{\bar{\Omega}_1}{\bar{\Omega}_2}\left[\left(x'f_y - y'f_x\right)y' + \left(x'f_z - z'f_x\right)z'\right] \ , \tag{9.397a}$$

$$y'' = \frac{\bar{\Omega}_1}{\bar{\Omega}_2} \left[\left(y' f_x - x' f_y \right) x' + \left(y' f_z - z' f_y \right) z' \right] , \tag{9.397b}$$

$$z'' = \frac{\bar{\Omega}_1}{\bar{\Omega}_2} \left[\left(z' f_x - x' f_z \right) x' + \left(z' f_y - y' f_z \right) y' \right] . \tag{9.397c}$$

It is usually difficult to explicitly solve the geodesic equations. However, there are some cases where the problem can be simplified to only computing integrals. This is described in the following.

9.7.2.5 Geodesics on Clairaut Surfaces

Let $\mathcal{S}$ be a regular surface and consider an orthogonal parametrization $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}^3_{\mathrm{p}}$ on that surface for which

$$\boxed{\mathrm{F}_{\mathbf{r}} = 0 \quad , \quad \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} = 0 \quad , \quad \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^2} = 0 \, .} \tag{9.398}$$

This is called *t^1−Clairaut parametrization*. In a similar fashion, a *t^2−Clairaut patch* is a patch $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}^3_{\mathrm{p}}$ on $\mathcal{S}$ whose metric coefficients satisfy

$$\boxed{\mathrm{F}_{\mathbf{r}} = 0 \quad , \quad \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^1} = 0 \quad , \quad \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1} = 0 \, .} \tag{9.399}$$

Attention is now focused on this case for which the Christoffel symbols $\Gamma^{\alpha}_{\beta\gamma}$, given in (9.235a)-(9.235c), render

$$\begin{bmatrix} \Gamma^1_{11} \\ \Gamma^2_{11} \end{bmatrix} = \frac{1}{2\mathrm{G}_{\mathbf{r}}} \begin{bmatrix} 0 \\ -\partial \mathrm{E}_{\mathbf{r}} / \partial t^2 \end{bmatrix} \quad , \quad \begin{bmatrix} \Gamma^1_{12} \\ \Gamma^2_{12} \end{bmatrix} = \frac{1}{2\mathrm{E}_{\mathbf{r}}} \begin{bmatrix} \partial \mathrm{E}_{\mathbf{r}} / \partial t^2 \\ 0 \end{bmatrix} , \tag{9.400a}$$

$$\begin{bmatrix} \Gamma^1_{22} \\ \Gamma^2_{22} \end{bmatrix} = \frac{1}{2\mathrm{G}_{\mathbf{r}}} \begin{bmatrix} 0 \\ \partial \mathrm{G}_{\mathbf{r}} / \partial t^2 \end{bmatrix} . \tag{9.400b}$$

Consequently, the geodesic equations are reduced to

$$\frac{d^2 t^1}{ds^2} + \frac{1}{\mathrm{E}_{\mathbf{r}}} \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} \frac{dt^1}{ds} \frac{dt^2}{ds} = 0 \, , \tag{9.401a}$$

$$\frac{d^2 t^2}{ds^2} - \frac{1}{2\mathrm{G}_{\mathbf{r}}} \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} \frac{dt^1}{ds} \frac{dt^1}{ds} + \frac{1}{2\mathrm{G}_{\mathbf{r}}} \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^2} \frac{dt^2}{ds} \frac{dt^2}{ds} = 0 \, . \tag{9.401b}$$

Let $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}^3_{\mathrm{p}}$ be a t^2-Clairaut patch on a surface $\mathcal{S}$. Introducing $(9.399)_3$ into $(9.372\mathrm{b})_1$ then yields $\kappa^{\mathrm{g}}_{t^2} = 0$. And $\kappa^{\mathrm{g}}_{t^1}$ in $(9.372\mathrm{a})_1$ vanishes if and only if $\partial \mathrm{E}_{\mathbf{r}} / \partial t^2$ vanishes. This leads to the following lemma (Gray et al. [5]).

Lemma A

All t^2-curves of a t^2-Clairaut patch on a surface $\mathcal{S}$ are geodesics and a t^1-curve on that patch is a geodesic if and only if $\partial \mathrm{E_r}/\partial t^2$ vanishes along that curve.

In what follows, the aim is to introduce an important theorem due to Clairaut.

Theorem B. Let $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}_{\mathrm{p}}^3$ be a t^2-Clairaut patch embedding a unit-speed curve $\mathcal{C}$ described by the parametrization $\mathbf{x}\left(t^1(s), t^2(s)\right) = \hat{\mathbf{x}}^{\mathrm{c}}(s)$. Suppose that $\mathcal{C}$ is a geodesic on that patch and let θ be the angle between $d\hat{\mathbf{x}}^{\mathrm{c}}/ds$ and $\partial \mathbf{x}/\partial t^1$. Then,

$$\sqrt{\mathrm{E_r}} \cos\theta := c_{\mathrm{sla}} \ , \qquad \longleftarrow \text{this is known as the } \textit{Clairaut relation} \tag{9.402}$$

is constant along $\mathcal{C}$. This constant, c_{sla}, is known as the *slant* of $\mathcal{C}$. Notice that $\mathcal{C}$ cannot leave the domain where $\mathrm{E_r} \geq c_{\mathrm{sla}}^2$. Moreover, c_{sla} along with $\mathrm{E_r}$ determines the angle θ.

Proof. By (9.399)$_2$ and (9.401a), one will have

$$\frac{d}{ds}\left(\mathrm{E_r}\frac{dt^1}{ds}\right) = \left(0 + \frac{\partial \mathrm{E_r}}{\partial t^2}\frac{dt^2}{ds}\right)\frac{dt^1}{ds} + \mathrm{E_r}\frac{d^2t^1}{ds^2} = 0 \ .$$

As a result, there exists a constant c_{sla} such that

$$\mathrm{E_r}\frac{dt^1}{ds} = c_{\mathrm{sla}} \ . \tag{9.403}$$

Finally,

$$\cos\theta = \underbrace{\left|\frac{\partial \mathbf{x}}{\partial t^1}\right|^{-1} \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} \cdot \frac{\partial \mathbf{x}}{\partial t^1}}_{= \frac{1}{\sqrt{\mathbf{a}_1 \cdot \mathbf{a}_1}}\left(\frac{dt^1}{ds}\mathbf{a}_1 + \frac{dt^2}{ds}\mathbf{a}_2\right) \cdot \mathbf{a}_1} = \underbrace{\frac{\mathrm{E_r}}{\sqrt{\mathrm{E_r}}}\frac{dt^1}{ds}}_{= \sqrt{\mathrm{E_r}}\frac{dt^1}{ds}} \quad \text{or} \quad \sqrt{\mathrm{E_r}}\cos\theta = c_{\mathrm{sla}} \ .$$

The constant c_{sla} helps characterize the derivative of the Gaussian coordinates with respect to the arc length parameter. This is described below.

Lemma B. Let $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}_{\mathrm{p}}^3$ be a t^2-Clairaut patch embedding a unit-speed curve $\mathcal{C}$ parametrically defined by $\mathbf{x}\left(t^1(s), t^2(s)\right) = \hat{\mathbf{x}}^{\mathrm{c}}(s)$. Suppose that $\mathcal{C}$ is a geodesic on that patch. Then,

$$\frac{dt^1}{ds} = \frac{c_{\mathrm{sla}}}{\mathrm{E_r}} \quad , \quad \frac{dt^2}{ds} = \pm\sqrt{\frac{\mathrm{E_r} - c_{\mathrm{sla}}^2}{\mathrm{E_r}\mathrm{G_r}}} \ . \tag{9.404}$$

Conversely, if these relations hold, then $\mathcal{C}$ is a geodesic for either $dt^2/ds = 0$ or $dt^2/ds \neq 0$.

Proof. The relation $(9.404)_1$ is trivially obtained from (9.403). To verify $(9.404)_2$, consider $d\hat{\mathbf{x}}^c/ds = \mathbf{a}_\alpha dt^\alpha/ds$ with $\left|d\hat{\mathbf{x}}^c/ds\right| = 1$ and $\mathrm{F_r} = 0$. Then,

$$\frac{dt^\alpha}{ds}\mathbf{a}_\alpha \cdot \frac{dt^\beta}{ds}\mathbf{a}_\beta = 1 \quad \text{gives} \quad \mathrm{E_r}\frac{dt^1}{ds}\frac{dt^1}{ds} + \mathrm{G_r}\frac{dt^2}{ds}\frac{dt^2}{ds} = 1 \; ,$$

or

$$\frac{c_{\mathrm{sla}}^2}{\mathrm{E_r}} + \mathrm{G_r}\frac{dt^2}{ds}\frac{dt^2}{ds} = 1 \quad \text{or} \quad \left(\frac{dt^2}{ds}\right)^2 = \frac{\mathrm{E_r} - c_{\mathrm{sla}}^2}{\mathrm{E_r G_r}} \; .$$

To prove the converse, suppose that the relation $(9.404)_1$ holds. Then, the geodesic equation (9.401a) is satisfied because

$$\text{from} \quad \frac{d}{ds}\left(\frac{dt^1}{ds} - \frac{c_{\mathrm{sla}}}{\mathrm{E_r}}\right) = 0 \quad \text{one obtains} \quad \underbrace{\frac{d^2t^1}{ds^2} + \frac{c_{\mathrm{sla}}}{\mathrm{E_r^2}}\frac{\partial \mathrm{E_r}}{\partial t^2}\frac{dt^2}{ds} = 0}_{\text{or } \frac{d^2t^1}{ds^2} + \frac{1}{\mathrm{E_r}}\frac{\partial \mathrm{E_r}}{\partial t^2}\frac{dt^1}{ds}\frac{dt^2}{ds} = 0} \; .$$

It only remains to show that $(9.404)_{1-2}$ imply (9.401b). Consider

$$\mathrm{E_r}\frac{dt^1}{ds}\frac{dt^1}{ds} + \mathrm{G_r}\frac{dt^2}{ds}\frac{dt^2}{ds} = 1 \quad \text{whose derivative, that is,} \quad \frac{d}{ds}\left[\mathrm{E_r}\left(\frac{dt^1}{ds}\right)^2 + \mathrm{G_r}\left(\frac{dt^2}{ds}\right)^2\right] = 0 \, ,$$

yields

$$\underbrace{\frac{\partial \mathrm{E_r}}{\partial t^2}\frac{dt^2}{ds}\left(\frac{dt^1}{ds}\right)^2 + 2\mathrm{E_r}\frac{dt^1}{ds}\frac{d^2t^1}{ds^2} + \frac{\partial \mathrm{G_r}}{\partial t^2}\left(\frac{dt^2}{ds}\right)^3 + 2\mathrm{G_r}\frac{dt^2}{ds}\frac{d^2t^2}{ds^2} = 0}_{\text{or } \frac{dt^2}{ds}\left[\frac{\partial \mathrm{E_r}}{\partial t^2}\left(\frac{dt^1}{ds}\right)^2 - 2\frac{\partial \mathrm{E_r}}{\partial t^2}\left(\frac{dt^1}{ds}\right)^2 + \frac{\partial \mathrm{G_r}}{\partial t^2}\left(\frac{dt^2}{ds}\right)^2 + 2\mathrm{G_r}\frac{d^2t^2}{ds^2}\right] = 0} \; .$$

If $dt^2/ds \neq 0$, the relation (9.401b) is then satisfied. And if $dt^2/ds = 0$, the equation $(9.404)_2$ implies that $\mathrm{E_r}$ is a constant. As a result, (9.401b) is again satisfied. And this completes the proof. ➛

The slant of a geodesic curve embedded in a t^2-Clairaut patch also helps characterize dt^1/dt^2 when that curve is defined by $t^1 = t^1\left(t^2\right)$. This is described in the following.

Proposition. Let $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}_\mathrm{p}^3$ be a t^2-Clairaut patch embedding a curve $\mathcal{C}$ of the form

$$\boldsymbol{\alpha}\left(t^2\right) = \mathbf{x}\left(t^1\left(t^2\right), t^2\right) \; . \tag{9.405}$$

This curve is then a geodesic if and only if

$$\frac{dt^1}{dt^2} = \pm c_{\text{sla}} \sqrt{\frac{\mathrm{G_r}}{\mathrm{E_r}\left(\mathrm{E_r} - c_{\text{sla}}^2\right)}} \,. \tag{9.406}$$

Proof. In principle, there exists a unit-speed geodesic curve, $\boldsymbol{\beta}$, which can reparameterize $\boldsymbol{\alpha}$. This is given by

$$\boldsymbol{\beta}(s) = \boldsymbol{\alpha}\left(t^2(s)\right) = \mathbf{x}\left(t^1\left(t^2(s)\right), t^2(s)\right) \,.$$

Then, using (9.404)$_{1-2}$ along with the chain rule of differentiation,

$$\frac{dt^1}{ds} = \frac{dt^1}{dt^2}\frac{dt^2}{ds} \quad \text{gives} \quad \frac{dt^1}{dt^2} = \frac{\frac{c_{\text{sla}}}{\mathrm{E_r}}}{\pm\sqrt{\frac{\mathrm{E_r} - c_{\text{sla}}^2}{\mathrm{E_r}\mathrm{G_r}}}} = \pm c_{\text{sla}} \sqrt{\frac{\mathrm{G_r}}{\mathrm{E_r}\left(\mathrm{E_r} - c_{\text{sla}}^2\right)}} \,.$$

Conversely, suppose that the relation (9.406) holds. One can then define

$$\frac{dt^1}{ds} = \pm c_{\text{sla}} \sqrt{\frac{\mathrm{G_r}}{\mathrm{E_r}\left(\mathrm{E_r} - c_{\text{sla}}^2\right)}} \frac{dt^2}{ds} \quad \text{and} \quad \mathrm{E_r}\frac{dt^1}{ds}\frac{dt^1}{ds} + \mathrm{G_r}\frac{dt^2}{ds}\frac{dt^2}{ds} = 1 \,,$$

which simply imply (9.404)$_{1-2}$. Thus, $\boldsymbol{\beta}$ is a unit-speed geodesic curve. And this completes the proof.

9.7.2.6 Geodesics on Surfaces of Revolution

Consider a **surface of revolution** parametrically described by

$$\underbrace{x_1 = f\left(t^2\right)\cos t^1 \quad , \quad x_2 = f\left(t^2\right)\sin t^1 \quad , \quad x_3 = g\left(t^2\right)}_{\text{these coordinates satisfy the implicit relation } \left(x_1/f\left(t^2\right)\right)^2 + \left(x_2/f\left(t^2\right)\right)^2 = x_3/g\left(t^2\right)} \quad , \tag{9.407}$$

where $0 \le t^1 < 2\pi$ and $a < t^2 < b$. This is illustrated in Fig. 9.26. A t^1-curve on this regular surface $\mathcal{S}$ is known as a *parallel* (note that all parallels on $\mathcal{S}$ are circles). And a t^2-curve on $\mathcal{S}$ is called a *meridian*. The rotation of the parametric curve $t^2 \to \left(f\left(t^2\right), g\left(t^2\right)\right)$ in the x_1x_3-plane about the x_3-axis basically generates $\mathcal{S}$. The function $f\left(t^2\right)$ can be viewed as the radius of a parallel. Thus, it makes sense to assume that such a function is positive. And this makes sure that the meridians do not intersect their axis of rotation.

The parametric equations (9.407) represent a t^2-Clairaut parametrization because

$$\mathbf{a}_1 = -f\left(t^2\right)\sin t^1 \widehat{\mathbf{e}}_1 + f\left(t^2\right)\cos t^1 \widehat{\mathbf{e}}_2 \,, \tag{9.408a}$$

$$\mathbf{a}_2 = \frac{df}{dt^2}\cos t^1 \widehat{\mathbf{e}}_1 + \frac{df}{dt^2}\sin t^1 \widehat{\mathbf{e}}_2 + \frac{dg}{dt^2}\widehat{\mathbf{e}}_3 \,, \tag{9.408b}$$

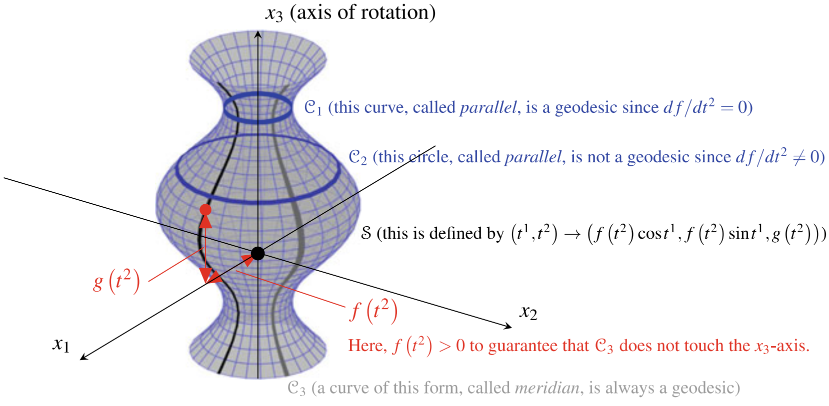


Fig. 9.26 Surface of revolution

and, hence,

$$\mathrm{E}_{\mathbf{r}} = \mathbf{a}_1 \cdot \mathbf{a}_1 = \left[f\left(t^2\right)\right]^2 \quad \Rightarrow \quad \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^1} = 0 \,, \tag{9.409a}$$

$$\mathrm{F}_{\mathbf{r}} = \mathbf{a}_1 \cdot \mathbf{a}_2 = 0 \,, \tag{9.409b}$$

$$\mathrm{G}_{\mathbf{r}} = \mathbf{a}_2 \cdot \mathbf{a}_2 = \left[\frac{df\left(t^2\right)}{dt^2}\right]^2 + \left[\frac{dg\left(t^2\right)}{dt^2}\right]^2 \quad \Rightarrow \quad \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1} = 0 \,. \tag{9.409c}$$

Thus, according to Lemma A on Sect. 9.7.2.5, any meridian on $\mathcal{S}$ is a geodesic and a parallel on that surface is a geodesic if and only if $df/dt^2 = 0$. Guided by (9.406), the following integral

$$\boxed{t^1 = \pm \int \frac{c_{\text{sla}}}{f\sqrt{f^2 - c_{\text{sla}}^2}} \sqrt{\left[\frac{df}{dt^2}\right]^2 + \left[\frac{dg}{dt^2}\right]^2}\, dt^2 + \text{constant} \,,} \tag{9.410}$$

should be computed to find the other geodesics.

The above considerations help infer that the **prime meridian** and **equator** of a sphere are geodesics. They are eventually the great circles on that surface. And by spherical symmetry, any great circle on sphere is also a geodesic.

In what follows, the goal is to find the geodesics on regular surfaces generalizing the surfaces of revolution. ☆

9.7.2.7 Geodesics on Liouville Surfaces

Consider a regular parametrized surface $\mathbf{x} = \hat{\mathbf{x}}^{s}\left(t^{1}, t^{2}\right)$ with the following coefficients of the first fundamental form

$$\boxed{\mathrm{E_r} = \mathrm{G_r} = U\left(t^{1}\right) + V\left(t^{2}\right) \quad , \quad \mathrm{F_r} = 0\,,} \quad \longleftarrow \text{see (9.797)} \tag{9.411}$$

where U and V are smooth functions of only one variable. Such a surface with the quadratic length of the line element

$$\underbrace{ds^{2} = (U + V)\left(dt^{1}dt^{1} + dt^{2}dt^{2}\right)}_{\text{or } \frac{1}{U+V} = \frac{dt^{1}}{ds}\frac{dt^{1}}{ds} + \frac{dt^{2}}{ds}\frac{dt^{2}}{ds}}\,, \tag{9.412}$$

is called the *Liouville surface*. Notice that the orthogonal parametrization (9.411) delivers the Clairaut parametrization (9.398) ((9.399)) when V (U) vanishes.

For the Liouville surface, the mean curvature $(9.257)_2$ and the Gaussian curvature $(9.497\text{a})_2$ become (see Exercise 9.8)

$$\overline{\mathrm{H}} = \frac{\Delta\mathbf{x}\cdot\widehat{\mathbf{n}}}{2\,(U+V)} \qquad \text{where} \qquad \Delta\mathbf{x} = \frac{\partial^{2}\mathbf{x}}{\partial t^{\alpha}\partial t^{\alpha}} = \frac{\partial\mathbf{a}_{\alpha}}{\partial t^{\alpha}}\,, \tag{9.413a}$$

$$\overline{\mathrm{K}} = -\frac{\Delta\log\left(U+V\right)}{2\,(U+V)} \qquad \text{where} \qquad \Delta\log\left(U+V\right) = \frac{\partial^{2}\log\left(U+V\right)}{\partial t^{\alpha}\partial t^{\alpha}}\,. \tag{9.413b}$$

And the geodesic equations (9.376) are reduced to

$$\frac{d^{2}t^{1}}{ds^{2}} + \frac{dU/dt^{1}}{2\,(U+V)}\frac{dt^{1}}{ds}\frac{dt^{1}}{ds} + \frac{dV/dt^{2}}{U+V}\frac{dt^{1}}{ds}\frac{dt^{2}}{ds} - \frac{dU/dt^{1}}{2\,(U+V)}\frac{dt^{2}}{ds}\frac{dt^{2}}{ds} = 0\,, \tag{9.414a}$$

$$\frac{d^{2}t^{2}}{ds^{2}} - \frac{dV/dt^{2}}{2\,(U+V)}\frac{dt^{1}}{ds}\frac{dt^{1}}{ds} + \frac{dU/dt^{1}}{U+V}\frac{dt^{1}}{ds}\frac{dt^{2}}{ds} + \frac{dV/dt^{2}}{2\,(U+V)}\frac{dt^{2}}{ds}\frac{dt^{2}}{ds} = 0\,. \tag{9.414b}$$

Multiplying (9.414a) with $-2V dt^{1}/ds$ and (9.414b) with $+2U dt^{2}/ds$ and then adding the resulting equations will lead to

$$\frac{d}{ds}\left[U\,(U+V)\frac{dt^{2}}{ds}\frac{dt^{2}}{ds} - V\,(U+V)\frac{dt^{1}}{ds}\frac{dt^{1}}{ds}\right] = 0\,. \tag{9.415}$$

Consequently,

$$U\frac{dt^{2}}{ds}\frac{dt^{2}}{ds} - V\frac{dt^{1}}{ds}\frac{dt^{1}}{ds} = \frac{\bar{c}_{1}}{U+V}\,, \tag{9.416}$$

where $\bar{c}_1$ is a constant. From (9.412) and (9.416), it follows that

$$\frac{1}{V+\bar{c}_1}\frac{dt^2}{ds}\frac{dt^2}{ds} = \frac{1}{U-\bar{c}_1}\frac{dt^1}{ds}\frac{dt^1}{ds} \ . \tag{9.417}$$

It is then easy to see that

$$\boxed{\int \frac{dt^2}{\sqrt{V+\bar{c}_1}} = \pm \int \frac{dt^1}{\sqrt{U-\bar{c}_1}} + \bar{c}_2 \ .} \tag{9.418}$$

where $\bar{c}_2$ presents another constant. The constants $\bar{c}_1$ and $\bar{c}_2$ are determined from the initial conditions. ✫

At the end, consider an orthogonal parametrization for a regular surface $\mathcal{S}$, i.e. at each point on that surface $\mathrm{F_r} = 0$ holds. Using (9.235a)-(9.235c) and (9.376), the geodesic equations then render

$$\frac{d^2t^1}{ds^2} + \frac{1}{2\mathrm{E_r}}\frac{\partial \mathrm{E_r}}{\partial t^1}\frac{dt^1}{ds}\frac{dt^1}{ds} + \frac{1}{\mathrm{E_r}}\frac{\partial \mathrm{E_r}}{\partial t^2}\frac{dt^1}{ds}\frac{dt^2}{ds} - \frac{1}{2\mathrm{E_r}}\frac{\partial \mathrm{G_r}}{\partial t^1}\frac{dt^2}{ds}\frac{dt^2}{ds} = 0 \ , \tag{9.419a}$$

$$\frac{d^2t^2}{ds^2} - \frac{1}{2\mathrm{G_r}}\frac{\partial \mathrm{E_r}}{\partial t^2}\frac{dt^1}{ds}\frac{dt^1}{ds} + \frac{1}{\mathrm{G_r}}\frac{\partial \mathrm{G_r}}{\partial t^1}\frac{dt^1}{ds}\frac{dt^2}{ds} + \frac{1}{2\mathrm{G_r}}\frac{\partial \mathrm{G_r}}{\partial t^2}\frac{dt^2}{ds}\frac{dt^2}{ds} = 0 \ . \tag{9.419b}$$

These two second-order differential equations can be rephrased as a system of four first-order differential equations. To do so, let

$$y_1 = t^1 \quad , \quad y_2 = \frac{dt^1}{ds} \quad , \quad y_3 = t^2 \quad , \quad y_4 = \frac{dt^2}{ds} \ , \tag{9.420}$$

and, consequently,

$$y_1' = y_2 \quad , \quad y_2' = \frac{d^2t^1}{ds^2} \quad , \quad y_3' = y_4 \quad , \quad y_4' = \frac{d^2t^2}{ds^2} \ . \tag{9.421}$$

One then arrives at

$$y_1' = y_2 \ , \tag{9.422a}$$

$$y_2' = -\frac{1}{2\mathrm{E_r}}\frac{\partial \mathrm{E_r}}{\partial t^1}y_2^2 - \frac{1}{\mathrm{E_r}}\frac{\partial \mathrm{E_r}}{\partial t^2}y_2y_4 + \frac{1}{2\mathrm{E_r}}\frac{\partial \mathrm{G_r}}{\partial t^1}y_4^2 \ , \tag{9.422b}$$

$$y_3' = y_4 \ , \tag{9.422c}$$

$$y_4' = \frac{1}{2\mathrm{G_r}}\frac{\partial \mathrm{E_r}}{\partial t^2}y_2^2 - \frac{1}{\mathrm{G_r}}\frac{\partial \mathrm{G_r}}{\partial t^1}y_2y_4 - \frac{1}{2\mathrm{G_r}}\frac{\partial \mathrm{G_r}}{\partial t^2}y_4^2 \ . \tag{9.422d}$$

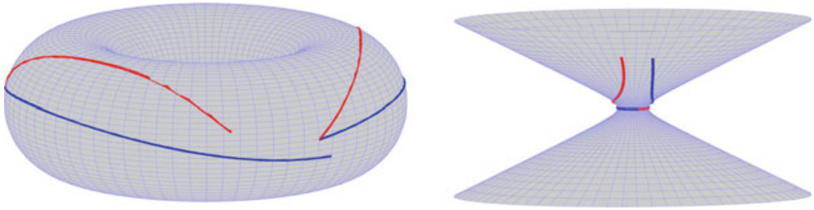


Fig. 9.27 Geodesics on torus and one-sheeted hyperboloid

This system of nonlinear equations can be solved numerically to deliver the geodesic curves. The results for a torus and hyperboloid of one sheet are plotted in Fig. 9.27 (see Exercise 9.15).

9.7.3 *Principal Curvatures*

The *principal curvatures* of a surface at a given point are the maximum and minimum values of the normal curvatures of embedded curves passing through that point. They are a measure of the local shape of the surface under consideration in the neighborhood of a point. The goal here is thus to obtain the extremum of the normal curvature and find the corresponding principal directions. At the end, a well-known relation, called *Euler formula*, will be introduced.

A question that naturally arises is in which directions the normal curvature attains its largest and smallest possible values. And this basically requires solving an eigenvalue problem. To begin with, one needs to rewrite $(9.270)_2$ as

$$\left(\underline{b}_{\alpha\beta} - \kappa^{\mathrm{n}} a_{\alpha\beta} \right) dt^{\alpha} dt^{\beta} = 0 \; . \tag{9.423}$$

Knowing that the necessary condition for κ^{n} to be extremal is $\partial \kappa^{\mathrm{n}} / \partial t^{\gamma} = 0$, one can then arrive at

$$\left(\underline{b}_{\alpha\gamma} - \kappa^{\mathrm{n}} a_{\alpha\gamma} \right) dt^{\alpha} = 0 \;\; \text{or, by index juggling,} \;\; \left(\underline{b}_{\alpha}^{\cdot\beta} - \kappa^{\mathrm{n}} \delta_{\alpha}^{\beta} \right) dt^{\alpha} = 0 \; . \tag{9.424}$$

This shows that the greatest and least possible values of the normal curvature are the eigenvalues of the surface mixed curvature tensor. Recall from $(9.103)_1$ and $(9.104)_1$ that the mean and Gaussian curvatures were the principal scalar invariants of this symmetric second-order tensor. Denoting by κ_1 and κ_2 the maximum and minimum values of κ^{n}, respectively, one can easily write

$$\overline{\mathrm{H}} = \frac{1}{2} \left(\kappa_1 + \kappa_2 \right) \; , \quad \longleftarrow \text{this is two-dimensional version of (4.14a)} \tag{9.425a}$$

$$\overline{\mathrm{K}} = \kappa_1 \kappa_2 \; , \quad \longleftarrow \text{this is two-dimensional version of (4.14c)} \tag{9.425b}$$

and, accordingly,

$$\boxed{\kappa_1 = \overline{\mathrm{H}} + \sqrt{\overline{\mathrm{H}}^2 - \overline{\mathrm{K}}}\,,} \tag{9.426a}$$

$$\boxed{\kappa_2 = \overline{\mathrm{H}} - \sqrt{\overline{\mathrm{H}}^2 - \overline{\mathrm{K}}}\,.} \tag{9.426b}$$

Hint: It is worthwhile to point out that by changing the positive sense of surface according to $\overline{\widehat{\mathbf{n}}} = -\widehat{\mathbf{n}}$, one can deduce that $\overline{\mathrm{II}}_{\mathbf{r}} = -\mathrm{II}_{\mathbf{r}}$ and consequently $\overline{\underline{b}_{\alpha}^{\cdot\beta}} = -\underline{b}_{\alpha}^{\cdot\beta}$. This leads to $\overline{\kappa}_1 = -\kappa_1$ and $\overline{\kappa}_2 = -\kappa_2$. As a result, $\overline{\overline{\mathrm{H}}} = -\overline{\mathrm{H}}$ and $\overline{\overline{\mathrm{K}}} = +\overline{\mathrm{K}}$ in accordance with (9.259).

In the following, it will be shown that the eigenvalues of the surface mixed curvature tensor can only be **real** numbers. ✦

Using $(9.232)_1$ and (9.254), the equation $(9.424)_1$ can be written in the convenient form

$$\begin{bmatrix} \mathrm{e_r} - \kappa^{\mathrm{n}}\mathrm{E_r} & \mathrm{f_r} - \kappa^{\mathrm{n}}\mathrm{F_r} \\ \mathrm{f_r} - \kappa^{\mathrm{n}}\mathrm{F_r} & \mathrm{g_r} - \kappa^{\mathrm{n}}\mathrm{G_r} \end{bmatrix} \begin{bmatrix} dt^1 \\ dt^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} . \tag{9.427}$$

This renders a homogeneous system of linear equations whose nontrivial solution is implied if and only if

$$\det \begin{bmatrix} \mathrm{e_r} - \kappa^{\mathrm{n}}\mathrm{E_r} & \mathrm{f_r} - \kappa^{\mathrm{n}}\mathrm{F_r} \\ \mathrm{f_r} - \kappa^{\mathrm{n}}\mathrm{F_r} & \mathrm{g_r} - \kappa^{\mathrm{n}}\mathrm{G_r} \end{bmatrix} = 0 . \tag{9.428}$$

Expanding this relation yields

$$\left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right) \kappa^{\mathrm{n}\,2} - \left(\mathrm{g_r E_r} - 2\mathrm{f_r F_r} + \mathrm{e_r G_r}\right) \kappa^{\mathrm{n}} + \left(\mathrm{e_r g_r} - \mathrm{f_r^2}\right) = 0 . \tag{9.429}$$

Knowing that $\left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right) > 0$ (and also $\mathrm{E_r}, \mathrm{G_r} > 0$) for a regular surface, the discriminant of the above quadratic equation in κ^{n} according to

$$\begin{aligned} \Delta_{\mathbf{r}} = {} & 4\mathrm{E_r^{-2}} \left(\mathrm{E_r G_r} - \mathrm{F_r^2}\right) \left(\mathrm{f_r E_r} - \mathrm{e_r F_r}\right)^2 \\ & + \left[\mathrm{g_r E_r} - \mathrm{e_r G_r} - 2\mathrm{E_r^{-1} F_r} \left(\mathrm{f_r E_r} - \mathrm{e_r F_r}\right)\right]^2 \geq 0 , \end{aligned} \tag{9.430}$$

reveals the fact that the characteristic values of the mixed curvature tensor belong to the set R of all real numbers. ✦

A point where $\Delta_{\mathbf{r}} = 0$, is called *umbilic*. This special point on the surface is locally **spherical**. And this means that the amount of bending is identical in all directions. A well-known example in this context regards sphere whose every point is umbilic. At umbilical points, there are double roots, i.e. $\kappa_1 = \kappa_2$. From (9.430), it then follows that

$$\boxed{\mathrm{f_r E_r} = \mathrm{e_r F_r} \quad , \quad \mathrm{g_r E_r} = \mathrm{e_r G_r}\ .} \tag{9.431}$$

These equations hold true if and only if there exists a proportionality constant $\bar{\kappa}$ such that

$$\mathrm{f_r} = \bar{\kappa}\,\mathrm{F_r} \quad , \quad \mathrm{e_r} = \bar{\kappa}\,\mathrm{E_r} \quad , \quad \mathrm{g_r} = \bar{\kappa}\,\mathrm{G_r} \; . \tag{9.432}$$

And these expressions immediately imply the extra condition

$$\boxed{\mathrm{g_r F_r} = \mathrm{f_r G_r} \; .} \tag{9.433}$$

A point where $\Delta_\mathrm{r} > 0$, is then referred to as *non-umbilic*. At non-umbilical points, there are two distinct roots. Accordingly, there exists two principal directions which, will be shown that they, are **orthogonal**. But at umbilical points, every direction is a principal direction.

Let $\hat{\lambda} = dt^2/dt^1$. Further, let $\hat{\lambda}_1$ ($\hat{\lambda}_2$) be the principal direction associated with the principal curvature κ_1 (κ_2). Using (9.427), these principal directions are then given by

$$\boxed{\hat{\lambda}_1 = -\frac{\mathrm{e_r} - \kappa_1 \mathrm{E_r}}{\mathrm{f_r} - \kappa_1 \mathrm{F_r}} \quad , \quad \hat{\lambda}_2 = -\frac{\mathrm{e_r} - \kappa_2 \mathrm{E_r}}{\mathrm{f_r} - \kappa_2 \mathrm{F_r}} \; ,} \tag{9.434}$$

or

$$\boxed{\hat{\lambda}_1 = -\frac{\mathrm{f_r} - \kappa_1 \mathrm{F_r}}{\mathrm{g_r} - \kappa_1 \mathrm{G_r}} \quad , \quad \hat{\lambda}_2 = -\frac{\mathrm{f_r} - \kappa_2 \mathrm{F_r}}{\mathrm{g_r} - \kappa_2 \mathrm{G_r}} \; .} \tag{9.435}$$

The goal here is to show that any direction at umbilical points represent a principal direction. It will also be verified that the principal directions at non-umbilical points are orthogonal. ❀

The extreme values of κ^n in (9.354) are obtained by computing $d\kappa^\mathrm{n}/d\hat{\lambda} = 0$. This leads to

$$\left[\mathrm{E_r} + 2\mathrm{F_r}\hat{\lambda} + \mathrm{G_r}\hat{\lambda}^2\right]\left(\mathrm{f_r} + \mathrm{g_r}\hat{\lambda}\right) - \left[\mathrm{e_r} + 2\mathrm{f_r}\hat{\lambda} + \mathrm{g_r}\hat{\lambda}^2\right]\left(\mathrm{F_r} + \mathrm{G_r}\hat{\lambda}\right) = 0 \; .$$

$$\text{or} \quad \left[\left(\mathrm{E_r} + \mathrm{F_r}\hat{\lambda}\right) + \hat{\lambda}\left(\mathrm{F_r} + \mathrm{G_r}\hat{\lambda}\right)\right]\left(\mathrm{f_r} + \mathrm{g_r}\hat{\lambda}\right) - \left[\left(\mathrm{e_r} + \mathrm{f_r}\hat{\lambda}\right) + \hat{\lambda}\left(\mathrm{f_r} + \mathrm{g_r}\hat{\lambda}\right)\right]\left(\mathrm{F_r} + \mathrm{G_r}\hat{\lambda}\right) = 0$$

$$\text{or} \quad \left(\mathrm{E_r} + \mathrm{F_r}\hat{\lambda}\right)\left(\mathrm{f_r} + \mathrm{g_r}\hat{\lambda}\right) = \left(\mathrm{e_r} + \mathrm{f_r}\hat{\lambda}\right)\left(\mathrm{F_r} + \mathrm{G_r}\hat{\lambda}\right)$$

Thus,

$$\boxed{\kappa^\mathrm{n} = \frac{\mathrm{e_r} + 2\mathrm{f_r}\hat{\lambda} + \mathrm{g_r}\hat{\lambda}^2}{\mathrm{E_r} + 2\mathrm{F_r}\hat{\lambda} + \mathrm{G_r}\hat{\lambda}^2} = \frac{\mathrm{f_r} + \mathrm{g_r}\hat{\lambda}}{\mathrm{F_r} + \mathrm{G_r}\hat{\lambda}} = \frac{\mathrm{e_r} + \mathrm{f_r}\hat{\lambda}}{\mathrm{E_r} + \mathrm{F_r}\hat{\lambda}} \; .} \tag{9.436}$$

Notice that one can simply use $(9.436)_{2-3}$ to arrive at (9.427). These relations can also help provide the quadratic equation

$$\left(\mathrm{g_r F_r} - \mathrm{f_r G_r}\right)\hat{\lambda}^2 + \left(\mathrm{g_r E_r} - \mathrm{e_r G_r}\right)\hat{\lambda} + \left(\mathrm{f_r E_r} - \mathrm{e_r F_r}\right) = 0 \; , \tag{9.437}$$

whose discriminant is identical to (9.430). This can also be represented by

$$\det \begin{bmatrix} \hat{\lambda}^2 & -\hat{\lambda} & 1 \\ \mathrm{E_r} & \mathrm{F_r} & \mathrm{G_r} \\ \mathrm{e_r} & \mathrm{f_r} & \mathrm{g_r} \end{bmatrix} = 0 \,. \tag{9.438}$$

Consider an umbilical point at which the expressions $(9.431)_{1-2}$ and (9.433) hold true. One can then deduce that any possible direction is a principal direction.

Consider now a non-umbilical point at which $\kappa_1 \neq \kappa_2$ and $\hat{\lambda}_1 \neq \hat{\lambda}_2$. The sum and product of the roots of (9.437) then render

$$\hat{\lambda}_1 + \hat{\lambda}_2 = -\frac{\mathrm{g_r E_r} - \mathrm{e_r G_r}}{\mathrm{g_r F_r} - \mathrm{f_r G_r}} \quad , \quad \hat{\lambda}_1 \hat{\lambda}_2 = \frac{\mathrm{f_r E_r} - \mathrm{e_r F_r}}{\mathrm{g_r F_r} - \mathrm{f_r G_r}} \,. \tag{9.439}$$

They finally satisfy

$$\mathrm{E_r} + \mathrm{F_r}\left(\hat{\lambda}_1 + \hat{\lambda}_2\right) + \mathrm{G_r}\left(\hat{\lambda}_1 \hat{\lambda}_2\right) = 0 \,. \tag{9.440}$$

And, guided by $(9.246)_2$, this proves the orthogonality of the principal directions at non-umbilical points. ❀

9.7.3.1 Classification of Points on Surface

Recall that the coefficients of the second fundamental form helped realize that a point on the surface can be either elliptic, parabolic, hyperbolic or flat. These points can also be recognized by means of the principal curvatures. With regard to this, a surface is said to have

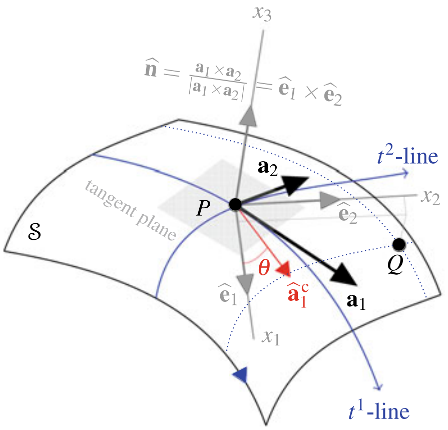
- an **elliptic point** if the principal curvatures are of the same sign. At such points, the Gaussian curvature is positive.
- a **parabolic point** if one of the principal curvatures vanishes. At such points, the Gaussian curvature identically becomes zero.
- a **hyperbolic** (or **saddle**) **point** if the principal curvatures are of different signs. At such points, the Gaussian curvature is negative.
- a **flat** (or **planar**) **point** if both of the principal curvatures become zero. At such points, the Gaussian curvature trivially vanishes.

The principal curvatures are important quantities which help characterize the geometry of a surface near a generic point. This is described below. ★

9.7.3.2 Local Shape of Surface

Let $\mathcal{S}$ be a regular surface. To characterize the local shape of $\mathcal{S}$, consider a point P corresponding to $(0, 0)$ and a sufficiently close point Q corresponding to $\left(t^1, t^2\right)$. Then, second-order (Taylor series) expansion of $\hat{\mathbf{x}}^s$ at $(0, 0)$ represents

Note that the principal curvatures of a regular surface at a generic point control the motion of such a surface perpendicular to its tangent plane at that point.



Let $\mathcal{S}$ be a regular surface with

$\mathbf{a}_\alpha = \frac{\partial \hat{\mathbf{x}}^{\mathrm{s}}}{\partial t^\alpha}$, $\frac{\partial \mathbf{a}_\alpha}{\partial t^\beta} = \Gamma^\gamma_{\alpha\beta}\mathbf{a}_\gamma + \underline{b}_{\alpha\beta}\widehat{\mathbf{n}}$.

Consider a non-umbilical point P at $\hat{\mathbf{x}}^{\mathrm{s}}(0,0)$ and a sufficiently close point Q at $\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1,t^2\right)$. Let $\left(\kappa_\gamma, \widehat{\mathbf{e}}_\gamma\right)$, $\gamma = 1, 2$, be the eigenpairs of $\underline{b}_{\alpha}^{\cdot\beta}$ at P.

A new ambient basis, $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{n}}\}$, can now be established. To characterize the local shape of $\mathcal{S}$ near P consider

$\hat{\mathbf{x}}^{\mathrm{s}}|_Q - \hat{\mathbf{x}}^{\mathrm{s}}|_P = x_1\widehat{\mathbf{e}}_1 + x_2\widehat{\mathbf{e}}_2 - x_3\widehat{\mathbf{n}}$

$\approx \mathbf{a}_\alpha|_P t^\alpha + \left(\Gamma^\gamma_{\alpha\beta}\mathbf{a}_\gamma + \underline{b}_{\alpha\beta}\widehat{\mathbf{n}}\right)\Big|_P \frac{t^\alpha t^\beta}{2}$.

After some manipulations, one can finally arrive at

$x_3 \approx -\frac{\kappa_1 x_1^2 + \kappa_2 x_2^2}{2}$.

Fig. 9.28 Surface local shape

$$\hat{\mathbf{x}}^{\mathrm{s}}\left(0 + t^1, 0 + t^2\right) \approx \hat{\mathbf{x}}^{\mathrm{s}}(0, 0) + \left(\frac{\partial \hat{\mathbf{x}}^{\mathrm{s}}}{\partial t^\alpha}\right)\Bigg|_{(0,0)} t^\alpha + \left(\frac{\partial^2 \hat{\mathbf{x}}^{\mathrm{s}}}{\partial t^\beta \partial t^\alpha}\right)\Bigg|_{(0,0)} \frac{t^\alpha t^\beta}{2} . \tag{9.441}$$

Guided by (9.10)$_1$ and (9.94), one can further write

$$\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right) - \hat{\mathbf{x}}^{\mathrm{s}}(0, 0) \approx \mathbf{a}_\alpha t^\alpha + \left(\Gamma^\gamma_{\alpha\beta}\mathbf{a}_\gamma + \underline{b}_{\alpha\beta}\widehat{\mathbf{n}}\right)\frac{t^\alpha t^\beta}{2} , \tag{9.442}$$

where all functions on the right hand side are taken at P. Suppose that P is a non-umbilical point at which the principal curvatures κ_1, κ_2 are distinct and the principal directions $\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2$ are orthogonal to each other (note that these are eventually the eigenpairs of the symmetric tensor $\underline{b}_{\alpha}^{\cdot\beta}$). This allows one to consider a Cartesian coordinate frame represented by the origin $\hat{\mathbf{x}}^{\mathrm{s}}(0, 0) = \mathbf{o}$ and the basis $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{n}}\}$ where $\widehat{\mathbf{n}}$ denotes the unit normal field to the surface. This has schematically been illustrated in Fig. 9.28. In this frame, the position vector of Q relative to the origin P can be written as

$$\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right) - \mathbf{o} = x_1\widehat{\mathbf{e}}_1 + x_2\widehat{\mathbf{e}}_2 - x_3\widehat{\mathbf{n}} . \tag{9.443}$$

The goal is now to find a functional relationship between the Cartesian components x_1, x_2 and x_3 through the principal curvatures. Note that

$$-x_3 = \left[\hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right) - \mathbf{o}\right] \cdot \widehat{\mathbf{n}} \approx \underline{b}_{\alpha\beta}\frac{t^\alpha t^\beta}{2} . \tag{9.444}$$

Let

$$\left.\begin{array}{l}\hat{\eta}^1_{\ 1} = \mathbf{a}^1 \cdot \widehat{\mathbf{e}}_1 \\ \hat{\eta}^2_{\ 1} = \mathbf{a}^2 \cdot \widehat{\mathbf{e}}_1\end{array}\right\} \quad , \quad \left.\begin{array}{l}\hat{\eta}^1_{\ 2} = \mathbf{a}^1 \cdot \widehat{\mathbf{e}}_2 \\ \hat{\eta}^2_{\ 2} = \mathbf{a}^2 \cdot \widehat{\mathbf{e}}_2\end{array}\right\} , \tag{9.445}$$

and

$$[\,M\,] = \begin{bmatrix} \hat{\eta}^1_{\ 1} & \hat{\eta}^1_{\ 2} \\ \hat{\eta}^2_{\ 1} & \hat{\eta}^2_{\ 2} \end{bmatrix} . \tag{9.446}$$

The principal directions $\widehat{\mathbf{e}}_1$ and $\widehat{\mathbf{e}}_2$ can then be expressed in terms of the surface basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ as

$$\underbrace{\widehat{\mathbf{e}}_1 = \hat{\eta}^1_{\ 1}\mathbf{a}_1 + \hat{\eta}^2_{\ 1}\mathbf{a}_2 \quad , \quad \widehat{\mathbf{e}}_2 = \hat{\eta}^1_{\ 2}\mathbf{a}_1 + \hat{\eta}^2_{\ 2}\mathbf{a}_2}_{\text{note that, e.g., } \widehat{\mathbf{e}}_1 = \mathbf{I}\widehat{\mathbf{e}}_1 = \left(\mathbf{a}_1 \otimes \mathbf{a}^1 + \mathbf{a}_2 \otimes \mathbf{a}^2 + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}\right)\widehat{\mathbf{e}}_1} \quad \text{or} \quad \widehat{\mathbf{e}}_\alpha = \hat{\eta}^\beta_{\ \alpha}\mathbf{a}_\beta . \tag{9.447}$$

This helps rewrite $\widehat{\mathbf{e}}_\alpha \cdot \widehat{\mathbf{e}}_\beta = \delta_{\alpha\beta}$ as

$$\delta_{\alpha\beta} = \hat{\eta}^\sigma_{\ \alpha} a_{\sigma\tau} \hat{\eta}^\tau_{\ \beta} = \begin{bmatrix} \hat{\eta}^1_{\ \alpha} & \hat{\eta}^2_{\ \alpha} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \hat{\eta}^1_{\ \beta} \\ \hat{\eta}^2_{\ \beta} \end{bmatrix} . \tag{9.448}$$

The eigenvalue problem (9.424)$_1$ can now be written as

$$\underline{b}_{\sigma\tau}\, \hat{\eta}^\tau_{\ \beta} = \kappa_\beta a_{\sigma\tau} \hat{\eta}^\tau_{\ \beta} \quad (\beta = 1, 2;\ \text{no sum}) , \tag{9.449}$$

or

$$\hat{\eta}^\sigma_{\ \alpha}\, \underline{b}_{\sigma\tau}\, \hat{\eta}^\tau_{\ \beta} = \kappa_\beta \hat{\eta}^\sigma_{\ \alpha} a_{\sigma\tau} \hat{\eta}^\tau_{\ \beta} = \kappa_\beta \delta_{\alpha\beta} \quad (\beta = 1, 2;\ \text{no sum}) , \tag{9.450}$$

or

$$\begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix} = [\,M\,]^{\mathrm{T}} \begin{bmatrix} \underline{b}_{11} & \underline{b}_{12} \\ \underline{b}_{12} & \underline{b}_{22} \end{bmatrix} [\,M\,] . \tag{9.451}$$

From (9.442), (9.443) and (9.447)$_{1-2}$, one can now arrive at

$$\begin{bmatrix} t^1 \\ t^2 \end{bmatrix} \approx [\,M\,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} , \tag{9.452}$$

where the higher-order terms involving the Christoffel symbols have been neglected. Substituting (9.452) into (9.444)$_2$ leads to

$$-x_3 \approx \frac{1}{2} \begin{bmatrix} t^1 & t^2 \end{bmatrix} \begin{bmatrix} \underline{b}_{11} & \underline{b}_{12} \\ \underline{b}_{12} & \underline{b}_{22} \end{bmatrix} \begin{bmatrix} t^1 \\ t^2 \end{bmatrix} \approx \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} [\,M\,]^{\mathrm{T}} \begin{bmatrix} \underline{b}_{11} & \underline{b}_{12} \\ \underline{b}_{12} & \underline{b}_{22} \end{bmatrix} [\,M\,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ,$$

which, by using (9.451), finally yields

$$\boxed{x_3 \approx -\frac{\kappa_1 x_1^2 + \kappa_2 x_2^2}{2} .} \tag{9.453}$$

One can thus infer that the local shape of a regular surface $\mathcal{S}$ in a neighborhood of a given point P has a **quadratic approximation** characterized by the principal curvatures at that point. ★

Guided by (9.350), one can write

$$\boxed{\kappa_1 = -\left.\widehat{\mathbf{n}}\right|_{\widehat{\mathbf{e}}_1} \cdot \widehat{\mathbf{e}}_1 \quad , \quad \kappa_2 = -\left.\widehat{\mathbf{n}}\right|_{\widehat{\mathbf{e}}_2} \cdot \widehat{\mathbf{e}}_2 \ ,} \tag{9.454}$$

since, for instance,

$$\begin{aligned}\kappa_1 &= -\left.\widehat{\mathbf{n}}\right|_{(\hat{\eta}^\sigma_{\ 1}\mathbf{a}_\sigma)} \cdot \widehat{\mathbf{e}}_1 = -\left.\widehat{\mathbf{n}}\right|_\sigma \hat{\eta}^\sigma_{\ 1} \cdot \widehat{\mathbf{e}}_1 \\ &= -\frac{\partial \widehat{\mathbf{n}}}{\partial t^\sigma} \hat{\eta}^\sigma_{\ 1} \cdot \widehat{\mathbf{e}}_1 = \underline{b}_{\sigma\tau} \hat{\eta}^\sigma_{\ 1} \mathbf{a}^\tau \cdot \hat{\eta}^\gamma_{\ 1} \mathbf{a}_\gamma \\ &= \underline{b}_{\sigma\tau} \hat{\eta}^\sigma_{\ 1} \delta^\tau_\gamma \hat{\eta}^\gamma_{\ 1} = \underline{b}_{\sigma\tau} \hat{\eta}^\sigma_{\ 1} \hat{\eta}^\tau_{\ 1} \ .\end{aligned}$$

In the following, the goal is to express the second fundamental form in the principal directions. This leads to a well-known relation due to Euler [28].

9.7.3.3 Euler Formula

Theorem C (Euler). Let $\mathcal{C}$ be a naturally represented regular surface curve whose nowhere vanishing unit tangent vector at a given point P is denoted by $\widehat{\mathbf{a}}^c_1$. Suppose that P is a non-umbilical point at which the principal curvatures κ_1, κ_2 are distinct and the principal directions $\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2$ are orthogonal to each other. Further, suppose that θ is the oriented angle from $\widehat{\mathbf{e}}_1$ to $\widehat{\mathbf{a}}^c_1$ (Fig. 9.28). Then, the normal curvature κ^n along the direction of $\widehat{\mathbf{a}}^c_1$ takes the form

$$\boxed{\kappa^n = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta \ .} \quad \longleftarrow \text{see (9.872)} \tag{9.455}$$

<u>Proof.</u> Referred to the surface covariant basis vectors $\mathbf{a}_1, \mathbf{a}_2$ and the principal directions $\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2$, the tensorial object $\widehat{\mathbf{a}}^c_1$ can be expressed as

$$\widehat{\mathbf{a}}^c_1 = \frac{dt^\alpha}{ds}\mathbf{a}_\alpha \quad , \quad \widehat{\mathbf{a}}^c_1 = \cos\theta\,\widehat{\mathbf{e}}_1 + \sin\theta\,\widehat{\mathbf{e}}_2 \ . \tag{9.456}$$

Recall from $(9.447)_{1-2}$ that $\widehat{\mathbf{e}}_1 = \hat{\eta}^\beta_{\ 1}\mathbf{a}_\beta$ and $\widehat{\mathbf{e}}_2 = \hat{\eta}^\beta_{\ 2}\mathbf{a}_\beta$. It then follows that

$$\frac{dt^\alpha}{ds} = \widehat{\mathbf{a}}^c_1 \cdot \mathbf{a}^\alpha = \hat{\eta}^\alpha_{\ 1}\cos\theta + \hat{\eta}^\alpha_{\ 2}\sin\theta \ . \tag{9.457}$$

Finally,

$$\begin{aligned}
\kappa^{\text{n}} &\xlongequal[(9.348)]{\text{from}} \underline{b}_{\alpha\beta} \frac{dt^\alpha}{ds} \frac{dt^\beta}{ds} \\
&\xlongequal[(9.457)]{\text{from}} \hat{\eta}^\alpha_1 \, \underline{b}_{\alpha\beta} \hat{\eta}^\beta_1 \cos^2\theta + \hat{\eta}^\alpha_2 \, \underline{b}_{\alpha\beta} \hat{\eta}^\beta_1 \sin\theta \cos\theta \\
&\quad + \hat{\eta}^\alpha_1 \, \underline{b}_{\alpha\beta} \hat{\eta}^\beta_2 \cos\theta \sin\theta + \hat{\eta}^\alpha_2 \, \underline{b}_{\alpha\beta} \hat{\eta}^\beta_2 \sin^2\theta \\
&\xlongequal[(9.450)]{\text{from}} \kappa_1 \cos^2\theta + 0 + 0 + \kappa_2 \sin^2\theta \ .
\end{aligned}$$

The theorem also holds true at an umbilical point where the principal directions are the same, $\kappa_1 = \kappa_2 = \kappa$, and every direction, $\widehat{\mathbf{e}} = \hat{\eta}^\alpha \mathbf{a}_\alpha$, is a principal direction. In this case, $\kappa^{\text{n}} = \kappa \left(\cos^2\theta + \sin^2\theta\right) = \kappa$ considering the fact that

$$\kappa^{\text{n}} = \hat{\eta}^\alpha \, \underline{b}_{\alpha\beta} \hat{\eta}^\beta = \kappa \hat{\eta}^\alpha a_{\alpha\beta} \hat{\eta}^\beta = \kappa \, \widehat{\mathbf{e}} \cdot \widehat{\mathbf{e}} = \kappa \ .$$

And this completes the proof.

Consider a non-umbilical point and suppose that the mean curvature $2\overline{\text{H}} = \kappa_1 + \kappa_2$ vanishes at that point. One then has $\kappa_1 = -\kappa_2$. From (9.455), the condition $\kappa^{\text{n}} = 0$ then leads to $\cos^2\theta = \sin^2\theta$ which results in $\theta_1 = \pi/4\,,\ \theta_2 = 3\pi/4$. Thus, the directions specified by $\widehat{\mathbf{u}} = \cos\theta_1 \, \widehat{\mathbf{e}}_1 + \sin\theta_1 \, \widehat{\mathbf{e}}_2$ and $\widehat{\mathbf{v}} = \cos\theta_2 \, \widehat{\mathbf{e}}_1 + \sin\theta_2 \, \widehat{\mathbf{e}}_2$ are orthogonal (since $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{v}} = 0$). This leads to the following lemma.

Lemma C

There are two orthogonal **asymptotic directions** at a non-umbilical point with vanishing mean curvature. ★

Consider now a pair of orthogonal directions specified by the unit vectors $\widehat{\mathbf{u}} = \cos\theta \, \widehat{\mathbf{e}}_1 + \sin\theta \, \widehat{\mathbf{e}}_2$ and $\widehat{\mathbf{v}} = \cos\left(\theta + \pi/2\right) \widehat{\mathbf{e}}_1 + \sin\left(\theta + \pi/2\right) \widehat{\mathbf{e}}_2$. Denoting by $\kappa^{\text{n}}_{\widehat{\mathbf{u}}}$ ($\kappa^{\text{n}}_{\widehat{\mathbf{v}}}$) the normal curvature in the direction of $\widehat{\mathbf{u}}$ ($\widehat{\mathbf{v}}$), one will have

$$\kappa^{\text{n}}_{\widehat{\mathbf{u}}} = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta \quad , \quad \kappa^{\text{n}}_{\widehat{\mathbf{v}}} = \kappa_1 \sin^2\theta + \kappa_2 \cos^2\theta \ .$$

The quantity

$$\boxed{\kappa^{\text{n}}_{\widehat{\mathbf{u}}} + \kappa^{\text{n}}_{\widehat{\mathbf{v}}} = \kappa_1 + \kappa_2 \ ,}$$

is thus **independent** of θ which is nothing but (twice) the mean curvature. This leads to the following lemma.

Lemma D

The sum of normal curvatures for every pair of orthogonal directions at a given point is **constant**. ★

9.7.4 Lines of Curvature and Geodesic Torsion

A regular curve $\mathcal{C}$ on a regular surface $\mathcal{S}$ is said to be a *line of curvature* if its tangent line at each point is directed along a principal direction of $\mathcal{S}$ at that point. The line of curvature basically satisfies the quadratic equation (9.437), i.e.

$$\underbrace{a_{\alpha\gamma}\varepsilon^{\gamma\delta}\,\underline{b}_{\delta\beta}dt^{\alpha}dt^{\beta} = 0}_{\text{or } \left(a_{11}\,\underline{b}_{21} - a_{12}\,\underline{b}_{11}\right)dt^1dt^1 + \left(a_{11}\,\underline{b}_{22} - a_{22}\,\underline{b}_{11}\right)dt^1dt^2 + \left(a_{21}\,\underline{b}_{22} - a_{22}\,\underline{b}_{12}\right)dt^2dt^2 = 0} \quad . \tag{9.458}$$

In what follows, the goal is to consider the necessary and sufficient conditions for the coordinate curves to be the lines of curvature. ✣

Theorem D. At a non-umbilical point on a regular surface, the parametric curves are the lines of curvature if and only if $a_{12} = \underline{b}_{12} = 0$.

Proof. Let t^1-curve (for which $dt^1 \neq 0$ and $dt^2 = 0$) be a line of curvature. The relation $(9.458)_2$ then implies that

$$a_{11}\,\underline{b}_{12} = a_{12}\,\underline{b}_{11} \quad \text{or} \quad a_{22}a_{11}\,\underline{b}_{12} = a_{22}a_{12}\,\underline{b}_{11} \; . \tag{9.459}$$

In a similar manner, on a t^2-curve,

$$a_{12}\,\underline{b}_{22} = a_{22}\,\underline{b}_{12} \quad \text{or} \quad a_{11}a_{12}\,\underline{b}_{22} = a_{11}a_{22}\,\underline{b}_{12} \; . \tag{9.460}$$

Consequently,

$$a_{12}\left(a_{22}\,\underline{b}_{11} - a_{11}\,\underline{b}_{22}\right) = 0 \quad \text{or, using (9.352)–(9.353),} \quad a_{12}\left(\kappa_1 - \kappa_2\right) = 0 \; . \tag{9.461}$$

One can thus infer that $a_{12} = 0$ (because $\kappa_1 - \kappa_2 = 0$ is not feasible by assumption). A similar procedure can be followed to deduce that $\underline{b}_{12} = 0$.
Conversely, suppose that $a_{12} = \underline{b}_{12} = 0$. The relation $(9.458)_2$ then reduces to

$$\left(a_{11}\,\underline{b}_{22} - a_{22}\,\underline{b}_{11}\right)dt^1dt^2 = 0 \quad \text{or} \quad \left(\kappa_2 - \kappa_1\right)dt^1dt^2 = 0 \; . \tag{9.462}$$

Consider the fact that $\kappa_1 \neq \kappa_2$ at a non-umbilical point. One then infers that either $dt^1 \neq 0, dt^2 = 0$ or $dt^1 = 0, dt^2 \neq 0$. In the former (latter) case, one can get a family of t^1-curves (t^2-curves). Thus, there are two families of curves which meet orthogonally on the surface and satisfy the condition $(9.458)_2$. ✣

Let the coordinate curves on a regular surface be the lines of curvature at a non-umbilical point. The principal curvatures κ_1 and κ_2 then satisfy

$$\boxed{\frac{\partial \kappa_1}{\partial t^2} = \frac{\partial a_{11}}{\partial t^2}\frac{\kappa_2 - \kappa_1}{2a_{11}}} \; , \quad \longleftarrow \text{the proof is given in Exercise 9.16} \tag{9.463a}$$

$$\boxed{\frac{\partial \kappa_2}{\partial t^1} = \frac{\partial a_{22}}{\partial t^1} \frac{\kappa_1 - \kappa_2}{2 a_{22}}} . \tag{9.463b}$$

In the following, the aim is to introduce an important relation characterizing the lines of curvature. ✪

Theorem E (Rodrigues). A regular curve $\mathcal{C}$ on a regular surface $\mathcal{S}$ is a line of curvature if and only if

$$\boxed{d\widehat{\mathbf{n}} + \kappa_p d\mathbf{x} = \mathbf{0}} , \quad \longleftarrow \text{this is known as the } \textit{Rodrigues formula}, \text{ see Exercise 9.17} \tag{9.464}$$

for some scalar κ_p. Here, κ_p presents a principal curvature. As a result, the vectors $d\widehat{\mathbf{n}}$ and $d\mathbf{x}$ will be parallel in every principal direction (see Rodrigues [29]).

Proof. Let $\mathcal{C}$ be a line of curvature on a regular surface $\mathcal{S}$. Then, using $(9.424)_2$, one will have $\underline{b}_{\alpha}^{\cdot\beta} dt^{\alpha} = \kappa_p dt^{\beta}$. With the aid of $(9.10)_1$, $(9.99)_1$ and the chain rule of differentiation, one can obtain

$$\frac{\partial \widehat{\mathbf{n}}}{\partial t^{\alpha}} dt^{\alpha} = -\underline{b}_{\alpha}^{\cdot\beta} dt^{\alpha} \mathbf{a}_{\beta} \quad \text{or} \quad d\widehat{\mathbf{n}} = -\kappa_p \frac{\partial \mathbf{x}}{\partial t^{\beta}} dt^{\beta} \quad \text{or} \quad d\widehat{\mathbf{n}} = -\kappa_p d\mathbf{x} .$$

Conversely, if $d\widehat{\mathbf{n}} = -\kappa_p d\mathbf{x}$ holds at every point of a regular surface curve, one can simply use the Weingarten formulas to arrive at the eigenvalue problem $(9.424)_2$. Thus, the tangent vector to that curve at a point is an eigenvector of $\underline{b}_{\alpha}^{\cdot\beta}$ at that point. ✪

Suppose one is given a line of curvature on a surface. This sometimes helps find the line of curvature on another surface. This is described below. ✳

Theorem F (Joachimstahl). Consider a curve $\mathcal{C}$ with the parametrization $\mathbf{x} = \mathbf{x}^{c}(t)$ which lies on the intersection of two regular surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$. Let $\widehat{\mathbf{n}}_1$ ($\widehat{\mathbf{n}}_2$) be the unit normal field to $\mathcal{S}_1$ ($\mathcal{S}_2$) and $\mathbf{a}_t = \mathbf{x}'$ be the tangent vector to $\mathcal{C}$ where $\bullet' = d\bullet/dt$. Now, suppose that $\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2 = \text{constant}$ along $\mathcal{C}$. Then, $\mathcal{C}$ is a line of curvature on $\mathcal{S}_1$ if and only if it is a line of curvature on $\mathcal{S}_2$ (see Joachimstahl [30]).

Proof. To begin with, consider $\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2 = \text{constant}$ which gives

$$\widehat{\mathbf{n}}_1' \cdot \widehat{\mathbf{n}}_2 + \widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2' = 0 . \tag{9.465}$$

Guided by (9.464), if $\mathcal{C}$ is a line of curvature on $\mathcal{S}_1$ then

$$\widehat{\mathbf{n}}_1' = -\kappa_p \, \mathbf{a}_t , \tag{9.466}$$

where κ_p denotes a principal curvature of $\mathcal{S}_1$. Note that $\mathbf{a}_t \cdot \widehat{\mathbf{n}}_1 = 0$ and $\mathbf{a}_t \cdot \widehat{\mathbf{n}}_2 = 0$. From (9.466), it then follows that $\widehat{\mathbf{n}}_1' \cdot \widehat{\mathbf{n}}_2 = 0$. This result along with (9.465) helps obtain $\widehat{\mathbf{n}}_2' \cdot \widehat{\mathbf{n}}_1 = 0$. Now, consider the fact that $\widehat{\mathbf{n}}_2 \cdot \widehat{\mathbf{n}}_2 = 1$ which immediately yields $\widehat{\mathbf{n}}_2' \cdot \widehat{\mathbf{n}}_2 = 0$. As can be seen, $\widehat{\mathbf{n}}_2'$ and $\mathbf{a}_t$ are orthogonal to both $\widehat{\mathbf{n}}_1$ and $\widehat{\mathbf{n}}_2$. Thus, there exists a scalar κ_q such that

$$\widehat{\mathbf{n}}_2' = -\kappa_q\, \mathbf{a}_t \ . \tag{9.467}$$

And this means that $\mathcal{C}$ is also a line of curvature on $\mathcal{S}_2$. The converse can be proved in a similar manner. ✳

As a consequence of the Joachimstahl theorem, the **meridians** and **parallels** of a surface of revolution $\mathcal{S}_1$ (Fig. 9.26) are lines of curvature.

To show this, consider a plane $\mathcal{S}_2$ passing through the axis of rotation of $\mathcal{S}_1$. There is (at least) one meridian $\mathcal{C}$ which lies on the intersection of $\mathcal{S}_1$ and $\mathcal{S}_2$. Notice that along this curve, one will have $\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{n}}_2 = 0$. Now, consider the fact that all points of a plane are planar at which the principal curvatures identically vanish and the unit normal fields are all the same. This means that all curves embedded in a plane are lines of curvature. Guided by the Joachimstahl theorem, one can thus infer that $\mathcal{C}$ should also be a line of curvature on $\mathcal{S}_1$. Similar arguments can finally be used to show that any parallel of a surface of revolution is a line of curvature.

A line of curvature can be characterized by means of an important attribute of a surface curve called the *geodesic torsion*. It is a measure of the tendency of a regular surface to twist about a curve embedded in that surface. Consider a regular surface $\mathcal{S}$ enveloping a curve $\mathcal{C}$. The geodesic torsion, τ^{g}, of $\mathcal{C}$ at a given point P is basically the torsion of a geodesic curve passing through P in the tangential direction of $\mathcal{C}$ at P (note that there is only one geodesic curve passing through an arbitrary point in a specific direction on a regular surface).

To characterize the geodesic torsion, consider a surface curve $\mathcal{C}$ and let $\widehat{\mathbf{a}}_2^{c} = -\widehat{\mathbf{n}}$ at a given point P. Accordingly, the torsion function in (9.305b) renders

$$\begin{aligned}
\tau^{g} &= \underbrace{-\frac{d\,\widehat{\mathbf{n}}}{ds} \cdot \widehat{\mathbf{a}}_3^{c}}_{\text{note that } \widehat{\mathbf{a}}_3^{c} = \widehat{\mathbf{a}}_1^{c} \times \widehat{\mathbf{a}}_2^{c} \text{ where } \widehat{\mathbf{a}}_1^{c} = d\mathbf{x}/ds} \\
&= \underbrace{\frac{d\,\widehat{\mathbf{n}}}{ds} \cdot \left(\frac{d\mathbf{x}}{ds} \times \widehat{\mathbf{n}} \right)}_{\text{note that } \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})} \\
&= \widehat{\mathbf{n}} \cdot \left(\frac{d\,\widehat{\mathbf{n}}}{ds} \times \frac{d\mathbf{x}}{ds} \right) .
\end{aligned} \tag{9.468}$$

One can further establish

$$\begin{aligned}
\tau^{g} &\xlongequal[\text{with applying the chain rule}]{\text{by using (9.468) along}} \widehat{\mathbf{n}} \cdot \left(\frac{\partial\,\widehat{\mathbf{n}}}{\partial t^{\alpha}} \times \frac{\partial \mathbf{x}}{\partial t^{\beta}} \right) \frac{dt^{\alpha}}{ds} \frac{dt^{\beta}}{ds} \\
&\xlongequal[\text{(9.10) and (9.99)}]{\text{by using}} \widehat{\mathbf{n}} \cdot \left(-\, \underline{b}_{\alpha}^{\cdot\gamma}\, \mathbf{a}_{\gamma} \times \mathbf{a}_{\beta} \right) \frac{dt^{\alpha}}{ds} \frac{dt^{\beta}}{ds}
\end{aligned}$$

$$\overset{\text{by using}}{\underset{(9.57)}{=\!=\!=}} \widehat{\mathbf{n}} \cdot \left(-\underline{b}_{\alpha}^{\cdot\gamma}\sqrt{a}\varepsilon_{\gamma\beta}\widehat{\mathbf{n}}\right)\frac{dt^{\alpha}}{ds}\frac{dt^{\beta}}{ds}$$

$$\overset{\text{by using}}{\underset{(9.12)}{=\!=\!=}} -\sqrt{a}\frac{dt^{\alpha}}{ds}\,\underline{b}_{\alpha}^{\cdot\gamma}\varepsilon_{\gamma\beta}\frac{dt^{\beta}}{ds}\,. \tag{9.469}$$

It should not be difficult now to see that

$$\tau^{\mathrm{g}} = \frac{\left[\mathrm{f_r E_r - e_r F_r}\right]\left(dt^1\right)^2 + \left[\mathrm{g_r E_r - e_r G_r}\right]dt^1 dt^2 + \left[\mathrm{g_r F_r - f_r G_r}\right]\left(dt^2\right)^2}{\sqrt{\mathrm{E_r G_r - F_r^2}}\left[\mathrm{E_r}\left(dt^1\right)^2 + 2\mathrm{F_r}dt^1dt^2 + \mathrm{G_r}\left(dt^2\right)^2\right]}\,. \tag{9.470}$$

From $(9.228)_{1-3}$, $(9.252\text{a})_1$, $(9.252\text{b})_1$, $(9.252\text{c})_1$, $(9.458)_2$ and (9.470), one can conclude that a curve $\mathcal{C}$ is a line of curvature on a regular surface $\mathcal{S}$ if and only if its geodesic torsion vanishes everywhere.

In addition to the above representations, the geodesic torsion admits some other forms. For instance,

$$\tau^{\mathrm{g}} = -\frac{\kappa^{\mathrm{c}}}{\kappa^{\mathrm{g}}}\widehat{\mathbf{a}}_2^{\mathrm{c}} \cdot \frac{d\widehat{\mathbf{n}}}{ds}\,, \tag{9.471}$$

because

$$\left(\widehat{\mathbf{a}}_2^{\mathrm{c}}\right)\cdot\frac{d\widehat{\mathbf{n}}}{ds} \overset{\text{from}}{\underset{(9.299)}{=\!=\!=}} \left(\widehat{\mathbf{a}}_3^{\mathrm{c}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right)\cdot\frac{d\widehat{\mathbf{n}}}{ds}$$

$$\overset{\text{from}}{\underset{(1.73)}{=\!=\!=}} \widehat{\mathbf{a}}_3^{\mathrm{c}} \cdot \left(\widehat{\mathbf{a}}_1^{\mathrm{c}} \times \frac{d\widehat{\mathbf{n}}}{ds}\right) \quad \longleftarrow \text{ note that both } \widehat{\mathbf{a}}_1^{\mathrm{c}} = \frac{d\mathbf{x}}{ds} \text{ and } \frac{d\widehat{\mathbf{n}}}{ds} \text{ lie in the tangent plane}$$

$$\overset{\text{from}}{\underset{(1.49\text{a}) \text{ and } (9.468)}{=\!=\!=\!=\!=\!=}} -\tau^{\mathrm{g}}\widehat{\mathbf{a}}_3^{\mathrm{c}} \cdot \widehat{\mathbf{n}}$$

$$\overset{\text{from}}{\underset{(9.347)}{=\!=\!=}} -\tau^{\mathrm{g}}\sin\phi$$

$$\overset{\text{from}}{\underset{(9.346)}{=\!=\!=}} -\tau^{\mathrm{g}}\frac{\kappa^{\mathrm{g}}}{\kappa^{\mathrm{c}}}\,.$$

Another form is

$$\tau^{\mathrm{g}} = \tau^{\mathrm{c}} + \frac{d\phi}{ds}\,, \quad \longleftarrow \text{ this is known as the } \textit{Bonnet formula} \text{ for the geodesic torsion, see (9.363)} \tag{9.472}$$

since

$$-\left(\sin\phi\right)\frac{d\phi}{ds} \overset{\text{by using (9.347) along}}{\underset{\text{with applying the product rule}}{=\!=\!=\!=\!=\!=\!=}} \frac{d\widehat{\mathbf{a}}_2^{\mathrm{c}}}{ds}\cdot\widehat{\mathbf{n}} + \widehat{\mathbf{a}}_2^{\mathrm{c}}\cdot\frac{d\widehat{\mathbf{n}}}{ds}$$

$$\overset{\text{by using}}{\underset{(9.305\text{b})}{=\!=\!=}} \left(\tau^{\mathrm{c}}\widehat{\mathbf{a}}_3^{\mathrm{c}} - \kappa^{\mathrm{c}}\widehat{\mathbf{a}}_1^{\mathrm{c}}\right)\cdot\widehat{\mathbf{n}} + \widehat{\mathbf{a}}_2^{\mathrm{c}}\cdot\frac{d\widehat{\mathbf{n}}}{ds} \quad \longleftarrow \text{ note that } \left\{\widehat{\mathbf{a}}_1^{\mathrm{c}}, \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}, \widehat{\mathbf{n}}\right\} \text{ is an orthonormal basis}$$

$$\overset{\text{by using}}{\underset{(9.346),\,(9.347) \text{ and } (9.471)}{=\!=\!=\!=\!=\!=\!=}} \left(\sin\phi\right)\tau^{\mathrm{c}} - \left(\sin\phi\right)\tau^{\mathrm{g}}\,.$$

Moreover,

$$\tau^{\mathrm{g}} = \frac{\kappa_2 - \kappa_1}{2} \sin 2\theta \ , \quad \longleftarrow \text{the proof is given in Exercise 9.20} \tag{9.473}$$

where θ denotes the angle between the principal direction $\widehat{\mathbf{e}}_1$ and the unit tangent vector $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ (Fig. 9.28).

At the end, the different curves introduced so far are listed. A parametrized curve, $\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{c}}(s)$,

* is called a **straight line** if its curvature κ^{c} vanishes everywhere.
* is called a **plane curve** if its torsion τ^{c} vanishes everywhere.
* embedded in a surface $\mathcal{S}$ is called an **asymptotic curve** if its normal curvature κ^{n} vanishes everywhere.
* embedded in a surface $\mathcal{S}$ is called a **geodesic curve** if its geodesic curvature κ^{g} vanishes everywhere.
* embedded in a surface $\mathcal{S}$ is called a **line of curvature** if its geodesic torsion τ^{g} vanishes everywhere.

9.7.5 Gaussian Curvature

The *Gaussian curvature* (or *total curvature*) was partially studied within the developments achieved so far. Recall that it was first introduced as the determinant of the surface mixed curvature tensor. Then, it was expressed in terms of the coefficients of the first and second fundamental forms. Finally, it turned out to be the product of the principal curvatures. This geometric object is an important characteristic of a surface. It represents an extension of the curvature of (one-dimensional) curves to (two-dimensional) surfaces. The main goal here is to show that the Gaussian curvature is an **intrinsic** quantity connecting the intrinsic and extrinsic perspectives. This is basically the essence of one of the greatest achievements of Gauss; called *remarkable theorem* (*Theorema Egregium*).

Recall that by prescribing the functions $\kappa^{\mathrm{c}}(s) > 0$ and $\tau^{\mathrm{c}}(s)$, there exists a regular s-parametrized space curve which is unique (up to a rigid body motion). This is the result of the fundamental theorem of space curves which guarantees existence and uniqueness of solutions of the Frenet formulas. The extension of this theorem to the case of curved surfaces requires more consideration. This is described below.

9.7.5.1 Integrability Conditions

Technically, by prescribing the coefficients of the first fundamental form (i.e. $\mathrm{E}_{\mathbf{r}}, \mathrm{G}_{\mathbf{r}}, \mathrm{F}_{\mathbf{r}}$ satisfying $\mathrm{E}_{\mathbf{r}} > 0$, $\mathrm{G}_{\mathbf{r}} > 0$, $\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^2 > 0$), and second fundamental form (i.e. $\mathrm{e}_{\mathbf{r}}, \mathrm{g}_{\mathbf{r}}, \mathrm{f}_{\mathbf{r}}$), one cannot determine a regular surface unless the following conditions

$$\boxed{\frac{\partial^2 \mathbf{a}_\alpha}{\partial t^\delta \partial t^\beta} = \frac{\partial^2 \mathbf{a}_\alpha}{\partial t^\beta \partial t^\delta}}\ , \tag{9.474}$$

called *integrability conditions*, are satisfied. They are also known as the *compatibility equations* of theory of surfaces. Suppose one is given the functions $\mathrm{E_r}$, $\mathrm{G_r}$, $\mathrm{F_r}$ ($\mathrm{E_r} > 0, \mathrm{G_r} > 0, \mathrm{E_r G_r} - \mathrm{F_r^2} > 0$) and $\mathrm{e_r}$, $\mathrm{g_r}$, $\mathrm{f_r}$. The equations of surface according to (9.94) and (9.99) then render 15 coupled **partial** differential equations while there are only 9 unknowns, i.e. the ambient basis $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}$ (note that a vector in the three-dimensional ambient space has apparently three components and also $\partial \mathbf{a}_1 / \partial t^2 = \partial \mathbf{a}_2 / \partial t^1$). In principle, 6 equations - being the integrability conditions - naturally need to be integrated in order to determine the trihedron $\{\mathbf{a}_1, \mathbf{a}_2, \widehat{\mathbf{n}}\}$ (apart from a rigid body motion). In practice, the equations (9.474) will reduce to

$$\boxed{\frac{\partial^2 \mathbf{a}_1}{\partial t^1 \partial t^2} = \frac{\partial^2 \mathbf{a}_1}{\partial t^2 \partial t^1}\ , \quad \frac{\partial^2 \mathbf{a}_2}{\partial t^1 \partial t^2} = \frac{\partial^2 \mathbf{a}_2}{\partial t^2 \partial t^1}}\ . \tag{9.475}$$

These integrability conditions do not allow one to choose the coefficients of the first and second fundamental forms arbitrary. Indeed, they represent a way of restricting such coefficients to avoid compatibility issues. The considerations above are summarized in a theorem called *fundamental theorem of surfaces*. It guaranties existence and uniqueness of solutions of the Gauss and Weingarten formulas assuming that the compatibility equations are satisfied (for a proof of this important theorem see, for example, do Carmo [8]).

In the following, it will be shown that the conditions of integrability lead to two independent equations which are of fundamental importance in differential geometry of surfaces.

9.7.5.2 Gauss Theorema Egregium and Mainardi-Codazzi Equations

Using $(9.92)_2$, (9.94), $(9.99)_1$ and $(9.132)_1$, the left hand side of (9.474) takes the following form

$$\begin{aligned}
\frac{\partial^2 \mathbf{a}_\alpha}{\partial t^\delta \partial t^\beta} &= \frac{\partial \Gamma^\gamma_{\alpha\beta}}{\partial t^\delta} \mathbf{a}_\gamma + \Gamma^\gamma_{\alpha\beta} \left[\Gamma^\rho_{\gamma\delta} \mathbf{a}_\rho + \underline{b}_{\gamma\delta} \widehat{\mathbf{n}} \right] + \frac{\partial\, \underline{b}_{\alpha\beta}}{\partial t^\delta} \widehat{\mathbf{n}} + \underline{b}_{\alpha\beta} \left[- \underline{b}_\delta^{\cdot\rho} \mathbf{a}_\rho \right] \\
&= \left[\frac{\partial \Gamma^\rho_{\alpha\beta}}{\partial t^\delta} + \Gamma^\rho_{\delta\gamma} \Gamma^\gamma_{\beta\alpha} - \underline{b}_{\alpha\beta}\, \underline{b}_\delta^{\cdot\rho} \right] \mathbf{a}_\rho + \left[\frac{\partial\, \underline{b}_{\alpha\beta}}{\partial t^\delta} + \Gamma^\gamma_{\alpha\beta}\, \underline{b}_{\gamma\delta} \right] \widehat{\mathbf{n}}\ .
\end{aligned} \tag{9.476}$$

In a similar manner, the right hand side of (9.474) renders

$$\frac{\partial^2 \mathbf{a}_\alpha}{\partial t^\beta \partial t^\delta} = \left[\frac{\partial \Gamma^\rho_{\alpha\delta}}{\partial t^\beta} + \Gamma^\rho_{\beta\gamma} \Gamma^\gamma_{\delta\alpha} - \underline{b}_{\alpha\delta}\, \underline{b}_\beta^{\cdot\rho} \right] \mathbf{a}_\rho + \left[\frac{\partial\, \underline{b}_{\alpha\delta}}{\partial t^\beta} + \Gamma^\gamma_{\alpha\delta}\, \underline{b}_{\gamma\beta} \right] \widehat{\mathbf{n}}\ . \tag{9.477}$$

Knowing that $\mathbf{a}_\rho \cdot \widehat{\mathbf{n}} = \mathbf{0}$, one then arrives at

$$\frac{\partial \underline{b}_{\alpha\delta}}{\partial t^\beta} - \frac{\partial \underline{b}_{\alpha\beta}}{\partial t^\delta} = \Gamma^\gamma_{\alpha\beta}\,\underline{b}_{\gamma\delta} - \Gamma^\gamma_{\alpha\delta}\,\underline{b}_{\gamma\beta}\ , \tag{9.478a}$$

$$\frac{\partial \Gamma^\rho_{\alpha\delta}}{\partial t^\beta} - \frac{\partial \Gamma^\rho_{\alpha\beta}}{\partial t^\delta} + \Gamma^\rho_{\beta\gamma}\Gamma^\gamma_{\delta\alpha} - \Gamma^\rho_{\delta\gamma}\Gamma^\gamma_{\beta\alpha} = \underline{b}_{\alpha\delta}\,\underline{b}_\beta^{\cdot\rho} - \underline{b}_{\alpha\beta}\,\underline{b}_\delta^{\cdot\rho}\ . \tag{9.478b}$$

By using (9.92)$_2$, (9.95)$_3$, (9.162d) and (9.199), these results can be rephrased as

$$\underline{b}_{\alpha\delta}\big|_\beta = \underline{b}_{\alpha\beta}\big|_\delta\ , \tag{9.479a}$$

$$\underline{\mathbb{R}}^{\rho\,\cdot\,\cdot\,\cdot}_{\cdot\,\alpha\beta\delta} = \underline{b}_{\alpha\delta}\,\underline{b}_\beta^{\cdot\rho} - \underline{b}_{\alpha\beta}\,\underline{b}_\delta^{\cdot\rho}\ . \tag{9.479b}$$

The expressions in (9.478a) or (9.479a) are called *Mainardi-Codazzi equations* (see Mainardi [31] and Codazzi [32]). They show that the object $\underline{b}_{\alpha\beta}\big|_\delta$ is symmetric with respect to its last two indices. And since $\underline{b}_{\alpha\beta}$ is symmetric, one can deduce that $\underline{b}_{\alpha\beta}\big|_\delta$ is totally symmetric with respect to all of its indices. As a result, in practice, the Mainardi-Codazzi formulas can be written as

$$\underline{b}_{11}\big|_2 = \underline{b}_{12}\big|_1 \quad , \quad \underline{b}_{22}\big|_1 = \underline{b}_{12}\big|_2\ . \tag{9.480}$$

The notable expressions in (9.478b) or (9.479b) are called the *Gauss equations of the surface*. By index juggling, it can be written as

$$\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \underline{b}_{\alpha\gamma}\,\underline{b}_{\beta\delta} - \underline{b}_{\alpha\delta}\,\underline{b}_{\beta\gamma}\ . \tag{9.481}$$

Recall that $\underline{\mathbb{R}}_{1212}$ was the only independent component of the surface covariant Riemann-Christoffel curvature tensor. Using (9.217) and (9.481), one then has

$$\underline{\mathbb{R}}_{1212} = \underline{b}_{11}\,\underline{b}_{22} - \underline{b}_{12}\,\underline{b}_{21} \quad \text{or} \quad \underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \left(\underline{b}_{11}\,\underline{b}_{22} - \underline{b}^2_{12}\right)\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta}\ . \tag{9.482}$$

From (9.255)$_2$, (9.258)$_3$ and (9.482)$_{1-2}$, one can finally arrive at

$$\boxed{\overline{\mathrm{K}} = \frac{\underline{\mathbb{R}}_{1212}}{a} \quad \text{or} \quad \underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \left(a\overline{\mathrm{K}}\right)\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta}\ .} \tag{9.483}$$

This celebrated result shows that the Gaussian curvature is an **intrinsic** object since the Riemann-Christoffel tensor can be obtained by means of the coefficients of the first fundamental form and their partial derivatives, see (9.199) and (9.235a)-(9.235c). In other words, although the objects $\underline{b}_{11}$, $\underline{b}_{22}$ and $\underline{b}_{12}$ are defined extrinsically, the combination $\underline{b}_{11}\,\underline{b}_{22} - \underline{b}^2_{12}$ is an intrinsic measure of curvature. That is why the Gaussian curvature can be viewed as an object connecting the intrinsic and extrinsic points of view. This is the essence of what is known as the *Theorema Egregium* (translated as the *remarkable theorem*) proposed by Gauss [33].

The remarkable theorem (9.483) has some alternative forms. The most common formulas for the Gaussian curvature are demonstrated in the following. ■

9.7.5.3 Some Alternative Forms of Gaussian Curvature

From (9.54)$_2$ and (9.483)$_1$, one will have $\underline{\mathbb{R}}_{1212} = \overline{\mathrm{K}} \left(a_{11} a_{22} - a_{12}^2 \right)$ having in mind that $a_{12} = a_{21}$. Then,

$$\boxed{\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \overline{\mathrm{K}} \left(a_{\alpha\gamma} a_{\beta\delta} - a_{\alpha\delta} a_{\beta\gamma} \right) \ .} \tag{9.484}$$

Accordingly, consider two vectors $\mathbf{u} = \underline{u}^\alpha \mathbf{a}_\alpha$, $\mathbf{v} = \underline{v}^\beta \mathbf{a}_\beta$ and let

$$R^\star (\mathbf{u}, \mathbf{v}, \mathbf{u}, \mathbf{v}) = \underline{\mathbb{R}}_{\alpha\beta\gamma\delta} \, \underline{u}^\alpha \, \underline{v}^\beta \, \underline{u}^\gamma \, \underline{v}^\delta \ . \tag{9.485}$$

Then, one can find that this quantity is nothing but the Gaussian curvature times the area squared of the parallelogram defined by those vectors:

$$\underbrace{R^\star (\mathbf{u}, \mathbf{v}, \mathbf{u}, \mathbf{v}) = \overline{\mathrm{K}} \det \begin{bmatrix} \mathbf{u} \cdot \mathbf{u} & \mathbf{u} \cdot \mathbf{v} \\ \mathbf{u} \cdot \mathbf{v} & \mathbf{v} \cdot \mathbf{v} \end{bmatrix}}_{\text{since } \overline{\mathrm{K}} \left(\underline{u}^\alpha a_{\alpha\gamma} \underline{u}^\gamma \underline{v}^\beta a_{\beta\delta} \underline{v}^\delta - \underline{u}^\alpha a_{\alpha\delta} \underline{v}^\delta \underline{u}^\gamma a_{\gamma\beta} \underline{v}^\beta \right) = \overline{\mathrm{K}} \left(a_{\alpha\gamma} a_{\beta\delta} - a_{\alpha\delta} a_{\beta\gamma} \right) \underline{u}^\alpha \underline{v}^\beta \underline{u}^\gamma \underline{v}^\delta} \ . \tag{9.486}$$

One can also establish

$$\boxed{\mathbf{a}_1|_{12} - \mathbf{a}_1|_{21} = \overline{\mathrm{K}} \left(\mathbf{a}_1 \times \mathbf{a}_2 \right) \times \mathbf{a}_1 \ ,} \tag{9.487}$$

because

$$\begin{aligned}
\left(\mathbf{a}_1|_1 \right)\big|_2 - \left(\mathbf{a}_1|_2 \right)\big|_1 \overset{\text{from (9.131)}}{\underset{\text{and (9.137c)}}{=\!=}} & \underbrace{\left(\Gamma^1_{11} \mathbf{a}_1 + \Gamma^2_{11} \mathbf{a}_2 \right)\Big|_2}_{= \frac{\partial \Gamma^1_{11}}{\partial t^2} \mathbf{a}_1 + \Gamma^1_{11} \Gamma^1_{12} \mathbf{a}_1 + \Gamma^1_{11} \Gamma^2_{12} \mathbf{a}_2 + \frac{\partial \Gamma^2_{11}}{\partial t^2} \mathbf{a}_2 + \Gamma^2_{11} \Gamma^1_{22} \mathbf{a}_1 + \Gamma^2_{11} \Gamma^2_{22} \mathbf{a}_2} \\
& - \underbrace{\left(\Gamma^1_{12} \mathbf{a}_1 + \Gamma^2_{12} \mathbf{a}_2 \right)\Big|_1}_{= \frac{\partial \Gamma^1_{12}}{\partial t^1} \mathbf{a}_1 + \Gamma^1_{12} \Gamma^1_{11} \mathbf{a}_1 + \Gamma^1_{12} \Gamma^2_{11} \mathbf{a}_2 + \frac{\partial \Gamma^2_{12}}{\partial t^1} \mathbf{a}_2 + \Gamma^2_{12} \Gamma^1_{12} \mathbf{a}_1 + \Gamma^2_{12} \Gamma^2_{12} \mathbf{a}_2} \\
\overset{\text{from (9.228),}}{\underset{\text{(9.493) and (9.494)}}{=\!=}} & - \underbrace{\left[\frac{\partial \Gamma^1_{12}}{\partial t^1} - \frac{\partial \Gamma^1_{11}}{\partial t^2} + \Gamma^1_{12} \Gamma^2_{12} - \Gamma^1_{22} \Gamma^2_{11} \right] \mathbf{a}_1}_{= \overline{\mathrm{K}} \, \mathrm{F}_\mathbf{r} \, \mathbf{a}_1 = \overline{\mathrm{K}} \, \left(\mathbf{a}_1 \cdot \mathbf{a}_2 \right) \mathbf{a}_1} \\
& + \underbrace{\left[\frac{\partial \Gamma^2_{11}}{\partial t^2} - \frac{\partial \Gamma^2_{12}}{\partial t^1} + \Gamma^1_{11} \Gamma^2_{12} + \Gamma^2_{11} \Gamma^2_{22} - \Gamma^1_{12} \Gamma^2_{11} - \Gamma^2_{12} \Gamma^2_{12} \right] \mathbf{a}_2}_{= \overline{\mathrm{K}} \, \mathrm{E}_\mathbf{r} \, \mathbf{a}_2 = \overline{\mathrm{K}} \, \left(\mathbf{a}_1 \cdot \mathbf{a}_1 \right) \mathbf{a}_2} \\
\overset{\text{from}}{\underset{\text{(1.72)}}{=\!=}} & \ \overline{\mathrm{K}} \left(\mathbf{a}_1 \times \mathbf{a}_2 \right) \times \mathbf{a}_1 \ .
\end{aligned}$$

By means of (9.38) and (9.483)$_2$, it follows that $\underline{\mathbb{R}}_{\alpha\beta\gamma\delta}\varepsilon^{\alpha\beta} = 2\left(a\overline{\mathrm{K}}\right)\varepsilon_{\gamma\delta}$. One then immediately finds out that

$$\overline{\mathrm{K}} = \frac{1}{4a}\,\underline{\mathbb{R}}_{\alpha\beta\gamma\delta}\varepsilon^{\alpha\beta}\varepsilon^{\gamma\delta} \;. \tag{9.488}$$

Moreover, having in mind the identity (9.39)$_1$, the relation (9.483)$_2$ by index juggling renders $\underline{\mathbb{R}}^{\alpha\beta}{}_{..\gamma\delta} = \overline{\mathrm{K}}\varepsilon^{\alpha\beta}\varepsilon_{\gamma\delta}$. Multiplying both sides of this result with δ^{γ}_{α} leads to $\underline{\mathbb{R}}^{\alpha\beta}{}_{..\alpha\delta} = \overline{\mathrm{K}}\varepsilon^{\alpha\beta}\varepsilon_{\alpha\delta}$, taking into account the replacement property of the Kronecker delta. Contracting now β with δ and using (9.38) yields

$$\boxed{\overline{\mathrm{K}} = \frac{1}{2}\,\underline{\mathbb{R}}^{\alpha\beta}{}_{..\alpha\beta} \;,} \tag{9.489}$$

or

$$\boxed{\overline{\mathrm{K}} = \frac{1}{2}\left(\underline{b}^{\alpha}{}_{.\alpha}\,\underline{b}^{\beta}{}_{.\beta} - \underline{b}^{\alpha}{}_{.\beta}\,\underline{b}^{\beta}{}_{.\alpha}\right) \;,} \tag{9.490}$$

which can also be represented by

$$\overline{\mathrm{K}} = \frac{1}{2}\left(\delta^{\alpha}_{\gamma}\delta^{\beta}_{\delta} - \delta^{\alpha}_{\delta}\delta^{\beta}_{\gamma}\right)\underline{b}^{\gamma}{}_{.\alpha}\,\underline{b}^{\delta}{}_{.\beta} \;, \tag{9.491}$$

or

$$\overline{\mathrm{K}} = \frac{1}{2}\varepsilon^{\alpha\beta}\varepsilon_{\gamma\delta}\,\underline{b}^{\gamma}{}_{.\alpha}\,\underline{b}^{\delta}{}_{.\beta} \;. \tag{9.492}$$

Recall that the Riemann-Christoffel tensor was expressed in terms of the Christoffel symbols which themselves were functions of the coefficients of the first fundamental form. This helps provide some elegant equivalents of the Gaussian curvature. For instance, such an important characteristic of the surface can be expressed as

$$\boxed{\overline{\mathrm{K}} = \mathrm{F}_{\mathbf{r}}^{-1}\left[\frac{\partial\Gamma^{1}_{12}}{\partial t^{1}} - \frac{\partial\Gamma^{1}_{11}}{\partial t^{2}} + \Gamma^{1}_{12}\Gamma^{2}_{12} - \Gamma^{1}_{22}\Gamma^{2}_{11}\right] \;,} \quad \longleftarrow \text{the proof is given in Exercise 9.21} \tag{9.493}$$

or

$$\boxed{\overline{\mathrm{K}} = \mathrm{E}_{\mathbf{r}}^{-1}\left[\frac{\partial\Gamma^{2}_{11}}{\partial t^{2}} - \frac{\partial\Gamma^{2}_{12}}{\partial t^{1}} + \Gamma^{1}_{11}\Gamma^{2}_{12} + \Gamma^{2}_{11}\Gamma^{2}_{22} - \Gamma^{1}_{12}\Gamma^{2}_{11} - \Gamma^{2}_{12}\Gamma^{2}_{12}\right] \;,} \tag{9.494}$$

or

$$\boxed{\overline{\mathrm{K}} = \mathrm{G}_{\mathbf{r}}^{-1}\left[\frac{\partial\Gamma^{1}_{22}}{\partial t^{1}} - \frac{\partial\Gamma^{1}_{12}}{\partial t^{2}} + \Gamma^{1}_{11}\Gamma^{1}_{22} + \Gamma^{1}_{12}\Gamma^{2}_{22} - \Gamma^{1}_{22}\Gamma^{2}_{12} - \Gamma^{1}_{12}\Gamma^{1}_{12}\right] \;,} \tag{9.495}$$

or

$$\boxed{\overline{\mathrm{K}} = \mathrm{F}_{\mathbf{r}}^{-1}\left[\frac{\partial\Gamma_{12}^{2}}{\partial t^{2}} - \frac{\partial\Gamma_{22}^{2}}{\partial t^{1}} + \Gamma_{12}^{1}\Gamma_{12}^{2} - \Gamma_{22}^{1}\Gamma_{11}^{2}\right]} \;. \tag{9.496}$$

Furthermore, one can represent the intrinsic formulas

$$\begin{aligned}
\overline{\mathrm{K}} &= \frac{1}{\sqrt{a}}\left\{-\frac{\partial}{\partial t^{1}}\left[\frac{\sqrt{a}}{a_{11}}\Gamma_{12}^{2}\right] + \frac{\partial}{\partial t^{2}}\left[\frac{\sqrt{a}}{a_{11}}\Gamma_{11}^{2}\right]\right\} \quad \longleftarrow \text{this is known as the } \textit{Bieberbach formula} \text{ (see [34])} \\
&= \frac{1}{2\sqrt{\mathrm{E_r G_r - F_r^2}}}\left\{-\frac{\partial}{\partial t^{1}}\left[\frac{\mathrm{E_r}\partial\mathrm{G_r}/\partial t^{1} - \mathrm{F_r}\partial\mathrm{E_r}/\partial t^{2}}{\mathrm{E_r}\sqrt{\mathrm{E_r G_r - F_r^2}}}\right]\right. \quad \nwarrow \text{the proof is given in Exercise 9.22} \\
&\quad \left. + \frac{\partial}{\partial t^{2}}\left[\frac{2\mathrm{E_r}\partial\mathrm{F_r}/\partial t^{1} - \mathrm{E_r}\partial\mathrm{E_r}/\partial t^{2} - \mathrm{F_r}\partial\mathrm{E_r}/\partial t^{1}}{\mathrm{E_r}\sqrt{\mathrm{E_r G_r - F_r^2}}}\right]\right\}\;,
\end{aligned} \tag{9.497a}$$

$$\begin{aligned}
\overline{\mathrm{K}} &= \frac{1}{\sqrt{a}}\left\{\frac{\partial}{\partial t^{1}}\left[\frac{\sqrt{a}}{a_{22}}\Gamma_{22}^{1}\right] - \frac{\partial}{\partial t^{2}}\left[\frac{\sqrt{a}}{a_{22}}\Gamma_{12}^{1}\right]\right\} \quad \longleftarrow \text{this is also known as the } \textit{Bieberbach formula} \\
&= \frac{1}{2\sqrt{\mathrm{E_r G_r - F_r^2}}}\left\{\frac{\partial}{\partial t^{1}}\left[\frac{2\mathrm{G_r}\partial\mathrm{F_r}/\partial t^{2} - \mathrm{G_r}\partial\mathrm{G_r}/\partial t^{1} - \mathrm{F_r}\partial\mathrm{G_r}/\partial t^{2}}{\mathrm{G_r}\sqrt{\mathrm{E_r G_r - F_r^2}}}\right]\right. \\
&\quad \left. - \frac{\partial}{\partial t^{2}}\left[\frac{\mathrm{G_r}\partial\mathrm{E_r}/\partial t^{2} - \mathrm{F_r}\partial\mathrm{G_r}/\partial t^{1}}{\mathrm{G_r}\sqrt{\mathrm{E_r G_r - F_r^2}}}\right]\right\}\;, \quad \nwarrow \text{the proof is given in Exercise 9.22}
\end{aligned} \tag{9.497b}$$

and

$$\overline{\mathrm{K}} = \frac{\overline{\mathrm{K}}^{\star} - \overline{\mathrm{K}}^{*}}{\left(\mathrm{E_r G_r - F_r^2}\right)^{2}}\;, \quad \longleftarrow \text{this is known as the } \textit{Brioschi formula} \text{ (see [35])} \tag{9.498}$$

where

$$\overline{\mathrm{K}}^{\star} = \det\begin{bmatrix} \frac{\partial^2\mathrm{F_r}}{\partial t^1\partial t^2} - \frac{\partial^2\mathrm{E_r}}{2\partial t^2\partial t^2} - \frac{\partial^2\mathrm{G_r}}{2\partial t^1\partial t^1} & \frac{\partial\mathrm{E_r}}{2\partial t^1} & \frac{\partial\mathrm{F_r}}{\partial t^1} - \frac{\partial\mathrm{E_r}}{2\partial t^2} \\ \frac{\partial\mathrm{F_r}}{\partial t^2} - \frac{\partial\mathrm{G_r}}{2\partial t^1} & \mathrm{E_r} & \mathrm{F_r} \\ \frac{\partial\mathrm{G_r}}{2\partial t^2} & \mathrm{F_r} & \mathrm{G_r} \end{bmatrix}\;, \tag{9.499a}$$

$$\overline{\mathrm{K}}^{*} = \det\begin{bmatrix} 0 & \frac{\partial\mathrm{E_r}}{2\partial t^2} & \frac{\partial\mathrm{G_r}}{2\partial t^1} \\ \frac{\partial\mathrm{E_r}}{2\partial t^2} & \mathrm{E_r} & \mathrm{F_r} \\ \frac{\partial\mathrm{G_r}}{2\partial t^1} & \mathrm{F_r} & \mathrm{G_r} \end{bmatrix}\;. \tag{9.499b}$$

See Hartmann [36] and Goldman [37] for further formulas on curvatures of curves and surfaces. ■

Lemma E (Hilbert). Consider a non-umbilical point P on a regular surface $\mathcal{S}$. Suppose that the principal curvature κ_1 (κ_2) has a local maximum (minimum) at P and $\kappa_1 > \kappa_2$. Then, the Gaussian curvature of $\mathcal{S}$ at P is nonpositive.

Proof. Recall that the principal directions at non-umbilical points were orthogonal. Thus, there exists an orthogonal parametrization $\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1, t^2\right)$ in a neighborhood of the non-umbilical point P such that

$$\mathrm{I}_{\mathbf{r}} = \mathrm{E}_{\mathbf{r}} \left(dt^1\right)^2 + \mathrm{G}_{\mathbf{r}} \left(dt^2\right)^2 \quad , \quad \mathrm{II}_{\mathbf{r}} = \mathrm{e}_{\mathbf{r}} \left(dt^1\right)^2 + \mathrm{g}_{\mathbf{r}} \left(dt^2\right)^2 \ . \tag{9.500}$$

Guided by the Theorem D on Sect. 9.7.4, the coordinate curves of this orthogonal patch are the lines of curvature. The relations (9.463a)–(9.463b) thus hold:

$$\frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} = -\frac{2\mathrm{E}_{\mathbf{r}}}{\kappa_1 - \kappa_2} \frac{\partial \kappa_1}{\partial t^2} \quad , \quad \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1} = \frac{2\mathrm{G}_{\mathbf{r}}}{\kappa_1 - \kappa_2} \frac{\partial \kappa_2}{\partial t^1} \ .$$

The fact that P is a critical point of κ_1 implies that $\partial \mathrm{E}_{\mathbf{r}} / \partial t^2$ vanishes. In a similar manner, one can conclude that $\partial \mathrm{G}_{\mathbf{r}} / \partial t^1 = 0$ at that point. Bearing this in mind, the Bieberbach formula $(9.497\mathrm{a})_2$ takes the form

$$\begin{aligned}
\overline{\mathrm{K}} &= -\frac{1}{2\sqrt{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}}}} \left[\frac{\partial}{\partial t^1} \left(\frac{1}{\sqrt{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}}}} \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1} \right) + \frac{\partial}{\partial t^2} \left(\frac{1}{\sqrt{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}}}} \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} \right) \right] \\
&= -\frac{1}{2\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}}} \left[\frac{\partial}{\partial t^1} \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1} + \frac{\partial}{\partial t^2} \frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} \right] \\
&= -\frac{1}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}}} \left[\frac{\mathrm{G}_{\mathbf{r}}}{\kappa_1 - \kappa_2} \frac{\partial^2 \kappa_2}{\partial t^1 \partial t^1} - \frac{\mathrm{E}_{\mathbf{r}}}{\kappa_1 - \kappa_2} \frac{\partial^2 \kappa_1}{\partial t^2 \partial t^2} \right] .
\end{aligned}$$

Recall from (9.19) that $\mathrm{E}_{\mathbf{r}} > 0$ and $\mathrm{G}_{\mathbf{r}} > 0$. Whereas $\kappa_1 - \kappa_2 > 0$, $\partial^2 \kappa_2 / \partial t^1 \partial t^1 \geq 0$ and $\partial^2 \kappa_1 / \partial t^2 \partial t^2 \leq 0$ by assumption. Thus, $\overline{\mathrm{K}} \leq 0$ at P (see Hilbert and Cohn-Vossen [38]).

It is important to note that the following condition

$$\boxed{\frac{\partial^2 \widehat{\mathbf{n}}}{\partial t^\delta \partial t^\beta} = \frac{\partial^2 \widehat{\mathbf{n}}}{\partial t^\beta \partial t^\delta} \ ,} \tag{9.501}$$

was not written in (9.474) since it does not encode any new information. This is described in the following. ✣

The left hand side of (9.501) can be decomposed with respect to the dual basis $\left\{\mathbf{a}^1, \mathbf{a}^2, \widehat{\mathbf{n}}\right\}$ as

$$\begin{aligned}
\frac{\partial^2 \widehat{\mathbf{n}}}{\partial t^\delta \partial t^\beta} &\overset{\text{by using (9.99)}}{=\!=\!=} \frac{\partial}{\partial t^\delta} \left[- \underline{b}_{\beta\alpha} \mathbf{a}^\alpha \right] \\
&\overset{\text{by using (9.157)}}{=\!=\!=} -\frac{\partial\, \underline{b}_{\beta\alpha}}{\partial t^\delta} \mathbf{a}^\alpha - \underline{b}_{\beta\alpha} \frac{\partial \mathbf{a}^\alpha}{\partial t^\delta} \\
&\overset{\text{by using (9.109)}}{=\!=\!=} -\frac{\partial\, \underline{b}_{\beta\alpha}}{\partial t^\delta} \mathbf{a}^\alpha + \underline{b}_{\beta\alpha} \Gamma^\alpha_{\delta\gamma} \mathbf{a}^\gamma - \underline{b}_{\beta\alpha}\, \underline{b}^{\alpha}_{.\delta}\, \widehat{\mathbf{n}}
\end{aligned}$$

$$\xlongequal[\text{the dummy indices}]{\text{by renaming}} -\frac{\partial\, \underline{b}_{\beta\alpha}}{\partial t^{\delta}} \mathbf{a}^{\alpha} + \underline{b}_{\beta\gamma} \Gamma^{\gamma}_{\delta\alpha}\, \mathbf{a}^{\alpha} - \underline{b}_{\beta\alpha}\, \underline{b}^{\alpha}_{.\delta}\, \widehat{\mathbf{n}}$$

$$\xlongequal[\text{(9.95) and (9.100)}]{\text{by using (9.92),}} \left[-\frac{\partial\, \underline{b}_{\alpha\beta}}{\partial t^{\delta}} + \Gamma^{\gamma}_{\alpha\delta}\, \underline{b}_{\gamma\beta} \right] \mathbf{a}^{\alpha} - \underline{b}_{\beta\alpha} a^{\alpha\gamma}\, \underline{b}_{\gamma\delta}\, \widehat{\mathbf{n}} \,. \tag{9.502}$$

In a similar fashion, the right hand side of (9.501) can be expressed with respect to the ambient contravariant basis $\{\mathbf{a}^1, \mathbf{a}^2, \widehat{\mathbf{n}}\}$ via the following relation

$$\underbrace{\frac{\partial^2 \widehat{\mathbf{n}}}{\partial t^{\beta} \partial t^{\delta}} = \left[-\frac{\partial\, \underline{b}_{\alpha\delta}}{\partial t^{\beta}} + \Gamma^{\gamma}_{\alpha\beta}\, \underline{b}_{\gamma\delta} \right] \mathbf{a}^{\alpha} - \underline{b}_{\delta\alpha} a^{\alpha\gamma}\, \underline{b}_{\gamma\beta}\, \widehat{\mathbf{n}}}_{\text{note that } \underline{b}_{\delta\alpha} a^{\alpha\gamma}\, \underline{b}_{\gamma\beta} = \underline{b}_{\beta\gamma} a^{\gamma\alpha}\, \underline{b}_{\alpha\delta} = \underline{b}_{\beta\alpha} a^{\alpha\gamma}\, \underline{b}_{\gamma\delta}} \,. \tag{9.503}$$

The results (9.502) and (9.503) then clearly deliver the Mainardi-Codazzi equations introduced in (9.478a). This may be viewed as an alternative derivation of these important relations. ✜

Hint: Notice that (9.501) basically renders the commutative property of the partial derivative for the unit normal vector to the surface. Recall that the surface unit normal field was an ambient object for which covariant differentiation reduced to partial differentiation, see $(9.185)_1$. As a result, the surface covariant derivative should also commute for this object with extrinsic attribute, that is,

$$\widehat{\mathbf{n}}|_{\beta\delta} = \widehat{\mathbf{n}}|_{\delta\beta} \ , \tag{9.504}$$

where

$$\widehat{\mathbf{n}}|_{\beta\delta} \xlongequal[\text{(9.179)}]{\text{from}} \left[\left(\underline{\widehat{n}}^{\,i}\Big|_{\beta} \right) \mathbf{g}_i + \underline{\widehat{n}}^{\,i} \left(\mathbf{g}_i|_{\beta} \right) \right]\Big|_{\delta}$$

$$\xlongequal[\text{(9.180) and (9.181)}]{\text{from}} \left[\underline{\widehat{n}}^{\,i}\Big|_j\, \overline{Z}^{\,j}_{\beta}\, \mathbf{g}_i \right]\Big|_{\delta}$$

$$\xlongequal[\text{(9.47) and (9.188b)}]{\text{from}} \underline{\widehat{n}}^{\,i}\Big|_{jk}\, \overline{Z}^{\,j}_{\beta} \overline{Z}^{\,k}_{\delta}\, \mathbf{g}_i + \underline{\widehat{n}}^{\,i}\Big|_j \left[\mathbf{g}^j \cdot \mathbf{a}_{\beta} \right]\Big|_{\delta}\, \mathbf{g}_i$$

$$\xlongequal[\text{(9.177a) and (9.183)}]{\text{from}} \underline{\widehat{n}}^{\,i}\Big|_{jk}\, \overline{Z}^{\,j}_{\beta} \overline{Z}^{\,k}_{\delta}\, \mathbf{g}_i + \underline{\widehat{n}}^{\,i}\Big|_j \left[\mathbf{g}^j \cdot \underline{b}_{\beta\delta} \widehat{\mathbf{n}} \right] \mathbf{g}_i$$

$$\xlongequal[\text{(5.65c)}]{\text{from}} \left(\underline{\widehat{n}}^{\,i}\Big|_{jk}\, \overline{Z}^{\,j}_{\beta} \overline{Z}^{\,k}_{\delta} + \underline{\widehat{n}}^{\,i}\Big|_j\, \underline{\widehat{n}}^{\,j}\, \underline{b}_{\beta\delta} \right) \mathbf{g}_i \ , \tag{9.505}$$

having in mind that any ambient vector $\underline{\widehat{n}}^{\,i}$ in the Euclidean space satisfies the property $\underline{\widehat{n}}^{\,i}\big|_{jk} = \underline{\widehat{n}}^{\,i}\big|_{kj}$. Consequently, the Mainardi-Codazzi relations in (9.479a) can be achieved once again since

$$\widehat{\mathbf{n}}|_{\beta\delta} \xlongequal[\text{(9.185)}]{\text{from}} \left[-\underline{b}_{\beta\alpha} \mathbf{a}^{\alpha} \right]\Big|_{\delta}$$

$$\xlongequal[\text{(9.176b)}]{\text{from}} \left[-\underline{b}_{\beta\alpha}\Big|_{\delta} \right] \mathbf{a}^{\alpha} - \underline{b}_{\beta\alpha} \left[\mathbf{a}^{\alpha}|_{\delta} \right]$$

$$\xrightarrow[\text{(9.95) and (9.177b)}]{\text{from}} \left[-\ \underline{b}_{\alpha\beta}\big|_{\delta}\right]\mathbf{a}^{\alpha} - \underline{b}_{\beta\alpha}\left[\underline{b}^{\alpha}_{.\delta}\,\widehat{\mathbf{n}}\right]$$

$$\xrightarrow[\text{(9.100)}]{\text{from}} \left[-\ \underline{b}_{\alpha\beta}\big|_{\delta}\right]\mathbf{a}^{\alpha} - \underline{b}_{\beta\alpha}a^{\alpha\gamma}\ \underline{b}_{\gamma\delta}\,\widehat{\mathbf{n}}\,, \tag{9.506}$$

and, in a similar fashion,

$$\underbrace{\widehat{\mathbf{n}}|_{\delta\beta} = \left[-\ \underline{b}_{\alpha\delta}\big|_{\beta}\right]\mathbf{a}^{\alpha} - \underline{b}_{\delta\alpha}a^{\alpha\gamma}\ \underline{b}_{\gamma\beta}\,\widehat{\mathbf{n}}}_{\text{note that }\ \underline{b}_{\delta\alpha}a^{\alpha\gamma}\ \underline{b}_{\gamma\beta} = \underline{b}_{\beta\gamma}a^{\gamma\alpha}\ \underline{b}_{\alpha\delta} = \underline{b}_{\beta\alpha}a^{\alpha\gamma}\ \underline{b}_{\gamma\delta}}\ . \tag{9.507}$$

Hint: The ambitious reader may want to obtain the Gauss and Mainardi-Codazzi equations of the surface in an alternative way. This relies on noncommutativity of the surface covariant differentiation for the surface covariant basis vectors. It follows that

$$\begin{aligned}
\underline{\mathbb{R}}^{\alpha}_{\ .\,\beta\gamma\delta}\ \mathbf{a}_{\alpha} &\xrightarrow[\text{(9.201b)}]{\text{from}} \mathbf{a}_{\beta}\big|_{\gamma\delta} - \mathbf{a}_{\beta}\big|_{\delta\gamma} \\
&\xrightarrow[\text{(9.177a)}]{\text{from}} \left[\underline{b}_{\beta\gamma}\,\widehat{\mathbf{n}}\right]\big|_{\delta} - \left[\underline{b}_{\beta\delta}\,\widehat{\mathbf{n}}\right]\big|_{\gamma} \\
&\xrightarrow[\text{the product rule}]{\text{from (9.185) and}} \underline{b}_{\beta\gamma}\big|_{\delta}\,\widehat{\mathbf{n}} + \underline{b}_{\beta\gamma}\left[-\,\underline{b}_{\delta\theta}\mathbf{a}^{\theta}\right] - \underline{b}_{\beta\delta}\big|_{\gamma}\,\widehat{\mathbf{n}} - \underline{b}_{\beta\delta}\left[-\,\underline{b}_{\gamma\theta}\mathbf{a}^{\theta}\right] \\
&\xrightarrow[\text{(9.28)}]{\text{from}} \left[\underline{b}_{\beta\delta}\,\underline{b}_{\gamma\theta} - \underline{b}_{\beta\gamma}\,\underline{b}_{\delta\theta}\right]a^{\theta\alpha}\mathbf{a}_{\alpha} + \left[\underline{b}_{\beta\gamma}\big|_{\delta} - \underline{b}_{\beta\delta}\big|_{\gamma}\right]\widehat{\mathbf{n}}\,.
\end{aligned}$$

Considering the fact that the three vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\widehat{\mathbf{n}}$ are linearly independent, one can deduce that

$$\underline{b}_{\beta\gamma}\big|_{\delta} = \underline{b}_{\beta\delta}\big|_{\gamma} \quad \text{and} \quad \underline{\mathbb{R}}^{\alpha}_{\ .\,\beta\gamma\delta} = a^{\alpha\theta}\left[\underline{b}_{\theta\gamma}\,\underline{b}_{\beta\delta} - \underline{b}_{\theta\delta}\,\underline{b}_{\beta\gamma}\right]\,,$$

or $\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \underline{b}_{\alpha\gamma}\,\underline{b}_{\beta\delta} - \underline{b}_{\alpha\delta}\,\underline{b}_{\beta\gamma}$, see (9.481).

9.7.6 Ricci Curvature Tensor and Scalar

The *Ricci curvature tensor* (or simply the *Ricci tensor*) and *Ricci scalar* (or *curvature scalar* or *curvature invariant*) are of crucial importance in general relativity. The Ricci tensor represents gravity in the general theory of relativity. It measures how the volume of a region changes when such a volume is parallel transported along geodesics in a curved space. The Ricci scalar is the trace of the Ricci tensor. When this quantity is positive (negative), the volume shrinks (expands). It basically describes how the volume of a small ball in curved space differs from that of standard ball in Euclidean space. These two objects construct the so-called *Einstein tensor* appearing in *Einstein's gravitational field equations* and basically belong to the intrinsic geometry of a manifold. See Günther and Müller [39] and Pais [40] for further considerations.

9.7.6.1 Ricci Curvature Tensor

The **first-kind** Ricci tensor is a **covariant** second-order tensor obtained by contracting the single contravariant index of the mixed Riemann-Christoffel curvature tensor with its second covariant index:

$$\begin{aligned}\underline{R}_{\beta\delta} &= \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\alpha\delta}^{\;\;\cdot\;\cdot\;\cdot} \\ &\xlongequal[\text{(9.92) and (9.199)}]{\text{from}} \frac{\partial \Gamma^{\alpha}_{\beta\delta}}{\partial t^{\alpha}} - \frac{\partial \Gamma^{\alpha}_{\beta\alpha}}{\partial t^{\delta}} + \Gamma^{\theta}_{\beta\delta}\Gamma^{\alpha}_{\theta\alpha} - \Gamma^{\theta}_{\beta\alpha}\Gamma^{\alpha}_{\theta\delta} \,. \end{aligned} \tag{9.508}$$

The first-kind Ricci tensor may also be written as

$$\begin{aligned}\underline{R}_{\beta\delta} &\xlongequal[\text{(9.115) and (9.508)}]{\text{by using}} \frac{\partial \Gamma^{\alpha}_{\beta\delta}}{\partial t^{\alpha}} - \frac{\partial\left[\partial \ln\sqrt{a}\,\right]}{\partial t^{\delta}\partial t^{\beta}} + \Gamma^{\theta}_{\beta\delta}\frac{\partial \ln\sqrt{a}}{\partial t^{\theta}} - \Gamma^{\theta}_{\beta\alpha}\Gamma^{\alpha}_{\theta\delta} \\ &\xlongequal[\text{the dummy indices}]{\text{by renaming}} \frac{1}{\sqrt{a}}\left(\sqrt{a}\frac{\partial \Gamma^{\alpha}_{\beta\delta}}{\partial t^{\alpha}} + \Gamma^{\alpha}_{\beta\delta}\frac{\partial\sqrt{a}}{\partial t^{\alpha}}\right) - \frac{\partial^2\left[\ln\sqrt{a}\,\right]}{\partial t^{\delta}\partial t^{\beta}} - \Gamma^{\alpha}_{\beta\theta}\Gamma^{\theta}_{\alpha\delta} \\ &\xlongequal[\text{the product rule}]{\text{by using}} \frac{1}{\sqrt{a}}\left(\frac{\partial}{\partial t^{\alpha}}\left[\sqrt{a}\,\Gamma^{\alpha}_{\beta\delta}\right]\right) - \left(\frac{\partial^2\left[\ln\sqrt{a}\,\right]}{\partial t^{\delta}\partial t^{\beta}} + \Gamma^{\alpha}_{\beta\theta}\Gamma^{\theta}_{\alpha\delta}\right) . \end{aligned} \tag{9.509}$$

It renders a symmetric tensor because

$$\begin{aligned}\underline{R}_{\delta\beta} &= \frac{1}{\sqrt{a}}\left(\frac{\partial}{\partial t^{\alpha}}\left[\sqrt{a}\,\Gamma^{\alpha}_{\delta\beta}\right]\right) - \left(\frac{\partial^2\left[\ln\sqrt{a}\,\right]}{\partial t^{\beta}\partial t^{\delta}} + \Gamma^{\alpha}_{\delta\theta}\Gamma^{\theta}_{\alpha\beta}\right) \\ &= \frac{1}{\sqrt{a}}\left(\frac{\partial}{\partial t^{\alpha}}\left[\sqrt{a}\,\Gamma^{\alpha}_{\beta\delta}\right]\right) - \left(\frac{\partial^2\left[\ln\sqrt{a}\,\right]}{\partial t^{\delta}\partial t^{\beta}} + \Gamma^{\theta}_{\delta\alpha}\Gamma^{\alpha}_{\theta\beta}\right) \\ &= \frac{1}{\sqrt{a}}\left(\frac{\partial}{\partial t^{\alpha}}\left[\sqrt{a}\,\Gamma^{\alpha}_{\beta\delta}\right]\right) - \left(\frac{\partial^2\left[\ln\sqrt{a}\,\right]}{\partial t^{\delta}\partial t^{\beta}} + \Gamma^{\theta}_{\alpha\delta}\Gamma^{\alpha}_{\beta\theta}\right) \\ &= \underline{R}_{\beta\delta} \,. \end{aligned} \tag{9.510}$$

Using (9.26)$_1$, (9.205)$_1$ and (9.508)$_1$ along with the replacement property of the Kronecker delta, one can obtain

$$\begin{aligned}\underline{R}_{\beta\delta} &= \underline{\mathbb{R}}^{1}{}_{.\,\beta 1\delta}^{\;\;\cdot\;\cdot\;\cdot} + \underline{\mathbb{R}}^{2}{}_{.\,\beta 2\delta}^{\;\;\cdot\;\cdot\;\cdot} = a^{1\alpha}\,\underline{\mathbb{R}}_{\alpha\beta 1\delta} + a^{2\alpha}\,\underline{\mathbb{R}}_{\alpha\beta 2\delta} \\ &= a^{11}\,\underline{\mathbb{R}}_{1\beta 1\delta} + a^{12}\,\underline{\mathbb{R}}_{2\beta 1\delta} + a^{21}\,\underline{\mathbb{R}}_{1\beta 2\delta} + a^{22}\,\underline{\mathbb{R}}_{2\beta 2\delta} \,. \end{aligned} \tag{9.511}$$

From (9.65)$_3$, (9.207)$_2$, (9.208), (9.216)$_1$, (9.483)$_1$ and (9.511)$_3$, it follows that

$$\underbrace{\underline{R}_{11} = a^{22}\, \underline{\mathbb{R}}_{2121} = a^{22}\, \underline{\mathbb{R}}_{1212} = \frac{a_{11}}{a}\, \underline{\mathbb{R}}_{1212} = a_{11}\overline{\mathrm{K}}}_{\text{and, in a similar manner, } \underline{R}_{12} = a_{12}\overline{\mathrm{K}} = a_{21}\overline{\mathrm{K}} = \underline{R}_{21} \;,\; \underline{R}_{22} = a_{22}\overline{\mathrm{K}}} \;.$$

Consequently,

$$\boxed{\underline{R}_{\beta\delta} = \overline{\mathrm{K}}\, a_{\beta\delta} \;,} \tag{9.512}$$

represents the proportionality between the fully covariant form of the Ricci curvature tensor and the surface covariant metric coefficients where the proportionality factor is nothing but the Gaussian curvature. It is worth mentioning that this result is only valid for the two-dimensional space under consideration.

From (9.273) and (9.512), one immediately obtains

$$\boxed{\underline{R}_{\beta\delta} = 2\overline{\mathrm{H}}\, \underline{b}_{\beta\delta} - \underline{b}_{\beta}^{\cdot\theta}\, \underline{b}_{\theta\delta} \;.} \tag{9.513}$$

The **second-kind** Ricci tensor is introduced as

$$\boxed{\underline{R}^{\beta}_{\cdot\delta} = a^{\beta\rho}\, \underline{R}_{\rho\delta} = \underline{R}_{\delta\rho} a^{\rho\beta} = \underline{R}_{\delta}^{\cdot\beta} \;.} \tag{9.514}$$

Using (9.509) and (9.512), it admits the following forms

$$\underline{R}^{\beta}_{\cdot\delta} = a^{\beta\rho} \left\{ \frac{1}{\sqrt{a}} \left(\frac{\partial}{\partial t^{\alpha}} \left[\sqrt{a}\, \Gamma^{\alpha}_{\rho\delta} \right] \right) - \left(\frac{\partial^2 \left[\ln \sqrt{a}\, \right]}{\partial t^{\delta} \partial t^{\rho}} + \Gamma^{\alpha}_{\rho\theta} \Gamma^{\theta}_{\alpha\delta} \right) \right\} \;, \tag{9.515a}$$

$$\underline{R}^{\beta}_{\cdot\delta} = \overline{\mathrm{K}}\, \delta^{\beta}_{\delta} \;. \tag{9.515b}$$

9.7.6.2 Ricci Scalar

The Ricci scalar is constructed by contracting the indices of the second-kind Ricci tensor:

$$\boxed{R = \underline{R}^{\beta}_{\cdot\beta} = a^{\beta\rho}\, \underline{R}_{\rho\beta} = \underline{R}_{\beta\rho} a^{\rho\beta} = \underline{R}_{\beta}^{\cdot\beta} \;.} \tag{9.516}$$

By means of (9.515a), it takes the form

$$R = \underline{R}^{\beta}_{\cdot\beta} = a^{\beta\rho} \left\{ \frac{1}{\sqrt{a}} \left(\frac{\partial}{\partial t^{\alpha}} \left[\sqrt{a}\, \Gamma^{\alpha}_{\rho\beta} \right] \right) - \left(\frac{\partial^2 \left[\ln \sqrt{a}\, \right]}{\partial t^{\beta} \partial t^{\rho}} + \Gamma^{\alpha}_{\rho\theta} \Gamma^{\theta}_{\alpha\beta} \right) \right\} \;. \tag{9.517}$$

From (9.512) and $(9.516)_2$, taking into account the identity $a^{\beta\rho} a_{\rho\beta} = 2$, one simply obtains

$$\boxed{R = 2\overline{\mathrm{K}} \;.} \tag{9.518}$$

Substituting (9.518) into $(9.483)_2$ and (9.484) then yields

$$\boxed{\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = \frac{aR}{2}\varepsilon_{\alpha\beta}\varepsilon_{\gamma\delta} = \frac{R}{2}\left(a_{\alpha\gamma}a_{\beta\delta} - a_{\alpha\delta}a_{\beta\gamma}\right) .} \tag{9.519}$$

This may also be represented by

$$\underline{\mathbb{R}}_{\alpha\beta\gamma\delta} = a_{\alpha\gamma}\,\underline{R}_{\beta\delta} - a_{\alpha\delta}\,\underline{R}_{\beta\gamma} \ . \tag{9.520}$$

The covariant derivative of the Ricci scalar renders

$$\begin{aligned}
&\xrightarrow[\text{(9.2.14)}]{\text{by using}} a^{\gamma\alpha}a^{\delta\beta}\left(\underline{\mathbb{R}}_{\alpha\beta\gamma\delta}\big|_{\rho} + \underline{\mathbb{R}}_{\alpha\beta\delta\rho}\big|_{\gamma} + \underline{\mathbb{R}}_{\alpha\beta\rho\gamma}\big|_{\delta}\right) = 0 \\
&\xrightarrow[\text{by using (9.164) and (9.208)}]{\text{in light of (9.83c) and}} a^{\delta\beta}\ \underline{\mathbb{R}}^{\gamma}_{\ .\ \beta\gamma\delta}\Big|_{\rho} - a^{\gamma\alpha}\ \underline{\mathbb{R}}^{\delta}_{\ .\ \alpha\delta\rho}\Big|_{\gamma} - a^{\delta\beta}\ \underline{\mathbb{R}}^{\gamma}_{\ .\ \beta\gamma\rho}\Big|_{\delta} = 0 \\
&\xrightarrow[\text{(9.508)}]{\text{by using}} a^{\delta\beta}\ \underline{R}_{\beta\delta}\big|_{\rho} - a^{\gamma\alpha}\ \underline{R}_{\alpha\rho}\big|_{\gamma} - a^{\delta\beta}\ \underline{R}_{\beta\rho}\big|_{\delta} = 0 \\
&\xrightarrow[\text{(9.164) and (9.516)}]{\text{by using (9.83c),}} R\big|_{\rho} - \underline{R}^{\gamma}_{\ .\rho}\big|_{\gamma} - \underline{R}^{\delta}_{\ .\rho}\big|_{\delta} = 0 \ .
\end{aligned}$$

This helps establish the so-called *contracted Bianchi identities* (see Voss [41])

$$\boxed{\underline{R}^{\beta}_{\ .\delta}\Big|_{\beta} = \frac{R|_{\delta}}{2} = \frac{\partial R}{2\partial t^{\delta}} \ .} \tag{9.521}$$

It is noteworthy that this result is not confined to two-dimensional spaces and generally remains valid for spaces with higher dimensions. For the two-dimensional space under consideration, it can also be easily obtained from the relations $(9.164)_2$, (9.515b) and (9.518).

From $(9.521)_1$, taking into consideration that the mixed Kronecker delta is covariantly constant and possess the replacement property, one can write

$$\left(\underline{R}^{\beta}_{\ .\delta} - \frac{R}{2}\delta^{\beta}_{\delta}\right)\Bigg|_{\beta} = 0 \ . \tag{9.522}$$

This helps establish

$$\boxed{\mathrm{div}^{s}\mathbf{G} = \mathbf{0}\ ,} \quad \longleftarrow \text{see (7.84) and (9.225)} \tag{9.523}$$

where the symmetric second-order tensor $\mathbf{G}$ with

$$\underline{G}^{\beta}_{\ .\delta} = \underline{R}^{\beta}_{\ .\delta} - \frac{R}{2}\delta^{\beta}_{\delta} \quad \text{or} \quad \underline{G}_{\beta\delta} = \underline{R}_{\beta\delta} - \frac{R}{2}a_{\beta\delta}\ , \tag{9.524}$$

is called the *Einstein tensor*. It renders a **traceless** tensor, i.e.

$$\underline{G}^{\beta}_{.\beta} = \underline{R}^{\beta}_{.\beta} - \frac{R}{2}\delta^{\beta}_{\beta} = R - \frac{R}{2}(1+1) = 0 , \tag{9.525}$$

and basically describes the curvature of spacetime in general relativity, see Hawking and Ellis [42] and Misner et al. [43].

9.7.6.3 Geometric Meaning of Ricci Tensor

The geodesics in a curved space can converge or spread apart specifically because of the curvature of that space. The volume of a ball surrounded by the geodesics thus changes as it moves along these curves. The Ricci tensor is an object by which that change in volume can be captured. Since the Ricci curvature tensor is constructed from the fourth-order Riemann-Christoffel tensor, the procedure to be followed here is similar to what that led to (9.199). This procedure relies on some basic assumptions discussed below (Robinson et al. [44]).

Consider a set of (dust) particles clumped together in some region of a space. And consider a set of parallel geodesic curves passing through these particles. It is assumed that the relative positions of the particles are initially fixed and therefore the relative velocity between any two of them is zero. In other words, the velocity vectors of the particles are all parallel at the beginning. But, the relative acceleration between any two points may not be zero. The problem can well be formulated by considering the motion of two infinitesimally close particles which follow their own geodesics in a very short period of time. In the following, it will be shown that how (in the limit) the change in area of a circle,[17] defined by the distance between these two points, is related to curvature of space.

Let $\mathcal{C}_1$ with the parametrization $t^{\alpha}(t)$ be a geodesic curve of a surface $\mathcal{S}$. Consider a point P at t, interpreted here as time, and an infinitesimally nearby point R corresponding to $t + \Delta t$. The tangent vector at P, interpreted here as velocity, is denoted by $\mathbf{v}_P$ (note that $t^{\alpha}(R) = t^{\alpha}(t) + \Delta t\, \underline{v}^{\alpha}_{P}$). Consider another vector sitting at P, $\mathbf{u}_P$, which helps identify a point Q on another geodesic curve $\mathcal{C}_2$ with $t^{\alpha}(Q) = t^{\alpha}(t) + \Delta t\, \underline{u}^{\alpha}_{P}$. It is important to point out that $\Delta t \mathbf{v}_P$, $\Delta t \mathbf{u}_P$ are infinitesimal vectors although $\mathbf{v}_P$, $\mathbf{u}_P$ can be quite finite. The velocity vector $\mathbf{v}_P$ describes flow along $\mathcal{C}_1$ while the separation vector $\mathbf{u}_P$ describes motion from $\mathcal{C}_1$ to $\mathcal{C}_2$. One can now parallel transport $\mathbf{v}_P$ a distance $\Delta t \mathbf{v}_P$ ($\Delta t \mathbf{u}_P$) to have $\mathbf{v}^{\parallel}_{P\to R}$ ($\mathbf{v}^{\parallel}_{P\to Q}$). Parallel transporting $\mathbf{u}_P$ along $\mathcal{C}_1$ from P to R helps identify another point S with $t^{\alpha}(S) = t^{\alpha}(R) + \Delta t\, \underline{u}^{\alpha\,\parallel}_{P\to R}$. Notice that $t^{\alpha}(S) = t^{\alpha}(Q) + \Delta t\, \underline{v}^{\alpha\,\parallel}_{P\to Q}$. One can thus have an infinitesimal closed loop $PQSR$ as shown in Fig. 9.29. Guided by the equations (9.197)-(9.198), the change in velocity vector after its parallel transport from R to S, i.e. $\mathbf{v}^{\parallel}_{R\to S}$, and from Q to S, i.e. $\mathbf{v}^{\parallel}_{Q\to S}$, is

$$\boxed{\Delta\, \underline{v}^{\alpha} = (\Delta t)^2\, \underline{\mathbb{R}}^{\alpha}_{.\,\beta\gamma\delta}\, \underline{v}^{\beta}\, \underline{v}^{\gamma}\, \underline{u}^{\delta} .} \tag{9.526}$$

[17] A circle (sphere) is also known as *1-sphere* (*2-sphere*) and the region enclosed by that circle (sphere) is called *2-ball* (*3-ball*).

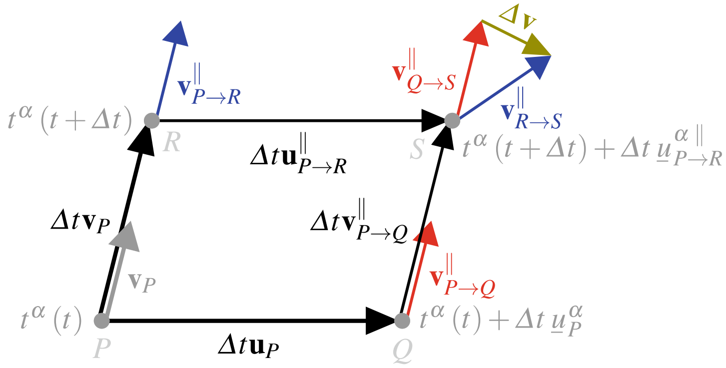


Fig. 9.29 Parallel transport of a vector along two paths defined by that vector and a generic one

This is eventually the relative velocity of the particles sitting at P and Q after the time increment Δt. The acceleration vector $\underline{a}^{\alpha}$ is then given by

$$\boxed{\underline{a}^{\alpha} = \lim_{\Delta t\to 0} \frac{\Delta\, \underline{v}^{\alpha}}{\Delta t}\ .} \tag{9.527}$$

It follows that

$$\lim_{\Delta t\to 0} \frac{\underline{a}^{\alpha}}{\Delta t} = \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}\ \underline{v}^{\beta}\ \underline{v}^{\gamma}\ \underline{u}^{\delta}\ . \tag{9.528}$$

This relation represents geodesic deviation which states that the two particles moving along their geodesics will accelerate with respect to each other.

At this stage, consider a circle of radius R_0 where $R_0 = t^{\alpha}(Q) - t^{\alpha}(P)$. After the time increment Δt, the new radius becomes $r^{\alpha}(\Delta t)$ where $r^{\alpha}(0) = R_0$ (see Fig. 9.30). Suppose that the starting points start the journey along their geodesics with zero velocity and nonzero acceleration, i.e.

$$\dot{r}^{\alpha}(0) = 0 \quad , \quad \ddot{r}^{\alpha}(0) = \underline{a}^{\alpha} \qquad \text{where} \qquad \dot{\bullet} = \frac{d\bullet}{d\Delta t}\ . \tag{9.529}$$

One then immediately obtains

$$r^{\alpha}(\Delta t) = R_0 + \frac{1}{2}\, \underline{a}^{\alpha}\, (\Delta t)^2\ . \tag{9.530}$$

It follows that

$$\lim_{\Delta t\to 0} \frac{\ddot{r}^{\alpha}(\Delta t)}{r^{\alpha}(\Delta t)} = \lim_{\Delta t\to 0} \left(\frac{\Delta t}{R_0}\, \frac{\underline{a}^{\alpha}}{\Delta t} \right) = \lim_{\Delta t\to 0} \left(\frac{\Delta t}{R_0}\, \underline{\mathbb{R}}^{\alpha}{}_{.\,\beta\gamma\delta}\ \underline{v}^{\beta}\ \underline{v}^{\gamma}\ \underline{u}^{\delta} \right)\ (\alpha = 1, 2;\ \text{no sum})\ .$$

Recall that the vector $\mathbf{u}_P$ was arbitrary. Thus, one can choose $\mathbf{u}_P = \mathbf{a}_{\alpha P}$. This immediately implies that $R_0 = \Delta t$. Consequently, the above relation becomes

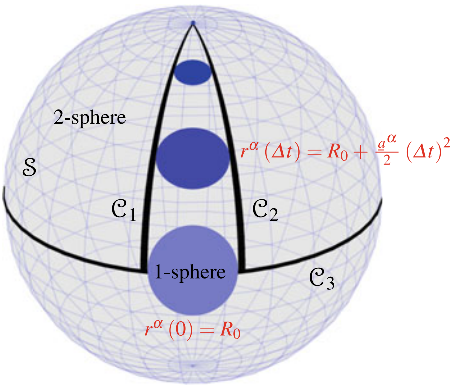


Fig. 9.30 Area change along geodesics on sphere

$$\lim_{\Delta t \to 0} \frac{\ddot{r}^{\alpha}(\Delta t)}{r^{\alpha}(\Delta t)} = \underline{\mathbb{R}}^{\alpha}{}_{.\beta\gamma\alpha}^{\;\;\cdot\;\cdot\;\cdot}\, \underline{v}^{\beta}\, \underline{v}^{\gamma} \quad (\alpha = 1, 2;\ \text{no sum}) \,. \tag{9.531}$$

Consider an ellipsoid of radii $r^{\alpha}(\Delta t)$ in a space of dimension 2. Its volume, $V_2(\Delta t)$, is then given by

$$V_2 = \pi r^1 r^2 \quad \text{and, consequently,} \quad \left.\begin{array}{l} \dot{V}_2 = \pi \dot{r}^1 r^2 + \pi r^1 \dot{r}^2 \,, \\ \ddot{V}_2 = \pi \ddot{r}^1 r^2 + 2\pi \dot{r}^1 \dot{r}^2 + \pi r^1 \ddot{r}^2 \end{array}\right\} . \tag{9.532}$$

Thus,

$$\frac{\ddot{V}_2}{V_2} = \frac{\ddot{r}^1}{r^1} + 2\frac{\dot{r}^1}{r^1}\frac{\dot{r}^2}{r^2} + \frac{\ddot{r}^2}{r^2} \,. \tag{9.533}$$

In the limit, one can arrive at

$$\lim_{\Delta t \to 0} \frac{\ddot{V}_2}{V_2} = \lim_{\Delta t \to 0} \sum_{\alpha=1}^{2} \frac{\ddot{r}^{\alpha}}{r^{\alpha}} = -\sum_{\alpha=1}^{2} \underline{\mathbb{R}}^{\alpha}{}_{.\beta\alpha\gamma}^{\;\;\cdot\;\cdot\;\cdot}\, \underline{v}^{\beta}\, \underline{v}^{\gamma} \,,$$

or, finally,

$$\boxed{\lim_{\Delta t \to 0} \frac{\ddot{V}_2}{V_2} = -\,\underline{R}_{\beta\gamma}\, \underline{v}^{\beta}\, \underline{v}^{\gamma} \,.} \tag{9.534}$$

That is why the Ricci tensor describes the change in a volume as it travels along geodesics in a curved space (see Exercise 9.23 for the geometric meaning of Ricci scalar).

Hint: The procedure outlined above can simply be extended to an n-dimensional space. In this case, consider an ellipsoid of radii $r^j(\Delta t)$ in a space of dimension n whose volume, $V_n(\Delta t)$, is given by

$$\underbrace{V_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} \prod_{j=1}^{n} r^j}_{\text{note that } n=2 \Rightarrow \Gamma(2)=1 \text{ and } n=3 \Rightarrow \Gamma\left(\frac{5}{2}\right)=\frac{3}{4}\sqrt{\pi}} \quad , \tag{9.535}$$

where Γ presents the *gamma function*. The first derivative of V_n takes the form

$$\dot{V}_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} \sum_{i=1}^{n} \frac{\dot{r}^i}{r^i} \prod_{j=1}^{n} r^j , \tag{9.536}$$

and its second derivative renders

$$\ddot{V}_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} \sum_{i=1}^{n} \left(\frac{\ddot{r}^i}{r^i} + \sum_{k=1}^{i-1} \frac{\dot{r}^i \dot{r}^k}{r^i r^k} \right) \prod_{j=1}^{n} r^j . \tag{9.537}$$

It is then easy to see that

$$\underbrace{\lim_{\Delta t \to 0} \frac{\ddot{V}_n}{V_n} = - \underline{R}_{ij}\, \underline{v}^i\, \underline{v}^j}_{\text{note that } \underline{v}^i \underline{R}_{ij} \underline{v}^j = \left[\left(\underline{v}^i \mathbf{g}_i\right) \otimes \left(\underline{v}^j \mathbf{g}_j\right)\right] : \left[\underline{R}_{kl} \mathbf{g}^k \otimes \mathbf{g}^l\right] \text{ is an invariant object}} \quad . \tag{9.538}$$

9.8 Lie Derivatives

The *Lie derivative* plays an important role in differential geometry of manifolds. It is also widely used in many branches of physics and engineering. Examples of which include general relativity, nonlinear solid mechanics and control theory. The technique of Lie derivative is an extension of the directional derivative which also remains invariant under a transformation from one coordinate system to another. It provides tensors out of tensors. Indeed, the Lie derivative of a tensor field is a tensor with the same order. Such a technique computes the change in a tensorial field variable as it moves along the flow of a vector field. In particular, the Lie derivative of a vector field in the direction of flow of a vector field measures how much the resulting flow curves fail to close. It basically indicates whether the coordinate curves of a coordinate system can be constructed from some families of curves or not. This motivates to completely devote this section to the study of Lie derivatives. The section begins with the geometrical description of the Lie derivative of a vector field. The result will then be generalized to tensors of higher ranks. The Lie derivative of the differential forms is also studied. At the end, this new differential operator is represented in a more sophisticated and abstract form by introducing what is known as *commutator*. This will help address more aspects of the problem. For a more detailed account on Lie derivatives, see, e.g., Helgason [45] and Fecko [46]. See also the older classic work of Cartan [47].

9.8.1 Lie Derivative of Vector Fields

Recall that the covariant derivative relied on the concept of parallel transport. The different tangent planes in this technique were related through the Christoffel symbols which themselves were determined from the metric coefficients. The covariant differentiation thus essentially relies on defining the metric coefficients. In contrary, there is no need to define the metric coefficients to compute the Lie derivative of a tensor field. It basically requires a simpler structure which relies on primary geometric objects associated with given vector fields. This is demonstrated in the following.

Any smooth vector field determines a family of curves called *flow lines* (or *field lines* or *streamlines* or *integral curves* or *trajectories*). It can be shown that such curves do exist and are unique. The flow lines of a vector field constitute a **congruence**. This means that there is only one curve passing through a point on a manifold. The vectors of a vector field are always tangent to the corresponding integral curves. Indeed, a vector field represents the velocity field of a particle moving along the resulting integral curves. Let $\mathbf{u} = \underline{u}^1\mathbf{a}_1 + \underline{u}^2\mathbf{a}_2$ be a given vector field. Further, let $\mathbf{x} = \hat{\mathbf{x}}^s\left(t^1(\lambda), t^2(\lambda)\right)$ be a parametrized surface curve whose tangent vector is written by $\mathbf{a}_\lambda = \left(dt^1/d\lambda\right)\mathbf{a}_1 + \left(dt^2/d\lambda\right)\mathbf{a}_2$. Then, the integral curves $t^\alpha(\lambda)$ of the vector field $\mathbf{u}$ are obtained by solving the following system of ordinary differential equations

$$\boxed{\frac{dt^1}{d\lambda} = \underline{u}^1\left(t^1, t^2\right) \quad , \quad \frac{dt^2}{d\lambda} = \underline{u}^2\left(t^1, t^2\right) \ .} \tag{9.539}$$

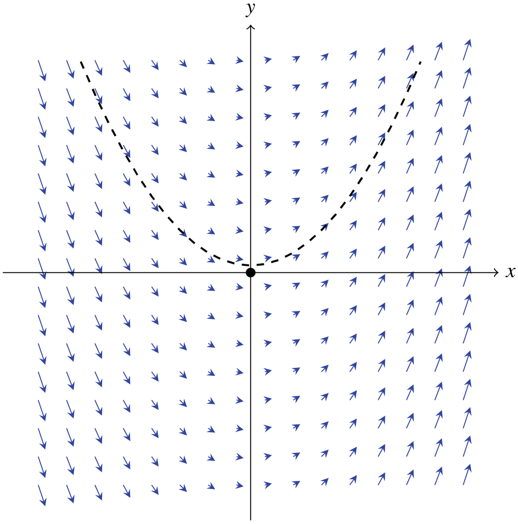
See Fig. 9.31 for a geometrical interpretation. The goal is now to examine that whether the flow lines of two given smooth vector fields on the tangent spaces of a two-dimensional manifold can properly form coordinate curves or not. This will be determined by the technique of Lie derivative described below.

The Lie derivative of a vector field is illustrated in Fig. 9.32. Let $\mathbf{u} = \underline{u}^\alpha\mathbf{a}_\alpha$ be a smooth vector field whose integral curves are denoted by $t^\theta(\lambda)$. Further, let $\hat{\mathbf{h}} = \hat{\underline{h}}^\alpha\mathbf{a}_\alpha$ be another smooth vector field with corresponding integral curves $t^\theta(\mu)$. Consider a point P whose coordinates are denoted by $\left(t^1_P, t^2_P\right)$. The tangent vectors passing through this point are denoted by $\underline{u}^\alpha_P := \underline{u}^\alpha\left(t^1_P, t^2_P\right)$ and $\hat{\underline{h}}^\alpha_P := \hat{\underline{h}}^\alpha\left(t^1_P, t^2_P\right)$. Suppose that a particle at P moves along the direction of flow of $\mathbf{u}$ to Q whose coordinates are given by

$$t^\theta_Q = t^\theta_P + \underline{u}^\theta_P \Delta\lambda \ . \tag{9.540}$$

In this expression, the vector $\underline{u}^\theta_P \Delta\lambda = t^\theta_Q - t^\theta_P$, relating the original position of that particle at P to its current position at Q, is called the *displacement vector*. The tangent vector $\hat{\underline{h}}^\alpha_Q := \hat{\underline{h}}^\alpha\left(t^1_Q, t^2_Q\right)$ passing through Q can be expressed in terms of $\hat{\underline{h}}^\alpha_P$ and its partial derivatives via the following first-order Taylor series expansion

$$\hat{\underline{h}}^\alpha_Q = \hat{\underline{h}}^\alpha_P + \left.\frac{\partial\, \hat{\underline{h}}^\alpha}{\partial t^\theta}\right|_P \left(\underline{u}^\theta_P \Delta\lambda\right) \ . \quad \leftarrow \text{see (6.24)–(6.25)} \tag{9.541}$$



Consider a two-dimensional Cartesian coordinate frame. The coordinates of an arbitrary point are written by $(x(\lambda), y(\lambda))$. Suppose one is given $\mathbf{u} = \widehat{\mathbf{e}}_x + x\widehat{\mathbf{e}}_y$. The solution of

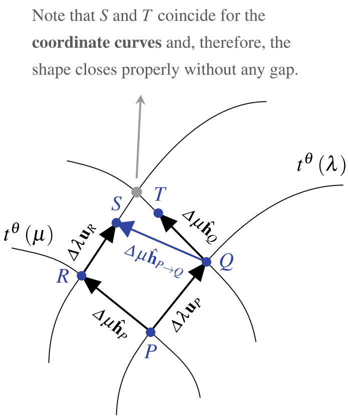
$$\frac{dx}{d\lambda} = 1 \,, \quad \frac{dy}{d\lambda} = x \,,$$

that is,

$$x = \lambda + c \,, \; y = \frac{1}{2}\lambda^2 + c\lambda + d \,,$$

will provide the flow lines of $\mathbf{u}$. In these relations, c and d are constants. The results have been plotted in this figure.

Fig. 9.31 Flow lines



Let $\mathbf{u} = \underline{u}^\alpha \mathbf{a}_\alpha$ be a smooth vector field which provides the integral curves $t^\theta(\lambda)$. Further, let $\hat{\mathbf{h}} = \hat{\underline{h}}^\alpha \mathbf{a}_\alpha$ be another smooth vector field with corresponding flow lines $t^\theta(\mu)$. One then has

$$\left.\begin{array}{l} t_Q^\theta = t_P^\theta + \underline{u}_P^\theta \Delta\lambda \\ t_R^\theta = t_P^\theta + \hat{\underline{h}}_P^\theta \Delta\mu \end{array}\right\} \,, \quad \left.\begin{array}{l} t_T^\theta = t_P^\theta + \underline{u}_P^\theta \Delta\lambda + \hat{\underline{h}}_Q^\theta \Delta\mu \\ t_S^\theta = t_P^\theta + \hat{\underline{h}}_P^\theta \Delta\mu + \underline{u}_R^\theta \Delta\lambda \end{array}\right\} \,,$$

and

$$\underline{u}_R^\alpha = \underline{u}_P^\alpha + \left.\frac{\partial\, \underline{u}^\alpha}{\partial t^\theta}\right|_P \hat{\underline{h}}_P^\theta \Delta\mu \,, \quad \left.\begin{array}{l} \hat{\underline{h}}_Q^\alpha = \hat{\underline{h}}_P^\alpha + \left.\frac{\partial\, \hat{\underline{h}}^\alpha}{\partial t^\theta}\right|_P \underline{u}_P^\theta \Delta\lambda \\ \hat{\underline{h}}_{P\to Q}^\alpha = \hat{\underline{h}}_P^\alpha + \left.\frac{\partial\, \underline{u}^\alpha}{\partial t^\theta}\right|_P \hat{\underline{h}}_P^\theta \Delta\lambda \end{array}\right\} \,.$$

The Lie derivative of $\hat{\mathbf{h}}$ with respect to $\mathbf{u}$ is defined by

$$\mathscr{L}_{\mathbf{u}} \hat{\mathbf{h}} := \lim_{\Delta\lambda \to 0} \frac{\hat{\mathbf{h}}_Q - \hat{\mathbf{h}}_{P\to Q}}{\Delta\lambda} \quad \text{or} \quad \mathscr{L}_{\mathbf{u}} \hat{\underline{h}}^\alpha := \lim_{\Delta\lambda \to 0} \frac{\hat{\underline{h}}_Q^\alpha - \hat{\underline{h}}_{P\to Q}^\alpha}{\Delta\lambda} \,,$$

which finally leads to

$$\mathscr{L}_{\mathbf{u}} \hat{\underline{h}}^\alpha = \underline{u}^\theta \frac{\partial\, \hat{\underline{h}}^\alpha}{\partial t^\theta} - \hat{\underline{h}}^\theta \frac{\partial\, \underline{u}^\alpha}{\partial t^\theta} \,.$$

Fig. 9.32 Lie derivative

At this stage, one can recognize another point T with

$$t_T^\theta = t_P^\theta + \underline{u}_P^\theta \Delta\lambda + \hat{\underline{h}}_Q^\theta \Delta\mu \,. \tag{9.542}$$

Suppose a particle at P is displaced in the direction of flow of $\hat{\mathbf{h}}$ to R whose coordinates are given by

$$t_R^\theta = t_P^\theta + \hat{\underline{h}}_P^\theta \Delta\mu \,. \tag{9.543}$$

Denoting by $\underline{u}_R^\alpha := \underline{u}^\alpha \left(t_R^1, t_R^2\right)$ the tangent vector passing through R, one can write

$$\underline{u}^{\alpha}_{R} = \underline{u}^{\alpha}_{P} + \left.\frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}}\right|_{P} \left(\underline{\hat{h}}^{\theta}_{P} \Delta\mu \right) . \tag{9.544}$$

Then, one can identify a point S with

$$t^{\theta}_{S} = t^{\theta}_{P} + \underline{\hat{h}}^{\theta}_{P} \Delta\mu + \underline{u}^{\theta}_{R} \Delta\lambda . \tag{9.545}$$

Notice that the tail of $\Delta\mu \hat{\mathbf{h}}_{P}$ has been displaced from P to Q while its tip has moved from R to S. Indeed, the given vector field $\mathbf{u}$ defines such a special way of transforming the other vector field $\hat{\mathbf{h}}$. The resulting vector, denoted here by $\hat{\mathbf{h}}_{P\to Q}$, basically represents the so-called *push-forward* of $\hat{\mathbf{h}}$ from P to Q. One then has

$$\underline{\hat{h}}^{\alpha}_{P\to Q} = \frac{t^{\alpha}_{S} - t^{\alpha}_{Q}}{\Delta\mu} = \frac{\underline{\hat{h}}^{\alpha}_{P} \Delta\mu + \underline{u}^{\alpha}_{R} \Delta\lambda - \underline{u}^{\alpha}_{P} \Delta\lambda}{\Delta\mu} . \tag{9.546}$$

or, using (9.544),[18]

$$\underline{\hat{h}}^{\alpha}_{P\to Q} = \underline{\hat{h}}^{\alpha}_{P} + \left.\frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}}\right|_{P} \left(\underline{\hat{h}}^{\theta}_{P} \Delta\lambda \right) . \tag{9.547}$$

It is worth mentioning that the subtraction $\hat{\mathbf{h}}_{Q} - \hat{\mathbf{h}}_{P}$ cannot simply be decomposed as $\left(\underline{\hat{h}}^{\alpha}_{Q} - \underline{\hat{h}}^{\alpha}_{P} \right) \mathbf{a}_{\alpha}$ due to the different bases of vectors. Indeed, the above procedure was followed to provide a consistent subtraction for differentiation. Notice that $\hat{\mathbf{h}}_{Q}$ and $\hat{\mathbf{h}}_{P\to Q}$ are now in the **same** tangent plane and, therefore, they can be expressed with respect to the basis vectors at Q. The vector difference of these two vectors is then given by

$$\hat{\mathbf{h}}_{Q} - \hat{\mathbf{h}}_{P\to Q} = \left(\underline{\hat{h}}^{\alpha}_{Q} - \underline{\hat{h}}^{\alpha}_{P\to Q} \right) \mathbf{a}_{\alpha} . \tag{9.548}$$

---

[18] The result (9.547) can be obtained in an alternative way using an **infinitesimal coordinate transformation**. This is demonstrated for the interested reader in the following.

Denoting by $\underline{\hat{h}}^{\alpha}\left(t^{1}, t^{2}\right)$ and $\bar{\underline{\hat{h}}}^{\alpha}\left(\bar{t}^{1}, \bar{t}^{2}\right)$ the old and new components of $\hat{\mathbf{h}}$, respectively, the vector transformation law reads

$$\bar{\underline{\hat{h}}}^{\alpha}\left(\bar{t}^{1}, \bar{t}^{2}\right) = \frac{\partial \bar{t}^{\alpha}}{\partial t^{\theta}} \underline{\hat{h}}^{\theta}\left(t^{1}, t^{2}\right) ,$$

where $\bar{t}^{\alpha} = t^{\alpha} + \underline{u}^{\alpha}\left(t^{1}, t^{2}\right) \Delta\lambda$. It follows that

$$\frac{\partial \bar{t}^{\alpha}}{\partial t^{\theta}} \underline{\hat{h}}^{\theta} = \frac{\partial \left(t^{\alpha} + \underline{u}^{\alpha} \Delta\lambda\right)}{\partial t^{\theta}} \underline{\hat{h}}^{\theta} = \delta^{\alpha}_{\theta} \underline{\hat{h}}^{\theta} + \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}} \Delta\lambda\, \underline{\hat{h}}^{\theta} = \underline{\hat{h}}^{\alpha} + \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}} \underline{\hat{h}}^{\theta} \Delta\lambda .$$

Thus,

$$\bar{\underline{\hat{h}}}^{\alpha}\left(\bar{t}^{1}, \bar{t}^{2}\right) = \underline{\hat{h}}^{\alpha}\left(t^{1}, t^{2}\right) + \frac{\partial\, \underline{u}^{\alpha}\left(t^{1}, t^{2}\right)}{\partial t^{\theta}} \underline{\hat{h}}^{\theta}\left(t^{1}, t^{2}\right) \Delta\lambda .$$

In comparison with (9.547), one should realize that $\bar{\underline{\hat{h}}}^{\alpha}\left(\bar{t}^{1}, \bar{t}^{2}\right)$ presents $\underline{\hat{h}}^{\alpha}_{P\to Q}\left(t^{1}_{Q}, t^{2}_{Q}\right)$ and $\underline{\hat{h}}^{\alpha}\left(t^{1}, t^{2}\right)$ is simply $\underline{\hat{h}}^{\alpha}\left(t^{1}_{P}, t^{2}_{P}\right)$.

This helps define the Lie derivative of $\hat{\mathbf{h}}$ with respect to $\mathbf{u}$ as

$$\boxed{\mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}} := \lim_{\Delta\lambda\to 0} \frac{\hat{\mathbf{h}}_Q - \hat{\mathbf{h}}_{P\to Q}}{\Delta\lambda} \quad \text{or} \quad \mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha} := \lim_{\Delta\lambda\to 0} \frac{\underline{\hat{h}}^{\alpha}_Q - \underline{\hat{h}}^{\alpha}_{P\to Q}}{\Delta\lambda}\ .} \tag{9.549}$$

Using (9.541), (9.547) and (9.549), the Lie derivative of a vector at an arbitrary point can finally be represented according to

$$\boxed{\mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}} = \left(\underline{u}^{\theta} \frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta} \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}} \right) \mathbf{a}_{\alpha} \ \text{or}\ \mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha} = \underline{u}^{\theta} \frac{\partial\, \underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta} \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}}\ .} \tag{9.550}$$

With the aid of $(9.92)_2$ and (9.128), this expression can be rephrased as

$$\boxed{\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha} = \underline{u}^{\theta} \left(\underline{\hat{h}}^{\alpha}\Big|_{\theta} \right) - \underline{\hat{h}}^{\theta} \left(\underline{u}^{\alpha}\big|_{\theta} \right)\ .} \tag{9.551}$$

As can be seen from (9.550) and (9.551), the partial derivative can simply be replaced by the covariant differentiation in the technique of Lie derivative. And this holds true for tensors of other ranks.

In the following, some important properties of the technique of Lie derivative are introduced. ✾

Let α and β be two arbitrary constants. Further, let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be three smooth vector fields. Then,

$$\boxed{\left.\begin{aligned} \mathscr{L}_{\mathbf{u}}\left(\alpha\mathbf{v} + \beta\mathbf{w}\right) &= \alpha\mathscr{L}_{\mathbf{u}}\mathbf{v} + \beta\mathscr{L}_{\mathbf{u}}\mathbf{w} \\ \mathscr{L}_{(\alpha\mathbf{u}+\beta\mathbf{v})}\mathbf{w} &= \alpha\mathscr{L}_{\mathbf{u}}\mathbf{w} + \beta\mathscr{L}_{\mathbf{v}}\mathbf{w} \end{aligned}\right\}\ .} \tag{9.552}$$

As can be seen, the Lie derivative is eventually a **bilinear** map.

The Lie derivative of a smooth vector field with respect to itself trivially vanishes:

$$\boxed{\mathscr{L}_{\mathbf{u}}\mathbf{u} = \mathbf{0}\ .} \tag{9.553}$$

The Lie derivative of a smooth vector field is **skew-symmetric**, that is,

$$\boxed{\mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}} = -\mathscr{L}_{\hat{\mathbf{h}}}\mathbf{u}\ .} \tag{9.554}$$

Any three smooth vector fields $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ satisfy the so-called *Jacobi identity*

$$\boxed{\mathscr{L}_{\mathbf{v}}\mathscr{L}_{\mathbf{w}}\,\underline{u}^{\alpha} + \mathscr{L}_{\mathbf{w}}\mathscr{L}_{\mathbf{u}}\,\underline{v}^{\alpha} + \mathscr{L}_{\mathbf{u}}\mathscr{L}_{\mathbf{v}}\,\underline{w}^{\alpha} = 0\ ,} \tag{9.555}$$

since, by definition,

$$\begin{aligned}
\mathscr{L}_{\mathbf{v}}\mathscr{L}_{\mathbf{w}}\,\underline{u}^{\alpha} &= \mathscr{L}_{\mathbf{v}}\left[\mathscr{L}_{\mathbf{w}}\,\underline{u}^{\alpha}\right] = \underline{v}^{\theta}\frac{\partial}{\partial t^{\theta}}\left[\mathscr{L}_{\mathbf{w}}\,\underline{u}^{\alpha}\right] - \left[\mathscr{L}_{\mathbf{w}}\,\underline{u}^{\theta}\right]\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\theta}} \\
&= \underline{v}^{\theta}\frac{\partial}{\partial t^{\theta}}\left[\underline{w}^{\rho}\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\rho}} - \underline{u}^{\rho}\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}}\right] - \left[\underline{w}^{\rho}\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\rho}} - \underline{u}^{\rho}\frac{\partial\,\underline{w}^{\theta}}{\partial t^{\rho}}\right]\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\theta}} \\
&= \underline{v}^{\theta}\frac{\partial\,\underline{w}^{\rho}}{\partial t^{\theta}}\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\rho}} + \underline{v}^{\theta}\,\underline{w}^{\rho}\frac{\partial^{2}\,\underline{u}^{\alpha}}{\partial t^{\theta}\partial t^{\rho}} - \underline{v}^{\theta}\frac{\partial\,\underline{u}^{\rho}}{\partial t^{\theta}}\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}} \\
&\quad - \underline{v}^{\theta}\,\underline{u}^{\rho}\frac{\partial^{2}\,\underline{w}^{\alpha}}{\partial t^{\theta}\partial t^{\rho}} - \underline{w}^{\rho}\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\rho}}\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\theta}} + \underline{u}^{\rho}\frac{\partial\,\underline{w}^{\theta}}{\partial t^{\rho}}\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\theta}}\,,
\end{aligned} \tag{9.556}$$

and, in a similar manner,

$$\begin{aligned}
\mathscr{L}_{\mathbf{w}}\mathscr{L}_{\mathbf{u}}\,\underline{v}^{\alpha} &= \underline{w}^{\rho}\frac{\partial}{\partial t^{\rho}}\left[\underline{u}^{\theta}\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\theta}} - \underline{v}^{\theta}\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}}\right] - \left[\underline{u}^{\theta}\frac{\partial\,\underline{v}^{\rho}}{\partial t^{\theta}} - \underline{v}^{\theta}\frac{\partial\,\underline{u}^{\rho}}{\partial t^{\theta}}\right]\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}} \\
&= \underline{w}^{\rho}\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\rho}}\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\theta}} + \underline{w}^{\rho}\,\underline{u}^{\theta}\frac{\partial^{2}\,\underline{v}^{\alpha}}{\partial t^{\rho}\partial t^{\theta}} - \underline{w}^{\rho}\frac{\partial\,\underline{v}^{\theta}}{\partial t^{\rho}}\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} \\
&\quad - \underline{w}^{\rho}\,\underline{v}^{\theta}\frac{\partial^{2}\,\underline{u}^{\alpha}}{\partial t^{\rho}\partial t^{\theta}} - \underline{u}^{\theta}\frac{\partial\,\underline{v}^{\rho}}{\partial t^{\theta}}\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}} + \underline{v}^{\theta}\frac{\partial\,\underline{u}^{\rho}}{\partial t^{\theta}}\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}}\,,
\end{aligned} \tag{9.557a}$$

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}}\mathscr{L}_{\mathbf{v}}\,\underline{w}^{\alpha} &= \underline{u}^{\theta}\frac{\partial}{\partial t^{\theta}}\left[\underline{v}^{\rho}\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}} - \underline{w}^{\rho}\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\rho}}\right] - \left[\underline{v}^{\rho}\frac{\partial\,\underline{w}^{\theta}}{\partial t^{\rho}} - \underline{w}^{\rho}\frac{\partial\,\underline{v}^{\theta}}{\partial t^{\rho}}\right]\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} \\
&= \underline{u}^{\theta}\frac{\partial\,\underline{v}^{\rho}}{\partial t^{\theta}}\frac{\partial\,\underline{w}^{\alpha}}{\partial t^{\rho}} + \underline{u}^{\theta}\,\underline{v}^{\rho}\frac{\partial^{2}\,\underline{w}^{\alpha}}{\partial t^{\theta}\partial t^{\rho}} - \underline{u}^{\theta}\frac{\partial\,\underline{w}^{\rho}}{\partial t^{\theta}}\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\rho}} \\
&\quad - \underline{u}^{\theta}\,\underline{w}^{\rho}\frac{\partial^{2}\,\underline{v}^{\alpha}}{\partial t^{\theta}\partial t^{\rho}} - \underline{v}^{\rho}\frac{\partial\,\underline{w}^{\theta}}{\partial t^{\rho}}\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} + \underline{w}^{\rho}\frac{\partial\,\underline{v}^{\theta}}{\partial t^{\rho}}\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}}\,,
\end{aligned} \tag{9.557b}$$

having in mind that the partial differentiation has the commutative property and the dummy indices can be renamed.

Similarly to the covariant derivative, the Lie derivative does not have the commutative property in general. Its **noncommutativity** is expressed as

$$\boxed{\mathscr{L}_{\mathbf{u}}\mathscr{L}_{\mathbf{v}}\,\underline{\hat{h}}^{\alpha} - \mathscr{L}_{\mathbf{v}}\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha} = \mathscr{L}_{\mathscr{L}_{\mathbf{u}}\mathbf{v}}\,\underline{\hat{h}}^{\alpha}\,,} \tag{9.558}$$

where

$$\mathscr{L}_{\mathscr{L}_{\mathbf{u}}\mathbf{v}}\,\underline{\hat{h}}^{\alpha} = \left[\mathscr{L}_{\mathbf{u}}\,\underline{v}^{\theta}\right]\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta}\,\frac{\partial}{\partial t^{\theta}}\left[\mathscr{L}_{\mathbf{u}}\,\underline{v}^{\alpha}\right]$$
$$= \underline{u}^{\rho}\,\frac{\partial\,\underline{v}^{\theta}}{\partial t^{\rho}}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{v}^{\rho}\,\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\rho}}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{u}^{\rho}}{\partial t^{\theta}}\,\frac{\partial\,\underline{v}^{\alpha}}{\partial t^{\rho}}$$
$$- \underline{\hat{h}}^{\theta}\,\underline{u}^{\rho}\,\frac{\partial^{2}\,\underline{v}^{\alpha}}{\partial t^{\theta}\partial t^{\rho}} + \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{v}^{\rho}}{\partial t^{\theta}}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\rho}} + \underline{\hat{h}}^{\theta}\,\underline{v}^{\rho}\,\frac{\partial^{2}\,\underline{u}^{\alpha}}{\partial t^{\theta}\partial t^{\rho}}\,. \tag{9.559}$$

Recall from (9.172a) that the object $\partial\,\underline{\hat{h}}^{\alpha}/\partial t^{\beta}$ was not a tensorial variable. By the definition (9.549), its Lie derivative represents

$$\mathscr{L}_{\mathbf{u}}\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}} = \frac{\partial\,\underline{u}^{\theta}}{\partial t^{\beta}}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} + \underline{u}^{\theta}\,\frac{\partial^{2}\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}\partial t^{\theta}}$$
$$- \frac{\partial\,\underline{\hat{h}}^{\theta}}{\partial t^{\beta}}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta}\,\frac{\partial^{2}\,\underline{u}^{\alpha}}{\partial t^{\beta}\partial t^{\theta}}\,. \tag{9.560}$$

↖ the proof is given in Exercise 9.24

This is exactly the partial differentiation of the Lie derivative of $\underline{\hat{h}}^{\alpha}$. And this helps establish the commutative property

$$\boxed{\mathscr{L}_{\mathbf{u}}\left(\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}\right) = \frac{\partial\left(\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha}\right)}{\partial t^{\beta}}\,.} \tag{9.561}$$

It is important to note that $\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha}$ is a vector but its partial differentiation, in general, is not a vector. ✾

Recall from $(9.140)_1$ that the parallel transport of $\hat{\mathbf{h}}$ along $\mathbf{u}$ was represented by $\hat{\mathbf{h}}\Big|_{\mathbf{u}} = \mathbf{0}$. In a similar manner, a smooth vector field $\underline{\hat{h}}^{\alpha}$ has been *Lie transported* (or *Lie dragged*) with respect to $\mathbf{u}$ if

$$\boxed{\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha} = 0\,.} \tag{9.562}$$

Consider a coordinate system whose coordinate curves are eventually the flow lines of its basis vectors. Suppose that $\mathbf{u} = \mathbf{a}_1$ and $\hat{\mathbf{h}} = \mathbf{a}_2$. One can then immediately deduce that

$$\boxed{\mathscr{L}_{\mathbf{a}_1}\mathbf{a}_2 = \mathbf{0}\,.} \tag{9.563}$$

This means that the tangent vectors of a coordinate system are always Lie transported along the coordinate curves. In other words, the coordinate lines of a coordinate system remain always closed.[19] This represents a major characteristic of a coordinate

[19] Suppose that the integral curves in Fig. 9.32 were the coordinate curves of a coordinate system. In this case, there will be no discrepancy between the points and, therefore, the points S and T

system. But $\mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}}$ does not vanish in general for arbitrary smooth vector fields. This means that the points S and T in Fig. 9.32 should not be identical in general. As a result, a gap may appear. When such a gap exists, the resulting integral curves will not be closed. That is why the Lie derivative of a vector field in the direction of flow of another vector field measures the failure of the resulting integral curves to be properly closed.

As a simple example, let $\hat{\mathbf{h}} = \mathbf{a}_\alpha$. Its Lie derivative with respect to $\mathbf{u}$ then takes the following form

$$\boxed{\mathscr{L}_{\mathbf{u}}\mathbf{a}_\alpha = -\frac{\partial\, \underline{u}^\beta}{\partial t^\alpha}\mathbf{a}_\beta \ .} \tag{9.564}$$

The Lie derivative of the smooth scalar function $\bar{h}\left(t^1, t^2\right)$ with respect to the vector field $\mathbf{u}$ is defined to be its directional derivative:

$$\mathscr{L}_{\mathbf{u}}\bar{h} = D_{\mathbf{u}}\bar{h} = \frac{\partial \bar{h}}{\partial \mathbf{x}} \cdot \mathbf{u} = \frac{\partial \bar{h}}{\partial t^\theta}\, \underline{u}^\theta = \frac{\partial \bar{h}}{\partial t^\theta}\frac{dt^\theta}{d\lambda} = \frac{d\bar{h}}{d\lambda} \ , \tag{9.565}$$

where $(6.11\text{b})_3$, $(9.33)_1$, $(9.80)_1$, $(9.220)_2$ and (9.539) along with the chain rule of differentiation have been used. Now, the Lie transport of the scalar function $\bar{h}$ is indicated by

$$\mathscr{L}_{\mathbf{u}}\bar{h} = 0 \ . \tag{9.566}$$

Let $\bar{h}_1$ and $\bar{h}_2$ be two smooth scalar functions. Then, one can readily verify that the Lie derivative of their product satisfies

$$\boxed{\mathscr{L}_{\mathbf{u}}\left(\bar{h}_1\bar{h}_2\right) = \left(\mathscr{L}_{\mathbf{u}}\bar{h}_1\right)\bar{h}_2 + \bar{h}_1\left(\mathscr{L}_{\mathbf{u}}\bar{h}_2\right) \ .} \quad \longleftarrow \text{see (9.585)} \tag{9.567}$$

Moreover, considering $\underline{\hat{h}}^\alpha$ as a scalar and $\mathbf{a}_\alpha$ as a vector, the Lie derivative of the object $\hat{\mathbf{h}} = \underline{\hat{h}}^\alpha \mathbf{a}_\alpha$ satisfies the product rule in the sense that

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}} &= \left(\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^\alpha\right)\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\left(\mathscr{L}_{\mathbf{u}}\mathbf{a}_\alpha\right) \\
&\xlongequal[\text{(9.565)}]{\text{from}} \left(\underline{u}^\theta\frac{\partial\,\underline{\hat{h}}^\alpha}{\partial t^\theta}\right)\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\left(\mathscr{L}_{\mathbf{u}}\mathbf{a}_\alpha\right) \\
&\xlongequal[\text{(9.564)}]{\text{from}} \underline{u}^\theta\frac{\partial\,\underline{\hat{h}}^\alpha}{\partial t^\theta}\mathbf{a}_\alpha + \underline{\hat{h}}^\alpha\left(-\frac{\partial\,\underline{u}^\beta}{\partial t^\alpha}\mathbf{a}_\beta\right) \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} \left(\underline{u}^\theta\frac{\partial\,\underline{\hat{h}}^\alpha}{\partial t^\theta} - \underline{\hat{h}}^\theta\frac{\partial\,\underline{u}^\alpha}{\partial t^\theta}\right)\mathbf{a}_\alpha \ . \quad \longleftarrow \text{see (9.138)}
\end{aligned} \tag{9.568}$$

should be identical. This means that a particle can freely move a distance $\Delta\lambda$ from P to Q followed by a distance $\Delta\mu$ to the point S. Now, it can move a distance $\Delta\lambda$ from S to R followed by a distance $\Delta\mu$ to arrive at its original position. Since there does not exist any gap between the points, the four points P, Q, S and R define a parallelogram.

Hint: It is worthwhile to mention that the result (9.550) can also be obtained by means of the so-called *pull-back operation*. In this case, one needs to transport the smooth vector $\hat{\mathbf{h}}$ from Q to P. Denoting by $\hat{\mathbf{h}}_{Q\to P}$ the *pull-back* of $\hat{\mathbf{h}}$ from Q to P, the Lie derivative $(9.549)_1$ now translates to

$$\boxed{\mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}} := \lim_{\Delta\lambda\to 0}\frac{\hat{\mathbf{h}}_{Q\to P}-\hat{\mathbf{h}}_P}{\Delta\lambda} \quad \text{or} \quad \mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha} := \lim_{\Delta\lambda\to 0}\frac{\underline{\hat{h}}^{\alpha}_{Q\to P}-\underline{\hat{h}}^{\alpha}_{P}}{\Delta\lambda}\ .} \tag{9.569}$$

In the following, the goal is to compute the Lie derivative of a covector. ❤

9.8.2 Lie Derivative of Covector Fields

By constructing the scalar function $\bar{h} = \underline{\hat{h}}^{\alpha}\,\underline{v}_{\alpha}$, one can use $(9.565)_3$ along with product rule of differentiation to represent

$$\mathscr{L}_{\mathbf{u}}\bar{h} = \frac{\partial\left(\underline{\hat{h}}^{\alpha}\,\underline{v}_{\alpha}\right)}{\partial t^{\theta}}\,\underline{u}^{\theta} = \underbrace{\underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}}}_{\longleftrightarrow}\,\underline{v}_{\alpha} + \underline{u}^{\theta}\,\frac{\partial\,\underline{v}_{\alpha}}{\partial t^{\theta}}\,\underline{\hat{h}}^{\alpha}\ . \tag{9.570}$$

Having in mind $(9.550)_2$, the demand for satisfying the product rule then implies that

$$\begin{aligned}\mathscr{L}_{\mathbf{u}}\bar{h} &= \left(\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha}\right)\underline{v}_{\alpha} + \underline{\hat{h}}^{\alpha}\left(\mathscr{L}_{\mathbf{u}}\,\underline{v}_{\alpha}\right)\\ &= \underbrace{\underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}}\,\underline{v}_{\alpha}}_{\longleftrightarrow} - \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}}\,\underline{v}_{\alpha} + \underline{\hat{h}}^{\alpha}\left(\mathscr{L}_{\mathbf{u}}\,\underline{v}_{\alpha}\right)\ .\end{aligned} \tag{9.571}$$

Comparing $(9.570)_2$ and $(9.571)_2$ now reveals

$$\left(\mathscr{L}_{\mathbf{u}}\,\underline{v}_{\alpha}\right)\underline{\hat{h}}^{\alpha} = \underline{u}^{\theta}\,\frac{\partial\,\underline{v}_{\alpha}}{\partial t^{\theta}}\,\underline{\hat{h}}^{\alpha} + \underline{v}_{\theta}\,\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\alpha}}\,\underline{\hat{h}}^{\alpha}\ .$$

The fact that $\underline{\hat{h}}^{\alpha}$ is entirely arbitrary finally leads to the desired result

$$\boxed{\underbrace{\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}_{\alpha} = \underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}_{\alpha}}{\partial t^{\theta}} + \underline{\hat{h}}_{\theta}\,\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\alpha}}}_{\text{note that } \mathscr{L}_{\mathbf{u}}\hat{\mathbf{h}} = \left(\underline{u}^{\theta}\frac{\partial\,\underline{\hat{h}}_{\alpha}}{\partial t^{\theta}} + \underline{\hat{h}}_{\theta}\frac{\partial\,\underline{u}^{\theta}}{\partial t^{\alpha}}\right)\mathbf{a}^{\alpha}}\ ,} \tag{9.572}$$

which can be rephrased, using (9.156), as

$$\boxed{\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}_{\alpha} = \underline{u}^{\theta}\left(\underline{\hat{h}}_{\alpha}\Big|_{\theta}\right) + \underline{\hat{h}}_{\theta}\left(\underline{u}^{\theta}\Big|_{\alpha}\right)\ .} \quad \longleftarrow \text{see (9.551)} \quad ❤ \tag{9.573}$$

Let $\hat{\mathbf{h}} = \mathbf{a}^{\alpha}$. Then, the Lie derivative of a surface contravariant basis vector with respect to an arbitrary smooth vector field $\mathbf{u}$ takes the form

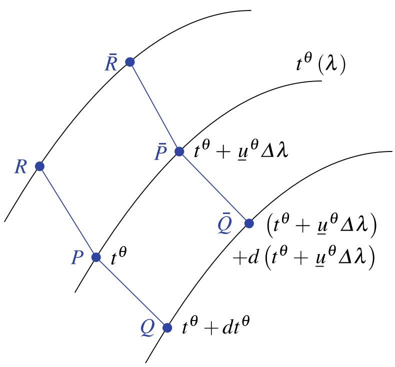
$$\boxed{\mathscr{L}_{\mathbf{u}} \mathbf{a}^{\alpha} = \frac{\partial \, \underline{u}^{\alpha}}{\partial t^{\beta}} \mathbf{a}^{\beta} \, .} \qquad \longleftarrow \text{see (9.564)} \tag{9.574}$$

Consistent with (9.568), one will have

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}} \hat{\mathbf{h}} &= \left(\mathscr{L}_{\mathbf{u}} \, \underline{\hat{h}}_{\alpha} \right) \mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha} \left(\mathscr{L}_{\mathbf{u}} \mathbf{a}^{\alpha} \right) \\
&\underset{\text{(9.565)}}{\overset{\text{from}}{=\!=\!=}} \left(\underline{u}^{\theta} \frac{\partial \, \underline{\hat{h}}_{\alpha}}{\partial t^{\theta}} \right) \mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha} \left(\mathscr{L}_{\mathbf{u}} \mathbf{a}^{\alpha} \right) \\
&\underset{\text{(9.574)}}{\overset{\text{from}}{=\!=\!=}} \underline{u}^{\theta} \frac{\partial \, \underline{\hat{h}}_{\alpha}}{\partial t^{\theta}} \mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha} \left(\frac{\partial \, \underline{u}^{\alpha}}{\partial t^{\beta}} \mathbf{a}^{\beta} \right) \\
&\underset{\text{the dummy indices}}{\overset{\text{by renaming}}{=\!=\!=\!=\!=}} \left(\underline{u}^{\theta} \frac{\partial \, \underline{\hat{h}}_{\alpha}}{\partial t^{\theta}} + \underline{\hat{h}}_{\theta} \frac{\partial \, \underline{u}^{\theta}}{\partial t^{\alpha}} \right) \mathbf{a}^{\alpha} \, .
\end{aligned} \tag{9.575}$$

9.8.3 Space Symmetry and Killing Vector

Let $\mathbf{u}$ be a given smooth vector field which produces the integral curves $t^{\theta}(\lambda)$ as shown in Fig. 9.33. Consider a point P at t^{θ} and an infinitesimally nearby point Q corresponding to $t^{\theta} + dt^{\theta}$. Now, suppose that P flows along $\mathbf{u}$ to $\bar{P}$ with the



Let $\mathbf{u} = \underline{u}^{\theta} \mathbf{a}_{\theta}$ be a smooth vector field which results in the flow lines $t^{\theta}(\lambda)$. Further, let $a_{\alpha\beta}$ be the metric of space. Consider the infinitely near points P and Q which have been displaced, respectively, to the points $\bar{P}$ and $\bar{Q}$. Now, the space under consideration is said to be **symmetric** under translations along the integral curves of $\mathbf{u}$ if, in the limit when $\Delta\lambda \to 0$, $PQ = \bar{P}\bar{Q}$ or

$ds^2 = a_{\alpha\beta}(t^{\theta})\, dt^{\alpha} dt^{\beta} = d\bar{s}^2$

$= a_{\alpha\beta}(t^{\theta} + \underline{u}^{\theta} \Delta\lambda)\, d(t^{\alpha} + \underline{u}^{\alpha} \Delta\lambda)\, d\left(t^{\beta} + \underline{u}^{\beta} \Delta\lambda\right)$,

which finally leads to

$\mathscr{L}_{\mathbf{u}} a_{\alpha\beta} = 0$.

And the vector satisfying this expression is referred to as the **killing vector**.

Fig. 9.33 Space symmetry

coordinates $t^\theta + \underline{u}^\theta \Delta\lambda$. Similarly, suppose that Q flows along this vector field to $\bar{Q}$ with the coordinates $\left(t^\theta + \underline{u}^\theta \Delta\lambda\right) + d\left(t^\theta + \underline{u}^\theta \Delta\lambda\right)$. The (square of) distance between P and Q is then given by

$$ds^2 = a_{\alpha\beta}\left(t^\theta\right) dt^\alpha dt^\beta \ . \tag{9.576}$$

In a similar manner, the distance between $\bar{P}$ and $\bar{Q}$ takes the form

$$d\bar{s}^2 = a_{\alpha\beta}\left(t^\theta + \underline{u}^\theta \Delta\lambda\right) d\left(t^\alpha + \underline{u}^\alpha \Delta\lambda\right) d\left(t^\beta + \underline{u}^\beta \Delta\lambda\right) \ . \tag{9.577}$$

Space is said to be *symmetric* under the action of a smooth vector field if any network of distances remains unchanged when points move along the flow lines of that vector field. This is indicated by

$$\begin{aligned}
\cancel{a_{\alpha\beta} dt^\alpha dt^\beta} &\xlongequal[\text{(9.576) and (9.577)}]{\text{by using}} a_{\alpha\beta}\left(t^\theta + \underline{u}^\theta \Delta\lambda\right) d\left(t^\alpha + \underline{u}^\alpha \Delta\lambda\right) d\left(t^\beta + \underline{u}^\beta \Delta\lambda\right) \\
&\xlongequal[\text{Taylor series expansion}]{\text{by using the first-order}} \left[a_{\alpha\beta} + \frac{\partial a_{\alpha\beta}}{\partial t^\theta}\underline{u}^\theta \Delta\lambda\right]\left[dt^\alpha \right. \\
&\quad \left. + \frac{\partial\,\underline{u}^\alpha}{\partial t^\rho} dt^\rho \Delta\lambda\right]\left[dt^\beta + \frac{\partial\,\underline{u}^\beta}{\partial t^\phi} dt^\phi \Delta\lambda\right] \\
&\xlongequal[\text{the higher-order terms}]{\text{by neglecting}} a_{\alpha\beta} dt^\alpha dt^\beta + \left[\underline{u}^\theta \frac{\partial a_{\alpha\beta}}{\partial t^\theta} dt^\alpha dt^\beta \right. \\
&\quad \left. + a_{\alpha\beta}\frac{\partial\,\underline{u}^\beta}{\partial t^\phi} dt^\alpha dt^\phi + a_{\alpha\beta}\frac{\partial\,\underline{u}^\alpha}{\partial t^\rho} dt^\rho dt^\beta\right]\Delta\lambda \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} \cancel{a_{\alpha\beta} dt^\alpha dt^\beta} + \left[\underline{u}^\theta \frac{\partial a_{\alpha\beta}}{\partial t^\theta} \right. \\
&\quad \left. + a_{\alpha\theta}\frac{\partial\,\underline{u}^\theta}{\partial t^\beta} + a_{\theta\beta}\frac{\partial\,\underline{u}^\theta}{\partial t^\alpha}\right] dt^\alpha dt^\beta \Delta\lambda \ .
\end{aligned}$$

The fact that $\Delta\lambda$ is arbitrary along with $dt^\alpha \neq 0$ then implies that

$$\boxed{\underline{u}^\theta \frac{\partial a_{\alpha\beta}}{\partial t^\theta} + a_{\alpha\theta}\frac{\partial\,\underline{u}^\theta}{\partial t^\beta} + a_{\theta\beta}\frac{\partial\,\underline{u}^\theta}{\partial t^\alpha} = 0 \ \text{ or, by (9.586b), } \ \mathscr{L}_{\mathbf{u}} a_{\alpha\beta} = 0 \ .} \tag{9.578}$$

This important result states that when the Lie derivative of a metric with respect to a smooth vector field vanishes, the geometry described by that metric is **symmetric** along the integral curves of such a vector field. And the special vector which defines such a symmetry is referred to as the *killing vector*. In the literature, it is often denoted by $\mathbf{k}$.

Suppose that the killing vector is a basis vector, say $\mathbf{k} = \mathbf{a}_1$. The symmetry condition (9.578) then takes the following form

$$\mathscr{L}_{\mathbf{a}_1} a_{\alpha\beta} = \frac{\partial a_{\alpha\beta}}{\partial t^1} = 0 \,. \tag{9.579}$$

And this helps conclude that when the covariant metric coefficients are independent of a coordinate variable, the corresponding basis vector then represents a killing vector field. But, one should note that if the covariant metric coefficients do depend on one of the coordinates, the space can still be symmetric under infinitesimal translations along the corresponding coordinate lines.

9.8.4 Lie Derivative of Higher-Order Tensor Fields

The main goal here is to compute the Lie derivative of a second-order tensor field. The results will then be extended to provide the Lie derivative of the surface fourth-order Riemann-Christoffel curvature tensor. The Lie derivative of the Christoffel symbols will also be addressed. And this helps establish some important properties.

To compute the Lie derivative of the tensor $\underline{\tilde{H}}^{\alpha\beta}$, one may construct the scalar function $\bar{h} = \underline{v}_\alpha \, \underline{\tilde{H}}^{\alpha\beta} \, \underline{w}_\beta$. Using (9.565)$_3$ along with the product rule of differentiation, the Lie derivative of this scalar function becomes

$$\begin{aligned} \mathscr{L}_{\mathbf{u}} \bar{h} &= \frac{\partial \left(\underline{v}_\alpha \, \underline{\tilde{H}}^{\alpha\beta} \, \underline{w}_\beta \right)}{\partial t^\theta} \, \underline{u}^\theta \\ &= \underline{u}^\theta \frac{\partial \underline{v}_\alpha}{\partial t^\theta} \, \underline{\tilde{H}}^{\alpha\beta} \, \underline{w}_\beta + \underline{u}^\theta \frac{\partial \, \underline{\tilde{H}}^{\alpha\beta}}{\partial t^\theta} \, \underline{v}_\alpha \, \underline{w}_\beta + \underline{u}^\theta \frac{\partial \underline{w}_\beta}{\partial t^\theta} \, \underline{v}_\alpha \, \underline{\tilde{H}}^{\alpha\beta} \,. \end{aligned} \tag{9.580}$$

Having in mind (9.550)$_2$ and (9.572)$_1$, the demand for satisfying the product rule then helps represent

$$\begin{aligned} \mathscr{L}_{\mathbf{u}} \bar{h} &= \left(\mathscr{L}_{\mathbf{u}} \, \underline{v}_\alpha \right) \underline{\tilde{H}}^{\alpha\beta} \, \underline{w}_\beta + \left(\mathscr{L}_{\mathbf{u}} \, \underline{\tilde{H}}^{\alpha\beta} \right) \underline{v}_\alpha \, \underline{w}_\beta + \left(\mathscr{L}_{\mathbf{u}} \, \underline{w}_\beta \right) \underline{v}_\alpha \, \underline{\tilde{H}}^{\alpha\beta} \\ &= \underline{u}^\theta \frac{\partial \underline{v}_\alpha}{\partial t^\theta} \, \underline{\tilde{H}}^{\alpha\beta} \, \underline{w}_\beta + \underline{v}_\theta \frac{\partial \, \underline{u}^\theta}{\partial t^\alpha} \, \underline{\tilde{H}}^{\alpha\beta} \, \underline{w}_\beta + \left(\mathscr{L}_{\mathbf{u}} \, \underline{\tilde{H}}^{\alpha\beta} \right) \underline{v}_\alpha \, \underline{w}_\beta \\ &\quad + \underline{u}^\theta \frac{\partial \underline{w}_\beta}{\partial t^\theta} \, \underline{v}_\alpha \, \underline{\tilde{H}}^{\alpha\beta} + \underline{w}_\theta \frac{\partial \, \underline{u}^\theta}{\partial t^\beta} \, \underline{v}_\alpha \, \underline{\tilde{H}}^{\alpha\beta} \,. \end{aligned} \tag{9.581}$$

From (9.580)$_2$ and (9.581)$_2$, it follows that

$$\left(\mathscr{L}_{\mathbf{u}} \, \underline{\tilde{H}}^{\alpha\beta} \right) \underline{v}_\alpha \, \underline{w}_\beta = \underline{u}^\theta \frac{\partial \, \underline{\tilde{H}}^{\alpha\beta}}{\partial t^\theta} \, \underline{v}_\alpha \, \underline{w}_\beta - \underline{\tilde{H}}^{\alpha\theta} \frac{\partial \, \underline{u}^\beta}{\partial t^\theta} \, \underline{v}_\alpha \, \underline{w}_\beta - \underline{\tilde{H}}^{\theta\beta} \frac{\partial \, \underline{u}^\alpha}{\partial t^\theta} \, \underline{v}_\alpha \, \underline{w}_\beta \,.$$

Considering the fact that $\underline{v}_\alpha$ and $\underline{w}_\beta$ are arbitrary chosen, one can finally obtain

$$\boxed{\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}^{\alpha\beta} = \underline{u}^\theta \frac{\partial\,\underline{\tilde{H}}^{\alpha\beta}}{\partial t^\theta} - \underline{\tilde{H}}^{\alpha\theta} \frac{\partial\,\underline{u}^\beta}{\partial t^\theta} - \underline{\tilde{H}}^{\theta\beta} \frac{\partial\,\underline{u}^\alpha}{\partial t^\theta}\ .} \tag{9.582}$$

Note that this result could also be attained by constructing $\underline{v}^\alpha = \underline{\tilde{H}}^{\alpha\beta}\,\underline{w}_\beta$. This is left as an exercise to be undertaken by the ambitious reader.

By following similar procedures which led to (9.582), one can arrive at

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}^\alpha_{\ .\beta} = \underline{u}^\theta \frac{\partial\,\underline{\tilde{H}}^\alpha_{\ .\beta}}{\partial t^\theta} + \underline{\tilde{H}}^\alpha_{\ .\theta} \frac{\partial\,\underline{u}^\theta}{\partial t^\beta} - \underline{\tilde{H}}^\theta_{\ .\beta} \frac{\partial\,\underline{u}^\alpha}{\partial t^\theta}\ , \tag{9.583a}$$

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}_\alpha^{\ .\beta} = \underline{u}^\theta \frac{\partial\,\underline{\tilde{H}}_\alpha^{\ .\beta}}{\partial t^\theta} - \underline{\tilde{H}}_\alpha^{\ .\theta} \frac{\partial\,\underline{u}^\beta}{\partial t^\theta} + \underline{\tilde{H}}_\theta^{\ .\beta} \frac{\partial\,\underline{u}^\theta}{\partial t^\alpha}\ , \tag{9.583b}$$

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}_{\alpha\beta} = \underline{u}^\theta \frac{\partial\,\underline{\tilde{H}}_{\alpha\beta}}{\partial t^\theta} + \underline{\tilde{H}}_{\alpha\theta} \frac{\partial\,\underline{u}^\theta}{\partial t^\beta} + \underline{\tilde{H}}_{\theta\beta} \frac{\partial\,\underline{u}^\theta}{\partial t^\alpha}\ . \tag{9.583c}$$

It is now easy to see that

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}^{\alpha\beta} = \underline{u}^\theta \left(\underline{\tilde{H}}^{\alpha\beta}\Big|_\theta \right) - \underline{\tilde{H}}^{\alpha\theta} \left(\underline{u}^\beta\big|_\theta \right) - \underline{\tilde{H}}^{\theta\beta} \left(\underline{u}^\alpha\big|_\theta \right)\ , \tag{9.584a}$$

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}^\alpha_{\ .\beta} = \underline{u}^\theta \left(\underline{\tilde{H}}^\alpha_{\ .\beta}\Big|_\theta \right) + \underline{\tilde{H}}^\alpha_{\ .\theta} \left(\underline{u}^\theta\big|_\beta \right) - \underline{\tilde{H}}^\theta_{\ .\beta} \left(\underline{u}^\alpha\big|_\theta \right)\ , \tag{9.584b}$$

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}_\alpha^{\ .\beta} = \underline{u}^\theta \left(\underline{\tilde{H}}_\alpha^{\ .\beta}\Big|_\theta \right) - \underline{\tilde{H}}_\alpha^{\ .\theta} \left(\underline{u}^\beta\big|_\theta \right) + \underline{\tilde{H}}_\theta^{\ .\beta} \left(\underline{u}^\theta\big|_\alpha \right)\ . \tag{9.584c}$$

$$\mathscr{L}_{\mathbf{u}}\,\underline{\tilde{H}}_{\alpha\beta} = \underline{u}^\theta \left(\underline{\tilde{H}}_{\alpha\beta}\Big|_\theta \right) + \underline{\tilde{H}}_{\alpha\theta} \left(\underline{u}^\theta\big|_\beta \right) + \underline{\tilde{H}}_{\theta\beta} \left(\underline{u}^\theta\big|_\alpha \right)\ . \tag{9.584d}$$

Suppose that the tensor product $\mathbf{v} \otimes \mathbf{w}$ is known in its contravariant components. Then, using (9.582),

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}} \left(\underline{v}^\alpha\,\underline{w}^\beta \right) &= \underline{u}^\theta \frac{\partial\,\underline{v}^\alpha}{\partial t^\theta}\,\underline{w}^\beta + \underline{u}^\theta\,\underline{v}^\alpha \frac{\partial\,\underline{w}^\beta}{\partial t^\theta} - \underline{v}^\alpha\,\underline{w}^\theta \frac{\partial\,\underline{u}^\beta}{\partial t^\theta} - \underline{v}^\theta\,\underline{w}^\beta \frac{\partial\,\underline{u}^\alpha}{\partial t^\theta} \\
&= \left(\underline{u}^\theta \frac{\partial\,\underline{v}^\alpha}{\partial t^\theta} - \underline{v}^\theta \frac{\partial\,\underline{u}^\alpha}{\partial t^\theta} \right) \underline{w}^\beta + \underline{v}^\alpha \left(\underline{u}^\theta \frac{\partial\,\underline{w}^\beta}{\partial t^\theta} - \underline{w}^\theta \frac{\partial\,\underline{u}^\beta}{\partial t^\theta} \right) \\
&= \left(\mathscr{L}_{\mathbf{u}}\,\underline{v}^\alpha \right) \underline{w}^\beta + \underline{v}^\alpha \left(\mathscr{L}_{\mathbf{u}}\,\underline{w}^\beta \right)\ .
\end{aligned}$$

Thus,

$$\boxed{\mathscr{L}_{\mathbf{u}} \left(\mathbf{v} \otimes \mathbf{w} \right) = \left(\mathscr{L}_{\mathbf{u}} \mathbf{v} \right) \otimes \mathbf{w} + \mathbf{v} \otimes \left(\mathscr{L}_{\mathbf{u}} \mathbf{w} \right)\ .} \tag{9.585}$$

$\longleftarrow$ see (9.594)

Using $(9.17)_3$, $(9.24)_3$, $(9.160)_1$, $(9.164)_1$, $(9.164)_3$, $(9.165)_1$, (9.582), (9.583c), (9.584a) and (9.584d), the Lie derivative of the surface metric coefficients will represent

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}} a^{\alpha\beta} &= \underline{u}^{\theta} \frac{\partial a^{\alpha\beta}}{\partial t^{\theta}} - a^{\alpha\theta} \frac{\partial \underline{u}^{\beta}}{\partial t^{\theta}} - a^{\theta\beta} \frac{\partial \underline{u}^{\alpha}}{\partial t^{\theta}} \\
&= -a^{\alpha\theta} \left(\underline{u}^{\beta}\big|_{\theta} \right) - a^{\theta\beta} \left(\underline{u}^{\alpha}\big|_{\theta} \right) \\
&= \boxed{- \underline{u}^{\alpha}\big|^{\beta} - \underline{u}^{\beta}\big|^{\alpha} \ ,}
\end{aligned} \tag{9.586a}$$

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}} a_{\alpha\beta} &= \underline{u}^{\theta} \frac{\partial a_{\alpha\beta}}{\partial t^{\theta}} + a_{\alpha\theta} \frac{\partial \underline{u}^{\theta}}{\partial t^{\beta}} + a_{\theta\beta} \frac{\partial \underline{u}^{\theta}}{\partial t^{\alpha}} \\
&= a_{\alpha\theta} \left(\underline{u}^{\theta}\big|_{\beta} \right) + a_{\theta\beta} \left(\underline{u}^{\theta}\big|_{\alpha} \right) \\
&= \boxed{\underline{u}_{\alpha}\big|_{\beta} + \underline{u}_{\beta}\big|_{\alpha} \ .}
\end{aligned} \tag{9.586b}$$

It is not then difficult to see that

$$\boxed{\mathscr{L}_{\mathbf{u}} \delta^{\alpha}_{\beta} = 0 \ .} \tag{9.587}$$

In accord with (9.557b), the second-order Lie derivative of a tensor is given by

$$\begin{aligned}
\mathscr{L}_{\mathbf{u}} \mathscr{L}_{\mathbf{v}} \underline{\tilde{H}}^{\alpha\beta} &= \underline{u}^{\theta} \frac{\partial}{\partial t^{\theta}} \left[\mathscr{L}_{\mathbf{v}} \underline{\tilde{H}}^{\alpha\beta} \right] - \left[\mathscr{L}_{\mathbf{v}} \underline{\tilde{H}}^{\alpha\theta} \right] \frac{\partial \underline{u}^{\beta}}{\partial t^{\theta}} - \left[\mathscr{L}_{\mathbf{v}} \underline{\tilde{H}}^{\theta\beta} \right] \frac{\partial \underline{u}^{\alpha}}{\partial t^{\theta}} \\
&= \underline{u}^{\theta} \frac{\partial \underline{v}^{\rho}}{\partial t^{\theta}} \frac{\partial \underline{\tilde{H}}^{\alpha\beta}}{\partial t^{\rho}} + \underline{u}^{\theta} \underline{v}^{\rho} \frac{\partial^2 \underline{\tilde{H}}^{\alpha\beta}}{\partial t^{\theta} \partial t^{\rho}} - \underline{u}^{\theta} \frac{\partial \underline{\tilde{H}}^{\alpha\rho}}{\partial t^{\theta}} \frac{\partial \underline{v}^{\beta}}{\partial t^{\rho}} \\
&\quad - \underline{u}^{\theta} \underline{\tilde{H}}^{\alpha\rho} \frac{\partial^2 \underline{v}^{\beta}}{\partial t^{\theta} \partial t^{\rho}} - \underline{u}^{\theta} \frac{\partial \underline{\tilde{H}}^{\rho\beta}}{\partial t^{\theta}} \frac{\partial \underline{v}^{\alpha}}{\partial t^{\rho}} - \underline{u}^{\theta} \underline{\tilde{H}}^{\rho\beta} \frac{\partial^2 \underline{v}^{\alpha}}{\partial t^{\theta} \partial t^{\rho}} \\
&\quad - \underline{v}^{\rho} \frac{\partial \underline{\tilde{H}}^{\alpha\theta}}{\partial t^{\rho}} \frac{\partial \underline{u}^{\beta}}{\partial t^{\theta}} + \underline{\tilde{H}}^{\alpha\rho} \frac{\partial \underline{v}^{\theta}}{\partial t^{\rho}} \frac{\partial \underline{u}^{\beta}}{\partial t^{\theta}} + \underline{\tilde{H}}^{\rho\theta} \frac{\partial \underline{v}^{\alpha}}{\partial t^{\rho}} \frac{\partial \underline{u}^{\beta}}{\partial t^{\theta}} \\
&\quad - \underline{v}^{\rho} \frac{\partial \underline{\tilde{H}}^{\theta\beta}}{\partial t^{\rho}} \frac{\partial \underline{u}^{\alpha}}{\partial t^{\theta}} + \underline{\tilde{H}}^{\theta\rho} \frac{\partial \underline{v}^{\beta}}{\partial t^{\rho}} \frac{\partial \underline{u}^{\alpha}}{\partial t^{\theta}} + \underline{\tilde{H}}^{\rho\beta} \frac{\partial \underline{v}^{\theta}}{\partial t^{\rho}} \frac{\partial \underline{u}^{\alpha}}{\partial t^{\theta}} \ .
\end{aligned} \tag{9.588}$$

The identity (9.558) can now be extended to

$$\boxed{\mathscr{L}_{\mathbf{u}} \mathscr{L}_{\mathbf{v}} \underline{\tilde{H}}^{\alpha\beta} - \mathscr{L}_{\mathbf{v}} \mathscr{L}_{\mathbf{u}} \underline{\tilde{H}}^{\alpha\beta} = \mathscr{L}_{\mathscr{L}_{\mathbf{u}} \mathbf{v}} \underline{\tilde{H}}^{\alpha\beta} \ ,} \tag{9.589}$$

where

$$\begin{aligned}
\mathscr{L}_{\mathscr{L}_{\mathbf{u}} \mathbf{v}} \underline{\tilde{H}}^{\alpha\beta} &= \left[\mathscr{L}_{\mathbf{u}} \underline{v}^{\theta} \right] \frac{\partial \underline{\tilde{H}}^{\alpha\beta}}{\partial t^{\theta}} - \underline{\tilde{H}}^{\alpha\theta} \frac{\partial}{\partial t^{\theta}} \left[\mathscr{L}_{\mathbf{u}} \underline{v}^{\beta} \right] - \underline{\tilde{H}}^{\theta\beta} \frac{\partial}{\partial t^{\theta}} \left[\mathscr{L}_{\mathbf{u}} \underline{v}^{\alpha} \right] \\
&= \underline{u}^{\rho} \frac{\partial \underline{v}^{\theta}}{\partial t^{\rho}} \frac{\partial \underline{\tilde{H}}^{\alpha\beta}}{\partial t^{\theta}} - \underline{v}^{\rho} \frac{\partial \underline{u}^{\theta}}{\partial t^{\rho}} \frac{\partial \underline{\tilde{H}}^{\alpha\beta}}{\partial t^{\theta}} - \underline{\tilde{H}}^{\alpha\theta} \frac{\partial \underline{u}^{\rho}}{\partial t^{\theta}} \frac{\partial \underline{v}^{\beta}}{\partial t^{\rho}} \\
&\quad - \underline{\tilde{H}}^{\alpha\theta} \underline{u}^{\rho} \frac{\partial^2 \underline{v}^{\beta}}{\partial t^{\theta} \partial t^{\rho}} + \underline{\tilde{H}}^{\alpha\theta} \frac{\partial \underline{v}^{\rho}}{\partial t^{\theta}} \frac{\partial \underline{u}^{\beta}}{\partial t^{\rho}} + \underline{\tilde{H}}^{\alpha\theta} \underline{v}^{\rho} \frac{\partial^2 \underline{u}^{\beta}}{\partial t^{\theta} \partial t^{\rho}}
\end{aligned}$$

$$- \underline{\tilde{H}}^{\theta\beta} \frac{\partial\, \underline{u}^{\rho}}{\partial t^{\theta}} \frac{\partial\, \underline{v}^{\alpha}}{\partial t^{\rho}} - \underline{\tilde{H}}^{\theta\beta}\, \underline{u}^{\rho} \frac{\partial^2\, \underline{v}^{\alpha}}{\partial t^{\theta} \partial t^{\rho}} + \underline{\tilde{H}}^{\theta\beta} \frac{\partial\, \underline{v}^{\rho}}{\partial t^{\theta}} \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\rho}} + \underline{\tilde{H}}^{\theta\beta}\, \underline{v}^{\rho} \frac{\partial^2\, \underline{u}^{\alpha}}{\partial t^{\theta} \partial t^{\rho}} \ . \tag{9.590}$$

An extension of (9.561) is

$$\boxed{\mathscr{L}_{\mathbf{u}} \left(\frac{\partial\, \underline{\tilde{H}}^{\alpha\beta}}{\partial t^{\theta}} \right) = \frac{\partial \left(\mathscr{L}_{\mathbf{u}}\, \underline{\tilde{H}}^{\alpha\beta} \right)}{\partial t^{\theta}} \ .} \tag{9.591}$$

The Lie derivative of a fourth-order tensor can be computed by extending the results (9.582)–(9.583c). For instance, the Lie derivative of the Riemann-Christoffel curvature tensor (9.199) is given by

$$\mathscr{L}_{\mathbf{u}}\, \underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\delta} = \underline{u}^{\theta} \frac{\partial\, \underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\delta}}{\partial t^{\theta}} + \underline{\mathbb{R}}^{\alpha}_{\ .\theta\gamma\delta} \frac{\partial\, \underline{u}^{\theta}}{\partial t^{\beta}} + \underline{\mathbb{R}}^{\alpha}_{\ .\beta\theta\delta} \frac{\partial\, \underline{u}^{\theta}}{\partial t^{\gamma}} + \underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\theta} \frac{\partial\, \underline{u}^{\theta}}{\partial t^{\delta}} - \underline{\mathbb{R}}^{\theta}_{\ .\beta\gamma\delta} \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}} \ , \tag{9.592}$$

or, using (9.128) and (9.215a),

$$\mathscr{L}_{\mathbf{u}}\, \underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\delta} = \underline{u}^{\theta} \left(\underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\delta} \big|_{\theta} \right) + \underline{\mathbb{R}}^{\alpha}_{\ .\theta\gamma\delta} \left(\underline{u}^{\theta} \big|_{\beta} \right) + \underline{\mathbb{R}}^{\alpha}_{\ .\beta\theta\delta} \left(\underline{u}^{\theta} \big|_{\gamma} \right) + \underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\theta} \left(\underline{u}^{\theta} \big|_{\delta} \right) - \underline{\mathbb{R}}^{\theta}_{\ .\beta\gamma\delta} \left(\underline{u}^{\alpha} \big|_{\theta} \right) \ . \tag{9.593}$$

It is then a simple exercise to verify that

$$\boxed{\mathscr{L}_{\mathbf{u}} \left(\mathbf{A} \otimes \mathbf{B} \right) = \left(\mathscr{L}_{\mathbf{u}} \mathbf{A} \right) \otimes \mathbf{B} + \mathbf{A} \otimes \left(\mathscr{L}_{\mathbf{u}} \mathbf{B} \right) \ .} \quad \longleftarrow \text{see (9.567)} \tag{9.594}$$

Although the Lie derivative of a tensor field absolutely delivers a tensor field, the Lie derivative of some special quantities which are nontensorial objects may still act as tensors. A well-known example will be the surface Christoffel symbols whose Lie derivative is given by

$$\mathscr{L}_{\mathbf{u}} \Gamma^{\alpha}_{\beta\gamma} = \frac{\partial^2\, \underline{u}^{\alpha}}{\partial t^{\beta} \partial t^{\gamma}} + \underline{u}^{\theta} \frac{\partial \Gamma^{\alpha}_{\beta\gamma}}{\partial t^{\theta}} + \Gamma^{\alpha}_{\beta\theta} \frac{\partial\, \underline{u}^{\theta}}{\partial t^{\gamma}} + \Gamma^{\alpha}_{\theta\gamma} \frac{\partial\, \underline{u}^{\theta}}{\partial t^{\beta}} - \Gamma^{\theta}_{\beta\gamma} \frac{\partial\, \underline{u}^{\alpha}}{\partial t^{\theta}} \ . \quad \nwarrow \text{the proof is given in Exercise 9.24} \tag{9.595}$$

Using (9.168a), (9.169a) and (9.199), this expression can be rewritten as

$$\boxed{\mathscr{L}_{\mathbf{u}} \Gamma^{\alpha}_{\beta\gamma} = \underline{u}^{\alpha} \big|_{\beta\gamma} - \underline{u}^{\theta}\, \underline{\mathbb{R}}^{\alpha}_{\ .\beta\gamma\theta} \ .} \tag{9.596}$$

Some identities can now be established. For instance,

$$\boxed{\underbrace{\mathscr{L}_{\mathbf{u}}\left(\underline{\hat{h}}^{\alpha}\Big|_{\beta}\right)-\left(\mathscr{L}_{\mathbf{u}}\,\underline{\hat{h}}^{\alpha}\right)\Big|_{\beta}}_{:=\,LHS}=\underbrace{\underline{\hat{h}}^{\theta}\,\mathscr{L}_{\mathbf{u}}\Gamma^{\alpha}_{\beta\theta}}_{:=\,RHS}\;,} \tag{9.597}$$

since

$$\begin{aligned}
LHS &\xlongequal[\text{along with (9.551) and (9.584b)}]{\text{by using the product rule}} \underline{u}^{\theta}\left(\underline{\hat{h}}^{\alpha}\Big|_{\beta\theta}\right)+\cancel{\underline{\hat{h}}^{\alpha}\Big|_{\theta}\left(\underline{u}^{\theta}\big|_{\beta}\right)}-\cancel{\underline{\hat{h}}^{\theta}\Big|_{\beta}\left(\underline{u}^{\alpha}\big|_{\theta}\right)} \\
&\quad -\cancel{\underline{u}^{\theta}\big|_{\beta}\left(\underline{\hat{h}}^{\alpha}\Big|_{\theta}\right)}-\underline{u}^{\theta}\left(\underline{\hat{h}}^{\alpha}\Big|_{\theta\beta}\right)+\cancel{\underline{\hat{h}}^{\theta}\Big|_{\beta}\left(\underline{u}^{\alpha}\big|_{\theta}\right)}+\underline{\hat{h}}^{\theta}\left(\underline{u}^{\alpha}\big|_{\theta\beta}\right) \\
&\xlongequal[\text{(9.596)}]{\text{from}} \underline{u}^{\theta}\left(\underline{\hat{h}}^{\alpha}\Big|_{\beta\theta}-\underline{\hat{h}}^{\alpha}\Big|_{\theta\beta}\right)+\underline{\hat{h}}^{\theta}\left(\mathscr{L}_{\mathbf{u}}\Gamma^{\alpha}_{\theta\beta}+\underline{u}^{\rho}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\theta\beta\rho}\right) \\
&\xlongequal[\text{(9.201a) and (9.207)}]{\text{from}} -\underline{\hat{h}}^{\rho}\,\underline{u}^{\theta}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\rho\beta\theta}+\underline{\hat{h}}^{\theta}\,\mathscr{L}_{\mathbf{u}}\Gamma^{\alpha}_{\theta\beta}+\underline{\hat{h}}^{\theta}\,\underline{u}^{\rho}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\theta\beta\rho} \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} \cancel{-\underline{\hat{h}}^{\theta}\,\underline{u}^{\rho}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\theta\beta\rho}}+\underline{\hat{h}}^{\theta}\,\mathscr{L}_{\mathbf{u}}\Gamma^{\alpha}_{\theta\beta}+\cancel{\underline{\hat{h}}^{\theta}\,\underline{u}^{\rho}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\theta\beta\rho}} \\
&\xlongequal[\text{(9.92)}]{\text{from}} RHS\,.
\end{aligned}$$

Moreover,

$$\boxed{\underbrace{\left(\mathscr{L}_{\mathbf{u}}\Gamma^{\alpha}_{\beta\delta}\right)\big|_{\gamma}-\left(\mathscr{L}_{\mathbf{u}}\Gamma^{\alpha}_{\beta\gamma}\right)\big|_{\delta}}_{:=\,LHS}=\underbrace{\mathscr{L}_{\mathbf{u}}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\gamma\delta}}_{:=\,RHS}\;,} \tag{9.598}$$

because

$$\begin{aligned}
LHS &\xlongequal[\text{along with (9.596)}]{\text{by using the product rule}} \left(\underline{u}^{\alpha}\big|_{\beta}\right)\Big|_{\delta\gamma}-\left(\underline{u}^{\alpha}\big|_{\beta}\right)\Big|_{\gamma\delta} \\
&\quad -\left[\underline{u}^{\theta}\big|_{\gamma}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\delta\theta}-\underline{u}^{\theta}\big|_{\delta}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\gamma\theta}\right]-\underline{u}^{\theta}\left[\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\delta\theta}\big|_{\gamma}-\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\gamma\theta}\big|_{\delta}\right] \\
&\xlongequal[\text{(9.207) and (9.212a)}]{\text{from}} \underline{u}^{\theta}\big|_{\beta}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\theta\gamma\delta}-\underline{u}^{\alpha}\big|_{\theta}\,\underline{\mathbb{R}}^{\theta}{}_{\cdot\beta\gamma\delta} \\
&\quad -\left[\underline{u}^{\theta}\big|_{\gamma}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\delta\theta}-\underline{u}^{\theta}\big|_{\delta}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\gamma\theta}\right]-\underline{u}^{\theta}\left[\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\delta\theta}\big|_{\gamma}+\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\theta\gamma}\big|_{\delta}\right] \\
&\xlongequal[\text{(9.207) and (9.214)}]{\text{from}} \underline{u}^{\theta}\big|_{\beta}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\theta\gamma\delta}-\underline{u}^{\alpha}\big|_{\theta}\,\underline{\mathbb{R}}^{\theta}{}_{\cdot\beta\gamma\delta} \\
&\quad +\underline{u}^{\theta}\big|_{\gamma}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\theta\delta}+\underline{u}^{\theta}\big|_{\delta}\,\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\gamma\theta}+\underline{\mathbb{R}}^{\alpha}{}_{\cdot\beta\gamma\delta}\big|_{\theta}\,\underline{u}^{\theta} \\
&\xlongequal[\text{(9.593)}]{\text{from}} RHS\,.
\end{aligned}$$

9.8.5 Lie Derivative of Differential Forms

Algebra and calculus of differential forms have briefly been studied in the previous chapter. As discussed, they have important applications in mathematics as well as mathematical physics and can be viewed as a complement to vector (or tensor) analysis. The goal here is to compute the Lie derivative of such useful mathematical creatures.

Recall that the simplest case of differential forms was a differential 0-form which represented a scalar function. For the two-dimensional **differentiable manifold** under consideration, it represents a function of the Gaussian coordinates.

Denoting by $\overset{0}{\omega} = \overset{0}{\omega}\left(t^1, t^2\right)$ a differential 0-form, its total differential is readily given by

$$\boxed{d\overset{0}{\omega} = \frac{\partial \overset{0}{\omega}}{\partial t^\alpha} dt^\alpha \ .} \quad \longleftarrow \text{see (8.78)} \tag{9.599}$$

The object $\partial \overset{0}{\omega} / \partial t^\alpha$ is basically a **covector**. In accord with (9.572), its Lie derivative is defined by

$$\boxed{\mathscr{L}_{\mathbf{u}} \frac{\partial \overset{0}{\omega}}{\partial t^\alpha} = \underline{u}^\theta \frac{\partial^2 \overset{0}{\omega}}{\partial t^\theta \partial t^\alpha} + \frac{\partial \overset{0}{\omega}}{\partial t^\theta} \frac{\partial \underline{u}^\theta}{\partial t^\alpha} \ ,} \tag{9.600}$$

or

$$\mathscr{L}_{\mathbf{u}} \left(d\overset{0}{\omega}\right) = \left(\underline{u}^\theta \frac{\partial^2 \overset{0}{\omega}}{\partial t^\theta \partial t^\alpha} + \frac{\partial \overset{0}{\omega}}{\partial t^\theta} \frac{\partial \underline{u}^\theta}{\partial t^\alpha} \right) dt^\alpha \ . \tag{9.601}$$

Consistent with (9.574), the Lie derivative of the basis $\{dt^\alpha\}$ takes the following form

$$\boxed{\mathscr{L}_{\mathbf{u}} \left(dt^\alpha\right) = \frac{\partial \underline{u}^\alpha}{\partial t^\beta} dt^\beta = d \, \underline{u}^\alpha \ .} \tag{9.602}$$

It is worth mentioning that the Lie derivative $\mathscr{L}_{\mathbf{u}}$ commutes with the exterior derivative d for the smooth scalar function under consideration:

$$\boxed{\mathscr{L}_{\mathbf{u}} \left(d\overset{0}{\omega}\right) = d \left(\mathscr{L}_{\mathbf{u}} \overset{0}{\omega}\right) \ ,} \tag{9.603}$$

because

$$d \left(\mathscr{L}_{\mathbf{u}} \overset{0}{\omega}\right) \xlongequal[\text{(9.565)}]{\text{by using}} d \left(\underline{u}^\theta \frac{\partial \overset{0}{\omega}}{\partial t^\theta} \right)$$

$$\xlongequal[\text{the product rule}]{\text{by using}} \frac{\partial \overset{0}{\omega}}{\partial t^\theta} d\, \underline{u}^\theta + \underline{u}^\theta d\left(\frac{\partial \overset{0}{\omega}}{\partial t^\theta}\right)$$

$$\xlongequal[\text{the chain rule}]{\text{by using}} \left(\frac{\partial \overset{0}{\omega}}{\partial t^\theta} \frac{\partial\, \underline{u}^\theta}{\partial t^\alpha} + \underline{u}^\theta \frac{\partial^2 \overset{0}{\omega}}{\partial t^\alpha \partial t^\theta}\right) dt^\alpha$$

$$\xlongequal[\text{partial derivative has commutative property}]{\text{by using (9.601) having in mind that the}} \mathscr{L}_{\mathbf{u}}\left(d\overset{0}{\omega}\right) .$$

Let $\overset{1}{\omega} = \overset{1}{\omega}_\alpha dt^\alpha$ be a differential 1-form. Its exterior derivative is then given by

$$\boxed{d\overset{1}{\omega} = d\overset{1}{\omega}_\alpha \wedge dt^\alpha = \frac{\partial \overset{1}{\omega}_\alpha}{\partial t^\beta} dt^\beta \wedge dt^\alpha = \frac{\partial \overset{1}{\omega}_\beta}{\partial t^\alpha} dt^\alpha \wedge dt^\beta\ .} \quad \leftarrow \text{see (8.80)} \tag{9.604}$$

Consistent with (9.583c), its Lie derivative is introduced as

$$\boxed{\mathscr{L}_{\mathbf{u}} \frac{\partial \overset{1}{\omega}_\beta}{\partial t^\alpha} = \underline{u}^\theta \frac{\partial^2 \overset{1}{\omega}_\beta}{\partial t^\theta \partial t^\alpha} + \frac{\partial \overset{1}{\omega}_\theta}{\partial t^\alpha} \frac{\partial\, \underline{u}^\theta}{\partial t^\beta} + \frac{\partial \overset{1}{\omega}_\beta}{\partial t^\theta} \frac{\partial\, \underline{u}^\theta}{\partial t^\alpha}\ ,} \tag{9.605}$$

or

$$\mathscr{L}_{\mathbf{u}}\left(d\overset{1}{\omega}\right) = \left(\underline{u}^\theta \frac{\partial^2 \overset{1}{\omega}_\beta}{\partial t^\theta \partial t^\alpha} + \frac{\partial \overset{1}{\omega}_\theta}{\partial t^\alpha} \frac{\partial\, \underline{u}^\theta}{\partial t^\beta} + \frac{\partial \overset{1}{\omega}_\beta}{\partial t^\theta} \frac{\partial\, \underline{u}^\theta}{\partial t^\alpha}\right) dt^\alpha \wedge dt^\beta\ . \tag{9.606}$$

In accord with (9.603), one can establish

$$\boxed{\mathscr{L}_{\mathbf{u}}\left(d\overset{1}{\omega}\right) = d\left(\mathscr{L}_{\mathbf{u}}\overset{1}{\omega}\right)\ ,} \tag{9.607}$$

since

$$d\left(\mathscr{L}_{\mathbf{u}}\overset{1}{\omega}\right) \xlongequal[\text{of (9.601)}]{\text{in light}} d\left(\left(\underline{u}^\theta \frac{\partial \overset{1}{\omega}_\alpha}{\partial t^\theta} + \overset{1}{\omega}_\theta \frac{\partial\, \underline{u}^\theta}{\partial t^\alpha}\right) dt^\alpha\right)$$

$$\xlongequal[\text{(9.604)}]{\text{from}} \frac{\partial}{\partial t^\beta}\left(\underline{u}^\theta \frac{\partial \overset{1}{\omega}_\alpha}{\partial t^\theta} + \overset{1}{\omega}_\theta \frac{\partial\, \underline{u}^\theta}{\partial t^\alpha}\right) dt^\beta \wedge dt^\alpha$$

$$\xlongequal[\text{the product and chain rules}]{\text{by using}} \left(\frac{\partial\, \underline{u}^\theta}{\partial t^\beta} \frac{\partial \overset{1}{\omega}_\alpha}{\partial t^\theta} + \underline{u}^\theta \frac{\partial^2 \overset{1}{\omega}_\alpha}{\partial t^\beta \partial t^\theta}\right.$$

$$
\begin{aligned}
&\left. + \frac{\partial \overset{1}{\omega}_{\theta}}{\partial t^{\beta}} \frac{\partial \underline{u}^{\theta}}{\partial t^{\alpha}} + \overset{1}{\omega}_{\theta} \frac{\partial^{2} \underline{u}^{\theta}}{\partial t^{\beta} \partial t^{\alpha}} \right) dt^{\beta} \wedge dt^{\alpha} \\
&\xlongequal[\text{of (2.79h)}]{\text{in light}} \left(\frac{\partial \underline{u}^{\theta}}{\partial t^{\beta}} \frac{\partial \overset{1}{\omega}_{\alpha}}{\partial t^{\theta}} + \underline{u}^{\theta} \frac{\partial^{2} \overset{1}{\omega}_{\alpha}}{\partial t^{\beta} \partial t^{\theta}} + \frac{\partial \overset{1}{\omega}_{\theta}}{\partial t^{\beta}} \frac{\partial \underline{u}^{\theta}}{\partial t^{\alpha}} \right) dt^{\beta} \wedge dt^{\alpha} \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} \left(\frac{\partial \underline{u}^{\theta}}{\partial t^{\alpha}} \frac{\partial \overset{1}{\omega}_{\beta}}{\partial t^{\theta}} + \underline{u}^{\theta} \frac{\partial^{2} \overset{1}{\omega}_{\beta}}{\partial t^{\alpha} \partial t^{\theta}} + \frac{\partial \overset{1}{\omega}_{\theta}}{\partial t^{\alpha}} \frac{\partial \underline{u}^{\theta}}{\partial t^{\beta}} \right) dt^{\alpha} \wedge dt^{\beta} \\
&\xlongequal[\text{partial derivative has commutative property}]{\text{by using (9.606) having in mind that the}} \mathscr{L}_{\mathbf{u}} \left(d\overset{1}{\omega} \right) . \qquad (9.608)
\end{aligned}
$$

Let $\overset{1}{\omega} = \overset{1}{\omega}_{\alpha} dt^{\alpha}$ be a differential 1-form and $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha} \mathbf{a}_{\alpha}$ be a vector field. The Lie derivative of their multiplication satisfies the product rule

$$
\boxed{\mathscr{L}_{\mathbf{u}} \left(\overset{1}{\omega} \hat{\mathbf{h}} \right) = \left(\mathscr{L}_{\mathbf{u}} \overset{1}{\omega} \right) \hat{\mathbf{h}} + \overset{1}{\omega} \left(\mathscr{L}_{\mathbf{u}} \hat{\mathbf{h}} \right) ,} \qquad (9.609)
$$

since

$$
\begin{aligned}
\mathscr{L}_{\mathbf{u}} \left(\overset{1}{\omega} \hat{\mathbf{h}} \right) &\xlongequal[\text{of (9.568) and (9.575)}]{\text{in light}} \left(\mathscr{L}_{\mathbf{u}} \overset{1}{\omega}_{\alpha} \right) dt^{\alpha} \hat{\mathbf{h}} + \overset{1}{\omega}_{\alpha} \left(\mathscr{L}_{\mathbf{u}} dt^{\alpha} \right) \hat{\mathbf{h}} \\
&\quad + \overset{1}{\omega} \left(\mathscr{L}_{\mathbf{u}} \underline{\hat{h}}^{\alpha} \right) \mathbf{a}_{\alpha} + \overset{1}{\omega} \underline{\hat{h}}^{\alpha} \left(\mathscr{L}_{\mathbf{u}} \mathbf{a}_{\alpha} \right) \\
&\xlongequal[\text{(9.564), (9.565) and (9.602)}]{\text{from}} \left(\underline{u}^{\theta} \frac{\partial \overset{1}{\omega}_{\alpha}}{\partial t^{\theta}} \right) dt^{\alpha} \hat{\mathbf{h}} + \overset{1}{\omega}_{\alpha} \left(\frac{\partial \underline{u}^{\alpha}}{\partial t^{\beta}} dt^{\beta} \right) \hat{\mathbf{h}} \\
&\quad + \overset{1}{\omega} \left(\underline{u}^{\theta} \frac{\partial \underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} \right) \mathbf{a}_{\alpha} + \overset{1}{\omega} \underline{\hat{h}}^{\alpha} \left(- \frac{\partial \underline{u}^{\beta}}{\partial t^{\alpha}} \mathbf{a}_{\beta} \right) \\
&\xlongequal[\text{the dummy indices}]{\text{by renaming}} \left(\left(\underline{u}^{\theta} \frac{\partial \overset{1}{\omega}_{\alpha}}{\partial t^{\theta}} + \overset{1}{\omega}_{\theta} \frac{\partial \underline{u}^{\theta}}{\partial t^{\alpha}} \right) dt^{\alpha} \right) \hat{\mathbf{h}} \\
&\quad + \overset{1}{\omega} \left(\left(\underline{u}^{\theta} \frac{\partial \underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta} \frac{\partial \underline{u}^{\alpha}}{\partial t^{\theta}} \right) \mathbf{a}_{\alpha} \right) \\
&\xlongequal[\text{in light of (9.601)}]{\text{from (9.550) and}} \left(\mathscr{L}_{\mathbf{u}} \overset{1}{\omega} \right) \hat{\mathbf{h}} + \overset{1}{\omega} \left(\mathscr{L}_{\mathbf{u}} \hat{\mathbf{h}} \right) .
\end{aligned}
$$

It is not then difficult to see that

$$
\boxed{\mathscr{L}_{\mathbf{u}} \left(\overset{1}{\omega} \tilde{\mathbf{H}} \right) = \left(\mathscr{L}_{\mathbf{u}} \overset{1}{\omega} \right) \tilde{\mathbf{H}} + \overset{1}{\omega} \left(\mathscr{L}_{\mathbf{u}} \tilde{\mathbf{H}} \right) .} \qquad (9.610)
$$

The above results can consistently be extended to any space of higher dimension. In this regard, let $\overset{k}{\varsigma}$ be a differential k-form in an n-dimensional space having in mind that $0 \leq k \leq n$. The Lie derivative of $\overset{k}{\varsigma}$ with respect o $\mathbf{u}$ is then defined by

$$\boxed{\mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma} = \left(\mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma}_{i_1 i_2 \cdots i_{k-1} i_k} \right) d\Theta^{i_1} \wedge d\Theta^{i_2} \wedge \cdots \wedge d\Theta^{i_{k-1}} \wedge d\Theta^{i_k} \ ,} \tag{9.611}$$

where

$$\begin{aligned} \mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma}_{i_1 i_2 \cdots i_{k-1} i_k} = \ & \underline{u}^m \frac{\partial \overset{k}{\varsigma}_{i_1 i_2 \cdots i_{k-1} i_k}}{\partial \Theta^m} + \overset{k}{\varsigma}_{m i_2 \cdots i_{k-1} i_k} \frac{\partial \, \underline{u}^m}{\partial \Theta^{i_1}} \\ & + \overset{k}{\varsigma}_{i_1 m \cdots i_{k-1} i_k} \frac{\partial \, \underline{u}^m}{\partial \Theta^{i_2}} + \cdots + \overset{k}{\varsigma}_{i_1 i_2 \cdots i_{k-1} m} \frac{\partial \, \underline{u}^m}{\partial \Theta^{i_k}} \ . \end{aligned} \tag{9.612}$$

Let $\overset{k}{\varsigma}$ and $\overset{k}{\pi}$ be two differential k-forms in a space of dimension n. Further, let $\mathbf{u}$, $\mathbf{v}$ be two smooth vector fields and α, β be two constants. Then, one can readily verify the following properties

$$\mathscr{L}_{\mathbf{u}} \left(\alpha \overset{k}{\varsigma} + \beta \overset{k}{\pi} \right) = \alpha \mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma} + \beta \mathscr{L}_{\mathbf{u}} \overset{k}{\pi} \ , \tag{9.613a}$$

$$\mathscr{L}_{\mathbf{u}} \left(\overset{k}{\varsigma} \wedge \overset{k}{\pi} \right) = \left(\mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma} \right) \wedge \overset{k}{\pi} + \overset{k}{\varsigma} \wedge \left(\mathscr{L}_{\mathbf{u}} \overset{k}{\pi} \right) \ , \tag{9.613b}$$

$$\mathscr{L}_{\mathbf{u}} \left(d \overset{k}{\varsigma} \right) = d \left(\mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma} \right) \ , \tag{9.613c}$$

$$\mathscr{L}_{\mathbf{u}} \mathscr{L}_{\mathbf{v}} \overset{k}{\varsigma} - \mathscr{L}_{\mathbf{v}} \mathscr{L}_{\mathbf{u}} \overset{k}{\varsigma} = \mathscr{L}_{\mathscr{L}_{\mathbf{u}} \mathbf{v}} \overset{k}{\varsigma} \ . \tag{9.613d}$$

9.8.6 Commutator

The Lie derivative of a vector field $\hat{\mathbf{h}}$ with respect to a vector field $\mathbf{u}$ is often introduced by the so-called *commutator* (or *Lie bracket*) of these vectors:

$$\left[\mathbf{u}, \hat{\mathbf{h}} \right] := \mathbf{u} \left(\hat{\mathbf{h}} \right) - \hat{\mathbf{h}} \left(\mathbf{u} \right) \ . \tag{9.614}$$

The key point to characterize the terms $\mathbf{u} \left(\hat{\mathbf{h}} \right)$ (read as $\mathbf{u}$ acting on $\hat{\mathbf{h}}$) and $\hat{\mathbf{h}} \left(\mathbf{u} \right)$ is that the vectors in differential geometry are basically **derivative operators**. In this context, the covariant basis vectors are defined by

$$\mathbf{a}_\alpha := \frac{\partial}{\partial t^\alpha} \ , \tag{9.615}$$

and, accordingly,

$$\mathbf{u} := \underline{u}^{\alpha}\,\frac{\partial}{\partial t^{\alpha}} \quad , \quad \hat{\mathbf{h}} := \underline{\hat{h}}^{\alpha}\,\frac{\partial}{\partial t^{\alpha}} \,. \tag{9.616}$$

The commutator of these two vectors then takes the form

$$\begin{aligned}\left[\mathbf{u}, \hat{\mathbf{h}}\right] &= \underline{u}^{\theta}\,\frac{\partial}{\partial t^{\theta}}\left(\underline{\hat{h}}^{\alpha}\,\frac{\partial}{\partial t^{\alpha}}\right) - \underline{\hat{h}}^{\theta}\,\frac{\partial}{\partial t^{\theta}}\left(\underline{u}^{\alpha}\,\frac{\partial}{\partial t^{\alpha}}\right) \\ &= \underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}}\,\frac{\partial}{\partial t^{\alpha}} + \underbrace{\underline{u}^{\theta}\,\underline{\hat{h}}^{\alpha}\,\frac{\partial^{2}}{\partial t^{\theta}\partial t^{\alpha}}}_{=\,\underline{u}^{\theta}\,\underline{\hat{h}}^{\alpha}\,\frac{\partial^{2}}{\partial t^{\alpha}\partial t^{\theta}}} - \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}}\,\frac{\partial}{\partial t^{\alpha}} - \underbrace{\underline{\hat{h}}^{\theta}\,\underline{u}^{\alpha}\,\frac{\partial^{2}}{\partial t^{\theta}\partial t^{\alpha}}}_{=\,\underline{u}^{\theta}\,\underline{\hat{h}}^{\alpha}\,\frac{\partial^{2}}{\partial t^{\alpha}\partial t^{\theta}}} \\ &= \left(\underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}}\right)\frac{\partial}{\partial t^{\alpha}} \,,\end{aligned} \tag{9.617}$$

or

$$\boxed{\left[\mathbf{u}, \hat{\mathbf{h}}\right]^{\alpha} = \underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} \,.} \tag{9.618}$$

With the aid of $(9.92)_2$, (9.128), $(9.139)_3$ and (9.618), one can establish

$$\boxed{\left[\mathbf{u}, \hat{\mathbf{h}}\right] = \hat{\mathbf{h}}\Big|_{\mathbf{u}} - \mathbf{u}|_{\hat{\mathbf{h}}} \,.} \tag{9.619}$$

This result has schematically been illustrated in Fig. 9.34 (see also Fig. 9.32). The interested reader should consult Hehl and Obukhov [48] for more details.

As an example, the bilinearity property (9.552) in this context reads

$$\left.\begin{aligned}[\mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{w}] &= \alpha\,[\mathbf{u}, \mathbf{v}] + \beta\,[\mathbf{u}, \mathbf{w}] \\ [\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}] &= \alpha\,[\mathbf{u}, \mathbf{w}] + \beta\,[\mathbf{v}, \mathbf{w}]\end{aligned}\right\} \,. \tag{9.620}$$

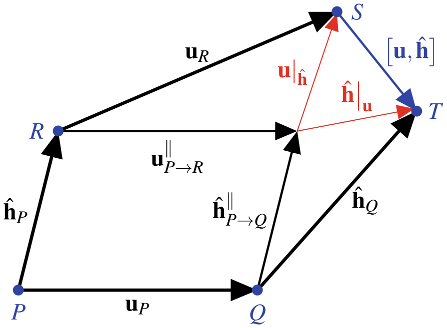
As another example, consider the skew-symmetry property (9.554) which can now be spelled out as

$$\left[\mathbf{u}, \hat{\mathbf{h}}\right] = -\left[\hat{\mathbf{h}}, \mathbf{u}\right] \quad \text{and this immediately implies that} \quad [\mathbf{u}, \mathbf{u}] = \mathbf{0} \,. \tag{9.621}$$

The last example here regards the Jacobi identity (9.555) which now takes the following form

$$[\mathbf{v}, [\mathbf{w}, \mathbf{u}]] + [\mathbf{w}, [\mathbf{u}, \mathbf{v}]] + [\mathbf{u}, [\mathbf{v}, \mathbf{w}]] = \mathbf{0} \,. \tag{9.622}$$

Note that there will be no separation between the parallel transported vectors when the space under consideration is **torsion-free**. Thus, $\hat{\mathbf{h}}^{\parallel}_{P\to Q}$ and $\mathbf{u}^{\parallel}_{P\to R}$ will respectively carry Q and R to the same point.



Let $\mathbf{u}$ and $\hat{\mathbf{h}}$ be two given smooth vector fields. The *commutator* (or *Lie bracket*) of these two vectors is defined as

$$\left[\mathbf{u},\hat{\mathbf{h}}\right] = \hat{\mathbf{h}}\big|_{\mathbf{u}} - \mathbf{u}\big|_{\hat{\mathbf{h}}} = \left(\underline{u}^{\theta}\,\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} - \underline{\hat{h}}^{\theta}\,\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} \right)\mathbf{a}_{\alpha} \ .$$

In this figure,

$$\hat{\mathbf{h}}_P = \underline{\hat{h}}^{\alpha}\mathbf{a}_{\alpha}\ ,\ \hat{\mathbf{h}}^{\parallel}_{P\to Q} = \left(\underline{\hat{h}}^{\alpha} - \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\beta}\,\underline{u}^{\theta}\right)\mathbf{a}_{\alpha}\ ,$$

$$\mathbf{u}_P = \underline{u}^{\alpha}\mathbf{a}_{\alpha}\ ,\ \mathbf{u}^{\parallel}_{P\to R} = \left(\underline{u}^{\alpha} - \Gamma^{\alpha}_{\beta\theta}\,\underline{u}^{\beta}\,\underline{\hat{h}}^{\theta}\right)\mathbf{a}_{\alpha}\ ,$$

$$\hat{\mathbf{h}}\big|_{\mathbf{u}} = \left(\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}} + \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\beta}\right)\underline{u}^{\theta}\mathbf{a}_{\alpha}\ ,$$

$$\mathbf{u}\big|_{\hat{\mathbf{h}}} = \left(\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}} + \Gamma^{\alpha}_{\beta\theta}\,\underline{u}^{\beta}\right)\underline{\hat{h}}^{\theta}\mathbf{a}_{\alpha}\ ,$$

$$\hat{\mathbf{h}}_Q = \left(\underline{\hat{h}}^{\alpha} + \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\theta}}\,\underline{u}^{\theta}\right)\mathbf{a}_{\alpha}\ ,$$

$$\mathbf{u}_R = \left(\underline{u}^{\alpha} + \frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\theta}}\,\underline{\hat{h}}^{\theta}\right)\mathbf{a}_{\alpha}\ .$$

Fig. 9.34 Commutator

9.9 Calculus of Moving Surfaces

So far the calculus of **stationary** curved surfaces has been studied wherein the surface covariant derivative was the central differential operator which helped achieve invariance (recall that the surface partial derivative was not an invariant operator since, in general, it did not preserve the tensor property of its operands). However, there are many applications in science and engineering, such as shape optimization and mechanics of biological membranes, where the two-dimensional manifolds basically represent **deformable** bodies which are moving in the ambient space. An extension of tensor analysis to these dynamically deforming surfaces is thus required. Indeed, the new discipline which characterizes the motion of surfaces is known as the *calculus of moving surfaces*. The surface coordinates and an additional **time** (or, in general, **time-like**) variable constitute the arguments of moving surface quantities. As expected, the main issue in this new language is that the partial time derivative does not provide tensors out of tensors and, therefore, a new derivative operator needs to be established. For the sake of consistency, the ultimate desire is to establish a differential operator possessing all properties of the surface covariant derivative. Such an operator, called *invariant time derivative*, can be defined properly based on a geometric approach without making any reference to coordinate systems. And this guarantees its invariance. This section is thus devoted entirely to the analysis of moving surfaces.

The calculus of moving surfaces was originated by Hadamard [49] and then extended over the years by several authors such as Thomas [50], Grinfeld [51], Grinfeld [52] and Svintradze [53].

9.9.1 Mathematical Description of Moving Surfaces

A two-dimensional surface moving in the ambient space can be seen as a family of surfaces. It may be described by the sufficiently smooth point function

$$\boxed{\mathbf{x} = \hat{\mathbf{x}}\left(\Theta^1, \Theta^2, \Theta^3\right) = \hat{\mathbf{x}}^{\,\mathrm{ms}}\left(t, t^1, t^2\right) \quad \text{for which} \quad \Theta^i = \Theta^i\left(t, t^1, t^2\right) \ ,} \tag{9.623}$$

where Θ^i, $i = 1, 2, 3$, denote the ambient coordinates, t^α, $\alpha = 1, 2$, are the surface coordinates and t presents time (or time-like) variable.

The **ambient velocity** of a particle is then defined by

$$\mathbf{v} = \frac{\partial \mathbf{x}}{\partial t} = \lim_{h \to 0} \frac{\hat{\mathbf{x}}^{\,\mathrm{ms}}\left(t + h, t^1, t^2\right) - \hat{\mathbf{x}}^{\,\mathrm{ms}}\left(t, t^1, t^2\right)}{h} \ , \tag{9.624}$$

and, in alignment with (9.10), the surface natural basis vectors are given by

$$\mathbf{a}_\alpha = \frac{\partial \mathbf{x}}{\partial t^\alpha} = \lim_{h \to 0} \frac{\hat{\mathbf{x}}^{\,\mathrm{ms}}\left(t, t^1 + h\delta^1_\alpha, t^2 + h\delta^2_\alpha\right) - \hat{\mathbf{x}}^{\,\mathrm{ms}}\left(t, t^1, t^2\right)}{h} \ . \tag{9.625}$$

Consequently, an infinitesimal change in the position vector takes the form

$$d\mathbf{x} = \mathbf{v} dt + \mathbf{a}_\alpha dt^\alpha \ . \tag{9.626}$$

One can also write

$$\mathbf{v} \xlongequal[\text{(9.624)}]{\text{by using}} \frac{\partial \mathbf{x}}{\partial t} \xlongequal[\text{with applying the chain rule}]{\text{by using (9.623) along}} \frac{\partial \mathbf{x}}{\partial \Theta^i} \frac{\partial \Theta^i}{\partial t} \xlongequal[\text{(5.3)}]{\text{from}} \frac{\partial \Theta^i}{\partial t} \mathbf{g}_i \ . \tag{9.627}$$

Considering the decompositions (5.64a) and $(9.627)_3$, the **ambient contravariant** (or **natural**) **components** of the velocity vector render

$$\boxed{\underline{v}^i = \frac{\partial \Theta^i}{\partial t} \ .} \tag{9.628}$$

It is worthwhile to point out that the velocity is not a surface vector. The ambient object $\mathbf{v} = \underline{v}^j \mathbf{g}_j$ should thus be expressed with respect to the ambient natural basis $\{\mathbf{a}_i\}$ as

$$\boxed{\mathbf{v} = \underline{v}^{\alpha}\,\mathbf{a}_{\alpha} + \widehat{c}\,\widehat{\mathbf{n}} \quad \text{with} \quad \underline{v}^{\alpha} = \widetilde{Z}^{\alpha}_{j}\,\underline{v}^{j} \quad , \quad \widehat{c} = \underline{v}^{j}\,\underline{\widehat{n}}_{j} \; ,} \tag{9.629}$$

where $\underline{v}^{\alpha}$, $\alpha = 1, 2$, denote the **surface contravariant** (or **natural**) **components** of velocity, $\widetilde{Z}^{\alpha}_{j}$ is the shift tensor defined in (9.60) and $\widehat{c}$ represents the normal component of velocity. This quantity measures the instantaneous velocity of the interface in the normal direction and plays an important role in the calculus of moving surfaces (see Fig. 9.35 for a geometrical interpretation).

In the following, it will be shown that the operator $\partial/\partial t$ does not preserve the tensor property even for invariants. It will also be shown that the variant $\underline{v}^{i}$ does not obey the tensor transformation law while the normal component $\widehat{c}$ does obey (meaning that it is an **invariant** quantity).

To begin with, let $\bar{h} = \bar{h}\left(t, t^1, t^2\right)$ be a smooth scalar field which remains invariant under a change of surface coordinates, i.e.

$$\bar{\bar{h}}\left(t, \bar{t}^{\,1}, \bar{t}^{\,2}\right) = \bar{h}\left(t, t^1, t^2\right) \; . \tag{9.630}$$

In this context, it makes sense to consider **time-dependent** changes of surface coordinates:

$$\bar{t}^{\,\alpha} = \bar{t}^{\,\alpha}\left(t, t^1, t^2\right) \quad , \quad t^{\alpha} = t^{\alpha}\left(t, \bar{t}^{\,1}, \bar{t}^{\,2}\right) \; . \quad \leftarrow \text{see (9.40)} \tag{9.631}$$

Consequently,

$$\bar{\bar{h}}\left(t, \bar{t}^{\,1}, \bar{t}^{\,2}\right) = \bar{h}\left(t, t^1\left(t, \bar{t}^{\,1}, \bar{t}^{\,2}\right), t^2\left(t, \bar{t}^{\,1}, \bar{t}^{\,2}\right)\right) \; . \tag{9.632}$$

Non-invariance of the partial time derivative of the given scalar field is then characterized by the following relation

$$\frac{\partial \bar{\bar{h}}}{\partial t} = \underbrace{\frac{\partial \bar{h}}{\partial t}}_{\text{tensorial part}} + \underbrace{\frac{\partial \bar{h}}{\partial t^{\alpha}}\frac{\partial t^{\alpha}}{\partial t}}_{\text{nontensorial contribution}} \; . \tag{9.633}$$

One can now conclude that $\partial\bar{h}/\partial t$ only remains invariant under a **time-independent** change of surface coordinates for which $\partial t^{\alpha}/\partial t = 0$.

Next, let $\underline{v}^{i}$ be the natural components of $\mathbf{v}$ in an old ambient coordinate system. Further, let $\underline{\bar{v}}^{i}$ be its contravariant components in a new ambient coordinate system. They are given by

$$\underline{\bar{v}}^{i} = \frac{\partial \bar{\Theta}^{i}}{\partial t} \quad \text{where} \quad \bar{\Theta}^{i}\left(t, \bar{t}^{\,1}, \bar{t}^{\,2}\right) = \bar{\Theta}^{i}\left(\Theta^1\left(\Upsilon\right), \Theta^2\left(\Upsilon\right), \Theta^3\left(\Upsilon\right)\right) \; , \tag{9.634}$$

in which

$$\Theta^j(\Upsilon) = \Theta^j\left(t, t^1\left(t, \bar{t}^1, \bar{t}^2\right), t^2\left(t, \bar{t}^1, \bar{t}^2\right)\right) . \tag{9.635}$$

The fact that $\partial/\partial t$ is not an invariant operator implies that the object $\underline{v}^i$ cannot be a tensor. Indeed, it transforms nontensorially according to

$$\underbrace{\bar{\underline{v}}^i = \frac{\partial \bar{\Theta}^i}{\partial t} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial t} + \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial t} = \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\underline{v}^j + \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial t}}_{\text{note that the tensor property is not preserved due to the development of } \frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial t}} . \tag{9.636}$$

It is not then difficult to see that

$$\bar{\underline{v}}^\alpha = \underbrace{\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\underline{v}^\beta}_{\text{tensorial portion}} + \underbrace{\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\frac{\partial t^\beta}{\partial t}}_{\text{nontensorial part}} . \tag{9.637}$$

Finally, let $\widehat{\mathbf{n}} = \bar{\widehat{\underline{n}}}_i \bar{\mathbf{g}}^i = \widehat{\underline{n}}_i \mathbf{g}^i$ be the ambient unit normal field to the surface whose old and new components are tensorially related by

$$\bar{\widehat{\underline{n}}}_i \overset{\text{from}}{\underset{(5.105\text{b})}{=\!=}} \frac{\partial \Theta^k}{\partial \bar{\Theta}^i}\widehat{\underline{n}}_k . \tag{9.638}$$

By means of $(5.56)_4$, $(9.12)_1$, $(9.47)_1$, $(9.47)_3$, $(9.636)_3$ and (9.638) along with the replacement property of the Kronecker delta (and renaming the dummy index k to i), one will have

$$\begin{aligned} \bar{\underline{v}}^i \bar{\widehat{\underline{n}}}_i &= \underbrace{\frac{\partial \Theta^k}{\partial \bar{\Theta}^i}\frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\underline{v}^j \widehat{\underline{n}}_k}_{= \delta^k_j \underline{v}^j \widehat{\underline{n}}_k = \underline{v}^k \widehat{\underline{n}}_k} + \underbrace{\frac{\partial \Theta^k}{\partial \bar{\Theta}^i}\frac{\partial \bar{\Theta}^i}{\partial \Theta^j}\widehat{\underline{n}}_k \frac{\partial \Theta^j}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial t}}_{= \delta^k_j \widehat{\underline{n}}_k Z^j_\alpha \frac{\partial t^\alpha}{\partial t} = \widehat{\underline{n}}_j \left(\mathbf{g}^j \cdot \mathbf{a}_\alpha\right)\frac{\partial t^\alpha}{\partial t} = \left(\widehat{\mathbf{n}} \cdot \mathbf{a}_\alpha\right)\frac{\partial t^\alpha}{\partial t} = 0} \\ &= \underline{v}^i \widehat{\underline{n}}_i . \end{aligned} \tag{9.639}$$

This result states that:

> The normal component of the ambient velocity, $\widehat{c} = \mathbf{v} \cdot \widehat{\mathbf{n}} = \underline{v}^i \widehat{\underline{n}}_i$, will remain **invariant** under a **time-dependent** change of surface coordinates.

9.9.2 *Geometric Approach to Invariant Time Derivative*

Consider a dynamically deforming surface $\mathcal{S}$. Let $\mathcal{S}_t$ be the position of that surface at time t and $\mathcal{S}_{t+h}$ be its position at nearby moment of time $t+h$ as illustrated in Fig. 9.35. Consider a point P on $\mathcal{S}_t$ with $\mathbf{x}(P) = \hat{\mathbf{x}}^{\text{ms}}\left(t, t^1, t^2\right)$. Further, consider a point S on $\mathcal{S}_{t+h}$ with the same surface coordinates as the point P, i.e. $\mathbf{x}(S) = \hat{\mathbf{x}}^{\text{ms}}\left(t+h, t^1, t^2\right)$. The relation $(9.624)_2$ can then be written as

$$\mathbf{x}(S) - \mathbf{x}(P) \approx h\mathbf{v} . \tag{9.640}$$

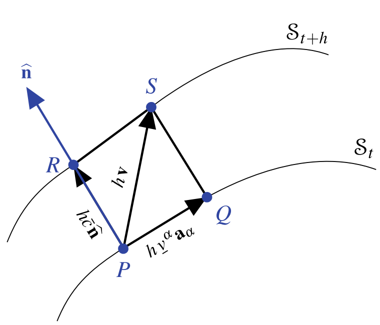
Now, with reference to Fig. 9.35, the goal is to identify two points, namely Q and R, which are crucially important in formulating the invariant time differentiation. When h is small enough, the projection of the ambient object $h\mathbf{v}$ onto the surface natural basis vectors $\mathbf{a}_\alpha$ intersects $\mathcal{S}_t$ at the point Q whose coordinates are given by

$$\mathbf{x}(Q) \approx \mathbf{x}(P) + h\,\underline{v}^\alpha \mathbf{a}_\alpha \quad \text{where} \quad \underline{v}^\alpha = \mathbf{v} \cdot \mathbf{a}^\alpha . \tag{9.641}$$

This result helps realize that Q corresponds to $\left(t, t^1 + h\,\underline{v}^1, t^2 + h\,\underline{v}^2\right)$. Notice that the angles PQS and PRS are approximately $\pi/2$ for sufficiently small amounts of time (or time-like) increment. Consequently,

$$\mathbf{x}(S) - \mathbf{x}(R) \approx \mathbf{x}(Q) - \mathbf{x}(P) , \tag{9.642}$$

The ambient object $\mathbf{v} = \underline{v}^i \mathbf{g}_i$ can also be expressed as $\mathbf{v} = \underline{v}^\alpha \mathbf{a}_\alpha + \widehat{c}\,\widehat{\mathbf{n}}$ where $\widehat{c} = \underline{v}^j\,\widehat{\underline{n}}_j$ measures the instantaneous velocity of the interface in the normal direction. This **invariant** quantity is defined by $\widehat{c} = \lim_{h\to 0} \dfrac{[\mathbf{x}(R) - \mathbf{x}(P)] \cdot \widehat{\mathbf{n}}}{h}$.



Consider a surface $\mathcal{S}$ at two nearby moments of time t and $t+h$ which are denoted by $\mathcal{S}_t$ and $\mathcal{S}_{t+h}$, respectively. Let P be a point on $\mathcal{S}_t$ with $\mathbf{x}(P) = \hat{\mathbf{x}}^{\text{ms}}\left(t, t^1, t^2\right)$ and S be a point on $\mathcal{S}_{t+h}$ with $\mathbf{x}(S) = \hat{\mathbf{x}}^{\text{ms}}\left(t+h, t^1, t^2\right)$. Then, the velocity vector is defined by

$$\mathbf{v} = \lim_{h\to 0} \frac{\mathbf{x}(S) - \mathbf{x}(P)}{h} .$$

This helps obtain $\mathbf{x}(S) - \mathbf{x}(P) \approx h\mathbf{v}$. When h is small enough, the projection of $h\mathbf{v}$ onto the surface covariant basis intersects $\mathcal{S}_t$ at Q with $\mathbf{x}(Q) \approx \mathbf{x}(P) + h\,\underline{v}^\alpha \mathbf{a}_\alpha$. For sufficiently small amounts of h, one will have $\mathbf{x}(S) - \mathbf{x}(R) \approx \mathbf{x}(Q) - \mathbf{x}(P)$ or $\mathbf{x}(R) \approx \mathbf{x}(S) - h\,\underline{v}^\alpha \mathbf{a}_\alpha$. Let $\bar{h} = \bar{h}\left(t, t^1, t^2\right)$ be a smooth scalar field. Its invariant time derivative is then defined by

$$\dot{\nabla}\bar{h} = \lim_{h\to 0} \frac{\bar{h}(R) - \bar{h}(P)}{h} .$$

And this finally leads to

$$\dot{\nabla}\bar{h} = \frac{\partial \bar{h}}{\partial t} - \bar{h}\big|_\alpha\,\underline{v}^\alpha .$$

Fig. 9.35 Invariant time derivative

or, using (9.641),

$$\mathbf{x}\,(R) \approx \mathbf{x}\,(S) - h\,\underline{v}^{\alpha}\mathbf{a}_{\alpha}\ . \tag{9.643}$$

The point R, where the unit normal vector $\widehat{\mathbf{n}}$ to $\mathcal{S}_t$ at P intersects $\mathcal{S}_{t+h}$, thus corresponds to $\left(t+h, t^1 - h\,\underline{v}^1, t^2 - h\,\underline{v}^2\right)$. The points P and R are eventually utilized to define the invariant time derivative. This relies on the following consideration.

Hint: The **invariant** instantaneous velocity of interface in the normal direction has precise geometric definition. According to Fig. 9.35, when h is small enough, the projection of the ambient object $h\mathbf{v}$ along the unit normal vector $\widehat{\mathbf{n}}$ to $\mathcal{S}_t$ at P intersects $\mathcal{S}_{t+h}$ at R with

$$\mathbf{x}\,(R) \approx \mathbf{x}\,(P) + h\widehat{c}\,\widehat{\mathbf{n}} \quad \text{where} \quad \widehat{c} = \mathbf{v}\cdot\widehat{\mathbf{n}}\ . \tag{9.644}$$

And this basically stems from

$$\widehat{c} = \lim_{h\to 0}\frac{\left[\mathbf{x}\,(R) - \mathbf{x}\,(P)\right]\cdot\widehat{\mathbf{n}}}{h}\ . \tag{9.645}$$

The invariant time derivative operator, denoted here by $\dot{\nabla}$, can now be defined properly. This will be demonstrated in the following.

Let $\bar{h}$ be a smooth scalar field of the time (or time-like) variable t and the Gaussian coordinates t^{α}, $\alpha = 1, 2$. The invariant time differentiation of this scalar-valued function is then defined by

$$\underbrace{\dot{\nabla}\bar{h} = \lim_{h\to 0}\frac{\bar{h}\left(t+h, t^1 - h\,\underline{v}^1, t^2 - h\,\underline{v}^2\right) - \bar{h}\left(t, t^1, t^2\right)}{h}}_{\text{note that } \bar{h}\left(t+h,\, t^1 - h\,\underline{v}^1,\, t^2 - h\,\underline{v}^2\right) = \bar{h}\,(R) \text{ and } \bar{h}\left(t,\, t^1,\, t^2\right) = \bar{h}\,(P)\,,\ \text{see Fig. 9.35}}\ . \tag{9.646}$$

It is important to point out that this definition also holds true for any smooth tensor field of arbitrary order. Moreover, this definition relies on a geometric approach in a consistent manner to guarantee its **invariance**, see (9.654).

Introducing the following first-order Taylor series expansion

$$\begin{aligned}\bar{h}\left(t+h, t^1 - h\,\underline{v}^1, t^2 - h\,\underline{v}^2\right) = \bar{h}\left(t, t^1, t^2\right) + \frac{\partial\bar{h}}{\partial t}h - \frac{\partial\bar{h}}{\partial t^1}h\,\underline{v}^1 \\ - \frac{\partial\bar{h}}{\partial t^2}h\,\underline{v}^2 + o\left(h, h\,\underline{v}^1, h\,\underline{v}^2\right)\ ,\end{aligned} \tag{9.647}$$

into (9.646) will lead to

$$\boxed{\dot{\nabla}\bar{h} = \frac{\partial\bar{h}}{\partial t} - \frac{\partial\bar{h}}{\partial t^{\alpha}}\,\underline{v}^{\alpha}\ .} \tag{9.648}$$

Using (9.129), this finally represents

$$\boxed{\dot{\nabla}\bar{h} = \frac{\partial \bar{h}}{\partial t} - \bar{h}\big|_{\alpha}\, \underline{v}^{\alpha}\ .} \tag{9.649}$$

As can be seen, this new time derivative operator is constructed based on a linear combination of the partial and covariant derivatives. Of interest here is to keep this structure for any other invariant field variable such as $\hat{\mathbf{h}}$ and $\tilde{\mathbf{H}}$. The fact that these well-established derivative operators satisfy the sum and product rules naturally implies that

$$\boxed{\left.\begin{aligned} \dot{\nabla}\left(\bar{h}_1 + \bar{h}_2\right) &= \dot{\nabla}\bar{h}_1 + \dot{\nabla}\bar{h}_2 \\ \dot{\nabla}\left(\bar{h}_1\bar{h}_2\right) &= \left(\dot{\nabla}\bar{h}_1\right)\bar{h}_2 + \bar{h}_1\left(\dot{\nabla}\bar{h}_2\right) \end{aligned}\right\}\ .} \tag{9.650}$$

It also satisfies the product rule regarding the single contraction between any two vector fields:[20]

[20] The proof is not difficult. Let $\hat{\mathbf{h}}_1 = \underline{\hat{h}}^i_1 \mathbf{g}_i$ and $\hat{\mathbf{h}}_2 = \underline{\hat{h}}^j_2 \mathbf{g}_j$ be two smooth ambient vector fields whose dot product renders $\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2 = \underline{\hat{h}}^i_1 g_{ij}\, \underline{\hat{h}}^j_2$. Recall from (9.136) or (9.191) that the surface covariant derivative satisfies the product rule when it applies to the scalar product between two vectors. To verify (9.651), one thus only needs to show that

$$\underbrace{\frac{\partial}{\partial t}\left(\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2\right)}_{:=\,LHS} = \underbrace{\frac{\partial \hat{\mathbf{h}}_1}{\partial t} \cdot \hat{\mathbf{h}}_2 + \hat{\mathbf{h}}_1 \cdot \frac{\partial \hat{\mathbf{h}}_2}{\partial t}}_{:=\,RHS}\ .$$

Guided by (9.681), one will have $\partial \hat{\mathbf{h}}_1/\partial t = \left(\partial\, \underline{\hat{h}}^i_1/\partial t + \underline{v}^k \Gamma^i_{km}\, \underline{\hat{h}}^m_1\right)\mathbf{g}_i$. Consequently, RHS takes the form

$$\begin{aligned} \frac{\partial \hat{\mathbf{h}}_1}{\partial t} \cdot \hat{\mathbf{h}}_2 + \hat{\mathbf{h}}_1 \cdot \frac{\partial \hat{\mathbf{h}}_2}{\partial t} &= \frac{\partial\, \underline{\hat{h}}^i_1}{\partial t} g_{ij}\, \underline{\hat{h}}^j_2 + \underline{v}^k \Gamma^i_{km}\, \underline{\hat{h}}^m_1 g_{ij}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 g_{ij} \frac{\partial\, \underline{\hat{h}}^j_2}{\partial t} + \underline{\hat{h}}^i_1 g_{ij}\, \underline{v}^k \Gamma^j_{km}\, \underline{\hat{h}}^m_2 \\ &= \frac{\partial\, \underline{\hat{h}}^i_1}{\partial t} g_{ij}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1\, \underline{v}^k \Gamma^m_{ki} g_{mj}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 g_{ij} \frac{\partial\, \underline{\hat{h}}^j_2}{\partial t} + \underline{\hat{h}}^i_1\, \underline{v}^k \Gamma^m_{kj} g_{im}\, \underline{\hat{h}}^j_2\ . \end{aligned}$$

Now, by using $(9.184)_3$, $(9.685)_3$ and (9.686d), one can arrive at the important result

$$\frac{\partial g_{ij}}{\partial t} = \underline{v}^k \Gamma^m_{ki} g_{mj} + \underline{v}^k \Gamma^m_{kj} g_{im}\ ,$$

which helps represent LHS as

$$\begin{aligned} \frac{\partial}{\partial t}\left(\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2\right) &= \frac{\partial}{\partial t}\left(\underline{\hat{h}}^i_1 g_{ij}\, \underline{\hat{h}}^j_2\right) \\ &= \frac{\partial\, \underline{\hat{h}}^i_1}{\partial t} g_{ij}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 \frac{\partial g_{ij}}{\partial t}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 g_{ij} \frac{\partial\, \underline{\hat{h}}^j_2}{\partial t} \\ &= \frac{\partial\, \underline{\hat{h}}^i_1}{\partial t} g_{ij}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1\, \underline{v}^k \Gamma^m_{ki} g_{mj}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1\, \underline{v}^k \Gamma^m_{kj} g_{im}\, \underline{\hat{h}}^j_2 + \underline{\hat{h}}^i_1 g_{ij} \frac{\partial\, \underline{\hat{h}}^j_2}{\partial t}\ . \end{aligned}$$

And this completes the proof.

$$\boxed{\dot{\nabla}\left(\hat{\mathbf{h}}_1 \cdot \hat{\mathbf{h}}_2\right) = \left(\dot{\nabla}\hat{\mathbf{h}}_1\right) \cdot \hat{\mathbf{h}}_2 + \hat{\mathbf{h}}_1 \cdot \left(\dot{\nabla}\hat{\mathbf{h}}_2\right) \ .} \tag{9.651}$$

Moreover, for any two tensor fields,

$$\boxed{\dot{\nabla}\left(\tilde{\mathbf{H}}_1 : \tilde{\mathbf{H}}_2\right) = \left(\dot{\nabla}\tilde{\mathbf{H}}_1\right) : \tilde{\mathbf{H}}_2 + \tilde{\mathbf{H}}_1 : \left(\dot{\nabla}\tilde{\mathbf{H}}_2\right) \ .} \tag{9.652}$$

The invariant time derivative of the position vector $\mathbf{x} - \mathbf{o}$ with the point function $\mathbf{x} = \hat{\mathbf{x}}\left(\Theta^1, \Theta^2, \Theta^3\right) = \hat{\mathbf{x}}^{\,\mathrm{ms}}\left(t, t^1, t^2\right)$ renders

$$\boxed{\dot{\nabla}\mathbf{x} = \widehat{c}\,\widehat{\mathbf{n}} \ ,} \tag{9.653}$$

because

$$\begin{aligned}
\dot{\nabla}\mathbf{x} &\xlongequal[\text{(9.646)}]{\text{by using}} \lim_{h \to 0} \frac{1}{h}\left[\mathbf{x}\left(t + h, t^1 - h\,\underline{v}^1, t^2 - h\,\underline{v}^2\right) - \mathbf{x}\left(t, t^1, t^2\right)\right] \\
&\xlongequal[\text{Taylor series expansion}]{\text{by using the first-order}} \lim_{h \to 0} \frac{1}{h}\left[\frac{\partial \mathbf{x}}{\partial t}(h) + \frac{\partial \mathbf{x}}{\partial t^\alpha}\left(-h\,\underline{v}^\alpha\right) + o\left(h, h\,\underline{v}^\alpha\right)\right] \\
&\xlongequal[\text{approaches zero faster than } h]{\text{by using the fact that } o} \frac{\partial \mathbf{x}}{\partial t} - \underline{v}^\alpha \frac{\partial \mathbf{x}}{\partial t^\alpha} \\
&\xlongequal[\text{(9.624) and (9.625)}]{\text{by using}} \mathbf{v} - \underline{v}^\alpha \mathbf{a}_\alpha \quad \longleftarrow \text{ note that, according to (9.629), } \mathbf{v} - \underline{v}^\alpha \mathbf{a}_\alpha = \widehat{c}\,\widehat{\mathbf{n}} \\
&\xlongequal[\text{(5.64a) and (9.46)}]{\text{by using}} \left(\underline{v}^i - \underline{v}^\alpha \overline{Z}^{\,i}_\alpha\right)\mathbf{g}_i \\
&\xlongequal[\text{(5.14)}]{\text{by using}} \left(\delta^i_j\,\underline{v}^j - \underline{v}^\alpha \overline{Z}^{\,i}_\alpha\right)\mathbf{g}_i \\
&\xlongequal[\text{(9.74)}]{\text{by using}} \left(\underline{\widehat{n}}^{\,i}\,\underline{\widehat{n}}_j\,\underline{v}^j + \overline{Z}^{\,i}_\alpha \widetilde{Z}^\alpha_j\,\underline{v}^j - \underline{v}^\alpha \overline{Z}^{\,i}_\alpha\right)\mathbf{g}_i \\
&\xlongequal[\text{(9.85a)}]{\text{by using}} \left(\underline{v}^j\,\underline{\widehat{n}}_j\right)\underline{\widehat{n}}^{\,i}\,\mathbf{g}_i \\
&\xlongequal[\text{(5.64a) and (9.629)}]{\text{by using}} \widehat{c}\,\widehat{\mathbf{n}} \ .
\end{aligned}$$

Hint: Recall from (9.633) that $\partial/\partial t$ did not preserve the tensor property even for an invariant of the form (9.632). But, invariance is achieved by the new established derivative operator $\dot{\nabla}$. This can be verified as follows:

$$\begin{aligned}
\dot{\nabla}\overset{=}{\bar{h}} &\xlongequal[\text{(9.648)}]{\text{by using}} \left(\frac{\partial \overset{=}{\bar{h}}}{\partial t}\right) - \frac{\partial \overset{=}{\bar{h}}}{\partial \bar{t}^{\,\alpha}}\left[\underline{\bar{v}}^\alpha\right] \\
&\xlongequal[\text{(9.633) and (9.637)}]{\text{by using (9.632),}} \left(\frac{\partial \bar{h}}{\partial t} + \frac{\partial \bar{h}}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial t}\right) - \frac{\partial \bar{h}}{\partial t^\gamma}\frac{\partial t^\gamma}{\partial \bar{t}^{\,\alpha}}\left[\frac{\partial \bar{t}^{\,\alpha}}{\partial t^\beta}\left\{\underline{v}^\beta + \frac{\partial t^\beta}{\partial t}\right\}\right]
\end{aligned}$$

$$\xlongequal[\text{indices and using the chain rule}]{\text{by renaming the dummy}} \frac{\partial \bar{h}}{\partial t} + \frac{\partial \bar{h}}{\cancel{\partial t^{\beta}}} \frac{\cancel{\partial t^{\beta}}}{\partial t} - \frac{\partial \bar{h}}{\partial t^{\alpha}} \underline{v}^{\alpha} - \frac{\partial \bar{h}}{\cancel{\partial t^{\beta}}} \frac{\cancel{\partial t^{\beta}}}{\partial t}$$
$$\xlongequal[\text{(9.648)}]{\text{by using}} \dot{\nabla} \bar{h} \ . \tag{9.654}$$

9.9.3 *Invariant Time Derivative of Objects with Surface Indices*

In accord with the partial and covariant derivatives, this new derivative operator is supposed to satisfy the product rule when it applies to an invariant combination of the surface components and basis vectors. This leads to its metrinilic property with respect to the surface metric coefficients (although this does not hold true for the surface basis vectors). The main goal here is thus to represent the invariant time derivative of the surface vector and tensor fields.

9.9.3.1 Partial Time Differentiation of Surface Basis Vectors and Metric Coefficients

The time rate of change of a surface covariant basis vector is given by

$$\frac{\partial \mathbf{a}_{\alpha}}{\partial t} \xlongequal[\text{with applying the product rule}]{\text{by using (9.46) along}} \frac{\partial \overline{Z}_{\alpha}^{i}}{\partial t} \mathbf{g}_i + \overline{Z}_{\alpha}^{i} \frac{\partial \mathbf{g}_i}{\partial t}$$
$$\xlongequal[\text{with applying the chain rule}]{\text{by using (9.47) along}} \left(\frac{\partial}{\partial t} \frac{\partial \Theta^i}{\partial t^{\alpha}} = \frac{\partial}{\partial t^{\alpha}} \frac{\partial \Theta^i}{\partial t} \right) \mathbf{g}_i + \overline{Z}_{\alpha}^{i} \left(\frac{\partial \mathbf{g}_i}{\partial \Theta^j} \right) \left(\frac{\partial \Theta^j}{\partial t} \right)$$
$$\xlongequal[\text{(7.8) and (9.628)}]{\text{by using}} \left(\frac{\partial \underline{v}^i}{\partial t^{\alpha}} \right) \mathbf{g}_i + \overline{Z}_{\alpha}^{i} \left(\Gamma_{ij}^{k} \mathbf{g}_k \right) \left(\underline{v}^j \right)$$
$$\xlongequal[\text{the dummy indices}]{\text{by renaming}} \left(\frac{\partial \underline{v}^i}{\partial t^{\alpha}} + \overline{Z}_{\alpha}^{j} \Gamma_{jm}^{i} \underline{v}^m \right) \mathbf{g}_i$$
$$\xlongequal[\text{(9.178)}]{\text{by using}} \mathbf{v}|_{\alpha} \ . \tag{9.655}$$

Using (9.176a), (9.177a), (9.179), $(9.185)_2$ and $(9.629)_1$, the result $(9.655)_5$ can further be written as

$$\boxed{\frac{\partial \mathbf{a}_{\alpha}}{\partial t} = \left(\dot{\Gamma}_{\alpha}^{\theta} \right) \mathbf{a}_{\theta} + \left(\underline{v}^{\theta} \, \underline{b}_{\theta\alpha} + \widehat{c}\,|_{\alpha} \right) \widehat{\mathbf{n}} \ ,} \tag{9.656}$$

where

$$\boxed{\dot{\Gamma}_{\alpha}^{\theta} := \underline{v}^{\theta}\big|_{\alpha} - \widehat{c} \, \underline{b}^{\theta}_{\ .\alpha} = \underline{v}^{\theta}\big|_{\alpha} - \widehat{c} \, \underline{b}_{\alpha}^{\ .\theta} \ .} \tag{9.657}$$

The relation (9.656) helps calculate $\partial a_{\alpha\beta}/\partial t = \partial\left(\mathbf{a}_\alpha \cdot \mathbf{a}_\beta\right)/\partial t$ by means of the product rule of differentiation. This basically requires dotting its both sides with $\mathbf{a}_\beta$. Having in mind $\widehat{\mathbf{n}} \cdot \mathbf{a}_\beta = 0$, one can write

$$\underbrace{\frac{\partial \mathbf{a}_\alpha}{\partial t} \cdot \mathbf{a}_\beta = \left(\dot{\Gamma}^{\theta}_{\alpha}\right) a_{\theta\beta}}_{\text{note that } \mathbf{a}_\alpha \cdot \frac{\partial \mathbf{a}_\beta}{\partial t} = \left(\dot{\Gamma}^{\theta}_{\beta}\right) a_{\theta\alpha}} \quad . \tag{9.658}$$

By means of (9.657)-(9.658), taking into account $a_{\theta\beta}\big|_\alpha = 0$, $\underline{v}_\beta = \underline{v}^\theta a_{\theta\beta}$ and $\underline{b}_{\alpha\beta} = \underline{b}_{\alpha}^{\cdot\theta} a_{\theta\beta} = \underline{b}_{\beta}^{\cdot\theta} a_{\theta\alpha}$, one can obtain the desired result

$$\boxed{\frac{\partial a_{\alpha\beta}}{\partial t} = \underline{v}_\alpha\big|_\beta + \underline{v}_\beta\big|_\alpha - 2\,\widehat{c}\;\underline{b}_{\alpha\beta}\ .} \tag{9.659}$$

Attention now will be on computing the time rate of change of the surface contravariant metric coefficients. The partial time derivative of $a^{\alpha\theta} a_{\theta\beta} = \delta^\alpha_\beta$, by applying the product rule of differentiation, gives the useful identity

$$\frac{\partial a^{\alpha\beta}}{\partial t} = -a^{\alpha\theta}\frac{\partial a_{\theta\rho}}{\partial t}a^{\rho\beta} \quad \text{where} \quad \frac{\partial a_{\theta\rho}}{\partial t} = \underline{v}_\theta\big|_\rho + \underline{v}_\rho\big|_\theta - 2\,\widehat{c}\;\underline{b}_{\theta\rho}\ . \tag{9.660}$$

It is then easy to see that

$$\boxed{\frac{\partial a^{\alpha\beta}}{\partial t} = 2\widehat{c}\;\underline{b}^{\alpha\beta} - \underline{v}^\alpha\big|^\beta - \underline{v}^\beta\big|^\alpha\ .} \tag{9.661}$$

The time rate of change of a surface contravariant basis vector is now formulated. This requires calculating $\partial\mathbf{a}^\alpha/\partial t = \partial\left(a^{\alpha\beta}\mathbf{a}_\beta\right)/\partial t$ upon use of the product rule of differentiation. By virtue of (9.656) and (9.661) along with index juggling, the desired result takes the form

$$\boxed{\frac{\partial \mathbf{a}^\alpha}{\partial t} = \left(-\dot{\Gamma}^{\alpha}_{\theta}\right)\mathbf{a}^\theta + \left(\underline{b}^{\alpha}_{\cdot\theta}\,\underline{v}^\theta + \widehat{c}\,|^\alpha\right)\widehat{\mathbf{n}}\ .} \tag{9.662}$$

In the following, the goal is to show that the time derivative operator $\dot{\nabla}$ is metrinilic with respect to the surface objects $\varepsilon^{\alpha\beta}/\sqrt{a}$ and $\sqrt{a}\varepsilon_{\alpha\beta}$.

To begin with, the sensitivity of the determinant of the covariant metric coefficients with respect to time is computed. By means of the chain rule, the partial time derivative of the determinant of $a_{\alpha\beta}$ becomes $\partial a/\partial t = \left(\partial a/\partial a_{\alpha\beta}\right)\left(\partial a_{\alpha\beta}/\partial t\right)$. In light of (6.20c) and (9.26) along with $a^{\beta\alpha} = a^{\alpha\beta}$, one can infer that $\partial a/\partial a_{\alpha\beta} = a\,a^{\alpha\beta}$. It is then easy to see that

$$\boxed{\frac{\partial a}{\partial t} = 2a\left(\underline{v}^{\gamma}\big|_{\gamma} - \widehat{c}\ \underline{b}^{\gamma}_{.\gamma} \right) = 2a\left(\dot{\Gamma}^{\gamma}_{\gamma}\right) .} \tag{9.663}$$

Consequently, one can trivially write

$$\frac{\partial}{\partial t}\left[\frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\right] = -\frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\left(\dot{\Gamma}^{\gamma}_{\gamma}\right) \quad , \quad \frac{\partial}{\partial t}\left[\sqrt{a}\varepsilon_{\alpha\beta}\right] = \sqrt{a}\varepsilon_{\alpha\beta}\left(\dot{\Gamma}^{\gamma}_{\gamma}\right) . \tag{9.664}$$

Recall from (9.57)$_2$ that $\mathbf{a}_\alpha \times \mathbf{a}_\beta = \sqrt{a}\varepsilon_{\alpha\beta}\widehat{\mathbf{n}}$ or $\sqrt{a}\varepsilon_{\alpha\beta} = \left(\mathbf{a}_\alpha \times \mathbf{a}_\beta\right) \cdot \widehat{\mathbf{n}}$ having in mind the identity $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$. The fact that $\mathbf{a}_\alpha \times \mathbf{a}_\beta$ points in the normal direction and $\partial\widehat{\mathbf{n}}/\partial t$, guided by (9.689)$_2$, lies in the tangent plane helps one to write

$$\frac{\partial}{\partial t}\left[\sqrt{a}\varepsilon_{\alpha\beta}\right] = \left(\frac{\partial \mathbf{a}_\alpha}{\partial t} \times \mathbf{a}_\beta\right) \cdot \widehat{\mathbf{n}} + \left(\mathbf{a}_\alpha \times \frac{\partial \mathbf{a}_\beta}{\partial t}\right) \cdot \widehat{\mathbf{n}} . \tag{9.665}$$

From (9.57)$_2$, (9.58) and (9.656), one then obtains

$$\frac{\partial \mathbf{a}_\alpha}{\partial t} \times \mathbf{a}_\beta = \left(\dot{\Gamma}^{\theta}_{\alpha}\right)\sqrt{a}\varepsilon_{\theta\beta}\widehat{\mathbf{n}} + \left(\underline{b}_{\alpha\theta}\ \underline{v}^{\theta} + \widehat{c}\,\big|_{\alpha}\right)\sqrt{a}\varepsilon_{\beta\rho}\mathbf{a}^{\rho} , \tag{9.666}$$

and, consequently,

$$\underbrace{\left(\frac{\partial \mathbf{a}_\alpha}{\partial t} \times \mathbf{a}_\beta\right) \cdot \widehat{\mathbf{n}} = \left(\dot{\Gamma}^{\theta}_{\alpha}\right)\sqrt{a}\varepsilon_{\theta\beta}}_{\text{in a similar manner, } \left(\mathbf{a}_\alpha \times \frac{\partial \mathbf{a}_\beta}{\partial t}\right) \cdot \widehat{\mathbf{n}} = \left(\dot{\Gamma}^{\theta}_{\beta}\right)\sqrt{a}\varepsilon_{\alpha\theta}} . \tag{9.667}$$

With the aid of (9.667)$_{1-2}$, the expression (9.665) takes the form

$$\frac{\partial}{\partial t}\left[\sqrt{a}\varepsilon_{\alpha\beta}\right] = \sqrt{a}\,\dot{\Gamma}^{\theta}_{\alpha}\varepsilon_{\theta\beta} + \sqrt{a}\,\dot{\Gamma}^{\theta}_{\beta}\varepsilon_{\alpha\theta} . \tag{9.668}$$

Comparing (9.664)$_2$ and (9.668) now reveals

$$\boxed{\varepsilon_{\alpha\beta}\dot{\Gamma}^{\gamma}_{\gamma} - \dot{\Gamma}^{\theta}_{\alpha}\varepsilon_{\theta\beta} - \dot{\Gamma}^{\theta}_{\beta}\varepsilon_{\alpha\theta} = 0 .} \tag{9.669}$$

Guided by (9.166)$_2$ and (9.676d), this result basically verifies that $\dot{\nabla}$ is **metrinilic** with respect to $\sqrt{a}\varepsilon_{\alpha\beta}$. Following similar procedures then lead to the metrinilic property of $\dot{\nabla}$ relative to the surface object $\varepsilon^{\alpha\beta}/\sqrt{a}$. Thus,

$$\boxed{\dot{\nabla}\left(\frac{\varepsilon^{\alpha\beta}}{\sqrt{a}}\right) = 0 \quad , \quad \dot{\nabla}\left(\sqrt{a}\varepsilon_{\alpha\beta}\right) = 0 .} \tag{9.670}$$

9.9.3.2 Invariant Time Differentiation of Surface Vector and Tensor Fields

The established relations regarding the partial time derivative of the surface basis vectors and metric coefficients are now utilized to represent the invariant time derivative of the surface vector and tensor fields as well as the basis vectors.

To begin with, consider an invariant surface object $\hat{\mathbf{h}} = \underline{\hat{h}}^{\alpha} \mathbf{a}_{\alpha}$ for which

$$\begin{aligned}
\dot{\nabla}\hat{\mathbf{h}} &\xlongequal[(9.649)]{\text{by using}} \frac{\partial \hat{\mathbf{h}}}{\partial t} - \left(\hat{\mathbf{h}} \Big|_{\theta} \right) \underline{v}^{\theta} \\
&\xlongequal[\text{the product rule}]{\text{by applying}} \frac{\partial \, \underline{\hat{h}}^{\alpha}}{\partial t} \mathbf{a}_{\alpha} + \underline{\hat{h}}^{\alpha} \frac{\partial \mathbf{a}_{\alpha}}{\partial t} - \left(\underline{\hat{h}}^{\alpha} \Big|_{\theta} \mathbf{a}_{\alpha} \right) \underline{v}^{\theta} - \underline{\hat{h}}^{\alpha} \left(\mathbf{a}_{\alpha}|_{\theta} \right) \underline{v}^{\theta} \\
&\xlongequal[(9.177a)]{\text{by using}} \left(\frac{\partial \, \underline{\hat{h}}^{\alpha}}{\partial t} - \underline{\hat{h}}^{\alpha} \Big|_{\theta} \underline{v}^{\theta} \right) \mathbf{a}_{\alpha} + \underline{\hat{h}}^{\alpha} \frac{\partial \mathbf{a}_{\alpha}}{\partial t} - \underline{\hat{h}}^{\alpha} \, \underline{b}_{\alpha\theta} \, \underline{v}^{\theta} \, \widehat{\mathbf{n}} \\
&\xlongequal[(9.656)]{\text{by using}} \left(\frac{\partial \, \underline{\hat{h}}^{\alpha}}{\partial t} - \underline{\hat{h}}^{\alpha} \Big|_{\theta} \underline{v}^{\theta} \right) \mathbf{a}_{\alpha} + \left(\dot{\Gamma}^{\theta}_{\alpha} \, \underline{\hat{h}}^{\alpha} \right) \mathbf{a}_{\theta} \\
&\quad + \left(\underline{v}^{\theta} \, \underline{b}_{\theta\alpha} \, \underline{\hat{h}}^{\alpha} + \widehat{c}|_{\alpha} \, \underline{\hat{h}}^{\alpha} - \underline{\hat{h}}^{\alpha} \, \underline{b}_{\alpha\theta} \, \underline{v}^{\theta} \right) \widehat{\mathbf{n}} \\
&\xlongequal[\text{renaming the dummy indices}]{\text{by using (9.95) along with}} \left(\frac{\partial \, \underline{\hat{h}}^{\alpha}}{\partial t} + \dot{\Gamma}^{\alpha}_{\theta} \, \underline{\hat{h}}^{\theta} - \underline{\hat{h}}^{\alpha} \Big|_{\theta} \underline{v}^{\theta} \right) \mathbf{a}_{\alpha} \\
&\quad + \widehat{c}|_{\alpha} \, \underline{\hat{h}}^{\alpha} \widehat{\mathbf{n}} \, .
\end{aligned} \tag{9.671}$$

The demand for satisfying the sum and product rules

$$\dot{\nabla}\hat{\mathbf{h}} = \left(\dot{\nabla} \, \underline{\hat{h}}^{\alpha} \right) \mathbf{a}_{\alpha} + \underline{\hat{h}}^{\alpha} \left(\dot{\nabla} \mathbf{a}_{\alpha} \right) , \quad \longleftarrow \text{see (9.176a)} \tag{9.672}$$

then helps establish

$$\boxed{\dot{\nabla} \, \underline{\hat{h}}^{\alpha} = \frac{\partial \, \underline{\hat{h}}^{\alpha}}{\partial t} + \dot{\Gamma}^{\alpha}_{\theta} \, \underline{\hat{h}}^{\theta} - \underline{\hat{h}}^{\alpha} \Big|_{\theta} \underline{v}^{\theta} ,} \quad \longleftarrow \text{see (9.128)} \tag{9.673a}$$

$$\boxed{\dot{\nabla} \mathbf{a}_{\alpha} = \widehat{c}|_{\alpha} \, \widehat{\mathbf{n}} \, .} \quad \longleftarrow \text{see (9.177a)} \tag{9.673b}$$

Next, consider an invariant of the form $\hat{\mathbf{h}} = \underline{\hat{h}}_{\alpha} \mathbf{a}^{\alpha}$ for which

$$\begin{aligned}
\dot{\nabla}\hat{\mathbf{h}} &= \frac{\partial \hat{\mathbf{h}}}{\partial t} - \left(\hat{\mathbf{h}} \Big|_{\theta} \right) \underline{v}^{\theta} \\
&= \frac{\partial \, \underline{\hat{h}}_{\alpha}}{\partial t} \mathbf{a}^{\alpha} + \underline{\hat{h}}_{\alpha} \frac{\partial \mathbf{a}^{\alpha}}{\partial t} - \left(\underline{\hat{h}}_{\alpha} \Big|_{\theta} \mathbf{a}^{\alpha} \right) \underline{v}^{\theta} - \underline{\hat{h}}_{\alpha} \left(\mathbf{a}^{\alpha}|_{\theta} \right) \underline{v}^{\theta}
\end{aligned}$$

$$= \left(\frac{\partial \, \underline{\hat{h}}_\alpha}{\partial t} - \underline{\hat{h}}_\alpha \Big|_\theta \, \underline{v}^\theta \right) \mathbf{a}^\alpha + \underline{\hat{h}}_\alpha \frac{\partial \mathbf{a}^\alpha}{\partial t} - \underline{\hat{h}}_\alpha \, \underline{b}^\alpha_{.\theta} \, \underline{v}^\theta \widehat{\mathbf{n}}$$
$$= \left(\frac{\partial \, \underline{\hat{h}}_\alpha}{\partial t} - \underline{\hat{h}}_\alpha \Big|_\theta \, \underline{v}^\theta \right) \mathbf{a}^\alpha + \left(-\dot{\Gamma}^\alpha_\theta \, \underline{\hat{h}}_\alpha \right) \mathbf{a}^\theta$$
$$+ \left(\underline{\hat{h}}_\alpha \, \underline{b}^\alpha_{.\theta} \, \underline{v}^\theta + \widehat{c}|^\alpha \, \underline{\hat{h}}_\alpha - \underline{\hat{h}}_\alpha \, \underline{b}^\alpha_{.\theta} \, \underline{v}^\theta \right) \widehat{\mathbf{n}}$$
$$= \left(\frac{\partial \, \underline{\hat{h}}_\alpha}{\partial t} - \dot{\Gamma}^\theta_\alpha \, \underline{\hat{h}}_\theta - \underline{\hat{h}}_\alpha \Big|_\theta \, \underline{v}^\theta \right) \mathbf{a}^\alpha + \widehat{c}|^\alpha \, \underline{\hat{h}}_\alpha \widehat{\mathbf{n}} \,, \tag{9.674}$$

implies that

$$\boxed{\dot{\nabla} \, \underline{\hat{h}}_\alpha = \frac{\partial \, \underline{\hat{h}}_\alpha}{\partial t} - \dot{\Gamma}^\theta_\alpha \, \underline{\hat{h}}_\theta - \underline{\hat{h}}_\alpha \Big|_\theta \, \underline{v}^\theta \,,} \quad \longleftarrow \text{see (9.156)} \tag{9.675a}$$
$$\boxed{\dot{\nabla} \mathbf{a}^\alpha = \widehat{c}|^\alpha \, \widehat{\mathbf{n}} \,.} \quad \longleftarrow \text{see (9.177b)} \tag{9.675b}$$

Finally, let $\tilde{\mathbf{H}}$ be a surface tensor field with the decompositions $(9.82)_{1-4}$. Then,

$$\dot{\nabla} \, \underline{\tilde{H}}^{\alpha\beta} = \frac{\partial \, \underline{\tilde{H}}^{\alpha\beta}}{\partial t} + \dot{\Gamma}^\alpha_\theta \, \underline{\tilde{H}}^{\theta\beta} + \dot{\Gamma}^\beta_\theta \, \underline{\tilde{H}}^{\alpha\theta} - \underline{\tilde{H}}^{\alpha\beta} \Big|_\theta \, \underline{v}^\theta \,, \quad \longleftarrow \text{see (9.162a)} \tag{9.676a}$$
$$\dot{\nabla} \, \underline{\tilde{H}}^\alpha_{\ .\beta} = \frac{\partial \, \underline{\tilde{H}}^\alpha_{\ .\beta}}{\partial t} + \dot{\Gamma}^\alpha_\theta \, \underline{\tilde{H}}^\theta_{\ .\beta} - \dot{\Gamma}^\theta_\beta \, \underline{\tilde{H}}^\alpha_{\ .\theta} - \underline{\tilde{H}}^\alpha_{\ .\beta} \Big|_\theta \, \underline{v}^\theta \,, \tag{9.676b}$$
$$\dot{\nabla} \, \underline{\tilde{H}}_\alpha^{\ .\beta} = \frac{\partial \, \underline{\tilde{H}}_\alpha^{\ .\beta}}{\partial t} - \dot{\Gamma}^\theta_\alpha \, \underline{\tilde{H}}_\theta^{\ .\beta} + \dot{\Gamma}^\beta_\theta \, \underline{\tilde{H}}_\alpha^{\ .\theta} - \underline{\tilde{H}}_\alpha^{\ .\beta} \Big|_\theta \, \underline{v}^\theta \,, \tag{9.676c}$$
$$\dot{\nabla} \, \underline{\tilde{H}}_{\alpha\beta} = \frac{\partial \, \underline{\tilde{H}}_{\alpha\beta}}{\partial t} - \dot{\Gamma}^\theta_\alpha \, \underline{\tilde{H}}_{\theta\beta} - \dot{\Gamma}^\theta_\beta \, \underline{\tilde{H}}_{\alpha\theta} - \underline{\tilde{H}}_{\alpha\beta} \Big|_\theta \, \underline{v}^\theta \,. \tag{9.676d}$$

Hint: The result (9.673b) can alternatively be obtained by formally applying $\dot{\nabla}$ to $\mathbf{a}_\alpha$ via (9.675a). With the aid of $(9.95)_3$, (9.177a) and (9.656), this is verified as follows:

$$\dot{\nabla} \mathbf{a}_\alpha = \underbrace{\frac{\partial \mathbf{a}_\alpha}{\partial t}}_{= \left(\dot{\Gamma}^\theta_\alpha \right) \mathbf{a}_\theta + \left(\underline{v}^\theta \, \underline{b}_{\theta\alpha} + \widehat{c}|_\alpha \right) \widehat{\mathbf{n}}} - \dot{\Gamma}^\theta_\alpha \mathbf{a}_\theta - \underbrace{\mathbf{a}_\alpha|_\theta \, \underline{v}^\theta}_{= \underline{b}_{\alpha\theta} \, \underline{v}^\theta \, \widehat{\mathbf{n}}}$$
$$= \left(\cancel{\dot{\Gamma}^\theta_\alpha} - \cancel{\dot{\Gamma}^\theta_\alpha} \right) \mathbf{a}_\theta + \left(\cancel{\underline{b}_{\alpha\theta} \, \underline{v}^\theta} + \widehat{c}|_\alpha - \cancel{\underline{b}_{\alpha\theta} \, \underline{v}^\theta} \right) \widehat{\mathbf{n}}$$
$$= \widehat{c}|_\alpha \, \widehat{\mathbf{n}} \,.$$

In a similar manner,

$$\dot{\nabla}\mathbf{a}^{\alpha} = \underbrace{\frac{\partial \mathbf{a}^{\alpha}}{\partial t}}_{= \left(-\dot{\Gamma}^{\alpha}_{\theta}\right)\mathbf{a}^{\theta} + \left(\underline{b}^{\alpha}_{\,.\theta}\, \underline{v}^{\theta} + \widehat{c}\,|^{\alpha}\right)\widehat{\mathbf{n}}} + \dot{\Gamma}^{\alpha}_{\theta}\mathbf{a}^{\theta} - \underbrace{\mathbf{a}^{\alpha}|_{\theta}\ \underline{v}^{\theta}}_{= \underline{b}^{\alpha}_{\,.\theta}\, \underline{v}^{\theta}\,\widehat{\mathbf{n}}}$$
$$= \left(-\dot{\Gamma}^{\alpha}_{\theta} + \dot{\Gamma}^{\alpha}_{\theta}\right)\mathbf{a}^{\theta} + \left(\underline{b}^{\alpha}_{\,.\theta}\, \underline{v}^{\theta} + \widehat{c}\,|^{\alpha} - \underline{b}^{\alpha}_{\,.\theta}\, \underline{v}^{\theta}\right)\widehat{\mathbf{n}}$$
$$= \widehat{c}\,|^{\alpha}\,\widehat{\mathbf{n}}\ .$$

Hint: In is worthwhile to point out that the invariant time derivative and surface covariant derivative do not commute. For instance, for any smooth scalar field,

$$\boxed{\dot{\nabla}\left(\bar{h}\big|_{\alpha}\right) - \left(\dot{\nabla}\bar{h}\right)\big|_{\alpha} = \widehat{c}\ \underline{b}_{\alpha}^{\ .\theta}\left(\bar{h}\big|_{\theta}\right)\ ,} \tag{9.677}$$

because,

$$\dot{\nabla}\left(\bar{h}\big|_{\alpha}\right) = \frac{\partial}{\partial t}\left(\bar{h}\big|_{\alpha}\right) - \left(\underline{v}^{\theta}\big|_{\alpha} - \widehat{c}\ \underline{b}^{\theta}_{\,.\alpha}\right)\left(\bar{h}\big|_{\theta}\right) - \left(\bar{h}\big|_{\alpha}\right)\big|_{\theta}\ \underline{v}^{\theta}$$
$$= \frac{\partial^2 \bar{h}}{\partial t\,\partial t^{\alpha}} - \underline{v}^{\theta}\big|_{\alpha}\ \bar{h}\big|_{\theta} + \widehat{c}\ \underline{b}_{\alpha}^{\ .\theta}\left(\bar{h}\big|_{\theta}\right) - \bar{h}\big|_{\alpha\theta}\ \underline{v}^{\theta}\ , \tag{9.678}$$

and,

$$\left(\dot{\nabla}\bar{h}\right)\big|_{\alpha} = \left(\frac{\partial \bar{h}}{\partial t} - \bar{h}\big|_{\theta}\ \underline{v}^{\theta}\right)\Big|_{\alpha}$$
$$= \underbrace{\frac{\partial^2 \bar{h}}{\partial t^{\alpha}\,\partial t}}_{= \frac{\partial^2 \bar{h}}{\partial t\,\partial t^{\alpha}}} - \underbrace{\bar{h}\big|_{\theta\alpha}}_{= \frac{\partial^2 \bar{h}}{\partial t^{\alpha}\,\partial t^{\theta}} - \Gamma^{\rho}_{\theta\alpha}\frac{\partial \bar{h}}{\partial t^{\rho}} = \bar{h}|_{\alpha\theta}}\ \underline{v}^{\theta} - \bar{h}\big|_{\theta}\ \underline{v}^{\theta}\big|_{\alpha}\ , \tag{9.679}$$

where $(9.92)_2$, $(9.100)_3$, (9.129), (9.156), (9.649), $(9.657)_1$ and (9.675a) along with the product rule of differentiation have been used.

9.9.3.3 Metrinilic Property of Invariant Time Differentiation with Respect to Surface Metric Coefficients

The results (9.673b) and (9.675b) clearly show that the invariant time derivative is not metrinilic with respect to the surface basis vectors. But, $\dot{\nabla}\mathbf{a}_{\alpha}$ and $\dot{\nabla}\mathbf{a}^{\alpha}$ point in the normal direction satisfying $\widehat{\mathbf{n}} \cdot \mathbf{a}_{\beta} = 0$ and $\widehat{\mathbf{n}} \cdot \mathbf{a}^{\beta} = 0$. Guided by (9.651), the following metrinilic property is thus implied:

$$\boxed{\dot{\nabla}a^{\alpha\beta} = 0\quad ,\quad \dot{\nabla}\delta^{\alpha}_{\beta} = 0\quad ,\quad \dot{\nabla}a_{\alpha\beta} = 0\ .} \quad \longleftarrow \text{see (9.164)} \tag{9.680}$$

Hint: Note that the important results (9.680) can also be attained from the expressions (9.676a)-(9.676d). For instance,

$$\begin{aligned}\dot{\nabla}a^{\alpha\beta} &= \underbrace{\frac{\partial a^{\alpha\beta}}{\partial t}}_{=2\widehat{c}\,b^{\alpha\beta} - \underline{v}^{\alpha}|^{\beta} - \underline{v}^{\beta}|^{\alpha}} + \underbrace{\dot{\Gamma}^{\alpha}_{\theta} a^{\theta\beta}}_{=\left(\underline{v}^{\alpha}|_{\theta} - \widehat{c}\,\underline{b}^{\alpha}_{.\theta}\right)a^{\theta\beta}} + \underbrace{\dot{\Gamma}^{\beta}_{\theta} a^{\alpha\theta}}_{=\left(\underline{v}^{\beta}|_{\theta} - \widehat{c}\,\underline{b}^{\beta}_{.\theta}\right)a^{\alpha\theta}} - \underbrace{a^{\alpha\beta}\big|_{\theta}}_{=0}\,\underline{v}^{\theta} \\ &= \cancel{2\widehat{c}\,\underline{b}^{\alpha\beta}} - \cancel{\underline{v}^{\alpha}|^{\beta}} - \cancel{\underline{v}^{\beta}|^{\alpha}} + \cancel{\underline{v}^{\alpha}|^{\beta}} - \cancel{\widehat{c}\,\underline{b}^{\alpha\beta}} + \cancel{\underline{v}^{\beta}|^{\alpha}} - \cancel{\widehat{c}\,\underline{b}^{\alpha\beta}} \\ &= 0\,,\end{aligned}$$

where $(9.24)_3$, $(9.83\text{d})_{1-2}$, $(9.100)_3$, $(9.160)_1$, $(9.164)_1$, (9.657), (9.661) and (9.676a) have been used.

9.9.4 Invariant Time Differentiation of Objects with Ambient Indices

As understood from previous considerations, the established time derivative operator properly satisfied the sum and product rules. The desire here is to keep these basic properties when it applies to an invariant combination of the ambient components and basis vectors. The main outcome is its metrinilic property with respect to the ambient basis vectors and, consequently, the ambient metric coefficients. This subsection also aims at characterizing the invariant time derivative of the unit normal vector to the surface as well as the shift tensors possessing both ambient and surface indices.

To begin with, suppose one is given an invariant of the form $\hat{\mathbf{h}} = \underline{\hat{h}}^i\,\mathbf{g}_i$. Then,

$$\begin{aligned}\dot{\nabla}\hat{\mathbf{h}} &\xlongequal[\text{(9.649)}]{\text{by using}} \frac{\partial \hat{\mathbf{h}}}{\partial t} - \left(\hat{\mathbf{h}}\Big|_{\alpha}\right)\underline{v}^{\alpha} \\ &\xlongequal[\text{with applying the product rule}]{\text{by using (9.181) along}} \frac{\partial\,\underline{\hat{h}}^i}{\partial t}\mathbf{g}_i + \underline{\hat{h}}^i\frac{\partial \mathbf{g}_i}{\partial t} - \left(\underline{\hat{h}}^i\Big|_{\alpha}\mathbf{g}_i\right)\underline{v}^{\alpha} \\ &\xlongequal[\text{rule of differentiation}]{\text{by applying the chain}} \frac{\partial\,\underline{\hat{h}}^i}{\partial t}\mathbf{g}_i + \underline{\hat{h}}^i\left(\frac{\partial \mathbf{g}_i}{\partial \Theta^j}\right)\left(\frac{\partial \Theta^j}{\partial t}\right) - \underline{\hat{h}}^i\Big|_{\alpha}\underline{v}^{\alpha}\mathbf{g}_i \\ &\xlongequal[\text{(7.7), (7.8) and (9.628)}]{\text{by using}} \frac{\partial\,\underline{\hat{h}}^i}{\partial t}\mathbf{g}_i + \underline{\hat{h}}^i\left(\Gamma^k_{ji}\mathbf{g}_k\right)\left(\underline{v}^j\right) - \underline{\hat{h}}^i\Big|_{\alpha}\underline{v}^{\alpha}\mathbf{g}_i \\ &\xlongequal[\text{the dummy indices}]{\text{by renaming}} \left(\frac{\partial\,\underline{\hat{h}}^i}{\partial t} + \underline{v}^j\,\Gamma^i_{jm}\,\underline{\hat{h}}^m - \underline{\hat{h}}^i\Big|_{\alpha}\underline{v}^{\alpha}\right)\mathbf{g}_i\,,\end{aligned} \tag{9.681}$$

along with the desire to satisfy the sum and product rules

$$\dot{\nabla}\hat{\mathbf{h}} = \left(\dot{\nabla}\,\underline{\hat{h}}^i\right)\mathbf{g}_i + \underline{\hat{h}}^i\left(\dot{\nabla}\mathbf{g}_i\right)\,, \quad \longleftarrow \text{see (9.672)} \tag{9.682}$$

helps establish

$$\boxed{\dot{\nabla}\,\underline{\hat{h}}^{i} = \frac{\partial\,\underline{\hat{h}}^{i}}{\partial t} + \underline{v}^{j}\,\Gamma^{i}_{jm}\,\underline{\hat{h}}^{m} - \underline{\hat{h}}^{i}\Big|_{\alpha}\,\underline{v}^{\alpha}\,,} \quad \longleftarrow \text{see (9.180)} \tag{9.683a}$$

$$\boxed{\dot{\nabla}\mathbf{g}_{i} = \mathbf{0}\,.} \quad \longleftarrow \text{see (9.181)} \tag{9.683b}$$

Next, for a generic ambient object $\hat{\mathbf{h}} = \underline{\hat{h}}_{i}\mathbf{g}^{i}$, one can similarly arrive at

$$\boxed{\dot{\nabla}\,\underline{\hat{h}}_{i} = \frac{\partial\,\underline{\hat{h}}_{i}}{\partial t} - \underline{v}^{j}\,\Gamma^{m}_{ji}\,\underline{\hat{h}}_{m} - \underline{\hat{h}}_{i}\Big|_{\alpha}\,\underline{v}^{\alpha}\,,} \quad \longleftarrow \text{see (9.182)} \tag{9.684a}$$

$$\boxed{\dot{\nabla}\mathbf{g}^{i} = \mathbf{0}\,.} \quad \longleftarrow \text{see (9.183)} \tag{9.684b}$$

Consistent with (7.36a), (7.36b), (9.181) and (9.183), the results (9.683b) and (9.684b) express the metrinilic property of the invariant time derivative with respect to the ambient basis vectors. Consequently, using (9.651), the metrinilic property of the invariant time derivative with respect to the ambient metric coefficients is implied:

$$\boxed{\dot{\nabla}g^{ij} = 0 \quad , \quad \dot{\nabla}\delta^{i}_{j} = 0 \quad , \quad \dot{\nabla}g_{ij} = 0\,.} \quad \longleftarrow \text{see (9.184)} \tag{9.685}$$

Finally, let $\tilde{\mathbf{H}} = \underline{\tilde{H}}^{ij}\mathbf{g}_{i}\otimes\mathbf{g}_{j} = \underline{\tilde{H}}^{i}_{\cdot j}\mathbf{g}_{i}\otimes\mathbf{g}^{j} = \underline{\tilde{H}}_{i}^{\cdot j}\mathbf{g}^{i}\otimes\mathbf{g}_{j} = \underline{\tilde{H}}_{ij}\mathbf{g}^{i}\otimes\mathbf{g}^{j}$ be an ambient second-order tensor field. It is then a simple exercise to show that

$$\dot{\nabla}\,\underline{\tilde{H}}^{ij} = \frac{\partial\,\underline{\tilde{H}}^{ij}}{\partial t} + \underline{v}^{k}\,\Gamma^{i}_{km}\,\underline{\tilde{H}}^{mj} + \underline{v}^{k}\,\Gamma^{j}_{km}\,\underline{\tilde{H}}^{im} - \underline{\tilde{H}}^{ij}\Big|_{\alpha}\,\underline{v}^{\alpha}\,, \quad \longleftarrow \text{see (9.188a)} \tag{9.686a}$$

$$\dot{\nabla}\,\underline{\tilde{H}}^{i}_{\cdot j} = \frac{\partial\,\underline{\tilde{H}}^{i}_{\cdot j}}{\partial t} + \underline{v}^{k}\,\Gamma^{i}_{km}\,\underline{\tilde{H}}^{m}_{\cdot j} - \underline{v}^{k}\,\Gamma^{m}_{kj}\,\underline{\tilde{H}}^{i}_{\cdot m} - \underline{\tilde{H}}^{i}_{\cdot j}\Big|_{\alpha}\,\underline{v}^{\alpha}\,, \tag{9.686b}$$

$$\dot{\nabla}\,\underline{\tilde{H}}_{i}^{\cdot j} = \frac{\partial\,\underline{\tilde{H}}_{i}^{\cdot j}}{\partial t} - \underline{v}^{k}\,\Gamma^{m}_{ki}\,\underline{\tilde{H}}_{m}^{\cdot j} + \underline{v}^{k}\,\Gamma^{j}_{km}\,\underline{\tilde{H}}_{i}^{\cdot m} - \underline{\tilde{H}}_{i}^{\cdot j}\Big|_{\alpha}\,\underline{v}^{\alpha}\,, \tag{9.686c}$$

$$\dot{\nabla}\,\underline{\tilde{H}}_{ij} = \frac{\partial\,\underline{\tilde{H}}_{ij}}{\partial t} - \underline{v}^{k}\,\Gamma^{m}_{ki}\,\underline{\tilde{H}}_{mj} - \underline{v}^{k}\,\Gamma^{m}_{kj}\,\underline{\tilde{H}}_{im} - \underline{\tilde{H}}_{ij}\Big|_{\alpha}\,\underline{v}^{\alpha}\,. \tag{9.686d}$$

Hint: The interested reader may want to obtain the properties $\dot{\nabla}\mathbf{g}_{i} = \mathbf{0}$ and $\dot{\nabla}\mathbf{g}^{i} = \mathbf{0}$ by formally applying the derivative operator $\dot{\nabla}$ to the ambient basis vectors $\mathbf{g}_{i}$ and $\mathbf{g}^{i}$ as follows:

$$\begin{aligned}
\dot{\nabla}\mathbf{g}_i &= \frac{\partial \mathbf{g}_i}{\partial t} - \underline{v}^j \Gamma^m_{ji}\mathbf{g}_m - \cancelto{\mathbf{0}}{\mathbf{g}_i|_\alpha}\ \underline{v}^\alpha \\
&= \left(\frac{\partial \mathbf{g}_i}{\partial \Theta^j}\right)\left(\frac{\partial \Theta^j}{\partial t}\right) - \underline{v}^j \Gamma^m_{ji}\mathbf{g}_m = \left(\Gamma^m_{ij}\mathbf{g}_m\right)\left(\underline{v}^j\right) - \underline{v}^j \Gamma^m_{ij}\mathbf{g}_m \\
&= \mathbf{0}\ , \\
\dot{\nabla}\mathbf{g}^i &= \frac{\partial \mathbf{g}^i}{\partial t} + \underline{v}^j \Gamma^i_{jm}\mathbf{g}^m - \cancelto{\mathbf{0}}{\mathbf{g}^i|_\alpha}\ \underline{v}^\alpha \\
&= \left(\frac{\partial \mathbf{g}^i}{\partial \Theta^j}\right)\left(\frac{\partial \Theta^j}{\partial t}\right) + \underline{v}^j \Gamma^i_{jm}\mathbf{g}^m = \left(-\Gamma^i_{jm}\mathbf{g}^m\right)\left(\underline{v}^j\right) + \underline{v}^j \Gamma^i_{jm}\mathbf{g}^m \\
&= \mathbf{0}\ .
\end{aligned}$$

Hint: The interested reader may also want to obtain the metrinilic property (9.685) from the relations (9.686a)-(9.686d). For instance, by using (7.16), $(9.184)_1$, (9.628) and (9.686a) along with applying the chain rule of differentiation (and renaming $l \to m$),

$$\begin{aligned}
\dot{\nabla} g^{ij} &= \underbrace{\frac{\partial g^{ij}}{\partial t}}_{= \frac{\partial g^{ij}}{\partial \Theta^k}\underline{v}^k} + \underline{v}^k \Gamma^i_{km} g^{mj} + \underline{v}^k \Gamma^j_{km} g^{im} - \underbrace{g^{ij}\big|_\alpha}_{=0}\ \underline{v}^\alpha \\
&= -\cancel{\Gamma^i_{kl} g^{lj} \underline{v}^k} - \cancel{\Gamma^j_{kl} g^{il} \underline{v}^k} + \cancel{\underline{v}^k \Gamma^i_{km} g^{mj}} + \cancel{\underline{v}^k \Gamma^j_{km} g^{im}} \\
&= 0\ .
\end{aligned}$$

Attention is now focused on representing the invariant time differentiation of the ambient unit normal vector to the surface, that is $\dot{\nabla}\widehat{\mathbf{n}} = \dot{\nabla}\,\underline{\widehat{n}}^i\,\mathbf{g}_i$. Having in mind $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$, the result $\dot{\nabla}\mathbf{a}_\alpha = \widehat{c}\,|_\alpha\,\widehat{\mathbf{n}}$, according to (9.673b), can now be rephrased as $\widehat{c}\,|_\alpha = \dot{\nabla}\mathbf{a}_\alpha \cdot \widehat{\mathbf{n}}$. Note that $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$ immediately gives $\dot{\nabla}\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 0$. And this helps infer that $\dot{\nabla}\widehat{\mathbf{n}} = \underline{w}_\alpha \mathbf{a}^\alpha$ where the object $\underline{w}_\alpha$ needs to be determined. On the other hand $\mathbf{a}_\alpha \cdot \widehat{\mathbf{n}} = 0$ helps establish $\dot{\nabla}\mathbf{a}_\alpha \cdot \widehat{\mathbf{n}} = -\mathbf{a}_\alpha \cdot \dot{\nabla}\widehat{\mathbf{n}} = -\mathbf{a}_\alpha \cdot \underline{w}_\beta \mathbf{a}^\beta = -\underline{w}_\alpha$. In other words, $\underline{w}_\alpha = -\widehat{c}\,|_\alpha$. Thus,

$$\boxed{\dot{\nabla}\widehat{\mathbf{n}} = -\,\widehat{c}\,|_\alpha\,\mathbf{a}^\alpha\ .} \tag{9.687}$$

By index juggling and using $\mathbf{a}_\alpha = \overline{Z}^i_\alpha \mathbf{g}_i$, according to $(9.46)_3$, the result (9.687) can further be written as

$$\boxed{\dot{\nabla}\widehat{\mathbf{n}} = -\,\widehat{c}\,|^\alpha\,\mathbf{a}_\alpha \quad \text{with} \quad \dot{\nabla}\,\underline{\widehat{n}}^i = -\,\widehat{c}\,|^\alpha\,\overline{Z}^i_\alpha\ .} \tag{9.688}$$

From (9.185)$_2$, (9.649) and (9.688)$_1$, it follows that[21]

$$\boxed{\frac{\partial \widehat{\mathbf{n}}}{\partial t} = \dot{\nabla}\widehat{\mathbf{n}} + \widehat{\mathbf{n}}|_\alpha \, \underline{v}^\alpha = -\left(\widehat{c}\,|^\beta + \underline{v}^\alpha \, \underline{b}_\alpha^{\cdot\beta}\right)\mathbf{a}_\beta \; .} \tag{9.689}$$

Attention here is focused on formulating the invariant time differentiation of the shift tensors $\overline{Z}_\alpha^i = \mathbf{g}^i \cdot \mathbf{a}_\alpha$ and $\widetilde{Z}_i^\alpha = \mathbf{g}_i \cdot \mathbf{a}^\alpha$ given in (9.47)$_1$ and (9.60)$_1$, respectively. Using (9.673b), (9.675b), (9.683b) and (9.684b) along with applying the product rule of differentiation, one will have

$$\boxed{\begin{array}{l} \dot{\nabla}\overline{Z}_\alpha^i = \mathbf{g}^i \cdot \dot{\nabla}\mathbf{a}_\alpha = \mathbf{g}^i \cdot \widehat{c}\,|_\alpha \widehat{\mathbf{n}} = \widehat{c}\,|_\alpha \; \underline{\widehat{n}}^i \; , \\ \text{or } \widehat{c}\,|_\alpha = \left(\dot{\nabla}\overline{Z}_\alpha^i\right) \underline{\widehat{n}}_i \text{ considering the identity } \underline{\widehat{n}}^i \, \underline{\widehat{n}}_i = 1 \end{array}} \tag{9.690}$$

and

$$\boxed{\begin{array}{l} \dot{\nabla}\widetilde{Z}_i^\alpha = \mathbf{g}_i \cdot \dot{\nabla}\mathbf{a}^\alpha = \mathbf{g}_i \cdot \widehat{c}\,|^\alpha \widehat{\mathbf{n}} = \widehat{c}\,|^\alpha \; \underline{\widehat{n}}_i \; . \\ \text{or } \widehat{c}\,|^\alpha = \left(\dot{\nabla}\widetilde{Z}_i^\alpha\right) \underline{\widehat{n}}^i \text{ considering the identity } \underline{\widehat{n}}_i \, \underline{\widehat{n}}^i = 1 \end{array}} \tag{9.691}$$

9.9.5 Invariant Time Differentiation of Surface Mixed Curvature Tensor and Its Principal Invariants

At the end of this section, the modern derivative operator $\dot{\nabla}$ will be applied to a crucially important quantity in the calculus of curved surfaces which is the extrinsic object $\underline{b}_\alpha^{\cdot\beta}$. The resulting expression then helps establish the invariant time derivative of the mean and Gaussian curvatures.

First, the invariant time derivative of the surface mixed curvature tensor is given by

$$\boxed{\dot{\nabla}\, \underline{b}_\alpha^{\cdot\beta} = \widehat{c}\,|_\alpha|^\beta + \widehat{c}\; \underline{b}_\alpha^{\cdot\theta}\; \underline{b}_\theta^{\cdot\beta} \; ,} \tag{9.692}$$

[21] The result (9.689) can be derived in an alternative way. This is demonstrated in the following. First, by using the expressions $\widehat{\mathbf{n}} \cdot \widehat{\mathbf{n}} = 1$ and $\widehat{\mathbf{n}} \cdot \mathbf{a}_\beta = 0$, one can write $\widehat{\mathbf{n}} \cdot \partial\widehat{\mathbf{n}}/\partial t = 0$ which reveals the fact that $\partial\widehat{\mathbf{n}}/\partial t = \underline{w}^\alpha \mathbf{a}_\alpha$ where $\underline{w}^\alpha$ are two unknown quantities to be determined. Then, making use of the equation (9.656), the partial time derivative of the relation $\widehat{\mathbf{n}} \cdot \mathbf{a}_\beta = 0$ helps conclude that $\partial\widehat{\mathbf{n}}/\partial t \cdot \mathbf{a}_\beta = -\left(\underline{v}^\theta \, \underline{b}_{\theta\beta} + \widehat{c}\,|_\beta\right)$ or $\underline{w}^\alpha a_{\alpha\beta} = -\left(\underline{v}^\theta \, \underline{b}_{\theta\beta} + \widehat{c}\,|_\beta\right)$. Finally, using the identity $a_{\alpha\beta} a^{\beta\rho} = \delta_\alpha^\rho$, one will arrive at $\underline{w}^\rho = -\left(\underline{v}^\theta \, \underline{b}_\theta^{\cdot\rho} + \widehat{c}\,|^\rho\right)$ or $\underline{w}^\beta = -\left(\widehat{c}\,|^\beta + \underline{v}^\alpha \, \underline{b}_\alpha^{\cdot\beta}\right)$.

because

$$\begin{aligned}
\dot{\nabla}\, \underline{b}_{\alpha}^{\cdot\beta} &\xlongequal[\text{(9.101) and (9.185)}]{\text{by using}} \dot{\nabla}\left(-\,\widehat{\mathbf{n}}|_{\alpha} \cdot \mathbf{a}^{\beta}\right) \\
&\xlongequal[\text{the product rule}]{\text{by applying}} -\dot{\nabla}\left(\widehat{\mathbf{n}}|_{\alpha}\right) \cdot \mathbf{a}^{\beta} - \widehat{\mathbf{n}}|_{\alpha} \cdot \left(\dot{\nabla}\mathbf{a}^{\beta}\right) \\
&\xlongequal[\text{(9.677)}]{\text{in light of}} -\left(\dot{\nabla}\,\widehat{\mathbf{n}}\right)\Big|_{\alpha} \cdot \mathbf{a}^{\beta} - \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\left(\widehat{\mathbf{n}}|_{\theta}\right) \cdot \mathbf{a}^{\beta} - \widehat{\mathbf{n}}|_{\alpha} \cdot \left(\dot{\nabla}\mathbf{a}^{\beta}\right) \\
&\xlongequal[\text{(9.675b) and (9.687)}]{\text{by using}} -\left(-\,\widehat{c}\,|_{\theta}\, \mathbf{a}^{\theta}\right)\Big|_{\alpha} \cdot \mathbf{a}^{\beta} - \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\left(\widehat{\mathbf{n}}|_{\theta}\right) \cdot \mathbf{a}^{\beta} - \widehat{\mathbf{n}}|_{\alpha} \cdot \left(\widehat{c}\,|^{\beta}\, \widehat{\mathbf{n}}\right) \\
&\xlongequal[\text{and noting that } \widehat{\mathbf{n}}|_{\alpha} \cdot \widehat{\mathbf{n}} = 0]{\text{by considering (9.12)}} \left(\widehat{c}\,|_{\theta}\, \mathbf{a}^{\theta}\right)\Big|_{\alpha} \cdot \mathbf{a}^{\beta} - \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\left(\widehat{\mathbf{n}}|_{\theta}\right) \cdot \mathbf{a}^{\beta} \\
&\xlongequal[\text{applying the product rule}]{\text{by using (9.185) and}} \widehat{c}\,|_{\theta\alpha}\, \mathbf{a}^{\theta} \cdot \mathbf{a}^{\beta} + \widehat{c}\,|_{\theta}\, \mathbf{a}^{\theta}\Big|_{\alpha} \cdot \mathbf{a}^{\beta} + \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\; \underline{b}_{\theta}^{\cdot\rho}\, \mathbf{a}_{\rho} \cdot \mathbf{a}^{\beta} \\
&\xlongequal[\text{(9.33) and (9.177b)}]{\text{by using (9.24),}} \widehat{c}\,|_{\theta\alpha}\, a^{\theta\beta} + \widehat{c}\,|_{\theta}\; \underline{b}^{\theta}_{\cdot\alpha}\, \widehat{\mathbf{n}} \cdot \mathbf{a}^{\beta} + \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\; \underline{b}_{\theta}^{\cdot\rho}\, \delta_{\rho}^{\beta} \\
&\xlongequal[\text{property of the mixed Kronecker delta}]{\text{by using (9.29) and the replacement}} \widehat{c}\,|_{\theta\alpha}\, a^{\theta\beta} + \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\; \underline{b}_{\theta}^{\cdot\beta} \\
&\xlongequal[\text{(9.129) and (9.156)}]{\text{by using (9.92),}} \widehat{c}\,|_{\alpha\theta}\, a^{\theta\beta} + \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\; \underline{b}_{\theta}^{\cdot\beta} \quad \longleftarrow \text{note that } \widehat{c}\,|_{\alpha\theta} = \frac{\partial^2 \widehat{c}}{\partial t^{\theta}\, \partial t^{\alpha}} - \Gamma_{\alpha\theta}^{\rho}\, \frac{\partial \widehat{c}}{\partial t^{\rho}} \\
&\xlongequal[\text{(9.160)}]{\text{by using}} \widehat{c}\,|_{\alpha}\,|^{\beta} + \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\; \underline{b}_{\theta}^{\cdot\beta}\;. \quad \longleftarrow \text{note that } \widehat{c}\,|_{\alpha}\,|^{\beta} = \widehat{c}\,|^{\beta}\Big|_{\alpha}
\end{aligned}$$

From (9.692), it simply follows that

$$\boxed{\dot{\nabla}\, \underline{b}_{\alpha}^{\cdot\alpha} = \widehat{c}\,|_{\alpha}\,|^{\alpha} + \widehat{c}\; \underline{b}_{\alpha}^{\cdot\theta}\; \underline{b}_{\theta}^{\cdot\alpha}\;.} \quad \longleftarrow \text{note that } \underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta}^{\cdot\alpha} = \underline{b}_{\alpha}^{\cdot\theta}\, \underline{b}_{\theta}^{\cdot\beta}\, \delta_{\beta}^{\alpha} = \text{tr}\left[\underline{b}_{\alpha}^{\cdot\beta}\right]^2 \tag{9.693}$$

Recall from $(9.103)_2$ that $\overline{\mathrm{H}} = \underline{b}_{\alpha}^{\cdot\alpha}/2$. Then, using $(9.276)_2$ and (9.693), the invariant time derivative of the mean curvature takes the following form

$$\boxed{\dot{\nabla}\overline{\mathrm{H}} = \frac{1}{2}\; \widehat{c}\,|_{\alpha}\,|^{\alpha} + \widehat{c}\left(2\overline{\mathrm{H}}^2 - \overline{\mathrm{K}}\right)\;.} \tag{9.694}$$

Finally, the invariant time derivative of the Gaussian curvature is given by

$$\boxed{\dot{\nabla}\overline{\mathrm{K}} = 2\; \widehat{c}\,|_{\alpha}\,|^{\alpha}\, \overline{\mathrm{H}} - \widehat{c}\,|_{\alpha}\,|^{\beta}\; \underline{b}_{\beta}^{\cdot\alpha} + 2\widehat{c}\,\overline{\mathrm{K}}\,\overline{\mathrm{H}}\;,} \tag{9.695}$$

since

$$\begin{aligned}
\dot{\nabla}\overline{\mathrm{K}} &\xlongequal[\text{(9.100) and (9.490)}]{\text{by using}} \frac{1}{2}\dot{\nabla}\left(\underline{b}_{\alpha}^{\cdot\alpha}\; \underline{b}_{\beta}^{\cdot\beta} - \underline{b}_{\alpha}^{\cdot\beta}\; \underline{b}_{\beta}^{\cdot\alpha}\right) \\
&\xlongequal[\text{and renaming the dummy indices}]{\text{by applying the product rule}} \left(\dot{\nabla}\, \underline{b}_{\alpha}^{\cdot\alpha}\right) \underline{b}_{\beta}^{\cdot\beta} - \frac{1}{2}\left(\dot{\nabla}\, \underline{b}_{\alpha}^{\cdot\beta}\right) \underline{b}_{\beta}^{\cdot\alpha} - \frac{1}{2}\, \underline{b}_{\alpha}^{\cdot\beta}\left(\dot{\nabla}\, \underline{b}_{\beta}^{\cdot\alpha}\right)
\end{aligned}$$

$$\xrightarrow[\text{(9.692) and (9.693)}]{\text{by using}} \left(\widehat{c}\,|_\alpha\,|^\alpha + \widehat{c}\; \underline{b}_\alpha^{\cdot\theta}\; \underline{b}_{\dot\theta}^{\;\alpha} \right) \underline{b}_{\dot\beta}^{\;\beta}$$
$$- \frac{1}{2} \left(\widehat{c}\,|_\alpha\,|^\beta\; \underline{b}_{\dot\beta}^{\;\alpha} + \widehat{c}\,|_\beta\,|^\alpha\; \underline{b}_\alpha^{\cdot\beta} \right) - \frac{1}{2} \left(\widehat{c}\; \underline{b}_\alpha^{\cdot\theta}\; \underline{b}_{\dot\theta}^{\;\beta}\; \underline{b}_{\dot\beta}^{\;\alpha} + \widehat{c}\; \underline{b}_{\dot\beta}^{\;\theta}\; \underline{b}_{\dot\theta}^{\;\alpha}\; \underline{b}_\alpha^{\cdot\beta} \right)$$
$$\xrightarrow[\text{the dummy indices}]{\text{by renaming}} \widehat{c}\,|_\alpha\,|^\alpha\; \underline{b}_{\dot\beta}^{\;\beta} + \widehat{c}\; \underline{b}_\alpha^{\cdot\theta}\; \underline{b}_{\dot\theta}^{\;\alpha}\; \underline{b}_{\dot\beta}^{\;\beta} - \widehat{c}\,|_\alpha\,|^\beta\; \underline{b}_{\dot\beta}^{\;\alpha} - \widehat{c}\; \underline{b}_\alpha^{\cdot\theta}\; \underline{b}_{\dot\theta}^{\;\rho}\; \underline{b}_{\dot\rho}^{\;\alpha}$$
$$\xrightarrow[\text{(2.89a)}]{\text{in light of}} \widehat{c}\,|_\alpha\,|^\alpha \operatorname{tr}\left[\underline{b}_\alpha^{\cdot\beta} \right] - \widehat{c}\,|_\alpha\,|^\beta\; \underline{b}_{\dot\beta}^{\;\alpha} + \widehat{c} \left(\operatorname{tr}\left[\underline{b}_\alpha^{\cdot\beta} \right]^2 \operatorname{tr}\left[\underline{b}_\alpha^{\cdot\beta} \right] - \operatorname{tr}\left[\underline{b}_\alpha^{\cdot\beta} \right]^3 \right)$$
$$\xrightarrow[\text{(9.276) and (9.277)}]{\text{by using (9.103),}} 2\; \widehat{c}\,|_\alpha\,|^\alpha\, \overline{\mathrm{H}} - \widehat{c}\,|_\alpha\,|^\beta\; \underline{b}_{\dot\beta}^{\;\alpha} + 2\widehat{c}\,\overline{\mathrm{K}}\,\overline{\mathrm{H}} \,.$$

9.10 Application to Shell Structures

Differential geometry of two-dimensional regular surfaces was studied in the previous sections. The aim here is to introduce an application of the surface theory in a basic **structural element** called *shell*. This initially curved element is subjected to some (mechanical, thermal, electrical, etc.) loads that cause **stretching**, **shearing** and/or **bending deformations**. Missiles and nanotubes in modern technology, leaves of trees and wings of insects in nature and red blood cells and arteries in human bodies are only a few examples of shells. For these three-dimensional creatures, the thickness is considerably small when compared with the other two dimensions. This particular geometric feature enables one to separate the thickness variable from the two in-plane ones. The geometry as well as deformation of a shell can thus be described by means of proper functions of two in-plane variables corresponding to the *midsurface* (i.e. the surface which bisects the shell thickness). The mechanical behavior of a shell under a given loading is studied through the *theory of elasticity*. This results in a boundary/initial value problem to be solved in order to finally have an **equilibrium state**. To achieve this ultimate goal, one only needs to find some unknown two-variable functions. Using the theory of elasticity in the context of structural mechanics, there is generally no closed-form solution for such functions and a shell problem should thus be treated numerically. Among all numerical procedures, finite element methods have gained much attention by the researchers. Numerous works have been published on finite element analysis of shells. For an overview see, e.g., Bathe et al. [54], Zienkiewicz et al. [55], Hughes and Tezduyar [56], Simo et al. [57], Ibrahimbegović et al. [58], Gruttmann et al. [59], Wriggers et al. [60], Betsch et al. [61], Bischoff and Ramm [62], Miehe [63], Hauptmann et al. [64], Cirak et al. [65] and Pimenta et al. [66].

9.10.1 Shell Geometry

The three-dimensional geometry of the shell is described by the vector function

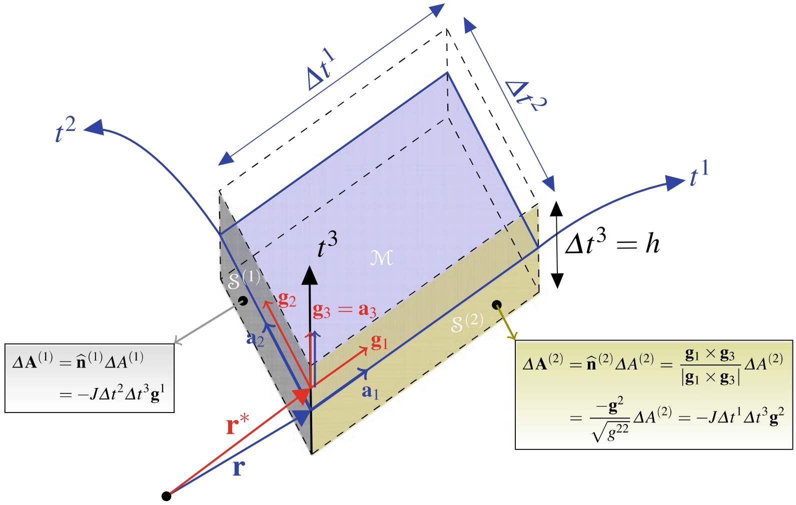


Fig. 9.36 Infinitesimal shell element

$$\boxed{\mathbf{r}^*\left(t^1, t^2, t^3\right) = \mathbf{r}\left(t^1, t^2\right) + t^3 \mathbf{a}_3\left(t^1, t^2\right) \quad \text{with} \quad -\frac{h}{2} \leq t^3 \leq +\frac{h}{2}} \ , \tag{9.696}$$

where $\mathbf{r}\left(t^1, t^2\right)$ denotes the position vector of a generic point on the midsurface $\mathcal{M}$, h presents the shell thickness and

$$\mathbf{a}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{|\mathbf{a}_1 \times \mathbf{a}_2|} = \widehat{\mathbf{n}} \quad \text{where} \quad \mathbf{a}_\alpha = \frac{\partial \mathbf{r}}{\partial t^\alpha} \ ,$$

represents the normal vector to the midsurface (see Fig. 9.36 for a geometrical interpretation). The sensitivity of $\mathbf{r}^*$ with respect to t^i, $i = 1, 2, 3$, will give the ambient covariant basis vectors

$$\mathbf{g}_\alpha = \frac{\partial\left(\mathbf{r} + t^3 \widehat{\mathbf{n}}\right)}{\partial t^\alpha} = \mathbf{a}_\alpha - t^3 \, \underline{b}_\alpha^{\cdot\beta} \mathbf{a}_\beta \ , \tag{9.697}$$

and

$$\mathbf{g}_3 = \frac{\partial\left(\mathbf{r} + t^3 \widehat{\mathbf{n}}\right)}{\partial t^3} = \widehat{\mathbf{n}} \ . \quad \leftarrow \text{note that } \mathbf{g}_3 = \mathbf{g}^3 = \mathbf{a}_3 = \mathbf{a}^3 = \widehat{\mathbf{n}} \text{ with } |\widehat{\mathbf{n}}| = 1 \tag{9.698}$$

The relation $(9.697)_2$ is sometimes rephrased as

$$\boxed{\mathbf{g}_\alpha = \mathbf{Z}\mathbf{a}_\alpha} \ , \tag{9.699}$$

where the symmetric second-order mixed contra-covariant tensor

$$\boxed{\mathbf{Z} = \left(\delta^{\alpha}_{\beta} - t^3\, \underline{b}^{\alpha}_{.\beta}\right) \mathbf{a}_{\alpha} \otimes \mathbf{a}^{\beta}\ ,} \tag{9.700}$$

is referred to as the *shell shifter* (see Bischoff et al. [67]). Making use of $(5.89b)_2$, $(9.103)_3$ and $(9.104)_2$, its determinant is given by

$$\boxed{Z = \det\left[\, \delta^{\alpha}_{\beta} - t^3\, \underline{b}^{\alpha}_{.\beta} \,\right] = 1 - 2t^3\overline{\mathrm{H}} + \left(t^3\right)^2 \overline{\mathrm{K}}\ .} \tag{9.701}$$

Recall from $(5.30)_3$ that $J = \mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3) > 0$ and consider $a = \det\left[\, a_{\alpha\beta} \,\right] > 0$ according to $(9.54)_1$. In this context, they are related through the following equation

$$\boxed{J = \sqrt{a}\, Z\ ,} \tag{9.702}$$

because

$$\begin{aligned}
J &\xrightarrow[\text{(5.30) and (9.697)}]{\text{from}} \left[\left(1 - t^3 b_{1}^{\,;1}\right) \mathbf{a}_1 - t^3 b_{1}^{\,;2} \mathbf{a}_2\right] \cdot \left[\left(1 - t^3 b_{2}^{\,;2}\right) \mathbf{a}_2 \times \widehat{\mathbf{n}} - t^3 b_{2}^{\,;1} \mathbf{a}_1 \times \widehat{\mathbf{n}}\right] \\
&\xrightarrow[\text{(1.149a) and (9.58)}]{\text{from}} \sqrt{a}\left[\left(1 - t^3 b_{1}^{\,;1}\right) \mathbf{a}_1 - t^3 b_{1}^{\,;2} \mathbf{a}_2\right] \cdot \left[\left(1 - t^3 b_{2}^{\,;2}\right) \mathbf{a}^1 + t^3 b_{2}^{\,;1} \mathbf{a}^2\right] \\
&\xrightarrow[\text{(9.33)}]{\text{from}} \sqrt{a}\left[\left(1 - t^3 b_{1}^{\,;1}\right)\left(1 - t^3 b_{2}^{\,;2}\right) - \left(t^3\right)^2 b_{1}^{\,;2} b_{2}^{\,;1}\right] \\
&\xrightarrow[\text{(9.100)}]{\text{from}} \sqrt{a}\left[\left(1 - t^3 b^{1}_{.1}\right)\left(1 - t^3 b^{2}_{.2}\right) - \left(t^3\right)^2 b^{2}_{.1} b^{1}_{.2}\right] \\
&\xrightarrow[\text{(9.701)}]{\text{from}} \sqrt{a}\, Z\ .
\end{aligned}$$

Consider an infinitesimal parallelepiped as illustrated in Fig. 9.36. The surface element $\Delta A^{(i)}$ ($\Delta A^{<i>}$) which is eventually a subset of the surface t^i = constant ($t^i + \Delta t^i$ = constant) is called the *negative face* (*positive face*). For instance, $\Delta A^{(1)}$ is characterized by

$$\widehat{\mathbf{n}}^{(1)} = \frac{\mathbf{g}_3 \times \mathbf{g}_2}{|\mathbf{g}_3 \times \mathbf{g}_2|} = \frac{-J\mathbf{g}^1}{\left|-J\mathbf{g}^1\right|} = -\frac{\mathbf{g}^1}{\sqrt{g^{11}}}\ . \tag{9.703}$$

Guided by (5.112) and consistent with $(8.2a)_1$, the above relation helps obtain

$$\left.\begin{aligned}
\Delta A^{(1)} &= \sqrt{g^{11}}\, J\, \Delta t^2 \Delta t^3 \\
\Delta \mathbf{A}^{(1)} &= \widehat{\mathbf{n}}^{(1)} \Delta A^{(1)} = -\left(J\mathbf{g}^1\right) \Delta t^2 \Delta t^3
\end{aligned}\right\}\ . \tag{9.704}$$

In a similar manner,

$$\Delta \mathbf{A}^{<1>} = +\left(J\mathbf{g}^1\right)\Big|_{t^1+\Delta t^1, t^2} \Delta t^2 \Delta t^3\ . \tag{9.705}$$

The unit normal vector in (9.703)$_3$ can also be written as

$$\widehat{\mathbf{n}}^{(1)} = -\frac{J}{\sqrt{g_{22}}}\mathbf{g}^1 \ , \tag{9.706}$$

because

$$\begin{aligned}
&\xrightarrow[(5.33)]{\text{from}} \quad \mathbf{g}_2 \times \mathbf{g}_3 = J\mathbf{g}^1 \\
&\xrightarrow[(5.46)]{\text{from}} \quad \mathbf{g}^1 \cdot \left[\mathbf{g}_2 \times \mathbf{g}_3\right] = Jg^{11} \\
&\xrightarrow[(5.43)]{\text{from}} \quad \mathbf{g}^1 \cdot \left[\left(g_{21}\mathbf{g}^1 + g_{22}\mathbf{g}^2 + g_{23}\mathbf{g}^3\right) \times \mathbf{g}_3\right] = Jg^{11} \\
&\xrightarrow[(1.51),\ (1.73)\text{ and }(9.698)]{\text{from}} \quad \mathbf{g}^1 \cdot \left[g_{22}\mathbf{g}^2 \times \mathbf{g}_3\right] = Jg^{11} \\
&\xrightarrow[(1.73)]{\text{from}} \quad \mathbf{g}_3 \cdot \left[\mathbf{g}^1 \times \mathbf{g}^2\right] = J\frac{g^{11}}{g_{22}} \\
&\xrightarrow[(5.35)]{\text{from}} \quad \mathbf{g}_3 \cdot \frac{\mathbf{g}_3}{J} = J\frac{g^{11}}{g_{22}} \\
&\xrightarrow[(9.698)]{\text{from}} \quad \boxed{g^{11} = J^{-2}g_{22} \ .}
\end{aligned}$$

9.10.2 External and Internal Forces and Moments

A shell can eventually be considered as a surface endowed with the **mechanical properties**. Any element of that deformable mechanical surface thus reacts to the applied surface as well as edge **forces** and **moments**. The results, to be predicted by a structural or mechanical engineer, will be the stretching, shearing and/or bending deformations.

In what follows, the goal is to characterize the external and internal forces and moments exerted on an infinitesimal element of a mechanical surface. Let

$$\mathbf{f}_{\text{ext}} = \underline{f}^{\,\alpha}_{\text{ext}}\mathbf{a}_\alpha + \underline{f}^{\,n}_{\text{ext}}\widehat{\mathbf{n}} \quad \text{and} \quad \mathbf{m}_{\text{ext}} = \underline{m}^{\alpha}_{\text{ext}}\widehat{\mathbf{n}} \times \mathbf{a}_\alpha \ , \tag{9.707}$$

be the external force and moment exerted on unit area of the shell midsurface, respectively. It has been assumed that the moment $\mathbf{m}_{\text{ext}}$ is **tangent** to the midsurface. The external force and couple exerted on $\mathcal{M}$, bounded by the curves $t^\alpha = \text{constant}$ and $t^\alpha + \Delta t^\alpha = \text{constant}$, are then given by

$$\left.\begin{aligned}
\mathbf{f}^{\text{tot}}_{\text{ext}} &= \int_{t^1}^{t^1+\Delta t^1}\int_{t^2}^{t^2+\Delta t^2} \mathbf{f}_{\text{ext}} dA = \int_{t^1}^{t^1+\Delta t^1}\int_{t^2}^{t^2+\Delta t^2} \mathbf{f}_{\text{ext}}\sqrt{a}\, dt^1 dt^2 \\
\mathbf{m}^{\text{tot}}_{\text{ext}} &= \int_{t^1}^{t^1+\Delta t^1}\int_{t^2}^{t^2+\Delta t^2} \mathbf{m}_{\text{ext}} dA = \int_{t^1}^{t^1+\Delta t^1}\int_{t^2}^{t^2+\Delta t^2} \mathbf{m}_{\text{ext}}\sqrt{a}\, dt^1 dt^2
\end{aligned}\right\} . \tag{9.708}$$

Let $\boldsymbol{\sigma} = \underline{\sigma}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j$ be the **Cauchy stress tensor**. By the **stress theorem of Cauchy**, consider the *Cauchy traction vector* $\mathbf{t} = \boldsymbol{\sigma}\,\widehat{\mathbf{n}}$. Then, the *resultant force* on $A^{(1)}$ can be written as

$$\begin{aligned}\int_{A^{(1)}} \mathbf{t}^{(1)} dA^{(1)} &= \int_{t^2}^{t^2+\Delta t^2} \int_{-h/2}^{h/2} \boldsymbol{\sigma}\left[\widehat{\mathbf{n}}^{(1)} dA^{(1)}\right] \\ &= -\int_{t^2}^{t^2+\Delta t^2} \left[\int_{-h/2}^{h/2} J\boldsymbol{\sigma}\,\mathbf{g}^1 dt^3\right] dt^2 \\ &= -\int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\left[\int_{-h/2}^{h/2} Z\boldsymbol{\sigma}\,\mathbf{g}^1 dt^3\right] dt^2 \\ &= -\int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\,\mathbf{f}^1 dt^2\ .\end{aligned}$$

In general,

$$\boxed{\int_{A^{(\alpha)}} \mathbf{t}^{(\alpha)} dA^{(\alpha)} = -\int_{t^\beta}^{t^\beta+\Delta t^\beta} \sqrt{a}\,\mathbf{f}^{\alpha} dt^{\beta}\ , \quad \alpha \neq \beta = 1, 2\ ,} \tag{9.709}$$

where the *stress resultant* $\mathbf{f}^{\alpha}$ is given by

$$\boxed{\mathbf{f}^{\alpha} = \int_{-h/2}^{h/2} Z\,\boldsymbol{\sigma}\,\mathbf{g}^{\alpha} dt^3\ .} \tag{9.710}$$

Now, the *resultant moment* on $A^{(1)}$ renders

$$\begin{aligned}\int_{A^{(1)}} \left[\mathbf{r}^*\right] \times \left[\mathbf{t}^{(1)} dA^{(1)}\right] &= \int_{t^2}^{t^2+\Delta t^2} \int_{-h/2}^{h/2} \left[\mathbf{r} + t^3\widehat{\mathbf{n}}\right] \times \left[\boldsymbol{\sigma}\left(\widehat{\mathbf{n}}^{(1)} dA^{(1)}\right)\right] \\ &= \int_{t^2}^{t^2+\Delta t^2} \int_{-h/2}^{h/2} \left[\mathbf{r} + t^3\widehat{\mathbf{n}}\right] \times \left[-\sqrt{a}\,Z\boldsymbol{\sigma}\,\mathbf{g}^1 dt^3 dt^2\right] \\ &= -\int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\,\mathbf{r} \times \left[\int_{-h/2}^{h/2} Z\boldsymbol{\sigma}\,\mathbf{g}^1 dt^3\right] dt^2 \\ &\quad -\int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\,\widehat{\mathbf{n}} \times \left[\int_{-h/2}^{h/2} Z\boldsymbol{\sigma}\,\mathbf{g}^1 t^3 dt^3\right] dt^2 \\ &= -\int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\,\left(\mathbf{r} \times \mathbf{f}^1 + \mathbf{m}^1\right) dt^2\ .\end{aligned}$$

In general,

$$\boxed{\int_{A^{(\alpha)}} \mathbf{r}^{*} \times \mathbf{t}^{(\alpha)} dA^{(\alpha)} = - \int_{t^{\beta}}^{t^{\beta}+\Delta t^{\beta}} \sqrt{a}\ (\mathbf{r} \times \mathbf{f}^{\alpha} + \mathbf{m}^{\alpha})\, dt^{\beta},\ \alpha \neq \beta = 1, 2\,,} \tag{9.711}$$

where

$$\boxed{\mathbf{m}^{\alpha} = \widehat{\mathbf{n}} \times \int_{-h/2}^{h/2} Z\, \boldsymbol{\sigma}\, \mathbf{g}^{\alpha} t^{3} dt^{3}\ .} \tag{9.712}$$

9.10.3 Equilibrium Equations

For a continuum body in equilibrium, the vector sum of all external and internal force variables acting on that body should vanish. And this should hold true not only on the entire body but also on any imaginary isolated element of that object. For the deformable surface element shown in Fig. 9.36, the force equilibrium condition is spelled out as

$$\begin{aligned} &- \int_{t^2}^{t^2+\Delta t^2} \left(\sqrt{a}\, \mathbf{f}^{1}\right)\Big|_{t^1,t^2} dt^2 - \int_{t^1}^{t^1+\Delta t^1} \left(\sqrt{a}\, \mathbf{f}^{2}\right)\Big|_{t^1,t^2} dt^1 \\ &+ \int_{t^2}^{t^2+\Delta t^2} \left(\sqrt{a}\, \mathbf{f}^{1}\right)\Big|_{t^1+\Delta t^1,t^2} dt^2 + \int_{t^1}^{t^1+\Delta t^1} \left(\sqrt{a}\, \mathbf{f}^{2}\right)\Big|_{t^1,t^2+\Delta t^2} dt^1 \\ &+ \int_{t^1}^{t^1+\Delta t^1} \int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\, \mathbf{f}_{\text{ext}}\, dt^1 dt^2 = \mathbf{0}\ . \end{aligned}$$

By using the first-order Taylor series expansion, one can arrive at

$$\begin{aligned} &\int_{t^2}^{t^2+\Delta t^2} \frac{\partial}{\partial t^1} \left(\sqrt{a}\, \mathbf{f}^{1}\right) \Delta t^1 dt^2 + \int_{t^1}^{t^1+\Delta t^1} \frac{\partial}{\partial t^2} \left(\sqrt{a}\, \mathbf{f}^{2}\right) \Delta t^2 dt^1 \\ &+ \int_{t^1}^{t^1+\Delta t^1} \int_{t^2}^{t^2+\Delta t^2} \sqrt{a}\, \mathbf{f}_{\text{ext}}\, dt^1 dt^2 = \mathbf{0}\ , \end{aligned}$$

which can be written in the useful form

$$\int_{t^1}^{t^1+\Delta t^1} \int_{t^2}^{t^2+\Delta t^2} \left[\frac{\partial}{\partial t^1} \left(\sqrt{a}\, \mathbf{f}^{1}\right) + \frac{\partial}{\partial t^2} \left(\sqrt{a}\, \mathbf{f}^{2}\right) + \sqrt{a}\, \mathbf{f}_{\text{ext}} \right] dt^1 dt^2 = \mathbf{0}\ . \tag{9.713}$$

In a similar manner, the moment equilibrium condition is given by

$$\int_{t^1}^{t^1+\Delta t^1}\int_{t^2}^{t^2+\Delta t^2}\left\{\frac{\partial}{\partial t^1}\left[\sqrt{a}\,\left(\mathbf{r}\times\mathbf{f}^1+\mathbf{m}^1\right)\right]+\frac{\partial}{\partial t^2}\left[\sqrt{a}\,\left(\mathbf{r}\times\mathbf{f}^2+\mathbf{m}^2\right)\right]\right.$$
$$\left.+\sqrt{a}\left[\mathbf{r}\times\mathbf{f}_{\text{ext}}+\mathbf{m}_{\text{ext}}\right]\right\}dt^1dt^2=\mathbf{0}\;. \tag{9.714}$$

Consider the fact that these integral expressions should hold for all shell elements. One can thus arrive at the local forms

$$\boxed{\frac{\partial}{\partial t^\alpha}\left(\sqrt{a}\,\mathbf{f}^{\,\alpha}\right)+\sqrt{a}\,\mathbf{f}_{\text{ext}}=\mathbf{0}\;,} \tag{9.715}$$

and

$$\boxed{\frac{\partial}{\partial t^\alpha}\left[\sqrt{a}\,\left(\mathbf{r}\times\mathbf{f}^{\,\alpha}+\mathbf{m}^\alpha\right)\right]+\sqrt{a}\,\left[\mathbf{r}\times\mathbf{f}_{\text{ext}}+\mathbf{m}_{\text{ext}}\right]=\mathbf{0}\;,} \tag{9.716}$$

called the *equilibrium equations*. Making use of (9.92)$_2$ and (9.115), one can further obtain

$$\boxed{\frac{\partial\mathbf{f}^{\,\alpha}}{\partial t^\alpha}+\Gamma^\beta_{\beta\alpha}\mathbf{f}^{\,\alpha}+\mathbf{f}_{\text{ext}}=\mathbf{0}\;,} \tag{9.717}$$

and

$$\boxed{\frac{\partial\mathbf{m}^\alpha}{\partial t^\alpha}+\Gamma^\beta_{\beta\alpha}\mathbf{m}^\alpha+\mathbf{a}_\alpha\times\mathbf{f}^{\,\alpha}+\mathbf{m}_{\text{ext}}=\mathbf{0}\;.} \tag{9.718}$$

Let

$$\mathbf{f}^{\,\alpha}=\underline{H}^{\alpha\beta}\mathbf{a}_\beta+\underline{h}^\alpha\widehat{\mathbf{n}}\quad,\quad\mathbf{m}^{\,\alpha}=\underline{M}^{\alpha\beta}\widehat{\mathbf{n}}\times\mathbf{a}_\beta\;. \tag{9.719}$$

One then has

$$\left(\underline{H}^{\alpha\beta}\big|_\alpha-\underline{h}^\alpha\,\underline{b}_\alpha^{\cdot\beta}+\underline{f}^{\beta}_{\text{ext}}\right)\mathbf{a}_\beta+\left(\underline{h}^\alpha\big|_\alpha+\underline{H}^{\alpha\beta}\,\underline{b}_{\alpha\beta}+\underline{f}^{\,n}_{\text{ext}}\right)\widehat{\mathbf{n}}=\mathbf{0}\;,$$

and

$$\left(\underline{M}^{\alpha\beta}\big|_\alpha-\underline{h}^\beta+\underline{m}^\beta_{\text{ext}}\right)\widehat{\mathbf{n}}\times\mathbf{a}_\beta+\sqrt{a}\left(\underline{H}^{\alpha\beta}-\underline{b}^\alpha_{\cdot\gamma}\,\underline{M}^{\gamma\beta}\right)\varepsilon_{\alpha\beta}\widehat{\mathbf{n}}=\mathbf{0}\;,$$

where $\widehat{\mathbf{n}}\times\mathbf{a}_\beta=\sqrt{a}\varepsilon_{\beta\gamma}\,\mathbf{a}^\gamma$, according to (9.58). Finally, consider the fact that the triplet of vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\widehat{\mathbf{n}}$ and also $\mathbf{a}^1$, $\mathbf{a}^2$, $\widehat{\mathbf{n}}$ are linearly independent. The equilibrium equations, in index notation, then render

$$\boxed{\left.\begin{aligned} \underline{H}^{\alpha\beta}\big|_{\alpha} - \underline{h}^{\alpha}\,\underline{b}_{\alpha}^{\cdot\beta} + \underline{f}^{\beta}_{\text{ext}} &= 0 \\ \underline{H}^{\alpha\beta}\,\underline{b}_{\alpha\beta} + \underline{h}^{\alpha}\big|_{\alpha} + \underline{f}^{n}_{\text{ext}} &= 0 \end{aligned}\right\} ,} \tag{9.720}$$

and

$$\boxed{\left.\begin{aligned} \underline{M}^{\alpha\beta}\big|_{\alpha} - \underline{h}^{\beta} + \underline{m}^{\beta}_{\text{ext}} &= 0 \\ \left(\underline{H}^{\alpha\beta} - \underline{b}^{\alpha}_{\cdot\gamma}\,\underline{M}^{\gamma\beta}\right)\varepsilon_{\alpha\beta} &= 0 \end{aligned}\right\} .} \tag{9.721}$$

Considering the identity $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$, according to $(9.36)_1$, the expression $(9.721)_2$ may be written as

$$\boxed{\underline{H}^{\alpha\beta} + \underline{b}^{\beta}_{\cdot\gamma}\,\underline{M}^{\gamma\alpha} = \underline{H}^{\beta\alpha} + \underline{b}^{\alpha}_{\cdot\gamma}\,\underline{M}^{\gamma\beta} \ , \quad \alpha, \beta = 1, 2, \ \alpha \neq \beta \ .} \tag{9.722}$$

To model (**small** or **finite**) **deformations** of a solid shell, a proper stress measure should be characterized for its material ingredients. For the purely mechanical problems, such a quantity is only a function of the strain tensor when the material under consideration is **elastic**, **isotropic** and **homogeneous**. See Exercises 3.4 and 6.16 for more discussions. In the following, the strain and stress measures as well as the displacement field will be characterized for a sophisticated shell model.

9.10.4 *Basic Shell Mathematical Model*

The goal here is to introduce the strain and stress measures of an elastic, isotropic and homogeneous shell in the **small strain regime** based on a model called *basic shell mathematical model* (see Lee and Bathe [68] and Chapelle and Bathe [69]).

In alignment with (9.696), the displacement field of a generic particle in this shell model is given by

$$\boxed{\mathbf{u}^*\left(t^1, t^2, t^3\right) = \mathbf{u}\left(t^1, t^2\right) + t^3\boldsymbol{\theta}\left(t^1, t^2\right) \ ,} \tag{9.723}$$

where $\mathbf{u}$ presents the **translational displacement** of a material point on the mid-surface $\mathcal{M}$ and $\boldsymbol{\theta}$ denotes the **infinitesimal rotation** of a material line perpendicular to that mechanical surface. Notice that $\mathbf{u}$ is an ambient vector while $\boldsymbol{\theta}$ is a surface vector. Referred to the dual basis $\left\{\mathbf{a}^1, \mathbf{a}^2, \widehat{\mathbf{n}}\right\}$, they are decomposed as

$$\boxed{\mathbf{u} = \underline{u}_{\alpha}\mathbf{a}^{\alpha} + \underline{u}^{n}\,\widehat{\mathbf{n}} \quad , \quad \boldsymbol{\theta} = \underline{\theta}_{\alpha}\mathbf{a}^{\alpha} \ .} \tag{9.724}$$

Here, $\underline{\theta}_1$ ($\underline{\theta}_2$) presents the rotation about $\mathbf{a}_2$ ($\mathbf{a}_1$) and the rotation about $\widehat{\mathbf{n}}$ is assumed to be zero. This is a well-known assumption due to *Reissner-Mindlin* (see Reissner [70], Mindlin [71] and Hencky [72]). The infinitesimal strain tensor for the given displacement field (9.723) renders

$$\boxed{\boldsymbol{\varepsilon} = \frac{1}{2}\left[\left(\frac{\partial \mathbf{u}^*}{\partial \mathbf{x}^*}\right) + \left(\frac{\partial \mathbf{u}^*}{\partial \mathbf{x}^*}\right)^{\mathrm{T}}\right] \quad \text{where} \quad \frac{\partial \mathbf{u}^*}{\partial \mathbf{x}^*} = \frac{\partial \mathbf{u}^*}{\partial t^k} \otimes \mathbf{g}^k \, .} \tag{9.725}$$

Referred to the ambient contravariant basis $\{\mathbf{g}^i\}$, the covariant components of this symmetric tensor render

$$\underline{\varepsilon}_{ij} = \frac{1}{2}\mathbf{g}_i \cdot \left[\frac{\partial \mathbf{u}^*}{\partial t^k} \otimes \mathbf{g}^k + \mathbf{g}^k \otimes \frac{\partial \mathbf{u}^*}{\partial t^k}\right]\mathbf{g}_j = \frac{1}{2}\left[\mathbf{g}_i \cdot \frac{\partial \mathbf{u}^*}{\partial t^j} + \frac{\partial \mathbf{u}^*}{\partial t^i} \cdot \mathbf{g}_j\right] . \tag{9.726}$$

It is then easy to see that

$$\begin{aligned} \underline{\varepsilon}_{\alpha\beta} = {} & \frac{1}{2}\left(\underline{u}_\alpha\big|_\beta + \underline{u}_\beta\big|_\alpha \right) - \underline{b}_{\alpha\beta}\, \underline{u}^n \\ & + t^3\left[\frac{1}{2}\left(\underline{\theta}_\alpha\big|_\beta + \underline{\theta}_\beta\big|_\alpha - \underline{b}_\alpha^{\cdot\gamma}\ \underline{u}_\gamma\big|_\beta - \underline{b}_\beta^{\cdot\gamma}\ \underline{u}_\gamma\big|_\alpha\right) + \underline{b}_\alpha^{\cdot\gamma}\, \underline{b}_{\gamma\beta}\, \underline{u}^n\right] \\ & - \left(t^3\right)^2\left[\frac{1}{2}\left(\underline{b}_\alpha^{\cdot\gamma}\ \underline{\theta}_\gamma\big|_\beta + \underline{b}_\beta^{\cdot\gamma}\ \underline{\theta}_\gamma\big|_\alpha\right)\right] , \end{aligned} \tag{9.727a}$$

$$\underline{\varepsilon}_{\alpha 3} = \frac{1}{2}\left(\underline{\theta}_\alpha + \frac{\partial\, \underline{u}^n}{\partial t^\alpha} + \underline{b}_\alpha^{\cdot\gamma}\, \underline{u}_\gamma\right) , \tag{9.727b}$$

$$\underline{\varepsilon}_{33} = 0 \, . \tag{9.727c}$$

As can be seen, the **inplane strains** $(\underline{\varepsilon}_{11},\ \underline{\varepsilon}_{22},\ \underline{\varepsilon}_{12})$ are quadratic through the shell thickness while the **transverse shear strains** $(\underline{\varepsilon}_{13},\ \underline{\varepsilon}_{23})$ are constant. Note that there will be no elongation through the shell thickness owing to the vanishing of the **transverse normal** $\underline{\varepsilon}_{33}$. Indeed, **inextensibility** of fibers normal to the shell midsurface is a well-known assumption usually made in shell theorems. With regard to this, the model under consideration may be called an *inextensible shell model*. See Sansour [73] and also Sansour [74] for some extensible shell models at finite deformations.

At the end, by means of the **plane stress** assumption, the contravariant components of the stress tensor for a linear elastic shell are given by

$$\underline{\sigma}^{\alpha\beta} = \frac{E}{2\,(1+\nu)}\left[g^{\alpha\gamma}g^{\beta\delta} + g^{\alpha\delta}g^{\beta\gamma} + \frac{2\nu}{1-\nu}g^{\alpha\beta}g^{\gamma\delta}\right]\underline{\varepsilon}_{\gamma\delta} \, , \tag{9.728a}$$

$$\underline{\sigma}^{\alpha 3} = \frac{E}{1+\nu}\left[g^{\alpha\beta}\right]\underline{\varepsilon}_{\beta 3} \, , \tag{9.728b}$$

$$\underline{\sigma}^{33} = 0 \, , \tag{9.728c}$$

where E denotes Young's modulus and ν presents Poisson's ratio.

9.11 Exercises

Exercise 9.1

Compute

(a) the covariant basis vectors $\mathbf{a}_1, \mathbf{a}_2$, according to (9.10)$_1$,
(b) the covariant metric coefficients $a_{\alpha\beta}$, according to (9.18),
(c) the area element dA, according to (9.56),
(d) the unit normal field $\widehat{\mathbf{n}}$, according to (9.31)$_1$,
(e) the covariant curvature tensor $\underline{b}_{\alpha\beta}$, according to (9.95)$_1$,
(f) the mixed curvature tensor $\underline{b}_{\alpha}^{\cdot\beta}$, according to (9.100)$_1$,
(g) the mean curvature $\overline{\mathrm{H}}$, according to (9.103)$_1$,
(h) the Gaussian curvature $\overline{\mathrm{K}}$, according to (9.104)$_1$, and
(i) the Christoffel symbols $\Gamma^{\alpha}_{\beta\gamma}$, given in (9.235a)-(9.235c),

for an **elliptic cylinder** defined by

$$\underbrace{x_1 = R\cos t^1 \quad , \quad x_2 = \overline{R}\sin t^1 \quad , \quad x_3 = t^2}_{\text{these coordinates satisfy the implicit relation } (x_1/R)^2 + \left(x_2/\overline{R}\right)^2 = 1} \quad , \tag{9.729}$$

where $R, \overline{R}$ are positive real numbers, $0 \leq t^1 < 2\pi$ and $-\infty < t^2 < \infty$ (see Fig. 9.17).

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = -R\sin t^1 \widehat{\mathbf{e}}_1 + \overline{R}\cos t^1 \widehat{\mathbf{e}}_2 \ , \tag{9.730a}$$

$$\mathbf{a}_2 = \widehat{\mathbf{e}}_3 \ . \tag{9.730b}$$

The covariant metric coefficients in matrix form:

$$\left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} R^2\sin^2 t^1 + \overline{R}^2\cos^2 t^1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{9.731}$$

The area element:

$$dA = \sqrt{R^2\sin^2 t^1 + \overline{R}^2\cos^2 t^1}\, dt^1 dt^2 \ . \tag{9.732}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \frac{\overline{R}\cos t^1 \widehat{\mathbf{e}}_1 + R\sin t^1 \widehat{\mathbf{e}}_2}{\sqrt{R^2\sin^2 t^1 + \overline{R}^2\cos^2 t^1}} \ . \tag{9.733}$$

The covariant curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha\beta} \right] = \frac{R\overline{R}}{\sqrt{R^2 \sin^2 t^1 + \overline{R}^2 \cos^2 t^1}} \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} . \tag{9.734}$$

The mixed curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha}^{\cdot\beta} \right] = \frac{R\overline{R}}{\left(R^2 \sin^2 t^1 + \overline{R}^2 \cos^2\right)^{3/2}} \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} . \tag{9.735}$$

The mean curvature:

$$\overline{\mathrm{H}} = -\frac{R\overline{R}}{2\left(R^2 \sin^2 t^1 + \overline{R}^2 \cos^2 t^1\right)^{3/2}} . \tag{9.736}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = 0 . \tag{9.737}$$

The only nonzero Christoffel symbols entry:

$$\Gamma^1_{11} = \frac{\left(R^2 - \overline{R}^2\right) \sin t^1 \cos t^1}{R^2 \sin^2 t^1 + \overline{R}^2 \cos^2 t^1} . \tag{9.738}$$

Exercise 9.2

Compute the desired quantities listed in Exercise 9.1 for the **sphere** (9.4), that is,

$$x_1 = R \sin t^1 \cos t^2 \quad , \quad x_2 = R \sin t^1 \sin t^2 \quad , \quad x_3 = R \cos t^1 \ .$$

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = R \cos t^1 \cos t^2 \widehat{\mathbf{e}}_1 + R \cos t^1 \sin t^2 \widehat{\mathbf{e}}_2 - R \sin t^1 \widehat{\mathbf{e}}_3 \ , \tag{9.739a}$$

$$\mathbf{a}_2 = -R \sin t^1 \sin t^2 \widehat{\mathbf{e}}_1 + R \sin t^1 \cos t^2 \widehat{\mathbf{e}}_2 \ . \tag{9.739b}$$

The covariant metric coefficients in matrix form:

$$\left[a_{\alpha\beta} \right] = \begin{bmatrix} R^2 & 0 \\ 0 & \left(R \sin t^1\right)^2 \end{bmatrix} . \tag{9.740}$$

The area element:

$$dA = R^2 \sin t^1 \, dt^1 dt^2 \ . \tag{9.741}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \sin t^1 \cos t^2 \widehat{\mathbf{e}}_1 + \sin t^1 \sin t^2 \widehat{\mathbf{e}}_2 + \cos t^1 \widehat{\mathbf{e}}_3 \ . \tag{9.742}$$

The covariant curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha\beta} \right] = -R \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 t^1 \end{bmatrix} . \tag{9.743}$$

The mixed curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha}^{\cdot\beta} \right] = -\frac{1}{R} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{9.744}$$

The mean curvature:

$$\overline{\mathrm{H}} = -\frac{1}{R} \ . \tag{9.745}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = \frac{1}{R^2} \ . \quad \leftarrow \text{note that the sphere is an object with constant positive Gaussian curvature} \tag{9.746}$$

The nonzero Christoffel symbols entries:

$$\Gamma^1_{22} = -\sin t^1 \cos t^1 \quad , \quad \Gamma^2_{12} = \cot t^1 \ . \tag{9.747}$$

Exercise 9.3

Compute the desired quantities listed in Exercise 9.1 for the **torus** (9.5), that is,

$$x_1 = \left(\hat{R} + \hat{r}\cos t^2\right)\cos t^1 \quad , \quad x_2 = \left(\hat{R} + \hat{r}\cos t^2\right)\sin t^1 \quad , \quad x_3 = \hat{r}\sin t^2 \ .$$

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = -\left(\hat{R} + \hat{r}\cos t^2\right)\sin t^1 \widehat{\mathbf{e}}_1 + \left(\hat{R} + \hat{r}\cos t^2\right)\cos t^1 \widehat{\mathbf{e}}_2 \ , \tag{9.748a}$$

$$\mathbf{a}_2 = -\hat{r}\sin t^2 \cos t^1 \widehat{\mathbf{e}}_1 - \hat{r}\sin t^2 \sin t^1 \widehat{\mathbf{e}}_2 + \hat{r}\cos t^2 \widehat{\mathbf{e}}_3 \ . \tag{9.748b}$$

The covariant metric coefficients in matrix form:

$$\left[a_{\alpha\beta} \right] = \begin{bmatrix} \left(\hat{R} + \hat{r}\cos t^2\right)^2 & 0 \\ 0 & (\hat{r})^2 \end{bmatrix} . \tag{9.749}$$

The area element:

$$dA = \hat{r}\left(\hat{R} + \hat{r}\cos t^2\right) dt^1 dt^2 \; . \tag{9.750}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \cos t^1 \cos t^2 \,\widehat{\mathbf{e}}_1 + \sin t^1 \cos t^2 \,\widehat{\mathbf{e}}_2 + \sin t^2 \,\widehat{\mathbf{e}}_3 \; . \tag{9.751}$$

The covariant curvature tensor in matrix form:

$$\left[\, \underline{b}_{\alpha\beta} \,\right] = \begin{bmatrix} -\left(\hat{R} + \hat{r}\cos t^2\right)\cos t^2 & 0 \\ 0 & -\hat{r} \end{bmatrix} . \tag{9.752}$$

The mixed curvature tensor in matrix form:

$$\left[\, \underline{b}_{\alpha}^{\,\cdot\beta} \,\right] = \begin{bmatrix} \frac{-\cos t^2}{\hat{R}+\hat{r}\cos t^2} & 0 \\ 0 & -\frac{1}{\hat{r}} \end{bmatrix} . \tag{9.753}$$

The mean curvature:

$$\overline{\mathrm{H}} = \frac{-\hat{R} - 2\hat{r}\cos t^2}{2\hat{r}\left(\hat{R} + \hat{r}\cos t^2\right)} \; . \tag{9.754}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = \frac{\cos t^2}{\hat{r}\left(\hat{R} + \hat{r}\cos t^2\right)} \; . \quad \longleftarrow \text{note that} \begin{cases} \overline{\mathrm{K}} > 0 & \text{on the outside} \\ \overline{\mathrm{K}} = 0 & \text{at the top and bottom circles} \\ \overline{\mathrm{K}} < 0 & \text{on the inside} \end{cases} \tag{9.755}$$

The nonzero Christoffel symbols entries:

$$\Gamma^1_{12} = -\frac{\hat{r}}{\hat{R} + \hat{r}\cos t^2}\sin t^2 \quad , \quad \Gamma^2_{11} = \frac{\hat{R} + \hat{r}\cos t^2}{\hat{r}}\sin t^2 \; . \tag{9.756}$$

Exercise 9.4

Compute the desired quantities listed in Exercise 9.1 for a **hyperboloid of revolution** (or **circular hyperboloid**) defined by

$$\underbrace{x_1 = R\cosh t^1 \cos t^2 \quad , \quad x_2 = R\cosh t^1 \sin t^2 \quad , \quad x_3 = \widehat{R}\sinh t^1}_{\text{note that the one-sheeted hyperboloid (9.7) is called the circular hyperboloid if } R = \overline{R}} \; . \tag{9.757}$$

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = R \sinh t^1 \cos t^2 \widehat{\mathbf{e}}_1 + R \sinh t^1 \sin t^2 \widehat{\mathbf{e}}_2 + \widehat{R} \cosh t^1 \widehat{\mathbf{e}}_3 \ , \tag{9.758a}$$

$$\mathbf{a}_2 = -R \cosh t^1 \sin t^2 \widehat{\mathbf{e}}_1 + R \cosh t^1 \cos t^2 \widehat{\mathbf{e}}_2 \ . \tag{9.758b}$$

The covariant metric coefficients in matrix form:

$$\left[a_{\alpha\beta} \right] = \begin{bmatrix} a^\star & 0 \\ 0 & R^2 \cosh^2 t^1 \end{bmatrix} \text{ where } a^\star = R^2 \sinh^2 t^1 + \widehat{R}^2 \cosh^2 t^1 \ . \tag{9.759}$$

The area element:

$$dA = R \cosh t^1 \sqrt{a^\star} dt^1 dt^2 \ . \tag{9.760}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \frac{-\widehat{R} \cosh t^1 \cos t^2 \widehat{\mathbf{e}}_1 - \widehat{R} \cosh t^1 \sin t^2 \widehat{\mathbf{e}}_2 + R \sinh t^1 \widehat{\mathbf{e}}_3}{\sqrt{a^\star}} \ . \tag{9.761}$$

The covariant curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha\beta} \right] = \frac{R\widehat{R}}{\sqrt{a^\star}} \begin{bmatrix} -1 & 0 \\ 0 & \cosh^2 t^1 \end{bmatrix} \ . \tag{9.762}$$

The mixed curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha}^{\cdot\beta} \right] = \frac{R\widehat{R}}{\sqrt{a^\star}} \begin{bmatrix} \frac{-1}{(a^\star)^2} & 0 \\ 0 & \frac{1}{R^2} \end{bmatrix} \ . \tag{9.763}$$

The mean curvature:

$$\overline{\mathrm{H}} = \frac{R^2 \widehat{R} \sinh^2 t^1 + \widehat{R}^3 \cosh^2 t^1 - R^2 \widehat{R}}{2R \left(a^\star\right)^{3/2}} \ . \tag{9.764}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = -\left(\frac{\widehat{R}}{a^\star} \right)^2 \ . \tag{9.765}$$

The nonzero Christoffel symbols entries:

$$\Gamma^1_{11} = \frac{\left(R^2 + \widehat{R}^2\right) \tanh t^1}{R^2 \tanh^2 t^1 + \widehat{R}^2} \ , \quad \Gamma^2_{12} = \tanh t^1 \ , \quad \Gamma^1_{22} = \frac{-R^2 \tanh t^1}{R^2 \tanh^2 t^1 + \widehat{R}^2} \ . \tag{9.766}$$

Exercise 9.5

Compute the desired quantities listed in Exercise 9.1 for the **hyperbolic paraboloid** (9.9), that is,

$$x_1 = Rt^1 \quad , \quad x_2 = \overline{R}t^2 \quad , \quad x_3 = \widehat{R}\left[\left(t^1\right)^2 - \left(t^2\right)^2\right] .$$

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = R\,\widehat{\mathbf{e}}_1 + 2\widehat{R}t^1\,\widehat{\mathbf{e}}_3 , \tag{9.767a}$$

$$\mathbf{a}_2 = \overline{R}\,\widehat{\mathbf{e}}_2 - 2\widehat{R}t^2\,\widehat{\mathbf{e}}_3 . \tag{9.767b}$$

The covariant metric coefficients in matrix form:

$$\left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} R^2 + \left(2\widehat{R}t^1\right)^2 & -4\widehat{R}^2 t^1 t^2 \\ -4\widehat{R}^2 t^1 t^2 & \overline{R}^2 + \left(2\widehat{R}t^2\right)^2 \end{bmatrix} . \tag{9.768}$$

The area element:

$$dA = \sqrt{a^{\text{hp}}}dt^1 dt^2 \quad \text{where} \quad a^{\text{hp}} = \left(2\overline{R}\widehat{R}t^1\right)^2 + \left(2R\widehat{R}t^2\right)^2 + \left(R\overline{R}\right)^2 . \tag{9.769}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \frac{-2\overline{R}\widehat{R}t^1\,\widehat{\mathbf{e}}_1 + 2R\widehat{R}t^2\,\widehat{\mathbf{e}}_2 + R\overline{R}\,\widehat{\mathbf{e}}_3}{\sqrt{a^{\text{hp}}}} . \tag{9.770}$$

The covariant curvature tensor in matrix form:

$$\left[\, \underline{b}_{\alpha\beta} \,\right] = \frac{2R\overline{R}\widehat{R}}{\sqrt{a^{\text{hp}}}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} . \tag{9.771}$$

The mixed curvature tensor in matrix form:

$$\left[\, \underline{b}_{\alpha}^{\cdot\beta} \,\right] = \frac{2R\overline{R}\widehat{R}}{\left(a^{\text{hp}}\right)^{3/2}} \begin{bmatrix} \overline{R}^2 + \left(2\widehat{R}t^2\right)^2 & 4\widehat{R}^2 t^1 t^2 \\ -4\widehat{R}^2 t^1 t^2 & -R^2 - \left(2\widehat{R}t^1\right)^2 \end{bmatrix} . \tag{9.772}$$

The mean curvature:

$$\overline{\mathrm{H}} = \frac{R\overline{R}\widehat{R}}{\left(a^{\text{hp}}\right)^{3/2}} \left\{ \left(\overline{R}^2 - R^2\right) - 4\widehat{R}^2 \left[\left(t^1\right)^2 - \left(t^2\right)^2\right] \right\} . \tag{9.773}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = -\left(\frac{2R\overline{R}\widehat{R}}{a^{\text{hp}}}\right)^2 . \tag{9.774}$$

The nonzero Christoffel symbols in matrix form:

$$\begin{bmatrix} \Gamma_{11}^1 \\ \Gamma_{11}^2 \end{bmatrix} = \frac{1}{a^{\text{hp}}} \begin{bmatrix} \left(2\overline{R}\widehat{R}\right)^2 t^1 \\ -\left(2R\widehat{R}\right)^2 t^2 \end{bmatrix} \quad , \quad \begin{bmatrix} \Gamma_{22}^1 \\ \Gamma_{22}^2 \end{bmatrix} = -\begin{bmatrix} \Gamma_{11}^1 \\ \Gamma_{11}^2 \end{bmatrix} . \tag{9.775}$$

Exercise 9.6

Compute the desired quantities listed in Exercise 9.1 for a **catenoid** (Fig. 9.37) parametrically described by

$$\underbrace{x_1 = c\cos t^1 \cosh\frac{t^2}{c} \quad , \quad x_2 = c\sin t^1 \cosh\frac{t^2}{c} \quad , \quad x_3 = t^2}_{\text{these coordinates satisfy the implicit relation } \sqrt{x_1^2 + x_2^2} = c\cosh\frac{x_3}{c}} \, , \tag{9.776}$$

where $c > 0$ is a constant, $0 \leq t^1 < 2\pi$ and $-\infty < t^2 < \infty$.

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = -c\sin t^1 \cosh\frac{t^2}{c}\,\widehat{\mathbf{e}}_1 + c\cos t^1 \cosh\frac{t^2}{c}\,\widehat{\mathbf{e}}_2 \, , \tag{9.777a}$$

$$\mathbf{a}_2 = \cos t^1 \sinh\frac{t^2}{c}\,\widehat{\mathbf{e}}_1 + \sin t^1 \sinh\frac{t^2}{c}\,\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3 \, . \tag{9.777b}$$

The covariant metric coefficients in matrix form:

$$\left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} c^2\cosh^2\frac{t^2}{c} & 0 \\ 0 & \cosh^2\frac{t^2}{c} \end{bmatrix} . \tag{9.778}$$

The area element:

$$dA = c\cosh^2\frac{t^2}{c}\, dt^1 dt^2 \, . \tag{9.779}$$

The unit normal field:

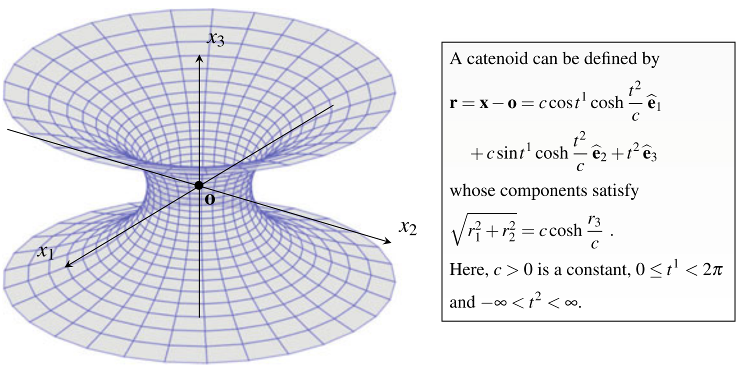


Fig. 9.37 Catenoid

$$\widehat{\mathbf{n}} = \mathrm{sech}\frac{t^2}{c}\left[\cos t^1\,\widehat{\mathbf{e}}_1 + \sin t^1\,\widehat{\mathbf{e}}_2 - \sinh\frac{t^2}{c}\,\widehat{\mathbf{e}}_3\right] . \tag{9.780}$$

The covariant curvature tensor in matrix form:

$$\left[\,\underline{b}_{\alpha\beta}\,\right] = \begin{bmatrix} -c & 0 \\ 0 & \frac{1}{c} \end{bmatrix} . \tag{9.781}$$

The mixed curvature tensor in matrix form:

$$\left[\,\underline{b}_{\alpha}^{\cdot\beta}\,\right] = \frac{1}{c\cosh^2\frac{t^2}{c}}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{9.782}$$

The mean curvature:

$$\overline{\mathrm{H}} = 0 . \tag{9.783}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = \frac{-1}{c^2\cosh^4\frac{t^2}{c}} . \tag{9.784}$$

The nonzero Christoffel symbols entries:

$$\Gamma^2_{11} = -c\tanh\frac{t^2}{c} \quad , \quad \Gamma^1_{12} = \frac{1}{c}\tanh\frac{t^2}{c} \quad , \quad \Gamma^2_{22} = \frac{1}{c}\tanh\frac{t^2}{c} . \tag{9.785}$$

Exercise 9.7

Compute the desired quantities listed in Exercise 9.1 for a surface defined by the explicit form

$$x_1 = t^1 \quad , \quad x_2 = t^2 \quad , \quad x_3 = f\left(t^1, t^2\right) , \tag{9.786}$$

noting that the **height function** f is a smooth function. This is known as the *Monge form* and the surface defined by this form is called the *Monge patch* (see the pioneering work of Monge [75]).

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = \widehat{\mathbf{e}}_1 + f_{t^1}\,\widehat{\mathbf{e}}_3 \quad \text{where} \quad f_{t^1} := \frac{\partial f}{\partial t^1} , \tag{9.787a}$$

$$\mathbf{a}_2 = \widehat{\mathbf{e}}_2 + f_{t^2}\,\widehat{\mathbf{e}}_3 \quad \text{where} \quad f_{t^2} := \frac{\partial f}{\partial t^2} . \tag{9.787b}$$

The covariant metric coefficients in matrix form:

$$\left[a_{\alpha\beta} \right] = \left[\begin{array}{cc} 1 + f_{t^1}^2 & f_{t^1} f_{t^2} \\ f_{t^1} f_{t^2} & 1 + f_{t^2}^2 \end{array} \right] . \tag{9.788}$$

The area element:

$$dA = \sqrt{1 + f_{t^1}^2 + f_{t^2}^2}\, dt^1 dt^2 . \tag{9.789}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \frac{-f_{t^1} \widehat{\mathbf{e}}_1 - f_{t^2} \widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3}{\sqrt{1 + f_{t^1}^2 + f_{t^2}^2}} . \tag{9.790}$$

The covariant curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha\beta} \right] = \frac{1}{\sqrt{1 + f_{t^1}^2 + f_{t^2}^2}} \left[\begin{array}{cc} f_{t^1t^1} & f_{t^1t^2} \\ f_{t^1t^2} & f_{t^2t^2} \end{array} \right] , \tag{9.791}$$

where

$$f_{t^1t^1} := \frac{\partial^2 f}{\partial t^1 \partial t^1} \quad , \quad f_{t^1t^2} := \frac{\partial^2 f}{\partial t^1 \partial t^2} \quad , \quad f_{t^2t^2} := \frac{\partial^2 f}{\partial t^2 \partial t^2} . \tag{9.792}$$

The mixed curvature tensor entries:

$$\underline{b}_1^{\cdot 1} = \frac{\left(1 + f_{t^2}^2\right) f_{t^1t^1} - f_{t^1} f_{t^2} f_{t^1t^2}}{\left(1 + f_{t^1}^2 + f_{t^2}^2\right)^{3/2}} , \tag{9.793a}$$

$$\underline{b}_1^{\cdot 2} = \frac{\left(1 + f_{t^1}^2\right) f_{t^1t^2} - f_{t^1} f_{t^2} f_{t^1t^1}}{\left(1 + f_{t^1}^2 + f_{t^2}^2\right)^{3/2}} , \tag{9.793b}$$

$$\underline{b}_2^{\cdot 1} = \frac{\left(1 + f_{t^2}^2\right) f_{t^1t^2} - f_{t^1} f_{t^2} f_{t^2t^2}}{\left(1 + f_{t^1}^2 + f_{t^2}^2\right)^{3/2}} . \tag{9.793c}$$

$$\underline{b}_2^{\cdot 2} = \frac{\left(1 + f_{t^1}^2\right) f_{t^2t^2} - f_{t^1} f_{t^2} f_{t^1t^2}}{\left(1 + f_{t^1}^2 + f_{t^2}^2\right)^{3/2}} . \tag{9.793d}$$

The mean curvature:

$$\overline{\mathrm{H}} = \frac{\left(1 + f_{t^2}^2\right) f_{t^1t^1} - 2 f_{t^1} f_{t^2} f_{t^1t^2} + \left(1 + f_{t^1}^2\right) f_{t^2t^2}}{2\left(1 + f_{t^1}^2 + f_{t^2}^2\right)^{3/2}} . \tag{9.794}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = \frac{f_{t^1t^1} f_{t^2t^2} - f_{t^1t^2}^2}{\left(1 + f_{t^1}^2 + f_{t^2}^2\right)^2} . \tag{9.795}$$

The nonzero Christoffel symbols in matrix form:

$$\begin{bmatrix} \Gamma^1_{11} \\ \Gamma^2_{11} \end{bmatrix} = \frac{f_{t^1t^1}}{1 + f^2_{t^1} + f^2_{t^2}} \begin{bmatrix} f_{t^1} \\ f_{t^2} \end{bmatrix} \ , \quad \begin{bmatrix} \Gamma^1_{12} \\ \Gamma^2_{12} \end{bmatrix} = \frac{f_{t^1t^2}}{1 + f^2_{t^1} + f^2_{t^2}} \begin{bmatrix} f_{t^1} \\ f_{t^2} \end{bmatrix} \ , \tag{9.796a}$$

$$\begin{bmatrix} \Gamma^1_{22} \\ \Gamma^2_{22} \end{bmatrix} = \frac{f_{t^2t^2}}{1 + f^2_{t^1} + f^2_{t^2}} \begin{bmatrix} f_{t^1} \\ f_{t^2} \end{bmatrix} \ . \tag{9.796b}$$

Exercise 9.8

Let $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}^3_\mathrm{p}$ be a parametrization of a regular surface $\mathcal{S}$ with

$$\mathrm{E}_\mathbf{r} = \mathrm{G}_\mathbf{r} = \lambda^2 \quad , \quad \mathrm{F}_\mathbf{r} = 0 \ . \tag{9.797}$$

This orthogonal parametrization is called the *isothermal* (or *isothermic*) parametrization. And λ^2 is known as the *scaling function* of the isothermal patch $\mathbf{x}$. Show that the mean and Gaussian curvatures of this patch can be written as

$$\overline{\mathrm{H}} = \frac{\Delta \mathbf{x} \cdot \widehat{\mathbf{n}}}{2\lambda^2} \quad \text{where} \quad \Delta \mathbf{x} = \frac{\partial^2 \mathbf{x}}{\partial t^\alpha \partial t^\alpha} = \frac{\partial \mathbf{a}_\alpha}{\partial t^\alpha} \ , \tag{9.798a}$$

$$\overline{\mathrm{K}} = -\frac{\Delta \log \lambda^2}{2\lambda^2} \quad \text{where} \quad \Delta \log \lambda^2 = \frac{\partial^2 \log \lambda^2}{\partial t^\alpha \partial t^\alpha} \ . \tag{9.798b}$$

Solution. The mean curvature $(9.257)_2$ for the problem at hand renders

$$\overline{\mathrm{H}} = \frac{\mathrm{e}_\mathbf{r} + \mathrm{g}_\mathbf{r}}{2\lambda^2} \ .$$

Using $(9.252a)_2$ and $(9.252b)_2$, one will have $\mathrm{e}_\mathbf{r} + \mathrm{g}_\mathbf{r} = \Delta \mathbf{x} \cdot \widehat{\mathbf{n}}$. Substituting this result into the above relation yields the desired result $(9.798a)_1$. This equation can also be written as

$$\Delta \mathbf{x} = 2\lambda^2 \overline{\mathrm{H}}\, \widehat{\mathbf{n}} \ , \tag{9.799}$$

which relies on the fact that $\Delta \mathbf{x}$ is a scalar multiple of $\widehat{\mathbf{n}}$. To show this, consider the equation (9.797), i.e. $\mathbf{a}_1 \cdot \mathbf{a}_1 = \mathbf{a}_2 \cdot \mathbf{a}_2 = \lambda^2$ and $\mathbf{a}_1 \cdot \mathbf{a}_2 = 0$, which implies that

$$\mathbf{a}_1 \cdot \frac{\partial \mathbf{a}_1}{\partial t^1} = \mathbf{a}_2 \cdot \frac{\partial \mathbf{a}_1}{\partial t^2} = -\mathbf{a}_1 \cdot \frac{\partial \mathbf{a}_2}{\partial t^2} \ ,$$

Consequently,

$$\mathbf{a}_1 \cdot \frac{\partial \mathbf{a}_\alpha}{\partial t^\alpha} = 0 \quad \text{and, in a similar manner,} \quad \mathbf{a}_2 \cdot \frac{\partial \mathbf{a}_\alpha}{\partial t^\alpha} = 0 \ .$$

And this means that $\partial \mathbf{a}_\alpha / \partial t^\alpha$ is parallel to $\widehat{\mathbf{n}}$. Thus, (9.798a)$_1$ and (9.799) imply each other.

For the isothermic surface $\mathcal{S}$, the required result (9.798b)$_1$ can be obtained from the Bieberbach formula (9.497a)$_2$ as follows:

$$\begin{aligned}\overline{\mathrm{K}} &= -\frac{1}{2\lambda^2}\left[\frac{\partial}{\partial t^1}\left(\frac{1}{\lambda^2}\frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1}\right) + \frac{\partial}{\partial t^2}\left(\frac{1}{\lambda^2}\frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2}\right)\right] \\ &= -\frac{1}{2\lambda^2}\left[\frac{\partial}{\partial t^1}\left(\frac{1}{\lambda^2}\frac{\partial \lambda^2}{\partial t^1}\right) + \frac{\partial}{\partial t^2}\left(\frac{1}{\lambda^2}\frac{\partial \lambda^2}{\partial t^2}\right)\right] \\ &= -\frac{1}{2\lambda^2}\left[\frac{\partial}{\partial t^1}\left(\frac{\partial}{\partial t^1}\log \lambda^2\right) + \frac{\partial}{\partial t^2}\left(\frac{\partial}{\partial t^2}\log \lambda^2\right)\right] \\ &= -\frac{1}{2\lambda^2}\left[\Delta \log \lambda^2\right] .\end{aligned}$$

Exercise 9.9

Compute

(i) the tangent vector $\mathbf{a}_t$, according to (9.282)$_1$, along with the covariant metric a_{tt}, according to (9.283)$_1$,
(ii) the arc length s, (from 0 to t) via (9.284)$_1$ in order to have the arc length parametrization,
(iii) the unit tangent vector $\widehat{\mathbf{a}}_1^{\mathrm{c}}$, according to (9.288)$_1$,
(iv) the curvature κ^{c}, according to (9.290)$_1$, along with the principal normal vector $\widehat{\mathbf{a}}_2^{\mathrm{c}}$, according to (9.291)$_2$,
(v) the binormal vector $\widehat{\mathbf{a}}_3^{\mathrm{c}}$, according to (9.299),
(vi) the torsion τ^{c}, according to (9.303), and
(vii) the Darboux vector $\mathbf{d}^{\mathrm{c}}$, given in (9.307),

for a **circular cylindrical helix** defined by

$$\underbrace{x_1 = R\cos t \quad , \quad x_2 = R\sin t \quad , \quad x_3 = \widehat{R}t}_{\text{note that the elliptic cylindrical helix (9.281) is called the circular cylindrical helix if } R = \overline{R}} \quad . \tag{9.800}$$

Solution. The tangent vector and the covariant metric tensor:

$$\mathbf{a}_t = -R\sin t\,\widehat{\mathbf{e}}_1 + R\cos t\,\widehat{\mathbf{e}}_2 + \widehat{R}\,\widehat{\mathbf{e}}_3 \ , \tag{9.801a}$$

$$a_{tt} = R^2 + \widehat{R}^2 \ . \tag{9.801b}$$

The natural parametrization:

$$\text{from} \quad s\,(t) = \sqrt{R^2 + \widehat{R}^2}\, t \quad \text{one obtains} \quad \left. \begin{aligned} x_1 &= R \cos \frac{s}{\sqrt{R^2 + \widehat{R}^2}} \\ x_2 &= R \sin \frac{s}{\sqrt{R^2 + \widehat{R}^2}} \\ x_3 &= \frac{\widehat{R} s}{\sqrt{R^2 + \widehat{R}^2}} \end{aligned} \right\} . \tag{9.802}$$

The unit tangent vector:

$$\underbrace{\widehat{\mathbf{a}}_1^{\mathrm{c}} = \frac{-R}{\sqrt{R^2 + \widehat{R}^2}} \left(\sin \frac{s}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_1 - \cos \frac{s}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_2 - \frac{\widehat{R}}{R}\, \widehat{\mathbf{e}}_3 \right)}_{\text{one then has } \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = \frac{-R}{R^2 + \widehat{R}^2} \left(\cos \frac{s}{\sqrt{R^2+\widehat{R}^2}}\, \widehat{\mathbf{e}}_1 + \sin \frac{s}{\sqrt{R^2+\widehat{R}^2}}\, \widehat{\mathbf{e}}_2 \right)} . \tag{9.803}$$

The curvature and the principal normal vector:

$$\kappa^{\mathrm{c}} = \frac{R}{R^2 + \widehat{R}^2} , \tag{9.804a}$$

$$\widehat{\mathbf{a}}_2^{\mathrm{c}} = -\cos \frac{s}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_1 - \sin \frac{s}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_2 . \tag{9.804b}$$

The binormal vector:

$$\underbrace{\widehat{\mathbf{a}}_3^{\mathrm{c}} = \frac{\widehat{R}}{\sqrt{R^2 + \widehat{R}^2}} \left(\sin \frac{s}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_1 - \cos \frac{s}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_2 + \frac{R}{\widehat{R}}\, \widehat{\mathbf{e}}_3 \right)}_{\text{and, consequently, } \frac{d\widehat{\mathbf{a}}_3^{\mathrm{c}}}{ds} = \frac{\widehat{R}}{R^2 + \widehat{R}^2} \left(\cos \frac{s}{\sqrt{R^2+\widehat{R}^2}}\, \widehat{\mathbf{e}}_1 + \sin \frac{s}{\sqrt{R^2+\widehat{R}^2}}\, \widehat{\mathbf{e}}_2 \right)} . \tag{9.805}$$

The torsion:

$$\tau^{\mathrm{c}} = \frac{\widehat{R}}{R^2 + \widehat{R}^2} . \tag{9.806}$$

The Darboux vector:

$$\mathbf{d}^{\mathrm{c}} = \frac{1}{\sqrt{R^2 + \widehat{R}^2}}\, \widehat{\mathbf{e}}_3 . \tag{9.807}$$

Exercise 9.10

Let $\mathbf{x}^{\mathrm{c}}(t) : \mathrm{I} \to \mathscr{E}_{\mathrm{p}}^{3}$ be a regular space curve with nowhere-vanishing curvature. It is parametrized by the general parameter $t \in \mathrm{I} \subset \mathrm{R}$ (not necessarily the arc length). Show that the curvature of $\mathbf{x}^{\mathrm{c}}(t)$ is governed by

$$\kappa^{\mathrm{c}} = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right|^{-3} \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \times \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} \right| , \tag{9.808}$$

and then express the corresponding moving trihedron $\left\{ \widehat{\mathbf{a}}_1^{\mathrm{c}}, \widehat{\mathbf{a}}_2^{\mathrm{c}}, \widehat{\mathbf{a}}_3^{\mathrm{c}} \right\}$ in terms of t. Moreover, show that its torsion can be represented by

$$\tau^{\mathrm{c}} = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \times \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} \right|^{-2} \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \cdot \left(\frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} \times \frac{d^3\mathbf{x}^{\mathrm{c}}}{dt^3} \right) , \tag{9.809}$$

and then reformulate the Frenet formulas (9.305a)-(9.305c) in terms of the time-like variable t. Finally, use these results to calculate the curvature and torsion of a **twisted cubic** (Fig. 9.38) defined by

$$\underbrace{x_1 = t \quad , \quad x_2 = t^2 \quad , \quad x_3 = t^3}_{\text{these coordinates satisfy the implicit relations } x_1^2 - x_2 = x_2^2 - x_1 x_3 = x_1 x_2 - x_3 = 0} \quad . \tag{9.810}$$

Solution. The given curve, in principle, can be parametrized by the arc length as written in $(9.287)_{1-2}$. By means of the chain rule of differentiation,

$$\underbrace{\frac{d\mathbf{x}^{\mathrm{c}}}{dt} = \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} \frac{ds}{dt} \quad \text{where} \quad \frac{ds}{dt} = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right|}_{\text{note that } \frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} = \frac{d\mathbf{x}^{\mathrm{c}}}{dt}\frac{dt}{ds} \text{ where } \frac{dt}{ds} = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right|^{-1}} . \tag{9.811}$$

Consequently,

$$\left(\frac{ds}{dt} \right)^2 = \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \cdot \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \quad \Longrightarrow \quad 2 \frac{ds}{dt} \frac{d^2 s}{dt^2} = 2 \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \cdot \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} ,$$

or

$$\frac{d^2 s}{dt^2} = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right|^{-1} \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \cdot \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} . \tag{9.812}$$

In a similar manner,

$$\frac{d^2 t}{ds^2} = - \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right|^{-4} \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \cdot \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} . \tag{9.813}$$

One then finds that

$$\begin{aligned}
\kappa^{\text{c}} &\xlongequal[\text{(9.290) and (9.811)}]{\text{by using}} \left| \frac{d}{ds}\left(\frac{d\mathbf{x}^{\text{c}}}{dt}\frac{dt}{ds}\right) \right| \\
&\xlongequal[\text{the product and chain rules}]{\text{by using}} \left| \left(\frac{dt}{ds}\right)^2 \frac{d^2\mathbf{x}^{\text{c}}}{dt^2} + \frac{d^2 t}{ds^2}\frac{d\mathbf{x}^{\text{c}}}{dt} \right| \\
&\xlongequal[\text{(9.811) and (9.813)}]{\text{by using}} \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\right|^{-4} \left| \left(\frac{d\mathbf{x}^{\text{c}}}{dt}\cdot\frac{d\mathbf{x}^{\text{c}}}{dt}\right)\frac{d^2\mathbf{x}^{\text{c}}}{dt^2} - \left(\frac{d\mathbf{x}^{\text{c}}}{dt}\cdot\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right)\frac{d\mathbf{x}^{\text{c}}}{dt} \right| \\
&\xlongequal[\text{(1.9a) and (1.72)}]{\text{by using}} \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\right|^{-4} \left| \left(\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right)\times\frac{d\mathbf{x}^{\text{c}}}{dt} \right| \\
&\xlongequal[\text{(1.50)}]{\text{by using}} \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\right|^{-4} \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right| \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\right| . \quad \nwarrow \text{ note that } \left(\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right)\cdot\frac{d\mathbf{x}^{\text{c}}}{dt} = 0
\end{aligned}$$

It should not be difficult now to see that

$$\widehat{\mathbf{a}}_1^{\text{c}} = \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\right|^{-1} \frac{d\mathbf{x}^{\text{c}}}{dt} , \tag{9.814a}$$

$$\widehat{\mathbf{a}}_2^{\text{c}} = \left| \left(\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right)\times\frac{d\mathbf{x}^{\text{c}}}{dt} \right|^{-1} \left(\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right)\times\frac{d\mathbf{x}^{\text{c}}}{dt} , \tag{9.814b}$$

$$\widehat{\mathbf{a}}_3^{\text{c}} = \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right|^{-1} \frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2} . \quad \diamond \tag{9.814c}$$

Next, the desired relation (9.809) is derived. Using $(9.290)_2$, (9.304) and (9.808), one can arrive at

$$\tau^{\text{c}} = \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\right|^{6} \left|\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2}\right|^{-2} \frac{d\hat{\mathbf{x}}^{\text{c}}}{ds}\cdot\left(\frac{d^2\hat{\mathbf{x}}^{\text{c}}}{ds^2}\times\frac{d^3\hat{\mathbf{x}}^{\text{c}}}{ds^3}\right) . \tag{9.815}$$

Thus, the above scalar triple product needs to expressed in terms of the time-like variable t. To do so, consider

$$\frac{d^2\mathbf{x}^{\text{c}}}{dt^2} = \frac{d^2 s}{dt^2}\frac{d\hat{\mathbf{x}}^{\text{c}}}{ds} + \left(\frac{ds}{dt}\right)^2 \frac{d^2\hat{\mathbf{x}}^{\text{c}}}{ds^2} , \quad \leftarrow \text{ note that } \frac{d\mathbf{x}^{\text{c}}}{dt} = \frac{ds}{dt}\frac{d\hat{\mathbf{x}}^{\text{c}}}{ds} \tag{9.816a}$$

$$\frac{d^3\mathbf{x}^{\text{c}}}{dt^3} = \frac{d^3 s}{dt^3}\frac{d\hat{\mathbf{x}}^{\text{c}}}{ds} + 3\frac{ds}{dt}\frac{d^2 s}{dt^2}\frac{d^2\hat{\mathbf{x}}^{\text{c}}}{ds^2} + \left(\frac{ds}{dt}\right)^3 \frac{d^3\hat{\mathbf{x}}^{\text{c}}}{ds^3} . \tag{9.816b}$$

One then obtains

$$\frac{d\mathbf{x}^{\text{c}}}{dt}\times\frac{d^2\mathbf{x}^{\text{c}}}{dt^2} = \left(\frac{ds}{dt}\right)^3 \frac{d\hat{\mathbf{x}}^{\text{c}}}{ds}\times\frac{d^2\hat{\mathbf{x}}^{\text{c}}}{ds^2} , \tag{9.817}$$

and

$$\underbrace{\frac{d^3\mathbf{x}^{\mathrm{c}}}{dt^3} \cdot \left(\frac{d\mathbf{x}^{\mathrm{c}}}{dt} \times \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} \right) = \left(\frac{ds}{dt} \right)^6 \frac{d^3\hat{\mathbf{x}}^{\mathrm{c}}}{ds^3} \cdot \left(\frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} \times \frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \right)}_{\text{note that } \mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) = \mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) \text{, according to (1.73)}} , \tag{9.818}$$

or

$$\frac{d\hat{\mathbf{x}}^{\mathrm{c}}}{ds} \cdot \left(\frac{d^2\hat{\mathbf{x}}^{\mathrm{c}}}{ds^2} \times \frac{d^3\hat{\mathbf{x}}^{\mathrm{c}}}{ds^3} \right) = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right|^{-6} \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \cdot \left(\frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} \times \frac{d^3\mathbf{x}^{\mathrm{c}}}{dt^3} \right) . \tag{9.819}$$

Introducing (9.819) into (9.815) then gives the desired result (9.809).

Recall that the triad $\widehat{\mathbf{a}}_1^{\mathrm{c}}(s)$, $\widehat{\mathbf{a}}_2^{\mathrm{c}}(s)$ and $\widehat{\mathbf{a}}_3^{\mathrm{c}}(s)$ were connected to their derivatives by using the curvature $\kappa^{\mathrm{c}}(s)$ and the torsion $\tau^{\mathrm{c}}(s)$. The results were called the Frenet formulas as demonstrated in (9.305a)-(9.305c). Here, these formulas are reformulated as

$$\frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{dt} = \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right| \kappa^{\mathrm{c}} \widehat{\mathbf{a}}_2^{\mathrm{c}} , \tag{9.820a}$$

$$\frac{d\widehat{\mathbf{a}}_2^{\mathrm{c}}}{dt} = - \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right| \kappa^{\mathrm{c}} \widehat{\mathbf{a}}_1^{\mathrm{c}} + \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right| \tau^{\mathrm{c}} \widehat{\mathbf{a}}_3^{\mathrm{c}} , \tag{9.820b}$$

$$\frac{d\widehat{\mathbf{a}}_3^{\mathrm{c}}}{dt} = - \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right| \tau^{\mathrm{c}} \widehat{\mathbf{a}}_2^{\mathrm{c}} , \tag{9.820c}$$

where the argument t of these functions has been omitted for convenience. The relation (9.820a) can readily be verified by setting $\widehat{\mathbf{a}}_1^{\mathrm{c}}(s) = \widehat{\mathbf{a}}_1^{\mathrm{c}}(t(s))$ and then using the chain rule of differentiation. The remaining results can be proved in a similar manner. ✧ ✧

Finally, the curvature and torsion of the twisted cubic $t \to \left(t, t^2, t^3\right)$ will be represented. For this curve,

$$\left.\begin{aligned} \frac{d\mathbf{x}^{\mathrm{c}}}{dt} &= \widehat{\mathbf{e}}_1 + 2t\,\widehat{\mathbf{e}}_2 + 3t^2\,\widehat{\mathbf{e}}_3 \\ \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} &= 2\,\widehat{\mathbf{e}}_2 + 6t\,\widehat{\mathbf{e}}_3 \\ \frac{d^3\mathbf{x}^{\mathrm{c}}}{dt^3} &= 6\,\widehat{\mathbf{e}}_3 \end{aligned}\right\} \quad , \quad \left.\begin{aligned} \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \right| &= \left(1 + 4t^2 + 9t^4\right)^{1/2} \\ \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \times \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} &= 6t^2\,\widehat{\mathbf{e}}_1 - 6t\,\widehat{\mathbf{e}}_2 + 2\,\widehat{\mathbf{e}}_3 \\ \left| \frac{d\mathbf{x}^{\mathrm{c}}}{dt} \times \frac{d^2\mathbf{x}^{\mathrm{c}}}{dt^2} \right| &= \left(4 + 36t^2 + 36t^4\right)^{1/2} \end{aligned}\right\} .$$

Hence,

$$\kappa^{\mathrm{c}} = \frac{\left(4 + 36t^2 + 36t^4\right)^{1/2}}{\left(1 + 4t^2 + 9t^4\right)^{3/2}} \quad , \quad \tau^{\mathrm{c}} = \frac{3}{1 + 9t^2 + 9t^4} .$$

These functions are plotted in Fig. 9.38. ✧ ✧ ✧

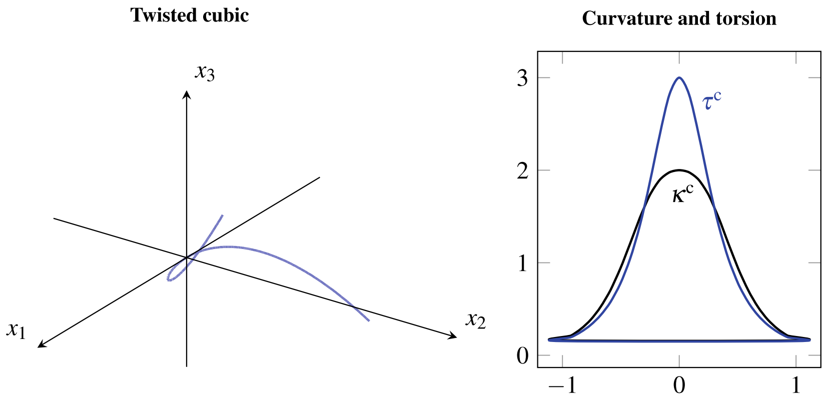


Fig. 9.38 Twisted cubic with its curvature and torsion

Exercise 9.11

Use the variational principles, followed in (9.378)-(9.390), to show that a regular surface is of minimum area if at every point in its domain the mean curvature vanishes.

Solution. A minimal surface is a surface whose area is minimum among a family of surfaces sharing the same boundary. The mean curvature of such a special surface vanishes everywhere. To show this, consider a regular surface $\mathbf{x} : U \subset \mathrm{R}^2 \to \mathscr{E}_{\mathrm{p}}^3$ and choose a bounded region $\tilde{D} \subset U$. Consider now a family of surfaces according to

$$\tilde{\mathbf{x}}\left(t^1, t^2, h\right) = \mathbf{x}\left(t^1, t^2\right) + h\eta\left(t^1, t^2\right)\widehat{\mathbf{n}}\left(t^1, t^2\right) \ , \tag{9.821}$$

where η presents an arbitrary differentiable function and $h \in (-\varepsilon, \varepsilon)$ (noting that $\varepsilon > 0$ is sufficiently small). The map $\tilde{\mathbf{x}} : \tilde{D} \times (-\varepsilon, \varepsilon) \to \mathscr{E}_{\mathrm{p}}^3$ is known as the *normal variation* of $\mathbf{x}\left(\tilde{D}\right)$. To characterize its area, one needs to have the basis vectors

$$\tilde{\mathbf{a}}_1 = \left(1 - h\eta\, \underline{b}_{1}{}^{1}\right)\mathbf{a}_1 + h\frac{\partial \eta}{\partial t^1}\widehat{\mathbf{n}} - h\eta\, \underline{b}_{1}{}^{2}\mathbf{a}_2 \ , \tag{9.822}$$

and

$$\tilde{\mathbf{a}}_2 = \left(1 - h\eta\, \underline{b}_{2}{}^{2}\right)\mathbf{a}_2 + h\frac{\partial \eta}{\partial t^2}\widehat{\mathbf{n}} - h\eta\, \underline{b}_{2}{}^{1}\mathbf{a}_1 \ , \tag{9.823}$$

which deliver the metric coefficients

$$\left.\begin{aligned} \tilde{a}_{11} &= a_{11} - 2h\eta\, \underline{b}_{11} + o\left(h\right) \\ \tilde{a}_{12} &= a_{12} - 2h\eta\, \underline{b}_{12} + o\left(h\right) \\ \tilde{a}_{22} &= a_{22} - 2h\eta\, \underline{b}_{22} + o\left(h\right) \end{aligned}\right\} \ . \tag{9.824}$$

It follows that

$$\tilde{a}_{11}\tilde{a}_{22} - (\tilde{a}_{12})^2 = \left[a_{11}a_{22} - (a_{12})^2\right]\left[1 - 4h\eta\overline{\mathrm{H}}\right] + o\,(h) \ ,$$

where

$$\overline{\mathrm{H}} = \frac{\underline{b}_{11}a_{22} - 2\,\underline{b}_{12}a_{12} + \underline{b}_{22}a_{11}}{2\left(a_{11}a_{22} - a_{12}^2\right)} \ .$$

As a result,

$$\begin{aligned}\tilde{A}\,(h) &= \int_{\tilde{D}} \sqrt{\tilde{a}_{11}\tilde{a}_{22} - (\tilde{a}_{12})^2}\,dt^1 dt^2 \\ &= \int_{\tilde{D}} \left[\left(1 - 2h\eta\overline{\mathrm{H}}\right)\sqrt{a_{11}a_{22} - (a_{12})^2} + o\,(h)\right] dt^1 dt^2 \ , \end{aligned} \tag{9.825}$$

whose derivative at $h = 0$ renders

$$\left.\frac{d\tilde{A}}{dh}\right|_{h=0} = \int_{\tilde{D}} \left(-2\eta\overline{\mathrm{H}}\right)\sqrt{a_{11}a_{22} - (a_{12})^2}\,dt^1 dt^2 \ .$$

Thus,

$$\delta A = \left.\frac{d\tilde{A}}{dh}\right|_{h=0} h = h\int_{\tilde{D}} \left(-2\eta\overline{\mathrm{H}}\right)\sqrt{a_{11}a_{22} - (a_{12})^2}\,dt^1 dt^2 \ . \tag{9.826}$$

The necessary condition for $\tilde{A}\,(h)$ to attain a minimum is that $\delta A = 0$. Having in mind that η is an arbitrary function, by the fundamental lemma of the variational calculus, one will have

$$\overline{\mathrm{H}}\sqrt{a_{11}a_{22} - (a_{12})^2} = 0 \ . \tag{9.827}$$

Recall from (9.20)$_3$ that $a_{11}a_{22} - (a_{12})^2 > 0$. And this implies the required result $\overline{\mathrm{H}} = 0$. That is why a surface with identically vanishing the mean curvature has least area.

Hint: A minimal surface is a special surface whose Gaussian curvature turns out to be either **negative** or **zero**, see the expressions (9.425a) and (9.425b).

Exercise 9.12

Verify (9.134).

Solution. The procedure used here to arrive at the desired result relies on the basic definition $(9.118)_1$, see Fig. 9.11. For the problem at hand, it renders

$$\frac{d\left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)}{dt} = \lim_{\Delta t \to 0} \frac{\hat{\mathbf{h}}^{\parallel}_{Q\to P} \cdot \mathbf{u}^{\parallel}_{Q\to P} - \hat{\mathbf{h}}_P \cdot \mathbf{u}_P}{\Delta t} . \tag{9.828}$$

Let these smooth vector fields be decomposed as $\hat{\mathbf{h}} = \underline{\hat{h}}^\alpha \mathbf{a}_\alpha$ and $\mathbf{u} = \underline{u}^\gamma \mathbf{a}_\gamma$. Then,

$$\frac{d\left(\hat{\mathbf{h}} \cdot \mathbf{u}\right)}{dt} = \lim_{\Delta t \to 0} \frac{\underline{\hat{h}}^\alpha_Q \, \underline{u}^\gamma_Q \, (\mathbf{a}_\alpha)^{\parallel}_{Q\to P} \cdot \left(\mathbf{a}_\gamma\right)^{\parallel}_{Q\to P} - \underline{\hat{h}}^\alpha_P \, \underline{u}^\gamma_P \, (\mathbf{a}_\alpha)_P \cdot \left(\mathbf{a}_\gamma\right)_P}{\Delta t} . \tag{9.829}$$

Taking this limit requires characterizing the tensor objects at Q and the parallel transported basis vectors sitting at P. Note that one only needs to expand these quantities up to the first-order. This is described in the following.

To begin with, recall from (9.121) that $(\mathbf{a}_\alpha)^{\parallel}_{Q\to P} = \xi^{\theta}_{\alpha\ P} \, (\mathbf{a}_\theta)_P + \varsigma_{\alpha\ P} \, (\widehat{\mathbf{n}})_P$. By using $(9.12)_{1-2}$, one then has

$$(\mathbf{a}_\alpha)^{\parallel}_{Q\to P} \cdot \left(\mathbf{a}_\gamma\right)^{\parallel}_{Q\to P} = \xi^\theta_\alpha \xi^\rho_\gamma a_{\theta\rho} + \varsigma_\alpha \varsigma_\gamma , \tag{9.830}$$

where the subscript P has been dropped for notational simplicity. For small enough Δt, one can write

$$\xi^\theta_\alpha = \delta^\theta_\alpha + \bar{\xi}^\theta_\alpha \Delta t \quad , \quad \varsigma_\alpha = (\mathbf{0})_\alpha + \bar{\varsigma}_\alpha \Delta t . \tag{9.831}$$

As a result,

$$(\mathbf{a}_\alpha)^{\parallel}_{Q\to P} \cdot \left(\mathbf{a}_\gamma\right)^{\parallel}_{Q\to P} = a_{\alpha\gamma} + \left(a_{\alpha\rho} \bar{\xi}^\rho_\gamma + \bar{\xi}^\theta_\alpha a_{\theta\gamma}\right) \Delta t . \tag{9.832}$$

Consider now $\underline{\hat{h}}^\alpha_Q = \underline{\hat{h}}^\alpha + \left(d \, \underline{\hat{h}}^\alpha / dt\right) \Delta t$, according to (9.120). Consequently,

$$\underline{\hat{h}}^\alpha_Q \, \underline{u}^\gamma_Q = \underline{\hat{h}}^\alpha \, \underline{u}^\gamma + \left(\frac{d \, \underline{\hat{h}}^\alpha}{dt} \, \underline{u}^\gamma + \underline{\hat{h}}^\alpha \frac{d \, \underline{u}^\gamma}{dt} \right) \Delta t . \tag{9.833}$$

Introducing (9.832) and (9.833) into (9.829) leads to

$$\frac{d\left(\hat{\mathbf{h}}\cdot\mathbf{u}\right)}{dt} = \frac{d\,\underline{\hat{h}}^{\alpha}}{dt}a_{\alpha\gamma}\,\underline{u}^{\gamma} + \underline{\hat{h}}^{\alpha}a_{\alpha\gamma}\frac{d\,\underline{u}^{\gamma}}{dt} + \underline{\hat{h}}^{\alpha}\,\underline{u}^{\gamma}a_{\alpha\rho}\lim_{\Delta t\to 0}\frac{\xi^{\rho}_{\gamma}-\delta^{\rho}_{\gamma}}{\Delta t} + \underline{\hat{h}}^{\alpha}\,\underline{u}^{\gamma}a_{\theta\gamma}\lim_{\Delta t\to 0}\frac{\xi^{\theta}_{\alpha}-\delta^{\theta}_{\alpha}}{\Delta t}\,. \tag{9.834}$$

At the end, suppose that t is one of the coordinate curves. By $(9.12)_{1-2}$, $(9.124)_1$, (9.128) and $(9.132)_3$ along with renaming the dummy indices, this expression can be rephrased as

$$\begin{aligned}\frac{\partial\left(\hat{\mathbf{h}}\cdot\mathbf{u}\right)}{\partial t^{\beta}} &= \frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}}a_{\alpha\gamma}\,\underline{u}^{\gamma} + \underline{\hat{h}}^{\alpha}\,\underline{u}^{\gamma}a_{\theta\gamma}\Gamma^{\theta}_{\beta\alpha} + \underline{\hat{h}}^{\alpha}a_{\alpha\gamma}\frac{\partial\,\underline{u}^{\gamma}}{\partial t^{\beta}} + \underline{\hat{h}}^{\alpha}\,\underline{u}^{\gamma}a_{\alpha\rho}\Gamma^{\rho}_{\beta\gamma}\\ &= \underbrace{\left(\frac{\partial\,\underline{\hat{h}}^{\alpha}}{\partial t^{\beta}} + \Gamma^{\alpha}_{\beta\theta}\,\underline{\hat{h}}^{\theta}\right)a_{\alpha\gamma}\,\underline{u}^{\gamma} + \underline{\hat{h}}^{\alpha}a_{\alpha\gamma}\left(\frac{\partial\,\underline{u}^{\gamma}}{\partial t^{\beta}} + \Gamma^{\gamma}_{\beta\rho}\,\underline{u}^{\rho}\right)}_{=\ \underline{\hat{h}}^{\alpha}\big|_{\beta}a_{\alpha\gamma}\,\underline{u}^{\gamma} + \left(\frac{\partial\,\underline{u}^{\alpha}}{\partial t^{\beta}} + \Gamma^{\alpha}_{\beta\theta}\,\underline{u}^{\theta}\right)a_{\alpha\gamma}\,\underline{\hat{h}}^{\gamma}\ =\ \underline{\hat{h}}^{\alpha}\big|_{\beta}a_{\alpha\gamma}\,\underline{u}^{\gamma} + \underline{u}^{\alpha}\big|_{\beta}a_{\alpha\gamma}\,\underline{\hat{h}}^{\gamma}}\\ &= \frac{\partial\hat{\mathbf{h}}}{\partial t^{\beta}}\cdot\mathbf{u} + \hat{\mathbf{h}}\cdot\frac{\partial\mathbf{u}}{\partial t^{\beta}}\,.\end{aligned}$$

Exercise 9.13

Consider the **unit** sphere

$$\left(t^1, t^2\right) \to \left(\sin t^1\cos t^2, \sin t^1\sin t^2, \cos t^1\right)\ ,$$

embedded in the three-dimensional Euclidean space whose nonzero Christoffel symbols, according to (9.747), are

$$\Gamma^{1}_{22} = -\sin t^1\cos t^1 \quad , \quad \Gamma^{2}_{12} = \Gamma^{2}_{21} = \cot t^1\,.$$

Consider now two curves of latitude defined by (see Fig. 9.39)

$$\begin{cases}\mathcal{C}_1:\ \ t\to\left(t^1=\frac{\pi}{2}, t^2=t\right)\\ \mathcal{C}_2:\ \ t\to\left(t^1=\frac{\pi}{4}, t^2=t\right)\end{cases}\ ,\quad 0\le t\le\frac{\pi}{2}\,,$$

on this planet. Let P be a point corresponding to $\left(t^1=\pi/2, t^2=0\right)$ on $\mathcal{C}_1$ and further let Q be a point corresponding to $\left(t^1=\pi/4, t^2=0\right)$ on $\mathcal{C}_2$. First, compute the covariant derivative of a surface vector field

$$\hat{\mathbf{h}} = \cos t^2 \mathbf{a}_1 + \sin t^2 \mathbf{a}_2 ,$$

along the given curves. Then, parallel transport an initial vector $\hat{\mathbf{h}}_0 = \alpha \mathbf{a}_1 + \beta \mathbf{a}_2$ sitting at P (Q) along $\mathcal{C}_1$ $(\mathcal{C}_2)$. Here, α and β are given constants.

Solution. Using (9.243)$_2$, the tangent vector to $\mathcal{C}_1$ simply renders

$$\mathbf{a}_t = \frac{dt^1}{dt} \mathbf{a}_1 + \frac{dt^2}{dt} \mathbf{a}_2 = \mathbf{a}_2 ,$$

which also holds true for $\mathcal{C}_2$. Then, guided by (9.127)$_2$ and (9.128), the covariant derivative of $\hat{\mathbf{h}}$ along $\mathcal{C}_1$ is given by

$$\begin{aligned}
\hat{\mathbf{h}}\Big|_{\mathbf{a}_2} &= \underline{\hat{h}}^1\Big|_2 \mathbf{a}_1 + \underline{\hat{h}}^2\Big|_2 \mathbf{a}_2 \\
&= \left(\frac{\partial \underline{\hat{h}}^1}{\partial t^2} + \Gamma^1_{21} \underline{\hat{h}}^1 + \Gamma^1_{22} \underline{\hat{h}}^2 \right) \mathbf{a}_1 + \left(\frac{\partial \underline{\hat{h}}^2}{\partial t^2} + \Gamma^2_{21} \underline{\hat{h}}^1 + \Gamma^2_{22} \underline{\hat{h}}^2 \right) \mathbf{a}_2 \\
&= \left(- \sin t^2 + 0 - \sin t^1 \cos t^1 \sin t^2 \right) \mathbf{a}_1 + \left(\cos t^2 + \cot t^1 \cos t^2 + 0 \right) \mathbf{a}_2 \\
&= - \sin t^2 \mathbf{a}_1 + \cos t^2 \mathbf{a}_2 .
\end{aligned}$$

In a similar manner, the object $\hat{\mathbf{h}}\Big|_{\mathbf{a}_2}$ along $\mathcal{C}_2$ delivers

$$\hat{\mathbf{h}}\Big|_{\mathbf{a}_2} = -\frac{3}{2} \sin t^2 \mathbf{a}_1 + 2 \cos t^2 \mathbf{a}_2 .$$

The parallel transport of a given vector along a curve results in a **vector field** whose components are the solution of a linear system of (generally coupled) first-order ordinary differential equations. For the first curve in this example, the condition (9.140)$_3$ now takes the form

$$\frac{\partial \underline{\hat{h}}^1}{\partial t^2} = 0 \quad , \quad \frac{\partial \underline{\hat{h}}^2}{\partial t^2} = 0 \quad \text{whose solution is} \quad \underline{\hat{h}}^1 = A \quad , \quad \underline{\hat{h}}^2 = B .$$

Considering the initial conditions $\underline{\hat{h}}^1 \left(t^2 = 0 \right) = \alpha$ and $\underline{\hat{h}}^2 \left(t^2 = 0 \right) = \beta$, one can infer that the parallel transport of $\hat{\mathbf{h}}_0 = \alpha \mathbf{a}_1 + \beta \mathbf{a}_2$ along $\mathcal{C}_1$ is nothing but itself. This is because the given curve represents a geodesic on the sphere.

To parallel transport $\hat{\mathbf{h}}_0 = \alpha \mathbf{a}_1 + \beta \mathbf{a}_2$ along $\mathcal{C}_2$, one needs to solve

$$\frac{\partial \underline{\hat{h}}^1}{\partial t^2} = \frac{1}{2} \underline{\hat{h}}^2 , \ \frac{\partial \underline{\hat{h}}^2}{\partial t^2} = - \underline{\hat{h}}^1 \quad \text{or, since } t^2 = t, \quad \frac{d \underline{\hat{h}}^1}{dt} = \frac{1}{2} \underline{\hat{h}}^2 , \ \frac{d \underline{\hat{h}}^2}{dt} = - \underline{\hat{h}}^1 .$$

The goal is now to decouple these differential equations. This can be done by differentiating these equations with respect to the time-like variable:

$$\frac{d^2\,\hat{\underline{h}}^1}{dt^2} = \frac{1}{2}\frac{d\,\hat{\underline{h}}^2}{dt} \quad , \quad \frac{d^2\,\hat{\underline{h}}^2}{dt^2} = -\frac{d\,\hat{\underline{h}}^1}{dt} \; .$$

And this helps obtain

$$\frac{d^2\,\hat{\underline{h}}^1}{dt^2} + \frac{1}{2}\,\hat{\underline{h}}^1 = 0 \quad , \quad \frac{d^2\,\hat{\underline{h}}^2}{dt^2} + \frac{1}{2}\,\hat{\underline{h}}^2 = 0 \; .$$

For solving the first differential equation, consider $\hat{\underline{h}}^1 = \exp(st)$. The characteristic equation then becomes $s^2 + 1/2 = 0$ whose roots are $s_1 = i\sqrt{2}/2$ and $s_2 = -i\sqrt{2}/2$ where i denotes the imaginary unit. By superposition, one can arrive at

$$\hat{\underline{h}}^1 = \bar{A}\exp\left(\frac{\sqrt{2}i}{2}t\right) + \bar{B}\exp\left(-\frac{\sqrt{2}i}{2}t\right) \; ,$$

or, using the Euler formulas $\exp(\pm\, i t) = \cos t \pm i \sin t$,

$$\hat{\underline{h}}^1 = A\cos\left(\frac{\sqrt{2}}{2}t\right) + B\sin\left(\frac{\sqrt{2}}{2}t\right) \; .$$

In a similar manner,

$$\hat{\underline{h}}^2 = C\cos\left(\frac{\sqrt{2}}{2}t\right) + D\sin\left(\frac{\sqrt{2}}{2}t\right) \; .$$

It only remains to determine these four constants. First, given that $\hat{\underline{h}}^1(0) = \alpha$ and $\hat{\underline{h}}^2(0) = \beta$,

$$\hat{\underline{h}}^1 = \alpha\cos\left(\frac{\sqrt{2}}{2}t\right) + B\sin\left(\frac{\sqrt{2}}{2}t\right) \; , \; \hat{\underline{h}}^2 = \beta\cos\left(\frac{\sqrt{2}}{2}t\right) + D\sin\left(\frac{\sqrt{2}}{2}t\right) \; .$$

Then, consider

$$\frac{d\,\hat{\underline{h}}^1}{dt} = \frac{1}{2}\,\hat{\underline{h}}^2 \quad \text{or} \quad \left(\beta - \sqrt{2}B\right)\cos\left(\frac{\sqrt{2}}{2}t\right) + \left(D + \sqrt{2}\alpha\right)\sin\left(\frac{\sqrt{2}}{2}t\right) = 0 \; .$$

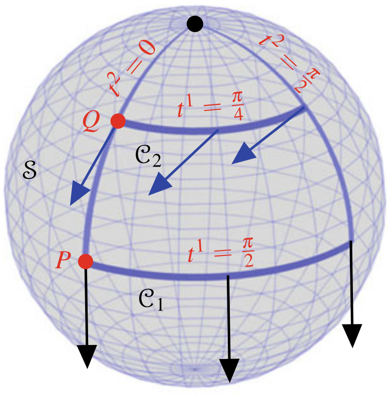
Notice that this result should hold for any t in $[0, \pi/2]$. As a result, one can finally arrive at

$$\hat{\mathbf{h}} = \left[\alpha \cos\left(\frac{\sqrt{2}}{2}t\right) + \frac{\sqrt{2}}{2}\beta \sin\left(\frac{\sqrt{2}}{2}t\right)\right]\mathbf{a}_1 + \left[\beta \cos\left(\frac{\sqrt{2}}{2}t\right) - \sqrt{2}\alpha \sin\left(\frac{\sqrt{2}}{2}t\right)\right]\mathbf{a}_2 \,.$$

Notice that $\mathcal{C}_2$ is not the straightest possible path on the planet (i.e. the acceleration vector has the tangential component). The initial vector $\hat{\mathbf{h}}_0 = \alpha\mathbf{a}_1 + \beta\mathbf{a}_2$ thus changes from point to point when it is parallel transported along $\mathcal{C}_2$.

Figure 9.39 illustrates the results when $\alpha = 1$ and $\beta = 0$. In this case, note that $\hat{\mathbf{h}}$ is a vector of unit length:

$$\begin{aligned}\hat{\mathbf{h}} \cdot \hat{\mathbf{h}} &= \cos^2\left(\frac{\sqrt{2}}{2}t\right) a_{11}\left(\frac{\pi}{4}, t\right) + 0 + 2\sin^2\left(\frac{\sqrt{2}}{2}t\right) a_{22}\left(\frac{\pi}{4}, t\right) \\ &= \cos^2\left(\frac{\sqrt{2}}{2}t\right)(1) + 2\sin^2\left(\frac{\sqrt{2}}{2}t\right)\left(\sin^2\frac{\pi}{4}\right) \\ &= 1 \,.\end{aligned}$$



Observe that how the unit vector $\mathbf{a}_1(Q)$ surfers a change in its orientation when it is parallel transported from point to point along $\mathcal{C}_2$.

Let $\mathcal{S}$ be a unit sphere defined by

$$\left(t^1, t^2\right) \to \left(\sin t^1 \cos t^2, \sin t^1 \sin t^2, \cos t^1\right) ,$$

and consider two points $P = (\pi/2, 0)$ and $Q = (\pi/4, 0)$ on this surface. Consider now two curves of latitude defined by

$$\begin{cases}\mathcal{C}_1: & t \to \left(t^1 = \frac{\pi}{2}, t^2 = t\right) \\ \mathcal{C}_2: & t \to \left(t^1 = \frac{\pi}{4}, t^2 = t\right)\end{cases} , \; 0 \le t \le \frac{\pi}{2} ,$$

on $\mathcal{S}$. The covariant derivative of $\hat{\mathbf{h}} = \mathbf{a}_1(P)$ along $\mathcal{C}_1$ is itself. But, the covariant derivative of $\hat{\mathbf{h}} = \mathbf{a}_1(Q)$ along $\mathcal{C}_2$ renders

$$\hat{\mathbf{h}} = \cos\left(\frac{\sqrt{2}}{2}t\right)\mathbf{a}_1 - \sqrt{2}\sin\left(\frac{\sqrt{2}}{2}t\right)\mathbf{a}_2 \,.$$

The results have been plotted in this figure.

Fig. 9.39 Parallel transport along two (geodesic and non-geodesic) curves on unit sphere

Exercise 9.14

Verify (9.360)-(9.363).

Solution. Consider a s-parametrized curve according to (9.336) whose first and second derivatives are given in (9.337) and (9.338). From these relations, one can obtain

$$\begin{aligned}\frac{d\hat{\mathbf{x}}^c}{ds}\times\frac{d^2\hat{\mathbf{x}}^c}{ds^2} &= (\mathbf{a}_1\times\mathbf{a}_2)\left(\frac{dt^1}{ds}\frac{d^2t^2}{ds^2}-\frac{d^2t^1}{ds^2}\frac{dt^2}{ds}\right)\\ &+\left(\mathbf{a}_1\times\frac{\partial\mathbf{a}_1}{\partial t^1}\right)\left(\frac{dt^1}{ds}\right)^3+\left(\mathbf{a}_2\times\frac{\partial\mathbf{a}_2}{\partial t^2}\right)\left(\frac{dt^2}{ds}\right)^3\\ &+\left(\mathbf{a}_1\times\frac{\partial\mathbf{a}_2}{\partial t^2}+2\mathbf{a}_2\times\frac{\partial\mathbf{a}_1}{\partial t^2}\right)\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2\\ &+\left(2\mathbf{a}_1\times\frac{\partial\mathbf{a}_1}{\partial t^2}+\mathbf{a}_2\times\frac{\partial\mathbf{a}_1}{\partial t^1}\right)\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}\,. \end{aligned}\tag{9.835}$$

Using (1.11), (1.78a), $(9.17)_1$, $(9.20)_3$, $(9.29)_3$, $(9.31)_1$, $(9.54)_2$, (9.94), (9.359) and (9.835), one will have

$$\begin{aligned}\kappa^g &= \left(\frac{d\hat{\mathbf{x}}^c}{ds}\times\frac{d^2\hat{\mathbf{x}}^c}{ds^2}\right)\cdot\frac{\mathbf{a}_1\times\mathbf{a}_2}{\sqrt{a}} \quad\longleftarrow\ \text{note that } \mathbf{a}_1\times\mathbf{a}_2=|\mathbf{a}_1\times\mathbf{a}_2|\,\widehat{\mathbf{n}} \text{ where } |\mathbf{a}_1\times\mathbf{a}_2|=\sqrt{a_{11}a_{22}-(a_{12})^2}=\sqrt{a}\\ &= \underbrace{\frac{1}{\sqrt{a}}(\mathbf{a}_1\times\mathbf{a}_2)\cdot(\mathbf{a}_1\times\mathbf{a}_2)\left(\frac{dt^1}{ds}\frac{d^2t^2}{ds^2}-\frac{d^2t^1}{ds^2}\frac{dt^2}{ds}\right)}_{=\frac{a}{\sqrt{a}}\left(\frac{dt^1}{ds}\frac{d^2t^2}{ds^2}-\frac{d^2t^1}{ds^2}\frac{dt^2}{ds}\right)=\sqrt{a}\left(\frac{dt^1}{ds}\frac{d^2t^2}{ds^2}-\frac{d^2t^1}{ds^2}\frac{dt^2}{ds}\right)}\\ &+ \underbrace{\frac{1}{\sqrt{a}}\left[\left(\mathbf{a}_1\times\frac{\partial\mathbf{a}_1}{\partial t^1}\right)\cdot(\mathbf{a}_1\times\mathbf{a}_2)\right]\left(\frac{dt^1}{ds}\right)^3}_{\substack{=\frac{1}{\sqrt{a}}\left[(\mathbf{a}_1\cdot\mathbf{a}_1)\left(\frac{\partial\mathbf{a}_1}{\partial t^1}\cdot\mathbf{a}_2\right)-(\mathbf{a}_1\cdot\mathbf{a}_2)\left(\frac{\partial\mathbf{a}_1}{\partial t^1}\cdot\mathbf{a}_1\right)\right]\left(\frac{dt^1}{ds}\right)^3\\ =\frac{1}{\sqrt{a}}\left[a_{11}\left(\Gamma^1_{11}a_{12}+\Gamma^2_{11}a_{22}\right)-a_{12}\left(\Gamma^1_{11}a_{11}+\Gamma^2_{11}a_{12}\right)\right]\left(\frac{dt^1}{ds}\right)^3=\left[\sqrt{a}\,\Gamma^2_{11}\right]\left(\frac{dt^1}{ds}\right)^3}}\\ &+ \underbrace{\frac{1}{\sqrt{a}}\left[\left(\mathbf{a}_2\times\frac{\partial\mathbf{a}_2}{\partial t^2}\right)\cdot(\mathbf{a}_1\times\mathbf{a}_2)\right]\left(\frac{dt^2}{ds}\right)^3}_{\substack{=\frac{1}{\sqrt{a}}\left[(\mathbf{a}_2\cdot\mathbf{a}_1)\left(\frac{\partial\mathbf{a}_2}{\partial t^2}\cdot\mathbf{a}_2\right)-(\mathbf{a}_2\cdot\mathbf{a}_2)\left(\frac{\partial\mathbf{a}_2}{\partial t^2}\cdot\mathbf{a}_1\right)\right]\left(\frac{dt^2}{ds}\right)^3\\ =\frac{1}{\sqrt{a}}\left[a_{12}\left(\Gamma^1_{22}a_{12}+\Gamma^2_{22}a_{22}\right)-a_{22}\left(\Gamma^1_{22}a_{11}+\Gamma^2_{22}a_{12}\right)\right]\left(\frac{dt^2}{ds}\right)^3=\left[-\sqrt{a}\,\Gamma^1_{22}\right]\left(\frac{dt^2}{ds}\right)^3}}\end{aligned}$$

$$+ \quad \underbrace{\frac{1}{\sqrt{a}}\left[\left(\mathbf{a}_1 \times \frac{\partial \mathbf{a}_2}{\partial t^2}\right)\cdot(\mathbf{a}_1 \times \mathbf{a}_2)\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2}$$

$$= \frac{1}{\sqrt{a}}\left[(\mathbf{a}_1\cdot\mathbf{a}_1)\left(\frac{\partial \mathbf{a}_2}{\partial t^2}\cdot\mathbf{a}_2\right) - (\mathbf{a}_1\cdot\mathbf{a}_2)\left(\frac{\partial \mathbf{a}_2}{\partial t^2}\cdot\mathbf{a}_1\right)\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2$$

$$= \frac{1}{\sqrt{a}}\left[a_{11}\left(\Gamma^1_{22}a_{12}+\Gamma^2_{22}a_{22}\right) - a_{12}\left(\Gamma^1_{22}a_{11}+\Gamma^2_{22}a_{12}\right)\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2 = \left[\sqrt{a}\,\Gamma^2_{22}\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2$$

$$+ \quad \underbrace{\frac{2}{\sqrt{a}}\left[\left(\mathbf{a}_2 \times \frac{\partial \mathbf{a}_1}{\partial t^2}\right)\cdot(\mathbf{a}_1 \times \mathbf{a}_2)\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2}$$

$$= \frac{2}{\sqrt{a}}\left[(\mathbf{a}_2\cdot\mathbf{a}_1)\left(\frac{\partial \mathbf{a}_1}{\partial t^2}\cdot\mathbf{a}_2\right) - (\mathbf{a}_2\cdot\mathbf{a}_2)\left(\frac{\partial \mathbf{a}_1}{\partial t^2}\cdot\mathbf{a}_1\right)\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2$$

$$= \frac{2}{\sqrt{a}}\left[a_{12}\left(\Gamma^1_{12}a_{12}+\Gamma^2_{12}a_{22}\right) - a_{22}\left(\Gamma^1_{12}a_{11}+\Gamma^2_{12}a_{12}\right)\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2 = \left[-2\sqrt{a}\,\Gamma^1_{12}\right]\frac{dt^1}{ds}\left(\frac{dt^2}{ds}\right)^2$$

$$+ \quad \underbrace{\frac{2}{\sqrt{a}}\left[\left(\mathbf{a}_1 \times \frac{\partial \mathbf{a}_1}{\partial t^2}\right)\cdot(\mathbf{a}_1 \times \mathbf{a}_2)\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}}$$

$$= \frac{2}{\sqrt{a}}\left[(\mathbf{a}_1\cdot\mathbf{a}_1)\left(\frac{\partial \mathbf{a}_1}{\partial t^2}\cdot\mathbf{a}_2\right) - (\mathbf{a}_1\cdot\mathbf{a}_2)\left(\frac{\partial \mathbf{a}_1}{\partial t^2}\cdot\mathbf{a}_1\right)\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}$$

$$= \frac{2}{\sqrt{a}}\left[a_{11}\left(\Gamma^1_{12}a_{12}+\Gamma^2_{12}a_{22}\right) - a_{12}\left(\Gamma^1_{12}a_{11}+\Gamma^2_{12}a_{12}\right)\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds} = \left[2\sqrt{a}\,\Gamma^2_{12}\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}$$

$$+ \quad \underbrace{\frac{1}{\sqrt{a}}\left[\left(\mathbf{a}_2 \times \frac{\partial \mathbf{a}_1}{\partial t^1}\right)\cdot(\mathbf{a}_1 \times \mathbf{a}_2)\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}} \quad .$$

$$= \frac{1}{\sqrt{a}}\left[(\mathbf{a}_2\cdot\mathbf{a}_1)\left(\frac{\partial \mathbf{a}_1}{\partial t^1}\cdot\mathbf{a}_2\right) - (\mathbf{a}_2\cdot\mathbf{a}_2)\left(\frac{\partial \mathbf{a}_1}{\partial t^1}\cdot\mathbf{a}_1\right)\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}$$

$$= \frac{1}{\sqrt{a}}\left[a_{12}\left(\Gamma^1_{11}a_{12}+\Gamma^2_{11}a_{22}\right) - a_{22}\left(\Gamma^1_{11}a_{11}+\Gamma^2_{11}a_{12}\right)\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds} = \left[-\sqrt{a}\,\Gamma^1_{11}\right]\left(\frac{dt^1}{ds}\right)^2\frac{dt^2}{ds}$$

which, by simplification, delivers the desired result (9.360). The elegant expression (9.361) then immediately follows.

The required result (9.362) can be shown in a similar fashion as follows:

$$\sqrt{a}\kappa^{\mathrm{g}} = (\mathbf{a}_1 \times \mathbf{a}_2)\cdot\left(\widehat{\mathbf{a}}^{\mathrm{c}}_1 \times \frac{d\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{ds}\right)$$

$$= \quad \underbrace{(\mathbf{a}_1 \times \mathbf{a}_2)\cdot\left(\widehat{\mathbf{a}}^{\mathrm{c}}_1 \times \frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^1}\right)\frac{dt^1}{ds}}$$

$$= \left(\frac{dt^1}{ds}\mathbf{a}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)\left(\mathbf{a}_2\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^1}\right) - \left(\frac{dt^1}{ds}\mathbf{a}_1\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^1}\right)\left(\mathbf{a}_2\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)$$

$$= \left(\widehat{\mathbf{a}}^{\mathrm{c}}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)\left(\mathbf{a}_2\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^1}\right) - \cancel{\left(\frac{dt^2}{ds}\mathbf{a}_2\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)\left(\mathbf{a}_2\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^1}\right)} - (0)\left(\mathbf{a}_2\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right) + \cancel{\left(\frac{dt^2}{ds}\mathbf{a}_2\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^1}\right)\left(\mathbf{a}_2\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)}$$

$$+ \quad \underbrace{(\mathbf{a}_1 \times \mathbf{a}_2)\cdot\left(\widehat{\mathbf{a}}^{\mathrm{c}}_1 \times \frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^2}\right)\frac{dt^2}{ds}}$$

$$= \left(\mathbf{a}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)\left(\frac{dt^2}{ds}\mathbf{a}_2\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^2}\right) - \left(\mathbf{a}_1\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^2}\right)\left(\frac{dt^2}{ds}\mathbf{a}_2\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)$$

$$= \left(\mathbf{a}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)(0) - \cancel{\left(\mathbf{a}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)\left(\frac{dt^1}{ds}\mathbf{a}_1\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^2}\right)} - \left(\mathbf{a}_1\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^2}\right)\left(\widehat{\mathbf{a}}^{\mathrm{c}}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right) + \cancel{\left(\mathbf{a}_1\cdot\frac{\partial\,\widehat{\mathbf{a}}^{\mathrm{c}}_1}{\partial t^2}\right)\left(\frac{dt^1}{ds}\mathbf{a}_1\cdot\widehat{\mathbf{a}}^{\mathrm{c}}_1\right)}$$

$$
\begin{aligned}
&= (1)\left(\mathbf{a}_2 \cdot \frac{\partial\, \widehat{\mathbf{a}}_1^{\mathrm{c}}}{\partial t^1} - \mathbf{a}_1 \cdot \frac{\partial\, \widehat{\mathbf{a}}_1^{\mathrm{c}}}{\partial t^2}\right) \\
&= \frac{\partial}{\partial t^1}\left(\mathbf{a}_2 \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}\right) - \cancel{\frac{\partial \mathbf{a}_2}{\partial t^1} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}} - \frac{\partial}{\partial t^2}\left(\mathbf{a}_1 \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}\right) + \cancel{\frac{\partial \mathbf{a}_1}{\partial t^2} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}} \; .
\end{aligned}
$$

To represent the geodesic curvature of $\varphi\left(t^1, t^2\right) = \text{constant}$, according to (9.363), consider

$$
\frac{\partial \varphi}{\partial t^1}\frac{dt^1}{ds} = -\frac{\partial \varphi}{\partial t^2}\frac{dt^2}{ds} \quad \text{and let} \quad \frac{dt^1}{ds} = \lambda \frac{\partial \varphi}{\partial t^2} \quad , \quad \frac{dt^2}{ds} = -\lambda \frac{\partial \varphi}{\partial t^1} \; .
$$

The first fundamental form $ds^2 = a_{11}\left(dt^1\right)^2 + 2a_{12}dt^1dt^2 + a_{22}\left(dt^2\right)^2$ then yields

$$
\lambda = \pm\frac{1}{\tilde{\Omega}} \quad \text{where} \quad \tilde{\Omega} := \sqrt{a_{11}\left(\frac{\partial \varphi}{\partial t^2}\right)^2 - 2a_{12}\frac{\partial \varphi}{\partial t^1}\frac{\partial \varphi}{\partial t^2} + a_{22}\left(\frac{\partial \varphi}{\partial t^1}\right)^2} \; .
$$

Consequently, the tangent vector to the curve under consideration takes the form

$$
\widehat{\mathbf{a}}_1^{\mathrm{c}} = \pm\frac{1}{\tilde{\Omega}}\frac{\partial \varphi}{\partial t^2}\mathbf{a}_1 \mp \frac{1}{\tilde{\Omega}}\frac{\partial \varphi}{\partial t^1}\mathbf{a}_2 \; .
$$

This helps obtain

$$
\mathbf{a}_2 \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} = \pm\frac{1}{\tilde{\Omega}}\frac{\partial \varphi}{\partial t^2}a_{12} \mp \frac{1}{\tilde{\Omega}}\frac{\partial \varphi}{\partial t^1}a_{22} \quad , \quad \mathbf{a}_1 \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} = \pm\frac{1}{\tilde{\Omega}}\frac{\partial \varphi}{\partial t^2}a_{11} \mp \frac{1}{\tilde{\Omega}}\frac{\partial \varphi}{\partial t^1}a_{12} \; .
$$

Introducing these expressions into (9.362) finally gives the desired result (9.363).

Exercise 9.15

Consider a regular surface $\mathcal{S}$ described by

$$
x_1 = Rt^2 \cos t^1 \quad , \quad x_2 = Rt^2 \sin t^1 \quad , \quad x_3 = \overline{R} \log t^2 \; , \tag{9.836}
$$

where $0 \leq t^1 < 2\pi, 0 < t^2 < \infty$ and $R, \overline{R}$ are positive real numbers. This is known as the **funnel surface** (Fig. 9.40). First, compute the desired quantities listed in Exercise 9.1 for this surface of revolution. Then, write a **computer program** to compute its geodesics.

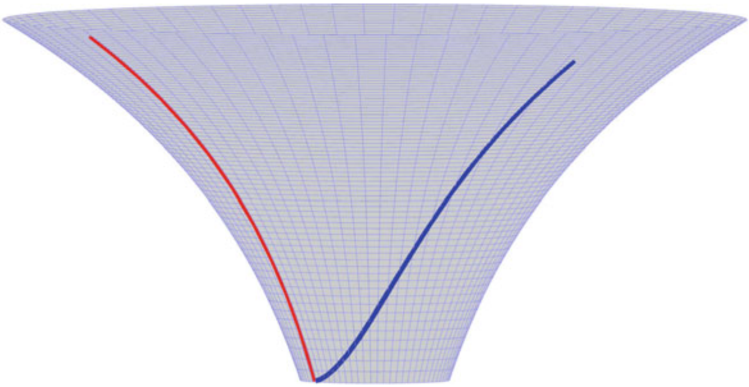


Fig. 9.40 Geodesics on funnel surface

Solution. The covariant basis vectors:

$$\mathbf{a}_1 = -Rt^2 \sin t^1 \widehat{\mathbf{e}}_1 + Rt^2 \cos t^1 \widehat{\mathbf{e}}_2 \ , \tag{9.837a}$$

$$\mathbf{a}_2 = R \cos t^1 \widehat{\mathbf{e}}_1 + R \sin t^1 \widehat{\mathbf{e}}_2 + \frac{\overline{R}}{t^2} \widehat{\mathbf{e}}_3 \ . \tag{9.837b}$$

The covariant metric coefficients in matrix form:

$$\left[a_{\alpha\beta} \right] = \begin{bmatrix} \left(Rt^2\right)^2 & 0 \\ 0 & R^2 + \left(\frac{\overline{R}}{t^2}\right)^2 \end{bmatrix} . \tag{9.838}$$

The area element:

$$dA = R\sqrt{\left(Rt^2\right)^2 + \overline{R}^2}\, dt^1 dt^2 \ . \tag{9.839}$$

The unit normal field:

$$\widehat{\mathbf{n}} = \frac{\overline{R} \cos t^1 \widehat{\mathbf{e}}_1 + \overline{R} \sin t^1 \widehat{\mathbf{e}}_2 - Rt^2 \widehat{\mathbf{e}}_3}{\sqrt{\left(Rt^2\right)^2 + \overline{R}^2}} \ . \tag{9.840}$$

The covariant curvature tensor in matrix form:

$$\left[\underline{b}_{\alpha\beta} \right] = \frac{R\,\overline{R}}{\sqrt{\left(Rt^2\right)^2 + \overline{R}^2}} \begin{bmatrix} -t^2 & 0 \\ 0 & \frac{1}{t^2} \end{bmatrix} . \tag{9.841}$$

The mixed curvature tensor in matrix form:

$$\left[\ \underline{b}_{\alpha}^{\cdot\beta}\ \right] = \frac{R\,\overline{R}}{\sqrt{\left(Rt^2\right)^2 + \overline{R}^2}} \begin{bmatrix} -\frac{1}{R^2t^2} & 0 \\ 0 & \frac{t^2}{\left(Rt^2\right)^2+\overline{R}^2} \end{bmatrix} . \tag{9.842}$$

The mean curvature:

$$\overline{\mathrm{H}} = -\frac{\overline{R}^3}{2Rt^2\left[\left(Rt^2\right)^2 + \overline{R}^2\right]^{3/2}} . \tag{9.843}$$

The Gaussian curvature:

$$\overline{\mathrm{K}} = -\frac{\overline{R}^2}{\left[\left(Rt^2\right)^2 + \overline{R}^2\right]^2} . \tag{9.844}$$

The nonzero Christoffel symbols entries:

$$\Gamma^2_{11} = -\frac{R^2\left(t^2\right)^3}{\left(Rt^2\right)^2 + \overline{R}^2} \quad , \quad \Gamma^1_{12} = \frac{1}{t^2} \quad , \quad \Gamma^2_{22} = -\frac{\overline{R}^2}{R^2\left(t^2\right)^3 + \overline{R}^2 t^2} . \tag{9.845}$$

Notice that $\mathrm{F}_{\mathbf{r}} = a_{12} = 0$ at each point of the funnel surface. And this means that the parametric equations (9.836) represent an orthogonal parametrization (note that (9.836) is eventually a t^2-Clairaut parametrization). One thus needs to solve the differential equations (9.422a)-(9.422d) to determine the geodesic curves on such a surface. The desired code can be downloaded for free from the website address https://data.uni-hannover.de/dataset/exercises-tensor-analysis.

Exercise 9.16

Verify (9.463a)-(9.463b).

Solution. Recall from Theorem D on Sect. 9.7.4 that when the coordinate curves are the lines of curvature at a non-umbilical point on a regular surface, the matrices $\left[\ a_{\alpha\beta}\ \right]$ and $\left[\ \underline{b}_{\alpha\beta}\ \right]$ have zero off-diagonal entries. Making use of the equations (9.162d), (9.235a)-(9.235c) and $(9.480)_1$, one then has

$$\frac{\partial\,\underline{b}_{11}}{\partial t^2} = \Gamma^1_{12}\,\underline{b}_{11} - \Gamma^2_{11}\,\underline{b}_{22} = \left(\frac{\underline{b}_{11}}{2a_{11}} + \frac{\underline{b}_{22}}{2a_{22}}\right)\frac{\partial a_{11}}{\partial t^2} , \tag{9.846}$$

or

$$\frac{\partial \mathrm{e}_{\mathbf{r}}}{\partial t^2} = \Gamma^1_{12}\mathrm{e}_{\mathbf{r}} - \Gamma^2_{11}\mathrm{g}_{\mathbf{r}} = \left(\frac{\mathrm{e}_{\mathbf{r}}}{2\mathrm{E}_{\mathbf{r}}} + \frac{\mathrm{g}_{\mathbf{r}}}{2\mathrm{G}_{\mathbf{r}}}\right)\frac{\partial \mathrm{E}_{\mathbf{r}}}{\partial t^2} . \tag{9.847}$$

In a similar manner,

$$\frac{\partial\, \underline{b}_{22}}{\partial t^1} = \Gamma^2_{12}\, \underline{b}_{22} - \Gamma^1_{22}\, \underline{b}_{11} = \left(\frac{\underline{b}_{22}}{2a_{22}} + \frac{\underline{b}_{11}}{2a_{11}} \right) \frac{\partial a_{22}}{\partial t^1} \ , \tag{9.848}$$

or

$$\frac{\partial \mathrm{g}_{\mathbf{r}}}{\partial t^1} = \Gamma^2_{12} \mathrm{g}_{\mathbf{r}} - \Gamma^1_{22} \mathrm{e}_{\mathbf{r}} = \left(\frac{\mathrm{g}_{\mathbf{r}}}{2\mathrm{G}_{\mathbf{r}}} + \frac{\mathrm{e}_{\mathbf{r}}}{2\mathrm{E}_{\mathbf{r}}} \right) \frac{\partial \mathrm{G}_{\mathbf{r}}}{\partial t^1} \ . \tag{9.849}$$

These relations help obtain

$$\underbrace{\frac{\partial}{\partial t^2} \frac{\underline{b}_{11}}{a_{11}} = \frac{\partial a_{11}}{2a_{11} \partial t^2} \left(\frac{\underline{b}_{22}}{a_{22}} - \frac{\underline{b}_{11}}{a_{11}} \right) \ , \ \frac{\partial}{\partial t^1} \frac{\underline{b}_{22}}{a_{22}} = \frac{\partial a_{22}}{2a_{22} \partial t^1} \left(\frac{\underline{b}_{11}}{a_{11}} - \frac{\underline{b}_{22}}{a_{22}} \right)}_{\text{or } \frac{\partial}{\partial t^2} \frac{\mathrm{e}_{\mathbf{r}}}{\mathrm{E}_{\mathbf{r}}} = \frac{\partial \mathrm{E}_{\mathbf{r}}}{2\mathrm{E}_{\mathbf{r}} \partial t^2} \left(\frac{\mathrm{g}_{\mathbf{r}}}{\mathrm{G}_{\mathbf{r}}} - \frac{\mathrm{e}_{\mathbf{r}}}{\mathrm{E}_{\mathbf{r}}} \right) \ , \ \frac{\partial}{\partial t^1} \frac{\mathrm{g}_{\mathbf{r}}}{\mathrm{G}_{\mathbf{r}}} = \frac{\partial \mathrm{G}_{\mathbf{r}}}{2\mathrm{G}_{\mathbf{r}} \partial t^1} \left(\frac{\mathrm{e}_{\mathbf{r}}}{\mathrm{E}_{\mathbf{r}}} - \frac{\mathrm{g}_{\mathbf{r}}}{\mathrm{G}_{\mathbf{r}}} \right)} \ ,$$

or, using (9.352)-(9.353),

$$\underbrace{\frac{\partial \kappa_1}{\partial t^2} = \frac{\partial a_{11}}{2a_{11} \partial t^2} \left(\kappa_2 - \kappa_1 \right) \ , \ \frac{\partial \kappa_2}{\partial t^1} = \frac{\partial a_{22}}{2a_{22} \partial t^1} \left(\kappa_1 - \kappa_2 \right)}_{\text{or } \frac{\partial \kappa_1}{\partial t^2} = \frac{\partial \mathrm{E}_{\mathbf{r}}}{2\mathrm{E}_{\mathbf{r}} \partial t^2} \left(\kappa_2 - \kappa_1 \right) \ , \ \frac{\partial \kappa_2}{\partial t^1} = \frac{\partial \mathrm{G}_{\mathbf{r}}}{2\mathrm{G}_{\mathbf{r}} \partial t^1} \left(\kappa_1 - \kappa_2 \right)} \ .$$

Exercise 9.17

A saddle-like surface can be resulted from the Monge patch (9.786) via the following parametrization

$$x_1 = t^1 \quad , \quad x_2 = t^2 \quad , \quad x_3 = \left(t^1 \right)^3 - 3t^1 \left(t^2 \right)^2 \ . \tag{9.850}$$

This is called **monkey saddle** (because the shape of this surface has room for two legs and also a tail, see Fig. 9.41). First, find the principal curvatures and directions of this surface at a given point P corresponding to (0, 1). Then, verify the Rodrigues formula (9.464) in the resulting directions.

Solution. The covariant basis vectors are

$$\mathbf{a}_1 = \widehat{\mathbf{e}}_1 + \left[3 \left(t^1 \right)^2 - 3 \left(t^2 \right)^2 \right] \widehat{\mathbf{e}}_3 \quad , \quad \mathbf{a}_1 \left(P \right) = \widehat{\mathbf{e}}_1 - 3 \widehat{\mathbf{e}}_3 \ , \tag{9.851a}$$

$$\mathbf{a}_2 = \widehat{\mathbf{e}}_2 + \left[-6t^1 t^2 \right] \widehat{\mathbf{e}}_3 \quad , \quad \mathbf{a}_2 \left(P \right) = \widehat{\mathbf{e}}_2 \ . \tag{9.851b}$$

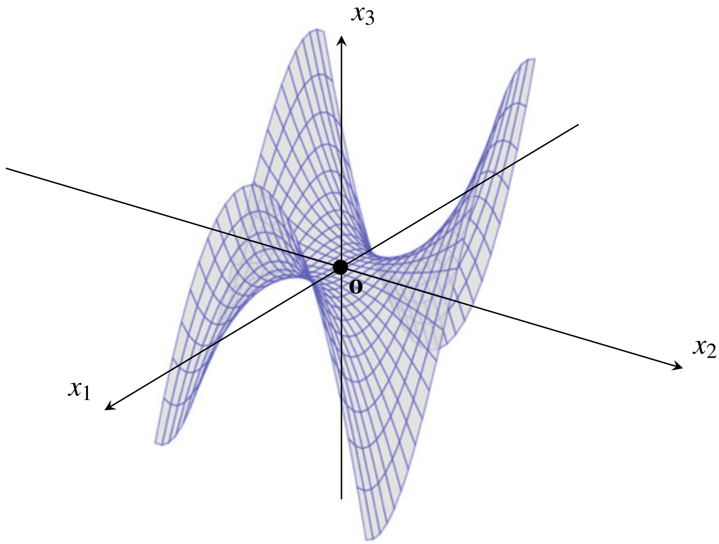


Fig. 9.41 Monkey saddle

The covariant metric coefficients in matrix form,

$$\left[\, a_{\alpha\beta} \,\right] = \begin{bmatrix} 1 + 9\left[\left(t^1\right)^2 - \left(t^2\right)^2\right]^2 & -18t^1t^2\left[\left(t^1\right)^2 - \left(t^2\right)^2\right] \\ -18t^1t^2\left[\left(t^1\right)^2 - \left(t^2\right)^2\right] & 1 + 36\left(t^1t^2\right)^2 \end{bmatrix} , \tag{9.852}$$

at the given point gives

$$\left[\, a_{\alpha\beta}\left(P\right) \,\right] = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix} .$$

The unit normal field,

$$\widehat{\mathbf{n}} = \frac{3\left[-\left(t^1\right)^2 + \left(t^2\right)^2\right]\widehat{\mathbf{e}}_1 + 6\left[t^1t^2\right]\widehat{\mathbf{e}}_2 + \widehat{\mathbf{e}}_3}{\sqrt{\widetilde{\Omega} := 1 + 9\left[\left(t^1\right)^2 + \left(t^2\right)^2\right]^2}} , \tag{9.853}$$

at P yields

$$\widehat{\mathbf{n}}\left(P\right) = \frac{3\sqrt{10}\,\widehat{\mathbf{e}}_1 + \sqrt{10}\,\widehat{\mathbf{e}}_3}{10} .$$

The covariant curvature tensor in matrix form renders

$$\left[\, \underline{b}_{\alpha\beta} \,\right] = \frac{6}{\sqrt{\widetilde{\Omega}}}\begin{bmatrix} t^1 & -t^2 \\ -t^2 & -t^1 \end{bmatrix} \quad , \quad \left[\, \underline{b}_{\alpha\beta}\left(P\right) \,\right] = -\frac{3\sqrt{10}}{5}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} . \tag{9.854}$$

At the given point, the mixed curvature tensor $\underline{b}_{\alpha}^{\cdot\beta} = \underline{b}_{\alpha\beta} a^{\alpha\beta}$ with the following matrix form

$$\left[\ \underline{b}_{\alpha}^{\cdot\beta}(P)\right] = -\frac{3\sqrt{10}}{5}\begin{bmatrix} 0 & 1 \\ \frac{1}{10} & 0 \end{bmatrix} \quad \text{has two distinct eigenvalues; namely,} \quad \left.\begin{array}{l} \kappa_1 = +\dfrac{3}{5} \\ \kappa_2 = -\dfrac{3}{5} \end{array}\right\} .$$

And this means that P is a non-umbilical point. From (9.434)$_{1-2}$, it then follows that

$$\hat{\lambda}_1 = -\frac{0 - \frac{3}{5}10}{-\frac{3\sqrt{10}}{5} - 0} = -\sqrt{10} \quad , \quad \hat{\lambda}_2 = -\frac{0 + \frac{3}{5}10}{-\frac{3\sqrt{10}}{5} - 0} = +\sqrt{10}$$

Accordingly, $dt_1^2 = -\sqrt{10}dt_1^1$ and $dt_2^2 = \sqrt{10}dt_2^1$. Now, the vectors $d\mathbf{x}_1 = dt_1^{\alpha}\mathbf{a}_{\alpha}$ and $d\mathbf{x}_2 = dt_2^{\alpha}\mathbf{a}_{\alpha}$ render

$$d\mathbf{x}_1 = \left(\widehat{\mathbf{e}}_1 - \sqrt{10}\,\widehat{\mathbf{e}}_2 - 3\,\widehat{\mathbf{e}}_3\right) dt_1^1 \quad , \quad d\mathbf{x}_2 = \left(\widehat{\mathbf{e}}_1 + \sqrt{10}\,\widehat{\mathbf{e}}_2 - 3\,\widehat{\mathbf{e}}_3\right) dt_2^1 \ .$$

And the partial derivatives

$$\frac{\partial \widehat{\mathbf{n}}}{\partial t^1} = -\,\underline{b}_1^{\cdot\beta}\mathbf{a}_{\beta} = \frac{3\sqrt{10}}{5}\widehat{\mathbf{e}}_2 \quad , \quad \frac{\partial \widehat{\mathbf{n}}}{\partial t^2} = -\,\underline{b}_2^{\cdot\beta}\mathbf{a}_{\beta} = \frac{3\sqrt{10}}{50}\widehat{\mathbf{e}}_1 - \frac{9\sqrt{10}}{50}\widehat{\mathbf{e}}_3 \ ,$$

help compute

$$d\,\widehat{\mathbf{n}}_1 = \frac{-3\,\widehat{\mathbf{e}}_1 + 3\sqrt{10}\,\widehat{\mathbf{e}}_2 + 9\,\widehat{\mathbf{e}}_3}{5} dt_1^1 \quad , \quad d\,\widehat{\mathbf{n}}_2 = \frac{3\,\widehat{\mathbf{e}}_1 + 3\sqrt{10}\,\widehat{\mathbf{e}}_2 - 9\,\widehat{\mathbf{e}}_3}{5} dt_2^1 \ .$$

Observe that $d\,\widehat{\mathbf{n}}_1 + \kappa_1 d\mathbf{x}_1 = \mathbf{0}$ and $d\,\widehat{\mathbf{n}}_2 + \kappa_2 d\mathbf{x}_2 = \mathbf{0}$. The Rodrigues formula is thus verified in each principal direction.

Exercise 9.18

Let $\mathcal{S}$ be a regular surface parametrized by $\mathbf{x} = \mathbf{x}\left(t^1, t^2\right)$ and consider a surface $\mathcal{S}^{\parallel}$ defined by

$$\mathbf{x}^{\parallel}\left(t^1, t^2\right) = \mathbf{x}\left(t^1, t^2\right) + h\,\widehat{\mathbf{n}}\left(t^1, t^2\right) \ . \tag{9.855}$$

This surface represents a *parallel surface* to $\mathcal{S}$ at (constant) distance h (Fig. 9.42). First, show that the mean curvature of $\mathcal{S}^{\parallel}$ can be written as

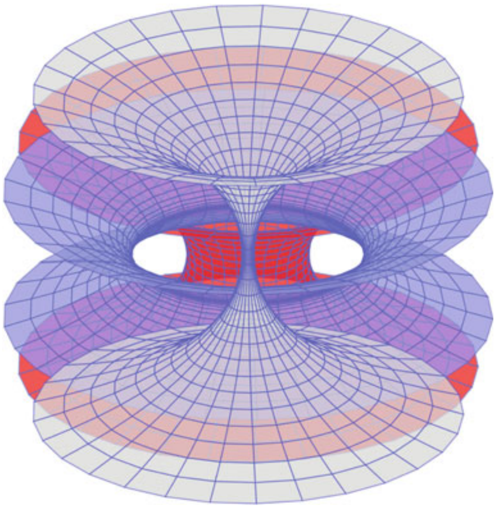


Fig. 9.42 Catenoid and its parallel surfaces

$$\overline{\mathrm{H}}^{\parallel} = \frac{\varpi\left(\overline{\mathrm{H}} - h\overline{\mathrm{K}}\right)}{1 - 2h\overline{\mathrm{H}} + h^2\overline{\mathrm{K}}} \quad \text{where} \quad \underbrace{\varpi = \operatorname{sgn}\left[(1 - h\kappa_1)(1 - h\kappa_2)\right]}_{\text{recall that } \kappa_1, \kappa_2 \text{ were the eigenvalues of } \underline{b}_{\alpha}^{\cdot\beta}} , \tag{9.856}$$

and subsequently obtain the following formula

$$\overline{\mathrm{K}}^{\parallel} = \frac{\overline{\mathrm{K}}}{1 - 2h\overline{\mathrm{H}} + h^2\overline{\mathrm{K}}} , \tag{9.857}$$

for its Gaussian curvature. Then, verify that the principal curvatures of such a surface are expressible as

$$\kappa_1^{\parallel} = \frac{\varpi\,\kappa_1}{1 - h\kappa_1} \quad , \quad \kappa_2^{\parallel} = \frac{\varpi\,\kappa_2}{1 - h\kappa_2} . \tag{9.858}$$

Finally, prove that

$$\left.\frac{d\mathrm{I}_{\mathbf{r}}^{\parallel}}{dh}\right|_{h=0} = -2\mathrm{II}_{\mathbf{r}} , \tag{9.859}$$

where $\mathrm{I}_{\mathbf{r}}^{\parallel}$ ($\mathrm{II}_{\mathbf{r}}$) denotes the first (second) fundamental form of $\mathcal{S}^{\parallel}$ ($\mathcal{S}$).

Solution. The covariant basis vectors of $\mathcal{S}^{\parallel}$ are given by

$$\mathbf{a}_{\alpha}^{\parallel} = \mathbf{a}_{\alpha} - h\,\underline{b}_{\alpha}^{\cdot\gamma}\mathbf{a}_{\gamma} , \tag{9.860}$$

or

$$\mathbf{a}_1^{\parallel} = \left(1 - h\,\underline{b}_{1}^{\cdot 1}\right)\mathbf{a}_1 - h\,\underline{b}_{1}^{\cdot 2}\mathbf{a}_2 \quad , \quad \mathbf{a}_2^{\parallel} = -h\,\underline{b}_{2}^{\cdot 1}\mathbf{a}_1 + \left(1 - h\,\underline{b}_{2}^{\cdot 2}\right)\mathbf{a}_2 . \tag{9.861}$$

In matrix form, they render

$$\begin{bmatrix} \vdots & \vdots \\ \mathbf{a}_1^{\parallel} & \mathbf{a}_2^{\parallel} \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots \\ \mathbf{a}_1 & \mathbf{a}_2 \\ \vdots & \vdots \end{bmatrix} \left[\, M \,\right] \quad \text{where} \quad [\, M\,] = \left[\, \delta_\alpha^\beta - h\, \underline{b}_\alpha^{\cdot\beta} \,\right]^{\mathrm{T}} \ . \tag{9.862}$$

Using (9.103)$_3$, (9.104)$_2$, (9.425a) and (9.425b), the determinant of $[\, M\,]$ takes the form

$$\det [\, M\,] = 1 - 2h\overline{\mathrm{H}} + h^2\overline{\mathrm{K}} = (1 - h\kappa_1)\,(1 - h\kappa_2) \ . \tag{9.863}$$

This helps establish

$$\mathbf{a}_1^{\parallel} \times \mathbf{a}_2^{\parallel} = (\det [\, M\,])\, \mathbf{a}_1 \times \mathbf{a}_2 \quad \text{and, consequently,} \quad \widehat{\mathbf{n}}^{\parallel} = \varpi\, \widehat{\mathbf{n}} \ , \tag{9.864}$$

where $\widehat{\mathbf{n}}^{\parallel}$ denotes the unit normal to the parallel surface and ϖ is the sign of $\det [\, M\,]$. As a result, $\partial\widehat{\mathbf{n}}^{\parallel}/\partial t^\alpha = \varpi\, \partial\widehat{\mathbf{n}}/\partial t^\alpha$. The Weingarten equations (9.99)$_1$ for $\mathcal{S}^{\parallel}$ can now be written as

$$\frac{\partial\widehat{\mathbf{n}}^{\parallel}}{\partial t^\alpha} = \left(-\, \underline{b}_\alpha^{\cdot\beta\,\parallel}\right) \mathbf{a}_\beta^{\parallel} \quad \text{or} \quad \left(\varpi\, \underline{b}_\alpha^{\cdot\beta}\right) \mathbf{a}_\beta = \left(\underline{b}_\alpha^{\cdot\beta\,\parallel}\right) \mathbf{a}_\beta^{\parallel} \ , \tag{9.865}$$

or

$$\varpi \begin{bmatrix} \vdots & \vdots \\ \mathbf{a}_1 & \mathbf{a}_2 \\ \vdots & \vdots \end{bmatrix} \left[\, \underline{b}_\alpha^{\cdot\beta} \,\right]^{\mathrm{T}} = \begin{bmatrix} \vdots & \vdots \\ \mathbf{a}_1^{\parallel} & \mathbf{a}_2^{\parallel} \\ \vdots & \vdots \end{bmatrix} \left[\, \underline{b}_\alpha^{\cdot\beta\,\parallel} \,\right]^{\mathrm{T}} \ . \tag{9.866}$$

It is then easy to see that

$$\left[\, \underline{b}_\alpha^{\cdot\beta\,\parallel} \,\right] = \varpi \left[\, \underline{b}_\alpha^{\cdot\beta} \,\right] \left[\, \delta_\alpha^\beta - h\, \underline{b}_\alpha^{\cdot\beta} \,\right]^{-1} \ . \tag{9.867}$$

The (half of) trace of this matrix and its determinant are the required results (9.856) and (9.857), respectively. ☆

The results (9.858)$_1$ and (9.858)$_2$ can be obtained from (9.426a) and (9.426b) in a straightforward manner. Considering $\left(\kappa_\gamma, \widehat{\mathbf{e}}_\gamma\right)$ as an eigenpair of $\left[\, \underline{b}_\alpha^{\cdot\beta} \,\right]$ and using the matrix expression (9.867), one can also arrive at the same results, i.e.

$$\left[\, \underline{b}_\alpha^{\cdot\beta\,\parallel} \,\right] \left[\, \widehat{\mathbf{e}}_\gamma \,\right] = \frac{\varpi\, \kappa_\gamma}{1 - h\kappa_\gamma} \left[\, \widehat{\mathbf{e}}_\gamma \,\right] \quad (\gamma = 1, 2;\ \text{no sum}) \ . \quad ☆☆ \tag{9.868}$$

Finally,

$$\left.\frac{d\mathrm{I}_{\mathbf{r}}^{\parallel}}{dh}\right|_{h=0} \overset{\text{by using (9.227) and}}{\underset{\text{applying the product rule}}{=\!=\!=}} \left[\frac{d\mathbf{a}_{\alpha}^{\parallel}}{dh}\cdot\mathbf{a}_{\beta}^{\parallel}+\mathbf{a}_{\alpha}^{\parallel}\cdot\frac{d\mathbf{a}_{\beta}^{\parallel}}{dh}\right]_{h=0} dt^{\alpha}dt^{\beta}$$

$$\overset{\text{by using}}{\underset{(9.860)}{=\!=\!=}} \left[-\,\underline{b}_{\alpha}^{\cdot\gamma}\mathbf{a}_{\gamma}\cdot\mathbf{a}_{\beta}-\mathbf{a}_{\alpha}\cdot\,\underline{b}_{\beta}^{\cdot\gamma}\mathbf{a}_{\gamma}\right]dt^{\alpha}dt^{\beta}$$

$$\overset{\text{by using}}{\underset{(9.17)}{=\!=\!=}} -\left[\,\underline{b}_{\alpha}^{\cdot\gamma}a_{\gamma\beta}+\,\underline{b}_{\beta}^{\cdot\gamma}a_{\gamma\alpha}\right]dt^{\alpha}dt^{\beta}$$

$$\overset{\text{by using}}{\underset{\text{(9.26) and (9.100)}}{=\!=\!=}} -\left[\,\underline{b}_{\alpha\beta}+\,\underline{b}_{\beta\alpha}\right]dt^{\alpha}dt^{\beta}$$

$$\overset{\text{by using}}{\underset{(9.95)}{=\!=\!=}} -2\,\underline{b}_{\alpha\beta}dt^{\alpha}dt^{\beta}$$

$$\overset{\text{by using}}{\underset{(9.251)}{=\!=\!=}} -2\mathrm{II}_{\mathbf{r}}\ . \qquad ☆☆☆$$

Exercise 9.19

Let $\boldsymbol{\beta}(s)$ be a naturally represented regular curve with the Frenet trihedron

$$\widehat{\mathbf{a}}_1^{\mathrm{c}}=\boldsymbol{\beta}' \quad , \quad \widehat{\mathbf{a}}_2^{\mathrm{c}}=\frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'' \quad , \quad \widehat{\mathbf{a}}_3^{\mathrm{c}}=\frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'\times\boldsymbol{\beta}'' \ , \qquad \leftarrow \text{see (9.288)}$$

and the Frenet formulas

$$\left(\widehat{\mathbf{a}}_1^{\mathrm{c}}\right)'=\boldsymbol{\beta}'' \quad , \quad \left(\widehat{\mathbf{a}}_2^{\mathrm{c}}\right)'=\frac{\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'\times\boldsymbol{\beta}''-\kappa^{\mathrm{c}}\boldsymbol{\beta}' \quad , \quad \left(\widehat{\mathbf{a}}_3^{\mathrm{c}}\right)'=-\frac{\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'' \ . \qquad \leftarrow \text{see (9.305a)}$$

Further, let $\mathcal{S}$ be the **principal normal surface** of $\boldsymbol{\beta}(s)$ parametrically described by $\hat{\mathbf{x}}^{\mathrm{s}}\left(s,t^2\right)=\boldsymbol{\beta}(s)+t^2\,\widehat{\mathbf{a}}_2^{\mathrm{c}}(s)$. Then, show that the mean curvature of $\mathcal{S}$ can be computed according to (9.334)-(9.335) and subsequently obtain the equations determining its principal curvatures and directions.

At the end, show that $\boldsymbol{\beta}(s)$ is a **geodesic** on its **binormal surface** parametrically described by $\hat{\mathbf{x}}^{\mathrm{s}}\left(s,t^2\right)=\boldsymbol{\beta}(s)+t^2\,\widehat{\mathbf{a}}_3^{\mathrm{c}}(s)$.

Solution. To this end, the argument s of the functions is dropped for simplifying the notation. To begin with, consider the fact that $\left(\boldsymbol{\beta}'\times\boldsymbol{\beta}''\right)\cdot\boldsymbol{\beta}'=\left(\boldsymbol{\beta}'\times\boldsymbol{\beta}''\right)\cdot\boldsymbol{\beta}''=0$ and $\boldsymbol{\beta}'\cdot\boldsymbol{\beta}''=0$. The distribution parameter (9.330) then takes the form

$$p=\frac{\left(\frac{\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'\times\boldsymbol{\beta}''-\kappa^{\mathrm{c}}\boldsymbol{\beta}'\right)\cdot\left[\boldsymbol{\beta}'\times\frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}''\right]}{\frac{\tau^{\mathrm{c}\,2}}{\kappa^{\mathrm{c}\,2}}\left|\boldsymbol{\beta}'\times\boldsymbol{\beta}''\right|^2+\kappa^{\mathrm{c}\,2}\left|\boldsymbol{\beta}'\right|^2}=\frac{\frac{\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}\,2}}\left|\boldsymbol{\beta}'\times\boldsymbol{\beta}''\right|^2}{\frac{\tau^{\mathrm{c}\,2}}{\kappa^{\mathrm{c}\,2}}\left|\boldsymbol{\beta}''\right|^2+\kappa^{\mathrm{c}\,2}}$$

$$=\frac{\tau^{\mathrm{c}}}{\tau^{\mathrm{c}\,2}+\kappa^{\mathrm{c}\,2}}\ .$$

The desired result (9.334) follows from (9.332). Thus, to compute

$$\overline{\mathrm{H}} = \frac{\left[\boldsymbol{\beta}'' + t^2\left(\tau^{\mathrm{c}}\widehat{\mathbf{a}}_3^{\mathrm{c}} - \kappa^{\mathrm{c}}\widehat{\mathbf{a}}_1^{\mathrm{c}}\right)'\right] \cdot \left\{\left[\boldsymbol{\beta}' + t^2\left(\frac{\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}' \times \boldsymbol{\beta}'' - \kappa^{\mathrm{c}}\boldsymbol{\beta}'\right)\right] \times \frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}''\right\}}{2\left[p^2 + \left(t^2\right)^2\right]^{3/2}\left(\tau^{\mathrm{c}\,2} + \kappa^{\mathrm{c}\,2}\right)^{3/2}} ,$$

one needs to have

$$\begin{aligned}\mathbf{u} &:= \boldsymbol{\beta}'' + t^2\left(\tau^{\mathrm{c}}\widehat{\mathbf{a}}_3^{\mathrm{c}} - \kappa^{\mathrm{c}}\widehat{\mathbf{a}}_1^{\mathrm{c}}\right)' \\ &= \left(1 - \frac{t^2\tau^{\mathrm{c}\,2}}{\kappa^{\mathrm{c}}} - t^2\kappa^{\mathrm{c}}\right)\boldsymbol{\beta}'' + \frac{t^2\tau^{\mathrm{c}\,\prime}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}' \times \boldsymbol{\beta}'' - t^2\kappa^{\mathrm{c}\,\prime}\boldsymbol{\beta}' ,\end{aligned}$$

and

$$\begin{aligned}\mathbf{v} &:= \left[\boldsymbol{\beta}' + t^2\left(\frac{\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}' \times \boldsymbol{\beta}'' - \kappa^{\mathrm{c}}\boldsymbol{\beta}'\right)\right] \times \frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'' \\ &= \left(\frac{1}{\kappa^{\mathrm{c}}} - t^2\right)\boldsymbol{\beta}' \times \boldsymbol{\beta}'' - \frac{t^2\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}\,2}}\left(\boldsymbol{\beta}'' \cdot \boldsymbol{\beta}''\right)\boldsymbol{\beta}' \\ &= \left(\frac{1}{\kappa^{\mathrm{c}}} - t^2\right)\boldsymbol{\beta}' \times \boldsymbol{\beta}'' - t^2\tau^{\mathrm{c}}\boldsymbol{\beta}' .\end{aligned}$$

Consequently,

$$\mathbf{u} \cdot \mathbf{v} = t^2\tau^{\mathrm{c}\,\prime}\left(1 - t^2\kappa^{\mathrm{c}}\right) + \left(t^2\right)^2\kappa^{\mathrm{c}\,\prime}\tau^{\mathrm{c}} .$$

Substituting this result into the above expression for the mean curvature leads to the required result.

Next, for the given surface $\hat{\mathbf{x}}^{\mathrm{s}}\left(s, t^2\right) = \boldsymbol{\beta}(s) + t^2\widehat{\mathbf{a}}_2^{\mathrm{c}}(s)$, consider

$$\left.\begin{aligned}\mathbf{a}_1 &= \left(1 - t^2\kappa^{\mathrm{c}}\right)\boldsymbol{\beta}' + \frac{t^2\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}' \times \boldsymbol{\beta}'' \\ \mathbf{a}_2 &= \frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}''\end{aligned}\right\} \quad \text{with} \quad \left.\begin{aligned}\mathrm{E}_{\mathbf{r}} &= \left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2 \\ \mathrm{G}_{\mathbf{r}} &= 1 \\ \mathrm{F}_{\mathbf{r}} &= 0\end{aligned}\right\} ,$$

and

$$\widehat{\mathbf{n}} = \frac{\left(\frac{1}{\kappa^{\mathrm{c}}} - t^2\right)\boldsymbol{\beta}' \times \boldsymbol{\beta}'' - t^2\tau^{\mathrm{c}}\boldsymbol{\beta}'}{\sqrt{\left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2}} .$$

It is then easy to see that

$$\mathrm{e}_{\mathbf{r}} = \frac{t^2\left[\tau^{\mathrm{c}\,\prime} + t^2\left(\kappa^{\mathrm{c}\,\prime}\tau^{\mathrm{c}} - \tau^{\mathrm{c}\,\prime}\kappa^{\mathrm{c}}\right)\right]}{\sqrt{\left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2}}\;,$$

$$\mathrm{g}_{\mathbf{r}} = 0\;,$$

$$\mathrm{f}_{\mathbf{r}} = \frac{\tau^{\mathrm{c}}}{\sqrt{\left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2}}\;.$$

Guided by (9.437), the principal directions are obtained via the following relation

$$\begin{aligned}
&+\tau^{\mathrm{c}}\left[\left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2\right] ds ds \\
&-t^2\left[\tau^{\mathrm{c}\,\prime} + t^2\left(\kappa^{\mathrm{c}\,\prime}\tau^{\mathrm{c}} - \tau^{\mathrm{c}\,\prime}\kappa^{\mathrm{c}}\right)\right] ds dt^2 \\
&-\tau^{\mathrm{c}} dt^2 dt^2 = 0\;.
\end{aligned}$$

The equation $\left(\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}_{\mathbf{r}}^2\right)\kappa^{\mathrm{n}\,2} - \left(\mathrm{g}_{\mathbf{r}}\mathrm{E}_{\mathbf{r}} - 2\mathrm{f}_{\mathbf{r}}\mathrm{F}_{\mathbf{r}} + \mathrm{e}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}}\right)\kappa^{\mathrm{n}} + \left(\mathrm{e}_{\mathbf{r}}\mathrm{g}_{\mathbf{r}} - \mathrm{f}_{\mathbf{r}}^2\right) = 0$, according to (9.429), for the problem at hand renders

$$\left[\left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2\right]^2 \kappa^{\mathrm{n}\,2} - \frac{t^2\left[\tau^{\mathrm{c}\,\prime} + t^2\left(\kappa^{\mathrm{c}\,\prime}\tau^{\mathrm{c}} - \tau^{\mathrm{c}\,\prime}\kappa^{\mathrm{c}}\right)\right]}{\left[\left(1 - t^2\kappa^{\mathrm{c}}\right)^2 + \left(t^2\tau^{\mathrm{c}}\right)^2\right]^{-1/2}}\kappa^{\mathrm{n}} - \tau^{\mathrm{c}\,2} = 0\;,$$

whose solution delivers the principal curvatures.

Finally, consider the binormal surface $\mathbf{x}\left(s, t^2\right) = \boldsymbol{\beta}\left(s\right) + \left(t^2/\kappa^{\mathrm{c}}\right)\boldsymbol{\beta}'\left(s\right) \times \boldsymbol{\beta}''\left(s\right)$ of $\boldsymbol{\beta}\left(s\right)$ with

$$\left.\begin{aligned}
\mathbf{a}_1 &= \boldsymbol{\beta}' - \frac{t^2\tau^{\mathrm{c}}}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'' \\
\mathbf{a}_2 &= \frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}' \times \boldsymbol{\beta}''
\end{aligned}\right\} \quad \text{and, consequently,} \quad \widehat{\mathbf{n}} = -\frac{\frac{1}{\kappa^{\mathrm{c}}}\boldsymbol{\beta}'' + t^2\tau^{\mathrm{c}}\boldsymbol{\beta}'}{\sqrt{1 + \left(t^2\tau^{\mathrm{c}}\right)^2}}\;.$$

The geodesic curvature (9.359) then becomes

$$\begin{aligned}
\kappa^{\mathrm{g}} &= \left(\boldsymbol{\beta}' \times \boldsymbol{\beta}''\right) \cdot \widehat{\mathbf{n}} \\
&= -\frac{1}{\kappa^{\mathrm{c}}}\frac{\left(\boldsymbol{\beta}' \times \boldsymbol{\beta}''\right) \cdot \boldsymbol{\beta}''}{\sqrt{1 + \left(t^2\tau^{\mathrm{c}}\right)^2}} - t^2\tau^{\mathrm{c}}\frac{\left(\boldsymbol{\beta}' \times \boldsymbol{\beta}''\right) \cdot \boldsymbol{\beta}'}{\sqrt{1 + \left(t^2\tau^{\mathrm{c}}\right)^2}} \\
&= 0\;.
\end{aligned}$$

This result apparently states that the base curve of a binormal surface is always a geodesic on that surface.

Exercise 9.20

First, verify (9.473).

Consider the Darboux trihedron $\widehat{\mathbf{a}}_1^{\mathrm{c}}$, $\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{n}}$ which was used to describe the geometry of a curve with respect to a surface embedding that curve (Fig. 9.25). Then, derive the derivative of such a triplet of vectors with respect to the arc length parameter.

Finally, find a useful Darboux vector along with the corresponding skew tensor and subsequently express the established system of ordinary differential equations in terms of these tensorial variables.

Solution. To begin with, consider a s-parametrized curve $\mathcal{C}$ embedded in a regular surface $\mathcal{S}$ according to (9.336), i.e. $\mathbf{x} = \hat{\mathbf{x}}^{\mathrm{s}}\left(t^1(s), t^2(s)\right) = \hat{\mathbf{x}}^{\mathrm{c}}(s)$. Let $\underline{b}_{\alpha\beta}$ be the symmetric covariant curvature tensor of $\mathcal{S}$ and further let $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ be the unit tangent vector to $\mathcal{C}$. Referred to the surface covariant basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and the principal directions $\widehat{\mathbf{e}}_1$, $\widehat{\mathbf{e}}_2$, this vector can then be expressed as

$$\widehat{\mathbf{a}}_1^{\mathrm{c}} = \frac{dt^\alpha}{ds}\mathbf{a}_\alpha \quad , \quad \widehat{\mathbf{a}}_1^{\mathrm{c}} = \cos\theta\,\widehat{\mathbf{e}}_1 + \sin\theta\,\widehat{\mathbf{e}}_2 \ ,$$

where θ denotes the inclination of $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ to $\widehat{\mathbf{e}}_1$ (Fig. 9.28). The derivative of the unit normal field to the surface with respect to the arc length parameter now renders

$$\frac{d\widehat{\mathbf{n}}}{ds} = -\kappa_1\cos\theta\,\widehat{\mathbf{e}}_1 - \kappa_2\sin\theta\,\widehat{\mathbf{e}}_2 \ , \tag{9.869}$$

because

$$\begin{aligned}
\frac{d\widehat{\mathbf{n}}}{ds} &\xlongequal[\text{of differentiation}]{\text{by using the chain rule}} \frac{dt^\alpha}{ds}\frac{\partial\widehat{\mathbf{n}}}{\partial t^\alpha} \\
&\xlongequal[(9.99)]{\text{by using}} -\frac{dt^\alpha}{ds}\,\underline{b}_{\alpha\beta}\mathbf{a}^\beta \\
&\xlongequal[(9.457)]{\text{by using}} -\cos\theta\,\hat{\eta}^\alpha_{\ 1}\,\underline{b}_{\alpha\beta}\mathbf{a}^\beta - \sin\theta\,\hat{\eta}^\alpha_{\ 2}\,\underline{b}_{\alpha\beta}\mathbf{a}^\beta \\
&\xlongequal[(9.449)]{\text{by using}} -\cos\theta\,\kappa_1\hat{\eta}^\alpha_{\ 1}a_{\alpha\beta}\mathbf{a}^\beta - \sin\theta\,\kappa_2\hat{\eta}^\alpha_{\ 2}a_{\alpha\beta}\mathbf{a}^\beta \\
&\xlongequal[(9.32)]{\text{by using}} -\cos\theta\,\kappa_1\hat{\eta}^\alpha_{\ 1}\mathbf{a}_\alpha - \sin\theta\,\kappa_2\hat{\eta}^\alpha_{\ 2}\mathbf{a}_\alpha \\
&\xlongequal[(9.447)]{\text{by using}} -\kappa_1\cos\theta\,\widehat{\mathbf{e}}_1 - \kappa_2\sin\theta\,\widehat{\mathbf{e}}_2 \ .
\end{aligned}$$

Considering the fact that $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{n}}\}$ is a right-handed orthonormal basis, the geodesic torsion $(9.468)_3$ takes the form

$$\begin{aligned}\tau^{\mathrm{g}} &= \widehat{\mathbf{n}} \cdot [(-\kappa_1 \cos\theta\, \widehat{\mathbf{e}}_1 - \kappa_2 \sin\theta\, \widehat{\mathbf{e}}_2) \times (\cos\theta\, \widehat{\mathbf{e}}_1 + \sin\theta\, \widehat{\mathbf{e}}_2)] \\ &= \widehat{\mathbf{n}} \cdot [-\kappa_1 \cos\theta \sin\theta\, \widehat{\mathbf{n}} + \kappa_2 \sin\theta \cos\theta\, \widehat{\mathbf{n}}] \\ &= (\kappa_2 - \kappa_1) \sin\theta \cos\theta = \frac{\kappa_2 - \kappa_1}{2} \sin 2\theta \ . \quad \blacktriangleleft\end{aligned}$$

Next, consider the trihedron $\{\widehat{\mathbf{a}}_1^{\mathrm{c}}, \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}, \widehat{\mathbf{n}}\}$ which also presents a right-handed orthonormal basis. One can write

$$\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} = -\sin\theta\, \widehat{\mathbf{e}}_1 + \cos\theta\, \widehat{\mathbf{e}}_2 \ , \tag{9.870}$$

which helps represent

$$\tau^{\mathrm{g}} = -\frac{d\, \widehat{\mathbf{n}}}{ds} \cdot \left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right) \ . \tag{9.871}$$

Having in mind the Euler formula (9.455), i.e. $\kappa^{\mathrm{n}} = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta$, one can also have

$$\kappa^{\mathrm{n}} = -\frac{d\, \widehat{\mathbf{n}}}{ds} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} \ . \tag{9.872}$$

It is then easy to see that $d\, \widehat{\mathbf{n}}/ds$ is a linear combination of $\widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}$, i.e.

$$\frac{d\, \widehat{\mathbf{n}}}{ds} = -\kappa^{\mathrm{n}} \widehat{\mathbf{a}}_1^{\mathrm{c}} - \tau^{\mathrm{g}} \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} \ .$$

Attention is now being focused on expressing $d\, \widehat{\mathbf{a}}_1^{\mathrm{c}}/ds$ in terms of $\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{n}}$. This is given by

$$\begin{aligned}\frac{d\, \widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} &\overset{\text{in light of}}{\underset{(1.26)}{=\!=\!=}} \underbrace{\left[\frac{d\, \widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \cdot \left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right)\right] \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} + \left[\frac{d\, \widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \cdot \widehat{\mathbf{n}}\right] \widehat{\mathbf{n}}}_{\text{note that } \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} = 0 \text{ since } \widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}} = 1} \\ &\overset{\text{by using}}{\underset{(9.358)}{=\!=\!=}} \kappa^{\mathrm{g}} \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} + \left[-\widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \frac{d\, \widehat{\mathbf{n}}}{ds}\right] \widehat{\mathbf{n}} \quad \longleftarrow \quad \begin{array}{l}\text{note that } \widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \widehat{\mathbf{n}} = 0 \text{ gives } \frac{d}{ds}\left[\widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \widehat{\mathbf{n}}\right] = 0 \\ \text{and, therefore, } \frac{d\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} \cdot \widehat{\mathbf{n}} = -\widehat{\mathbf{a}}_1^{\mathrm{c}} \cdot \frac{d\widehat{\mathbf{n}}}{ds}\end{array} \\ &\overset{\text{by using}}{\underset{(9.872)}{=\!=\!=}} \kappa^{\mathrm{g}} \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} + \kappa^{\mathrm{n}} \widehat{\mathbf{n}} \ .\end{aligned}$$

In a similar fashion,

$$\begin{aligned}\frac{d\left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right)}{ds} &= \left[\frac{d\left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right)}{ds} \cdot \widehat{\mathbf{a}}_1^{\mathrm{c}}\right] \widehat{\mathbf{a}}_1^{\mathrm{c}} + \left[\frac{d\left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right)}{ds} \cdot \widehat{\mathbf{n}}\right] \widehat{\mathbf{n}} \\ &= \left[-\left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right) \cdot \frac{d\, \widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds}\right] \widehat{\mathbf{a}}_1^{\mathrm{c}} + \left[-\left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right) \cdot \frac{d\, \widehat{\mathbf{n}}}{ds}\right] \widehat{\mathbf{n}} \\ &= -\kappa^{\mathrm{g}} \widehat{\mathbf{a}}_1^{\mathrm{c}} + \tau^{\mathrm{g}} \widehat{\mathbf{n}} \ .\end{aligned}$$

The above results are listed in the following:

$$\frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} = \kappa^{\mathrm{g}}\,\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} + \kappa^{\mathrm{n}}\,\widehat{\mathbf{n}}\,, \qquad \longleftarrow \text{see (9.305a)} \tag{9.873a}$$

$$\frac{d\left(\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\right)}{ds} = -\kappa^{\mathrm{g}}\,\widehat{\mathbf{a}}_1^{\mathrm{c}} + \tau^{\mathrm{g}}\,\widehat{\mathbf{n}}\,, \tag{9.873b}$$

$$\frac{d\,\widehat{\mathbf{n}}}{ds} = -\kappa^{\mathrm{n}}\,\widehat{\mathbf{a}}_1^{\mathrm{c}} - \tau^{\mathrm{g}}\,\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}\,. \tag{9.873c}$$

These equations are eventually the analogues of the Frenet formulas for the moving trihedron $\widehat{\mathbf{a}}_1^{\mathrm{c}}$, $\widehat{\mathbf{a}}_2^{\mathrm{c}}$ and $\widehat{\mathbf{a}}_3^{\mathrm{c}}$. In matrix notation, they render

$$\begin{bmatrix} \vdots & \vdots & \vdots \\ \frac{d\,\widehat{\mathbf{a}}_1^{\mathrm{c}}}{ds} & \frac{d\left(\widehat{\mathbf{n}}\times\widehat{\mathbf{a}}_1^{\mathrm{c}}\right)}{ds} & \frac{d\,\widehat{\mathbf{n}}}{ds} \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \widehat{\mathbf{a}}_1^{\mathrm{c}} & \widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}} & \widehat{\mathbf{n}} \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & -\kappa^{\mathrm{g}} & -\kappa^{\mathrm{n}} \\ \kappa^{\mathrm{g}} & 0 & -\tau^{\mathrm{g}} \\ \kappa^{\mathrm{n}} & \tau^{\mathrm{g}} & 0 \end{bmatrix}. \blacktriangleleft\blacktriangleleft \tag{9.874}$$

Finally, let the Darboux trihedron $\widehat{\mathbf{a}}_1^{\mathrm{c}}$, $\widehat{\mathbf{n}} \times \widehat{\mathbf{a}}_1^{\mathrm{c}}$ and $\widehat{\mathbf{n}}$ be denoted by $\widehat{\mathbf{u}}_1$, $\widehat{\mathbf{u}}_2$ and $\widehat{\mathbf{u}}_3$, respectively. Then, by introducing the Darboux vector

$$\mathbf{d}^{\mathrm{c}} = \tau^{\mathrm{g}}\,\widehat{\mathbf{u}}_1 - \kappa^{\mathrm{n}}\,\widehat{\mathbf{u}}_2 + \kappa^{\mathrm{g}}\,\widehat{\mathbf{u}}_3\,, \qquad \longleftarrow \text{see (9.307)} \tag{9.875}$$

the formulas (9.873a)-(9.873c) can be represented in a more elegant way as follows:

$$\frac{d\,\widehat{\mathbf{u}}_i}{ds} = \mathbf{d}^{\mathrm{c}} \times \widehat{\mathbf{u}}_i\,, \quad i = 1, 2, 3\,. \tag{9.876}$$

They may also be written as

$$\frac{d\,\widehat{\mathbf{u}}_i}{ds} = \mathbf{W}_s^{\mathrm{c}}\,\widehat{\mathbf{u}}_i\,, \quad i = 1, 2, 3\,, \tag{9.877}$$

where the skew-symmetric tensor $\mathbf{W}_s^{\mathrm{c}}$ is given by

$$\mathbf{W}_s^{\mathrm{c}} = \tau^{\mathrm{g}}\left(\widehat{\mathbf{u}}_3 \otimes \widehat{\mathbf{u}}_2 - \widehat{\mathbf{u}}_2 \otimes \widehat{\mathbf{u}}_3\right) - \kappa^{\mathrm{n}}\left(\widehat{\mathbf{u}}_1 \otimes \widehat{\mathbf{u}}_3 - \widehat{\mathbf{u}}_3 \otimes \widehat{\mathbf{u}}_1\right) + \kappa^{\mathrm{g}}\left(\widehat{\mathbf{u}}_2 \otimes \widehat{\mathbf{u}}_1 - \widehat{\mathbf{u}}_1 \otimes \widehat{\mathbf{u}}_2\right)\,. \blacktriangleleft\blacktriangleleft\blacktriangleleft \tag{9.878}$$

Exercise 9.21

Verify (9.493) to (9.496).

Solution. Attention here is focused on verifying (9.493) and (9.494). The proof simply follows from $(9.475)_1$. By using the Gauss formulas (9.94) along with the product rule (9.132)$_2$, one can write

$$\begin{aligned}\frac{\partial^2 \mathbf{a}_1}{\partial t^2 \partial t^1} &= \frac{\partial}{\partial t^2}\left(\Gamma^1_{11}\mathbf{a}_1 + \Gamma^2_{11}\mathbf{a}_2 + \mathrm{e}_{\mathbf{r}}\widehat{\mathbf{n}}\right)\\
&= \frac{\partial \Gamma^1_{11}}{\partial t^2}\mathbf{a}_1 + \Gamma^1_{11}\left(\Gamma^1_{12}\mathbf{a}_1 + \Gamma^2_{12}\mathbf{a}_2 + \mathrm{f}_{\mathbf{r}}\widehat{\mathbf{n}}\right) + \frac{\partial \Gamma^2_{11}}{\partial t^2}\mathbf{a}_2\\
&\quad + \Gamma^2_{11}\left(\Gamma^1_{22}\mathbf{a}_1 + \Gamma^2_{22}\mathbf{a}_2 + \mathrm{g}_{\mathbf{r}}\widehat{\mathbf{n}}\right) + \frac{\partial \mathrm{e}_{\mathbf{r}}}{\partial t^2}\widehat{\mathbf{n}} + \mathrm{e}_{\mathbf{r}}\left(-\underline{b}_2{}^1\mathbf{a}_1 - \underline{b}_2{}^2\mathbf{a}_2\right)\\
&= \underbrace{\left[\frac{\partial \Gamma^1_{11}}{\partial t^2} + \Gamma^1_{11}\Gamma^1_{12} + \Gamma^1_{22}\Gamma^2_{11} - \frac{\mathrm{e}_{\mathbf{r}}\left(\mathrm{f}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{g}_{\mathbf{r}}\mathrm{F}_{\mathbf{r}}\right)}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}^2_{\mathbf{r}}}\right]}_{:=\overline{A}_1}\mathbf{a}_1\\
&\quad + \underbrace{\left[\Gamma^1_{11}\Gamma^2_{12} + \frac{\partial \Gamma^2_{11}}{\partial t^2} + \Gamma^2_{11}\Gamma^2_{22} - \frac{\mathrm{e}_{\mathbf{r}}\left(-\mathrm{f}_{\mathbf{r}}\mathrm{F}_{\mathbf{r}} + \mathrm{g}_{\mathbf{r}}\mathrm{E}_{\mathbf{r}}\right)}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}^2_{\mathbf{r}}}\right]}_{:=\overline{A}_2}\mathbf{a}_2\\
&\quad + \underbrace{\left[\Gamma^1_{11}\mathrm{f}_{\mathbf{r}} + \Gamma^2_{11}\mathrm{g}_{\mathbf{r}} + \frac{\partial \mathrm{e}_{\mathbf{r}}}{\partial t^2}\right]}_{:=\overline{A}_3}\widehat{\mathbf{n}}\,.\end{aligned}$$

In a similar manner,

$$\begin{aligned}\frac{\partial^2 \mathbf{a}_1}{\partial t^1 \partial t^2} &= \underbrace{\left[\frac{\partial \Gamma^1_{12}}{\partial t^1} + \Gamma^1_{11}\Gamma^1_{12} + \Gamma^1_{12}\Gamma^2_{12} - \frac{\mathrm{f}_{\mathbf{r}}\left(\mathrm{e}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{f}_{\mathbf{r}}\mathrm{F}_{\mathbf{r}}\right)}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}^2_{\mathbf{r}}}\right]}_{:=\overline{B}_1}\mathbf{a}_1\\
&\quad + \underbrace{\left[\Gamma^1_{12}\Gamma^2_{11} + \frac{\partial \Gamma^2_{12}}{\partial t^1} + \Gamma^2_{12}\Gamma^2_{12} - \frac{\mathrm{f}_{\mathbf{r}}\left(-\mathrm{e}_{\mathbf{r}}\mathrm{F}_{\mathbf{r}} + \mathrm{f}_{\mathbf{r}}\mathrm{E}_{\mathbf{r}}\right)}{\mathrm{E}_{\mathbf{r}}\mathrm{G}_{\mathbf{r}} - \mathrm{F}^2_{\mathbf{r}}}\right]}_{:=\overline{B}_2}\mathbf{a}_2\\
&\quad + \underbrace{\left[\Gamma^1_{12}\mathrm{e}_{\mathbf{r}} + \Gamma^2_{12}\mathrm{f}_{\mathbf{r}} + \frac{\partial \mathrm{f}_{\mathbf{r}}}{\partial t^1}\right]}_{:=\overline{B}_3}\widehat{\mathbf{n}}\,.\end{aligned}$$

Thus,

$$\frac{\partial^2 \mathbf{a}_1}{\partial t^1 \partial t^2} - \frac{\partial^2 \mathbf{a}_1}{\partial t^2 \partial t^1} = \mathbf{0} \implies \left(\overline{B}_1 - \overline{A}_1\right)\mathbf{a}_1 + \left(\overline{B}_2 - \overline{A}_2\right)\mathbf{a}_2 + \left(\overline{B}_3 - \overline{A}_3\right)\widehat{\mathbf{n}} = \mathbf{0}\,.$$

The fact that $\mathbf{a}_1, \mathbf{a}_2$ and $\widehat{\mathbf{n}}$ are three linearly independent vectors implies that $\overline{B}_1 = \overline{A}_1$, $\overline{B}_2 = \overline{A}_2$ and $\overline{B}_3 = \overline{A}_3$. Now, by using the identities

$$\frac{\mathrm{f_r}\left(\mathrm{e_r G_r} - \mathrm{f_r F_r}\right)}{\mathrm{E_r G_r} - \mathrm{F_r^2}} - \frac{\mathrm{e_r}\left(\mathrm{f_r G_r} - \mathrm{g_r F_r}\right)}{\mathrm{E_r G_r} - \mathrm{F_r^2}} = \frac{\mathrm{F_r}\left(\mathrm{e_r g_r} - \mathrm{f_r^2}\right)}{\mathrm{E_r G_r} - \mathrm{F_r^2}} = \mathrm{F_r}\overline{\mathrm{K}} \ ,$$

$$\frac{\mathrm{e_r}\left(-\mathrm{f_r F_r} + \mathrm{g_r E_r}\right)}{\mathrm{E_r G_r} - \mathrm{F_r^2}} - \frac{\mathrm{f_r}\left(-\mathrm{e_r F_r} + \mathrm{f_r E_r}\right)}{\mathrm{E_r G_r} - \mathrm{F_r^2}} = \frac{\mathrm{E_r}\left(\mathrm{e_r g_r} - \mathrm{f_r^2}\right)}{\mathrm{E_r G_r} - \mathrm{F_r^2}} = \mathrm{E_r}\overline{\mathrm{K}} \ ,$$

one can arrive at the desired results (9.493) and (9.494). At the end, it should not be difficult to derive the relations (9.495) and (9.496). This remains to be undertaken by the ambitious reader.

Exercise 9.22

Verify $(9.497\text{a})_1$ and $(9.497\text{b})_1$.

Solution. Consider a regular surface $\mathcal{S}$ embedded in the three-dimensional Euclidean space with the unit normal field $\widehat{\mathbf{n}} = (\mathbf{a}_1 \times \mathbf{a}_2) / |\mathbf{a}_1 \times \mathbf{a}_2|$ according to $(9.31)_1$. Let

$$\widehat{\mathbf{u}} := \frac{\mathbf{a}_1}{|\mathbf{a}_1|} = \frac{\mathbf{a}_1}{\sqrt{a_{11}}} \ . \tag{9.879}$$

This helps construct an orthonormal basis $\{\widehat{\mathbf{u}}, \widehat{\mathbf{v}}, \widehat{\mathbf{n}}\}$ in the sense that

$$\widehat{\mathbf{v}} = \widehat{\mathbf{n}} \times \widehat{\mathbf{u}} = \frac{\widehat{\mathbf{n}} \times \mathbf{a}_1}{\sqrt{a_{11}}} \underset{(9.58)}{\overset{\text{from}}{=\!=}} \frac{\sqrt{a}}{\sqrt{a_{11}}}\mathbf{a}^2 \quad , \quad \widehat{\mathbf{u}} = \widehat{\mathbf{v}} \times \widehat{\mathbf{n}} \quad , \quad \widehat{\mathbf{n}} = \widehat{\mathbf{u}} \times \widehat{\mathbf{v}} \ . \tag{9.880}$$

One then immediately obtains

$$\left.\begin{aligned} \frac{\partial \widehat{\mathbf{n}}}{\partial t^\alpha} \cdot \widehat{\mathbf{u}} &= -\widehat{\mathbf{n}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^\alpha} \\ \frac{\partial \widehat{\mathbf{n}}}{\partial t^\alpha} \cdot \widehat{\mathbf{v}} &= -\widehat{\mathbf{n}} \cdot \frac{\partial \widehat{\mathbf{v}}}{\partial t^\alpha} \\ \frac{\partial \widehat{\mathbf{u}}}{\partial t^\alpha} \cdot \widehat{\mathbf{u}} &= \frac{\partial \widehat{\mathbf{v}}}{\partial t^\alpha} \cdot \widehat{\mathbf{v}} = 0 \end{aligned}\right\} \quad , \quad \mathbf{I} = \widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}} + \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}} \ . \tag{9.881}$$

From (1.78a), (2.5), (2.13), (2.43), $(9.29)_3$, $(9.33)_1$, (9.94), (9.105) and (9.879) to (9.881), one can finally arrive at

$$\begin{aligned} \sqrt{a}\,\overline{\mathrm{K}} &= \left(\frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \times \frac{\partial \widehat{\mathbf{n}}}{\partial t^2}\right) \cdot (\widehat{\mathbf{u}} \times \widehat{\mathbf{v}}) \\ &= \underbrace{\left(\frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \cdot \widehat{\mathbf{u}}\right)\left(\frac{\partial \widehat{\mathbf{n}}}{\partial t^2} \cdot \widehat{\mathbf{v}}\right) - \left(\frac{\partial \widehat{\mathbf{n}}}{\partial t^1} \cdot \widehat{\mathbf{v}}\right)\left(\frac{\partial \widehat{\mathbf{n}}}{\partial t^2} \cdot \widehat{\mathbf{u}}\right)}_{= \left(\widehat{\mathbf{n}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^1}\right)\left(\widehat{\mathbf{n}} \cdot \frac{\partial \widehat{\mathbf{v}}}{\partial t^2}\right) - \left(\widehat{\mathbf{n}} \cdot \frac{\partial \widehat{\mathbf{v}}}{\partial t^1}\right)\left(\widehat{\mathbf{n}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^2}\right)} \end{aligned}$$

$$= \underbrace{\frac{\partial \widehat{\mathbf{u}}}{\partial t^1} \cdot (\widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}} + \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}) \frac{\partial \widehat{\mathbf{v}}}{\partial t^2}}_{= \frac{\partial \widehat{\mathbf{u}}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{v}}}{\partial t^2}} - \underbrace{\frac{\partial \widehat{\mathbf{v}}}{\partial t^1} \cdot (\widehat{\mathbf{u}} \otimes \widehat{\mathbf{u}} + \widehat{\mathbf{v}} \otimes \widehat{\mathbf{v}} + \widehat{\mathbf{n}} \otimes \widehat{\mathbf{n}}) \frac{\partial \widehat{\mathbf{u}}}{\partial t^2}}_{= \frac{\partial \widehat{\mathbf{v}}}{\partial t^1} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^2}}$$

$$= \underbrace{\frac{\partial}{\partial t^2} \left\{ \widehat{\mathbf{v}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^1} \right\}}_{= \frac{\partial}{\partial t^2} \left\{ \frac{\sqrt{a}}{\sqrt{a_{11}}} \mathbf{a}^2 \cdot \left[\frac{\partial}{\partial t^1} \left(\frac{1}{\sqrt{a_{11}}} \right) \mathbf{a}_1 + \frac{\Gamma^1_{11} \mathbf{a}_1 + \Gamma^2_{11} \mathbf{a}_2 + b_{11} \widehat{\mathbf{n}}}{\sqrt{a_{11}}} \right] \right\}} - \underbrace{\frac{\partial}{\partial t^1} \left\{ \widehat{\mathbf{v}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^2} \right\}}_{= \frac{\partial}{\partial t^1} \left\{ \frac{\sqrt{a}}{\sqrt{a_{11}}} \mathbf{a}^2 \cdot \left[\frac{\partial}{\partial t^2} \left(\frac{1}{\sqrt{a_{11}}} \right) \mathbf{a}_1 + \frac{\Gamma^1_{12} \mathbf{a}_1 + \Gamma^2_{12} \mathbf{a}_2 + b_{12} \widehat{\mathbf{n}}}{\sqrt{a_{11}}} \right] \right\}}$$

$$= \frac{\partial}{\partial t^2} \left\{ \frac{\sqrt{a}}{a_{11}} \Gamma^2_{11} \right\} - \frac{\partial}{\partial t^1} \left\{ \frac{\sqrt{a}}{a_{11}} \Gamma^2_{12} \right\} .$$

To verify (9.497b)$_1$, consider

$$\widehat{\mathbf{u}} := \frac{\mathbf{a}_2}{|\mathbf{a}_2|} = \frac{\mathbf{a}_2}{\sqrt{a_{22}}} \quad \text{and} \quad \left. \begin{aligned} \widehat{\mathbf{v}} &= \widehat{\mathbf{n}} \times \widehat{\mathbf{u}} = -\frac{\sqrt{a}}{\sqrt{a_{22}}} \mathbf{a}^1 \\ \widehat{\mathbf{u}} &= \widehat{\mathbf{v}} \times \widehat{\mathbf{n}} \\ \widehat{\mathbf{n}} &= \widehat{\mathbf{u}} \times \widehat{\mathbf{v}} \end{aligned} \right\} . \tag{9.882}$$

In this case, the identities in (9.881) also hold true. Thus,

$$\sqrt{a}\, \overline{\mathrm{K}} = \underbrace{\frac{\partial}{\partial t^2} \left\{ \widehat{\mathbf{v}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^1} \right\}}_{= \frac{\partial}{\partial t^2} \left\{ -\frac{\sqrt{a}}{\sqrt{a_{22}}} \mathbf{a}^1 \cdot \left[\frac{\partial}{\partial t^1} \left(\frac{1}{\sqrt{a_{22}}} \right) \mathbf{a}_2 + \frac{\Gamma^1_{21} \mathbf{a}_1 + \Gamma^2_{21} \mathbf{a}_2 + b_{21} \widehat{\mathbf{n}}}{\sqrt{a_{22}}} \right] \right\}} - \underbrace{\frac{\partial}{\partial t^1} \left\{ \widehat{\mathbf{v}} \cdot \frac{\partial \widehat{\mathbf{u}}}{\partial t^2} \right\}}_{= \frac{\partial}{\partial t^1} \left\{ -\frac{\sqrt{a}}{\sqrt{a_{22}}} \mathbf{a}^1 \cdot \left[\frac{\partial}{\partial t^2} \left(\frac{1}{\sqrt{a_{22}}} \right) \mathbf{a}_2 + \frac{\Gamma^1_{22} \mathbf{a}_1 + \Gamma^2_{22} \mathbf{a}_2 + b_{22} \widehat{\mathbf{n}}}{\sqrt{a_{22}}} \right] \right\}}$$

$$= \frac{\partial}{\partial t^1} \left\{ \frac{\sqrt{a}}{a_{22}} \Gamma^1_{22} \right\} - \frac{\partial}{\partial t^2} \left\{ \frac{\sqrt{a}}{a_{22}} \Gamma^1_{12} \right\} .$$

Exercise 9.23

Let $\mathcal{S}_f$ be a 2-ball of radius r in flat space. Further, let $\mathcal{S}_c$ be a spherical cap corresponding to $\mathcal{S}_f$ which sits at the top of a 3-ball $\mathcal{S}$ of radius R_0 as illustrated in Fig. 9.43. The area of $\mathcal{S}_f$ ($\mathcal{S}_c$) is denoted by A_f (A_c). Notice that when the radius (curvature) of $\mathcal{S}$ deceases (increases), the area of that upside-down bowl-shaped surface will increase.

Show that, up to the second-order, the deviation of A_c from its expected value A_f is governed by

$$\frac{A_c}{A_f} = 1 - \frac{R}{24}r^2 + \cdots \, , \tag{9.883}$$

where R presents the **Ricci scalar** of the sphere $\mathcal{S}$; given by,

$$R = \frac{2}{R_0^2} \, . \tag{9.884}$$

From (9.883), one can now conclude that:

> The Ricci scalar measures how much the area of a small ball deviates from its standard value in flat space.

Solution. To begin with, consider a sphere of radius R_0 with the Gaussian curvature $\overline{\mathrm{K}} = 1/R_0^2$, according to (9.746). Consequently, the Ricci scalar (9.518) becomes $R = 2/R_0^2$. The area of $\mathcal{S}_f$ is simply $A_f = \pi r^2$ while the area of $\mathcal{S}_c$ renders (see Fig. 9.43)

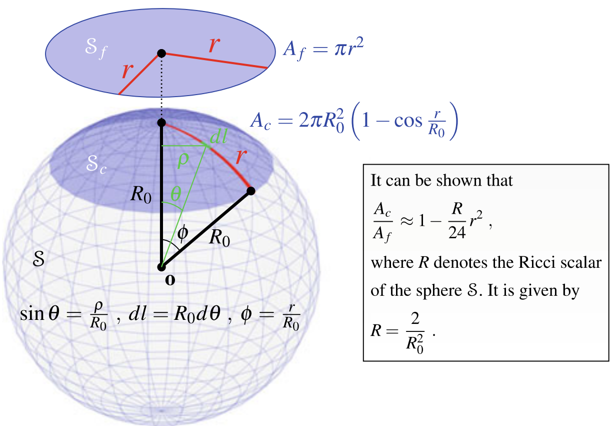


Fig. 9.43 Deviation of the area of a small ball in curved space from that of standard ball in flat space

$$A_c = \int_0^\phi 2\pi \rho \, dl = \int_0^\phi 2\pi R_0^2 \sin\theta \, d\theta = 2\pi R_0^2 (1 - \cos\phi) \ .$$

Using $\phi = r/R_0$, one then has

$$A_c = 2\pi R_0^2 \left(1 - \cos \frac{r}{R_0}\right) \ . \tag{9.885}$$

Consider the following Taylor series expansion

$$\begin{aligned} A_c &= 2\pi R_0^2 \left[1 - \left(1 - \frac{1}{2!}\frac{r^2}{R_0^2} + \frac{1}{4!}\frac{r^4}{R_0^4} + \cdots\right)\right] \\ &= 2\pi R_0^2 \left[\frac{r^2}{2R_0^2} - \frac{r^4}{24R_0^4} + \cdots\right] , \end{aligned}$$

which helps obtain the desired result

$$\frac{A_c}{A_f} \approx 1 - \frac{R}{24} r^2 \ .$$

The fact that a sphere is an object with the positive Gaussian curvature (or curvature invariant) implies that $A_c < A_f$ for a given radius r.

Exercise 9.24

Verify (9.560), i.e.

$$\mathscr{L}_{\mathbf{u}} \frac{\partial \, \underline{\hat{h}}^\alpha}{\partial t^\beta} = \frac{\partial \, \underline{u}^\theta}{\partial t^\beta} \frac{\partial \, \underline{\hat{h}}^\alpha}{\partial t^\theta} + \underline{u}^\theta \frac{\partial^2 \, \underline{\hat{h}}^\alpha}{\partial t^\beta \partial t^\theta} - \frac{\partial \, \underline{\hat{h}}^\theta}{\partial t^\beta} \frac{\partial \, \underline{u}^\alpha}{\partial t^\theta} - \underline{\hat{h}}^\theta \frac{\partial^2 \, \underline{u}^\alpha}{\partial t^\beta \partial t^\theta} \ .$$

Further, prove the important relation (9.595), i.e.

$$\mathscr{L}_{\mathbf{u}} \Gamma^\gamma_{\alpha\beta} = \frac{\partial^2 \, \underline{u}^\gamma}{\partial t^\alpha \partial t^\beta} + \underline{u}^\theta \frac{\partial \Gamma^\gamma_{\alpha\beta}}{\partial t^\theta} + \Gamma^\gamma_{\alpha\theta} \frac{\partial \, \underline{u}^\theta}{\partial t^\beta} + \Gamma^\gamma_{\theta\beta} \frac{\partial \, \underline{u}^\theta}{\partial t^\alpha} - \Gamma^\theta_{\alpha\beta} \frac{\partial \, \underline{u}^\gamma}{\partial t^\theta} \ .$$

Solution. These results may be attained by using an infinitesimal coordinate transformation as discussed in the footnote on Sect. 9.8.1. This procedure will be used in the following.

To begin with, let $\mathbf{u} = \underline{u}^\alpha \mathbf{a}_\alpha$ be a smooth vector field whose integral curves are denoted by $t^\theta(\lambda)$. Further, let $\hat{\mathbf{h}} = \underline{\hat{h}}^\alpha \mathbf{a}_\alpha$ be another smooth vector field with the corresponding flow lines $t^\theta(\mu)$, see Fig. 9.32. Now, consider two infinitesimally close points (t^1, t^2) and $(\bar{t}^1 = t^1 + \underline{u}^1 \Delta\lambda, \bar{t}^2 = t^2 + \underline{u}^2 \Delta\lambda)$. Then, guided by $(9.549)_2$,

$$\mathscr{L}_{\mathbf{u}} \frac{\partial \, \hat{\underline{h}}^{\alpha}}{\partial t^{\beta}} := \lim_{\Delta\lambda \to 0} \frac{1}{\Delta\lambda} \left[\frac{\partial \, \hat{\underline{h}}^{\alpha} \left(\bar{t}^{1}, \bar{t}^{2} \right)}{\partial t^{\beta}} - \frac{\partial \, \bar{\hat{\underline{h}}}^{\alpha} \left(\bar{t}^{1}, \bar{t}^{2} \right)}{\partial \bar{t}^{\beta}} \right] , \tag{9.886}$$

where

$$\underbrace{\frac{\partial \, \hat{\underline{h}}^{\alpha} \left(\bar{t}^{1}, \bar{t}^{2} \right)}{\partial t^{\beta}} = \frac{\partial \, \hat{\underline{h}}^{\alpha}}{\partial t^{\beta}} + \frac{\partial^{2} \, \hat{\underline{h}}^{\alpha}}{\partial t^{\gamma} \partial t^{\beta}} \underline{u}^{\gamma} \Delta\lambda + o\left(\Delta\lambda\right)}_{\text{note that the Landau order symbol } o(\Delta\lambda) \text{ satisfies } \lim_{\Delta\lambda \to 0} o(\Delta\lambda)/\Delta\lambda = 0} , \tag{9.887}$$

and $\partial \, \bar{\hat{\underline{h}}}^{\alpha} / \partial \bar{t}^{\beta}$ at $\left(\bar{t}^{1}, \bar{t}^{2} \right)$, according to (9.172a)$_2$, needs to be characterized for the problem at hand. To do so, one should have

$$\frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}} = \frac{\partial \left(t^{\alpha} + \underline{u}^{\alpha} \Delta\lambda \right)}{\partial t^{\gamma}} = \delta^{\alpha}_{\gamma} + \frac{\partial \, \underline{u}^{\alpha}}{\partial t^{\gamma}} \Delta\lambda , \tag{9.888}$$

and

$$\begin{aligned} \frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} &= \frac{\partial \left(\bar{t}^{\delta} - \underline{u}^{\delta} \Delta\lambda \right)}{\partial \bar{t}^{\beta}} = \delta^{\delta}_{\beta} - \frac{\partial \, \underline{u}^{\delta}}{\partial t^{\theta}} \left(\delta^{\theta}_{\beta} - \frac{\partial \, \underline{u}^{\theta}}{\partial \bar{t}^{\beta}} \Delta\lambda \right) \Delta\lambda \\ &= \delta^{\delta}_{\beta} - \frac{\partial \, \underline{u}^{\delta}}{\partial t^{\beta}} \Delta\lambda + o\left(\Delta\lambda\right) . \end{aligned} \tag{9.889}$$

Consequently,

$$\begin{aligned} \frac{\partial \, \bar{\hat{\underline{h}}}^{\alpha} \left(\bar{t}^{1}, \bar{t}^{2} \right)}{\partial \bar{t}^{\beta}} = & \underbrace{\frac{\partial \bar{t}^{\alpha}}{\partial t^{\gamma}} \frac{\partial \, \hat{\underline{h}}^{\gamma}}{\partial t^{\delta}} \frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}}}_{= \left[\delta^{\alpha}_{\gamma} + \frac{\partial \, \underline{u}^{\alpha}}{\partial t^{\gamma}} \Delta\lambda \right] \frac{\partial \, \hat{\underline{h}}^{\gamma}}{\partial t^{\delta}} \left[\delta^{\delta}_{\beta} - \frac{\partial \, \underline{u}^{\delta}}{\partial t^{\beta}} \Delta\lambda + o(\Delta\lambda) \right]} \\ & + \underbrace{\frac{\partial^{2} \bar{t}^{\alpha}}{\partial t^{\delta} \partial t^{\gamma}} \frac{\partial t^{\delta}}{\partial \bar{t}^{\beta}} \hat{\underline{h}}^{\gamma}}_{= \frac{\partial}{\partial t^{\delta}} \left[\delta^{\alpha}_{\gamma} + \frac{\partial \, \underline{u}^{\alpha}}{\partial t^{\gamma}} \Delta\lambda \right] \left[\delta^{\delta}_{\beta} - \frac{\partial \, \underline{u}^{\delta}}{\partial t^{\beta}} \Delta\lambda + o(\Delta\lambda) \right] \hat{\underline{h}}^{\gamma}} \\ = & \frac{\partial \, \hat{\underline{h}}^{\alpha}}{\partial t^{\beta}} + \left[\frac{\partial \, \underline{u}^{\alpha}}{\partial t^{\gamma}} \frac{\partial \, \hat{\underline{h}}^{\gamma}}{\partial t^{\beta}} - \frac{\partial \, \hat{\underline{h}}^{\alpha}}{\partial t^{\delta}} \frac{\partial \, \underline{u}^{\delta}}{\partial t^{\beta}} \right] \Delta\lambda \\ & + \frac{\partial^{2} \, \underline{u}^{\alpha}}{\partial t^{\beta} \partial t^{\gamma}} \hat{\underline{h}}^{\gamma} \Delta\lambda + o\left(\Delta\lambda\right) . \end{aligned} \tag{9.890}$$

With the aid of (9.886), (9.887) and (9.890), one can now provide the Lie derivative of the partial differentiation of a vector.

Next, consider the object $\Gamma^{\gamma}_{\alpha\beta}$ whose Lie derivative is defined by

$$\mathscr{L}_{\mathbf{u}}\Gamma^{\gamma}_{\alpha\beta} := \lim_{\Delta\lambda\to 0}\frac{\Gamma^{\gamma}_{\alpha\beta}\left(\bar{t}^{1},\bar{t}^{2}\right) - \bar{\Gamma}^{\gamma}_{\alpha\beta}\left(\bar{t}^{1},\bar{t}^{2}\right)}{\Delta\lambda} \ , \tag{9.891}$$

where

$$\Gamma^{\gamma}_{\alpha\beta}\left(\bar{t}^{1},\bar{t}^{2}\right) = \Gamma^{\gamma}_{\alpha\beta} + \frac{\partial\Gamma^{\gamma}_{\alpha\beta}}{\partial t^{\theta}}\,\underline{u}^{\theta}\,\Delta\lambda + o\left(\Delta\lambda\right) \ , \tag{9.892}$$

and $\bar{\Gamma}^{\gamma}_{\alpha\beta}\left(\bar{t}^{1},\bar{t}^{2}\right)$, according to (9.93), now takes the form

$$\begin{aligned}
\bar{\Gamma}^{\gamma}_{\alpha\beta}\left(\bar{t}^{1},\bar{t}^{2}\right) = & \underbrace{\frac{\partial t^{\delta}}{\partial\bar{t}^{\alpha}}\frac{\partial t^{\theta}}{\partial\bar{t}^{\beta}}\frac{\partial\bar{t}^{\gamma}}{\partial t^{\rho}}\Gamma^{\rho}_{\delta\theta}}_{=\left[\delta^{\delta}_{\alpha} - \frac{\partial\underline{u}^{\delta}}{\partial t^{\alpha}}\Delta\lambda + o(\Delta\lambda)\right]\left[\delta^{\theta}_{\beta} - \frac{\partial\underline{u}^{\theta}}{\partial t^{\beta}}\Delta\lambda + o(\Delta\lambda)\right]\left[\delta^{\gamma}_{\rho} + \frac{\partial\underline{u}^{\gamma}}{\partial t^{\rho}}\Delta\lambda\right]\Gamma^{\rho}_{\delta\theta}} \\
& + \underbrace{\frac{\partial^{2} t^{\rho}}{\partial\bar{t}^{\alpha}\partial\bar{t}^{\beta}}\frac{\partial\bar{t}^{\gamma}}{\partial t^{\rho}}}_{=\frac{\partial}{\partial\bar{t}^{\alpha}}\left[\delta^{\rho}_{\beta} - \frac{\partial\underline{u}^{\rho}}{\partial t^{\beta}}\Delta\lambda + o(\Delta\lambda)\right]\left[\delta^{\gamma}_{\rho} + \frac{\partial\underline{u}^{\gamma}}{\partial t^{\rho}}\Delta\lambda\right]} \\
= & \Gamma^{\gamma}_{\alpha\beta} + \left[\frac{\partial\,\underline{u}^{\gamma}}{\partial t^{\rho}}\Gamma^{\rho}_{\alpha\beta} - \frac{\partial\,\underline{u}^{\theta}}{\partial t^{\beta}}\Gamma^{\gamma}_{\alpha\theta} - \frac{\partial\,\underline{u}^{\delta}}{\partial t^{\alpha}}\Gamma^{\gamma}_{\delta\beta}\right]\Delta\lambda \\
& - \frac{\partial^{2}\,\underline{u}^{\gamma}}{\partial t^{\alpha}\partial t^{\beta}}\Delta\lambda + o\left(\Delta\lambda\right) \ .
\end{aligned} \tag{9.893}$$

It is then easy to arrive at the desired expression for the Lie derivative of the connection coefficients.

References

1. Weatherburn, C. E. (1955). *Differential geometry of three dimensions* (Vol. 1). Cambridge: Cambridge University Press.
2. Stoker, J. J. (1969). *Differential geometry*. Wiley.
3. Kreyszig, E. (1991). *Differential geometry*. New York: Dover Publications.
4. Spivak, M. (1999). *A comprehensive introduction to differential geometry* (3rd ed., Vol. 1). Publish or Perish, Inc.
5. Gray, A., Abbena, E., & Salamon, S. (2006). *Modern differential geometry of curves and surfaces with mathematica* (3rd ed.). Chapman & Hall/CRC.
6. O'Neill, B. (2006). *Elementary differential geometry* (Revised Second Edition). Academic.
7. Pressley, A. (2010). *Elementary differential geometry* (2nd ed.). London: Springer.
8. do Carmo, M. P. (2016). *Differential geometry of curves and surfaces: Revised & Updated* (2nd ed.). Mineola, New York: Dover Publications.
9. Banchoff, T., & Lovett, S. (2016). *Differential geometry of curves and surfaces* (2nd ed.). CRC Press.
10. Christoffel, E. B. (1869). Über die Transformation der homogenen Differentialausdrücke zweiten Grades. *Journal für die reine und angewandte Mathematik, 70*, 46–70.

11. Levi-Civita, T. (1925). *Lezioni di Calcolo Differenziale Assoluto*. Zanichelli. Reprinted as *The Absolute Differential Calculus* New York: Dover Publications (1977).
12. Steinmann, P. (2015). *Geometrical foundations of continuum mechanics: An application to first- and second-order elasticity and elasto-plasticity*. Lecture notes in applied mathematics and mechanicsBerlin, Heidelberg: Springer.
13. Cartan, É. (1986). *On manifolds with an affine connection and the theory of general relativity*. In *Monographs and Textbooks in Physical Science*, Bibliopolis, Napoli.
14. Naghdi, P. M. (1973). The theory of shells and plates. In C. Truesdell (Ed.), *Linear theories of elasticity and thermoelasticity.* Berlin, Heidelberg: Springer.
15. Grinfeld, P. (2013). *Introduction to tensor analysis and the calculus of moving surfaces*. New York, NY: Springer.
16. Carroll, S. M. (2003). *Spacetime and geometry: an introduction to general relativity*. Addison-Wesley.
17. Dalarsson, M., & Dalarsson, N. (2015). *Tensors, relativity, and cosmology* (2nd ed.). Academic.
18. Riemann, B. (1868). Über die Hypothesen, welche der Geometrie zu Grunde liegen. *Abhandlungen der Königlichen Gesellshaft der Wissenschaften zu Göttingen., 13*, 133–152.
19. Bianchi, L. (1902). Sui simboli a quattro indici e sulla curvatura di Riemann. *Atti dell'Accademia Nazionale dei Lincei, Rendiconti, 11*(5), 3–7.
20. Frenet, F. (1852). Sur les courbes à double courbure. *Journal de Mathématiques Pures et Appliquées,* 437–447.
21. Serret, J. A. (1851). Sur quelques formules relatives à la théorie des courbes à double courbure. *Journal de Mathématiques Pures et Appliquées,* 193–207.
22. Darboux, G. (1887–1896). *Leçons sur la théorie générale des surfaces* (Vol. I to IV). Paris: Gauthier-Villars.
23. Casey, J. (1996). *Exploring curvature*. Wiesbaden: Vieweg+Teubner Verlag.
24. Meusnier, J. B. (lu 1776) (1785). Mémoire sur la courbure des surfaces. Mémoires de Mathématique et de Physique présentés à l'Académie Royale des Sciences, par Divers Savants, & lûs dans ses Assemblés (Paris) *10*, 477–510.
25. Washizu, K. (1968). *Variational methods in elasticity and plasticity*. London: Pergamon Press.
26. Oden, J. T., & Reddy, J. N. (1983). *Variational methods in theoretical mechanics* (2nd ed.). Berlin, Heidelberg: Springer.
27. Jost, J., & Li-Jost, X. (1998). *Calculus of variations*. Cambridge: Cambridge University Press.
28. Euler, L. (1767). Recherches sur la courbure des surfaces. *Mémoires de l'Académie des Sciences de Berlin, 16*, 119–143.
29. Rodrigues, O. (1815). Recherches sur la théorie analytique des lignes et des rayons de courbure des surfaces. *Correspondance sur l'Ecole Royale Polytechnique, tome, III*, 162–182.
30. Joachimstahl, F. (1846). Demonstrationes theorematum ad superficies curvas spectantum. *Journal für die reine und angewandte Mathematik, 30*, 347–350.
31. Mainardi, G. (1856). Sulla teoria generale delle superficie. *Giornale dell'I. R. Istituto Lombardo, 9*, 385–398.
32. Codazzi, D. (1867). Sulle coordinate curvilinee d'una superficie e dello spazio. *Annali di Matematica Pura ed Applicata, 1*, 293–316.
33. Gauss, C. F. (1828). *Disquisitiones Generales Circa Superficies Curvas*. Typis Dieterichianis.
34. Bieberbach, L. (1933). *Differentialgeometrie*. Teubners Mathematische Leitfäden, Bd. 31, Teubner, Leipzig und Berlin 1932.
35. Brioschi, F. (1852). Intorno ad alcuni punti della teoria delle superficie. *Annali di Scienze Matematiche e Fisiche, 3*, 293–321.
36. Hartmann, E. (1999). On the curvature of curves and surfaces defined by normalforms. *Computer Aided Geometric Design, 16*, 355–376.
37. Goldman, R. (2005). Curvature formulas for implicit curves and surfaces. *Computer Aided Geometric Design, 22*, 632–658.
38. Hilbert, D., & Cohn-Vossen, S. (1952). *Geometry and the imagination*. AMS Chelsea Publishing.

39. Günther, H., & Müller, V. (2019). *The special theory of relativity: Einstein's world in new axiomatics*. Singapore: Springer.
40. Pais, A. (1982). *Subtle is the lord: The science and the life of Albert Einstein*. USA: Oxford University Press.
41. Voss, A. (1880). Zur Theorie der Transformation quadratischer Differentialausdrücke und der Krümmung höherer Mannigfaltigketien. *Mathematische Annalen, 16*, 129–179.
42. Hawking, S. W., & Ellis, G. F. R. (1973). *The large scale structure of space-time*. Cambridge: Cambridge University Press.
43. Misner, C. W., Thorne, K. S., & Wheeler, J. A. (1973). *Gravitation*. W. H. Freeman and Company.
44. Robinson, M. B., Ali, T., & Cleaver, G. B. (2009). *A simple introduction to particle physics: Part II - geometric foundations and relativity*. arXiv: High Energy Physics - Theory.
45. Helgason, S. (1978). *Differential geometry, lie groups, and symmetric spaces*. Academic.
46. Fecko, M. (2006). *Differential geometry and lie groups for physicists*. Cambridge: Cambridge University Press.
47. Cartan, É. (1922). *Leçons sur les invariants intégraux*. Paris: Hermann & Fils.
48. Hehl, F. W., & Obukhov, Y. N. (2007). Élie Cartan's torsion in geometry and in field theory, an essay. *Annales de la Fondation Louis de Broglie, 32*, 157–194.
49. Hadamard, J. (1908). *Mémoire sur le probléme d'analyse relatif à l'équilibre des plaques élastiques encastrées*. Paris: Imprimerie nationale.
50. Thomas, T. Y. (1961). *Concepts from tensor analysis and differential geometry*. Academic.
51. Grinfeld, M. A. (1991). *Thermodynamic methods in the theory of heterogeneous systems (Interaction of mechanics and mathematics)*. Wiley.
52. Grinfeld, P. (2012). A better calculus of moving surfaces. *Geometry and Symmetry in Physics, 26,* 61–69.
53. Svintradze, D. V. (2019). Shape dynamics of bouncing droplets. *Scientific Reports, 9,* 6105.
54. Bathe, K. J., Ramm, E., & Wilson, E. L. (1975). Finite element formulations for large deformation dynamic analysis. *International Journal for Numerical Methods in Engineering, 9*, 353–386.
55. Zienkiewicz, O. C., Bauer, J., Morgan, K., & Oãte, E. (1977). A simple and efficient element for axisymmetric shells. *International Journal for Numerical Methods in Engineering, 11*, 1545–1558.
56. Hughes, T. J. R., & Tezduyar, T. E. (1981). Finite elements based upon Mindlin plate theory with particular reference to the four-node bilinear isoparametric element. *Journal of Applied Mechanics, 48,* 587–596.
57. Simo, J. C., Fox, D., & Rifai, M. S. (1990). On a stress resultant geometrically exact shell model. Part III: Computational aspects of the nonlinear theory. *Applied Mechanics and Engineering, 79*, 21–70.
58. Ibrahimbegović, A., Taylor, R. L., & Wilson, E. L. (1990). A robust quadrilateral membrane finite element with drilling degrees of freedom. *International Journal for Numerical Methods in Engineering, 30*, 445–457.
59. Gruttmann, F., Wagner, W., & Wriggers, P. (1992). A nonlinear quadrilateral shell element with drilling degrees of freedom. *Archive of Applied Mechanics, 62*, 474–486.
60. Wriggers, P., Eberlein, R., & Reese, S. (1996). A comparison of three-dimensional continuum and shell elements for finite plasticity. *International Journal of Solids and Structures, 33*, 3309–3326.
61. Betsch, P., Gruttmann, F., & Stein, E. (1996). A 4-node finite shell element for the implementation of general hyperelastic 3D-elasticity at finite strains. *Computer Methods in Applied Mechanics and Engineering, 130*, 57–79.
62. Bischoff, M., & Ramm, E. (1997). Shear deformable shell elements for large strains and rotations. *International Journal for Numerical Methods in Engineering, 40*, 4427–4449.
63. Miehe, C. (1998). A theoretical and computational model for isotropic elastoplastic stress analysis in shells at large strains. *Computer Methods in Applied Mechanics and Engineering, 155*, 193–233.

64. Hauptmann, R., Schweizerhof, K., & Doll, S. (2000). Extension of the "solid-shell" concept for application to large elastic and large elastoplastic deformations. *International Journal for Numerical Methods in Engineering, 49*, 1121–1141.
65. Cirak, F., Ortiz, M., & Schröder, P. (2000). Subdivision surfaces: A new paradigm for thin-shell finite-element analysis. *International Journal for Numerical Methods in Engineering, 47*, 2039–2072.
66. Pimenta, P. M., Campello, E. M. B., & Wriggers, P. (2004). A fully nonlinear multi-parameter shell model with thickness variation and a triangular shell finite element. *Computational Mechanics, 34*, 181–193.
67. Bischoff, M., Wall, W. A., Bletzinger, K. U., & Ramm, E. (2004). *Models and Finite Elements for Thin-walled Structures*. In Stein, E., de Borst, R., & T. J. R. Hughes, T. J. R. (eds.) *Encyclopedia of Computational Mechanics*, Volume 2: Solids and Structures. Wiley.
68. Lee, P. S., & Bathe, K. J. (2005). Insight into finite element shell discretizations by use of the basic shell mathematical model. *Computers & Structures, 83*, 69–90.
69. Chapelle, D., & Bathe, K. J. (2011). *The finite element analysis of shells - fundamentals* (2nd ed.). Berlin, Heidelberg: Springer.
70. Reissner, E. (1945). The effect of transverse shear deformation on the bending of elastic plates. *Journal of Applied Mechanics, 12*, 69–77.
71. Mindlin, R. D. (1951). Influence of rotary inertia and shear on flexural motions of isotropic, elastic plates. *Journal of Applied Mechanics, 18*, 31–38.
72. Hencky, H. (1947). Über die Berücksichtigung der Schubverzerrung in ebenen Platten. *Ingenieur-Archiv, 16*, 72–76.
73. Sansour, C. (1995). A theory and finite element formulation of shells at finite deformations involving thickness change: Circumventing the use of a rotation tensor. *Archive of Applied Mechanics, 65*, 194–216.
74. Sansour, C. (1998). Large strain deformations of elastic shells constitutive modelling and finite element analysis. *Computer Methods in Applied Mechanics and Engineering, 161*, 1–18.
75. Monge, G. (1807). *Application De L'analyse À La Géométrie*. Paris: Perronneau.

Index

A
Actual difference, 232
Addition, 1, 4, 5, 9, 13, 15, 36, 37, 93, 99, 377
Additive decomposition, 48, 66, 429
Admissible, 525, 526
Ampère's circuital law, 422
Angle, 5, 7, 247, 435, 468, 486, 489
Anisotropic, 233, 236
Anti-commutative, 17
Antisymmetric, 48, 237, 330
Arc length, 499–501, 507, 525
Associative, 3, 4, 6, 21, 22, 380
Asymptotic
  curve, 494, 495, 549
  direction, 494

B
Backward, 230, 232
2-ball, 561, 654
3-ball, 561, 654
Base curve, 508, 512, 647
Basic invariants, 233, 237, 245, 247
Basis, 5, 40, 95, 104, 131, 132
  Cartesian, 2, 12, 40, 95, 142
  change of, 142
  contravariant, 175, 177, 179, 317
  covariant, 172, 177, 178, 211, 317
  -dependent, 320
  dual, 211, 219
  -free, 147
  functional, 234, 245, 250, 256, 276, 286
  general, 12, 54, 171, 174, 316, 434
  global, 174
  integrity, 233, 234, 237, 250, 256, 276, 277
  irregular, 172, 174
  local, 174
  minimal integrity, 233
  new, 67, 70
  old, 67, 70
  orthogonal, 149
  orthonormal, 7, 9, 20, 21, 65, 68
  positively oriented, 11
  rectangular, 95
  regular, 172
  standard, 12, 39, 54, 169, 211
  surface contravariant, 444
  surface covariant, 441
Bending, 435, 502, 504, 538, 604, 607
Bianchi identities, 349, 482
Bilinear, 6, 21, 36, 43, 50, 109, 164, 184, 199, 202, 491, 568
Bilinearity, 9, 22, 25, 39, 175, 376, 377, 380, 410, 447, 451, 584
Binormal vector, 503, 505, 509

C
Calculus
  of moving surfaces, 475, 585–587
  tensor, 112
  of variations, 525, 527
Catenoid, 459, 619
Cayley-Hamilton, 56, 139, 147, 233, 273, 278, 280, 281, 283, 496
Central, 230, 232
Characteristic(s), 1, 10, 42, 146, 177, 332, 434, 490, 549, 553, 570

S. Sahraee and P. Wriggers, *Tensor Calculus and Differential Geometry for Engineers*, https://doi.org/10.1007/978-3-031-33953-0

determinant, 133
equation, 133, 134, 139, 153, 160, 288, 633
polynomial(s), 133, 136, 144, 147
space(s), 135, 156, 157, 255
values, 131, 136, 538
Charge
-current 1-form, 423, 428
-current 3-form, 423
-current differential forms, 428
density, 422, 425
Christoffel symbols, 318, 455
of the first kind, 318
of the second kind, 317
Circle, 438
generating, 438
great, 524, 529, 534
kissing, 502
osculating, 502
Circulation, 371, 374
Clairaut, 531
parametrization, 530, 533, 535, 639
patch, 530–532
Closed, 435, 479, 481
curve, 370, 371, 414
surface, 198, 365, 366, 368, 370, 409, 410, 417
Closure, 64, 435
Coaxial, 157, 158, 241
Cofactor, 59, 80, 138, 206
Collinear, 520
Commutation theorem, 156
Commutative, 3, 6, 42, 86, 109, 178, 180, 556, 569, 570
Commutator, 469, 470, 564, 583, 584
Commute(s), 79, 88, 156, 157, 328, 526, 556, 580, 598
Compatibility equations, 550
Complex, 134
conjugate, 9
eigenvalues, 132
number, 9
vector, 9
vector space, 9
Components, 5, 15, 40, 51, 60, 67, 68, 70, 93, 131, 158, 191
ambient, 454, 599
ambient contravariant, 586
Cartesian, 1, 11, 95–97, 104, 106, 174
co-contravariant, 192
contra-covariant, 192
contravariant, 183, 187, 191, 270, 327
covariant, 183, 187, 192, 270
curvilinear, 193, 194, 229
fully contravariant, 194
fully covariant, 329
mixed, 329, 481
mixed co-contravariant, 187
mixed contra-covariant, 187
natural, 586, 587
normal, 484
rectangular, 11
spectral, 142, 144, 307
surface, 454, 593
surface contravariant, 587
Composition, 42, 83, 85, 86, 108, 109, 111, 112, 126, 189, 195, 206
Compressible neo-Hookean material, 310
Compression, 435
Computer program, 206, 310, 353, 637
Cone
elliptic, 499
generalized, 497, 511, 512
Configuration
current, 291, 293, 295, 308
deformed, 291
initial, 291
reference, 291, 293, 308
spatial, 291
undeformed, 291
Congruence, 565
Connection, 434, 446
Christoffel, 460
coefficients, 456, 465, 470, 657
Levi-Civita, 460
Riemann, 460
Conservation of electric charge, 365
Constitutive, 295
equations, 295
modeling, 140
relations, 297, 300
response, 217
Continuity equation, 424, 428
Continuum mechanics, 2, 56, 93, 96, 157, 169, 209, 290, 291, 315
Contracted, 281, 283
Bianchi identities, 560
Contraction
double, 50, 52, 94–96, 108, 113, 125, 206, 212, 238, 239, 279
quadruple, 111
simple, 44, 54, 56, 81, 95, 157, 206
single, 44, 591
Control theory, 131
Convection-diffusion, 344
Convention, 11, 198, 291, 366, 370, 386

Einstein summation, 14
summation, 14–16, 28, 40, 183, 215, 436
Coordinate(s), 11, 14, 169, 326, 343
axes, 11
Cartesian, 169, 171, 286, 308, 328, 333, 342, 344, 379, 386
a change of, 71, 97, 106, 177, 181
curves, 464, 495, 523, 545, 555, 565, 570, 639
curvilinear, 2, 169, 175, 177, 225, 316–318, 333, 343, 435
cylindrical, 169, 171, 199, 319, 335, 353, 385, 419
free, 36, 392, 411
global, 71, 169, 171
line, 11
local, 71, 169, 171, 395
representation, 21, 41, 106, 109, 121, 463
spherical, 169, 171, 199, 319, 335, 353, 385, 419
Coplanar, 289, 438
Coulomb's law for static electricity, 360
Covariantly constant, 324, 325, 330, 331, 472
Covector, 183, 382, 392, 406, 407, 481, 572, 580
Creatures, 365, 374, 580, 604
Curl, 315, 339–342, 350
Current
configuration, 291, 293, 295, 308
density, 422
Curved surfaces, 433, 456, 461, 462, 466, 497, 506, 549, 585, 602
Cyclic, 53, 349
Cylinder, 435, 494, 501
elliptic, 499, 613
generalized, 497, 511

D
Damage, 290
Deformation gradient, 291, 293, 302
Degree, 281–285
partial, 281, 282
total, 281, 284
Derivative
contravariant, 321, 322
covariant, 321, 323–325, 328, 330, 348, 349
directional, 213, 218, 280, 316, 347, 376, 465, 564, 571
partial, 215, 225, 428, 585
Determinant, 23, 28, 55, 56, 120, 206, 260, 384, 458, 549, 594
Diagonalizable, 131, 143, 144, 146, 158
Differential
0-form, 375, 377, 388, 580
1-form, 376, 377, 388, 393, 395, 581, 582
2-form, 377, 378, 389, 390, 407, 430
3-form, 380–382, 391, 427
calculus, 211, 387, 526
forms, 365, 374, 375, 382, 385, 387, 422, 425
total, 30, 212, 213, 216, 218, 221, 226, 264, 270, 580
Dimension, 5, 40, 95, 134, 241, 375–377
Direction(s), 292
cosines, 69
principal Eulerian, 292
principal Lagrangian, 292
principal material, 292
principal referential, 292
principal spatial, 292
Directrix, 508, 513
Distribution parameter, 513, 645
Distributive, 4, 15
Divergence, 315, 336–339, 350, 367, 368, 383
-free, 337
theorem, 198, 365, 368, 370, 405, 408
Dual, 142, 143, 145, 175, 177, 197, 200, 211, 317, 392, 442, 444, 555
Dyad, 45, 47, 54, 94, 175
Dyadic, 40

E
Eigenbasis, 132, 135, 136, 139, 144
Eigenpairs, 132, 140, 141, 148, 150, 243, 270, 541
Eigenprojections, 146, 149, 150
Eigenspace(s), 135, 136, 144, 146, 148, 154, 156, 157, 241
Eigenvalue(s), 61, 66, 131, 537, 538
complex, 132
decomposition, 140
distinct, 132, 135, 140, 141
identical, 132, 133, 138
multiple, 143, 144, 149
non-degenerate, 132
non-multiple, 142, 149, 241, 287
real, 241
repeated, 149, 253
simple, 136, 288
Eigenvectors, 131, 132, 141, 148
generalized, 144, 158, 159
left, 133, 142

linearly independent, 135, 136
normalized, 132, 154
orthonormal set of, 253
right, 132, 133
span of, 135
Einstein's
general theory of relativity, 434
gravitational field equations, 557
Elastic, 127, 611, 612
compliance tensor, 128
material, 231
stiffness tensor, 128
Elasticity tensor, 108, 128, 295, 308
compliance, 128
Eulerian, 296
first, 296, 301
Lagrangian, 296
material, 296, 305
referential, 296
second, 296, 303
spatial, 296
Electromagnetism, 315, 422
Element, 5, 11, 40, 54, 67
arc length, 486
area, 448, 613
diagonal, 62
generator, 237, 250
initially curved, 604
isolated, 609
line, 172, 196, 370, 535
off-diagonal, 288
redundant, 5
structural, 604
surface, 198, 407, 606, 609
volume, 196, 198, 366, 368
Elementary, 136, 144, 376, 379, 382
Ellipsoid, 437, 563
Elliptic, 494
geometry, 434, 437
paraboloid, 438
point, 494, 540
Embedded
curves, 537
surfaces, 433, 435
Engineering, 290
aerospace, 290
biomedical, 290
civil, 290
mechanical, 290
nuclear, 290
tissue, 233
Entities, 1, 35, 36, 386, 387, 435
Equator, 534
Equilibrium, 117, 604, 609
equations, 344, 610
Equilibrium condition
force, 609
moment, 610
Equivalent, 282, 285
Error, 230, 237, 310
round-off, 230
true, 230
truncation, 230
Euler equations, 528
Extension, 281–283
Exterior
calculus, 383, 392, 393, 396, 398, 419, 424, 429
derivative, 387, 388, 390, 419, 425, 426, 428, 430
Extremum, 527, 537
Extrinsic, 435, 457, 471, 486, 501, 515, 521, 549, 551, 556

F
Faraday 2-form, 423
Faraday's law of induction, 422
Field(s), 1, 9, 37
displacement, 611, 612
electric, 422
lines, 565
magnetic, 422
outward unit normal, 410, 414, 451
scalar, 315, 328, 333, 336, 347, 353, 419, 423, 485, 587, 590
tensor, 209, 368, 372, 374
vector, 209, 365, 369, 371
Finite, 210
deformations, 308, 612
differences, 230
element, 29, 67, 290, 604
rotation, 82
strains, 230
First
Bianchi identity, 331
-kind Ricci tensor, 558
-order covariant derivative, 321
-order Taylor series, 209
-order tensors, 1, 93
Flat, 486
intrinsically, 332, 435, 484, 497, 501, 504
manifold, 435
point, 495, 540
space(s), 328, 332, 434, 484, 654
Flow lines, 565, 570, 574, 655

Flux integral, 367, 368
Folding, 511
Formula(s)
  Bieberbach, 554, 555, 623
  binomial, 86
  Bonnet, 522, 548
  Brioschi, 554
  Cardano's, 134
  closed-form, 230
  Doyle-Ericksen, 309
  Euler, 537, 633, 649
  Euler-Rodrigues, 83, 85
  Frenet, 506, 507
  Frenet-Serret, 506
  Gauss, 456, 457, 465, 550, 650
  intrinsic, 554
  Liouville, 524
  Mainardi-Codazzi, 551
  recursive, 137
  representation, 257, 258
  Rodrigues, 546, 640, 642
  Rodrigues rotation, 82
  spectral, 255, 272, 304
  Sylvester's, 146
  Vieta's, 137
  Voss-Weyl, 336, 485
  Weingarten, 457, 546, 550
Forward, 230, 231
Frame, 11, 287
  Cartesian coordinate, 11, 12, 64, 169, 434
  Frenet, 504, 520
  new, 69, 288
  old, 71
  reference, 286, 287
  of reference, 70
Frobenius covariants, 146
Function, 30
  exponential tensor, 44, 86, 87, 89, 267
  gamma, 564
  linear, 35, 65, 129, 212, 249, 251, 255, 277, 417
  non-polynomial, 234
  point, 170, 333, 436, 497, 508, 586, 592
  polynomial, 234
  scaling, 622
  single-valued, 234, 286, 290
  stored-energy, 293, 301, 304
  strain-energy, 293, 294, 297
Functional, 527
Fundamental, 433, 459, 462
  differential quadratic form, 197
  lemma of variational calculus, 527
  theorem of calculus, 365
  theorem of space curves, 507, 549
  theorem of surfaces, 550
Fundamental form(s), 486, 496
  first, 486, 488, 490, 493, 495, 517, 549, 550
  second, 491, 495, 517, 540, 543, 549, 550
  third, 495

G

Gauss equations of the surface, 551
Gaussian
  coordinates, 436, 441, 456
  curvature, 433, 458, 492, 549
Gauss's
  law, 422
  law for magnetism, 422
  theorem, 365, 367
Generalized, 189
  cone, 497
  cylinder, 497
  eigenvectors, 158, 159
  Hooke's law, 128
  Rivlin's identities, 280
  Stokes' theorem, 365, 375, 405
General relativity, 433, 465, 469, 557, 561, 564
Geodesic, 524, 525, 529, 531, 532, 534
  curvature, 433, 515, 519, 520, 522, 524, 529, 549
  curvature vector, 516
  curve, 524, 529, 532, 547, 549, 561
  equations, 524, 525, 528–530, 536
  torsion, 521, 547–549, 648
Geometrical reasoning, 234
Geometry, 433, 472
  differential, 209, 433, 434, 478
  elliptic, 434
  Euclidean, 434
  Gauss-Bolyai-Lobachevsky, 434
  hyperbolic, 434
  intrinsic, 558
  non-Euclidean, 434
  Riemannian, 332
  spherical, 434
$\mathcal{G}$-invariant, 237
Gradient, 315, 332–335, 350, 383, 394, 424, 455, 484
  theorem for line integrals, 365, 406
Gram-Schmidt, 149
  orthogonalization, 149
  orthonormalization, 149
Green's theorem, 365, 372, 405

Groundform
first, 486, 487, 493
second, 491–493
third, 495
Growth conditions, 293

H
Heat conduction, 344
Helix, 401
circular cylindrical, 623
elliptic conical, 498
elliptic cylindrical, 499
Hemitropic, 236
Hessian, 344, 345
Hodge, 385
dual, 383, 387, 428
star operator, 383, 386
Homogeneous, 127, 315, 611
Hydrostatic, 66
Hyperbolic, 540
geometry, 434, 438, 439
paraboloid, 440, 617
point, 495, 540
Hyperboloid
circular, 616
of one sheet, 346, 509, 537
one-sheeted, 438, 495
of revolution, 616
single-sheeted, 509
two-sheeted, 439
Hyperelastic, 296, 310
materials, 108, 194, 295
solids, 116, 158
Hyperelasticity, 231, 293, 294, 296

I
Imaginary unit, 9
Index, 13
dummy, 14, 173, 183, 331, 588
free, 14, 44, 50, 173
juggling, 177, 330, 349, 350, 453, 484, 551, 553, 594, 601
live, 14
notation, 58, 610
subscript, 180
summation, 14
superscript, 179, 318, 460
Inextensibility, 612
Inextensible shell model, 612
Infinitesimal, 435
coordinate transformation, 567, 655
rotation, 74, 97, 611
strain tensor, 128, 611
strain theory, 128
Inplane strains, 612
Integrability conditions, 550
Integral, 367
curves, 565, 570, 571, 573, 574, 655
line, 370
surface, 367
theorems, 198, 365, 374, 375, 405, 411
volume, 366
Intrinsic, 435, 459, 465, 484, 486, 490, 515, 549, 551
Invariant, 131, 195
curvature, 557
object, 196, 465, 475, 519
time differentiation, 433, 589, 590, 601, 602
Inverse, 57, 79, 128, 145, 161, 174, 176, 179, 181, 189, 206, 445
additive, 198
square law, 360
Invertible, 56–58, 62, 80, 121, 132, 145, 158, 174, 217, 260, 291
Irreducible, 233, 234, 236, 237, 247, 286
Irrotational, 341
Isometric, 435
Isosurface, 333
Isotropic, 72, 97, 106, 107, 158, 296, 304, 309, 310, 611
hyperelasticity, 294
invariants, 235, 245, 254, 255, 286, 288–290
materials, 128, 293
relative, 236
scalar-valued function, 235, 245–250, 255, 289
tensor functions, 236, 281
tensors of order four, 106
tensors of order one, 72
tensors of order three, 97
tensor-valued function, 241, 242
tensors of order two, 73
vector-valued function, 249, 251, 253

J
Jacobian, 176–178, 181, 206, 318
matrix, 441
volumetric, 295
Jacobi identity, 568, 584
Joachimstahl theorem, 546, 547
Jordan
block, 158

canonical form, 158, 159, 162
normal form, 158, 159
vectors, 158

K
Kernel, 135
Klein bottle, 374
Kronecker delta, 14, 18, 19, 151
mixed, 173, 180, 182, 330, 560

L
Lagrange basis polynomials, 146
Lamé, 128, 310
Laplace equation, 344
Laplacian, 342–345, 399
Left, 15
-handed, 11, 20, 176
mapping, 44, 100, 101, 117
stretch tensor, 2, 291
Leibniz rule of differentiation, 467
Length, 5–7, 15, 49, 62, 69, 149, 179, 243, 435, 443, 463, 486
Lie
bracket, 470, 583
derivative, 433, 470, 481, 564, 565, 568, 580, 657
dragged, 570
transported, 570
Line, 347
of curvature, 545–549
element, 172
integral, 370–372
straight, 494, 499, 502, 504, 508, 529, 549
of striction, 513
tangent, 499, 502, 519, 545
Linear, 249, 252, 256, 258
algebra, 7, 15, 56, 131
algebraic equations, 240
approximation, 214
change, 67
combination, 5, 40, 67, 377, 434
elastic, 127, 612
elasticity, 115
equations, 469
function, 65, 129
map, 35, 45, 56, 94, 383
mapping, 35, 41, 93, 121, 129, 375
momentum, 365
operator, 232
ordinary differential equations, 44, 268
ordinary tensorial differential equations, 267
property, 67
relation, 35
structure, 5
subspace, 240
system of ordinary differential equations, 506
term, 213
transformation, 35, 36, 38, 58, 94, 131, 375, 449
Linearity, 6, 15, 35, 45, 46, 80, 93, 98, 151, 183, 266, 341
Linearly independent, 5, 55, 137, 174, 241, 441
Loads, 290, 291, 604
chemical, 290
electrical, 290
magnetic, 290
mechanical, 290
thermal, 290
tribological, 290

M
Mainardi-Codazzi equations, 551, 556, 557
Major, 14, 21, 36, 54, 462, 501
characteristic, 570
radius, 438
skew-symmetric, 116
symmetric, 115, 483
symmetries, 115, 119, 256, 257, 296, 482
Manifold, 433, 436, 480, 515, 558, 565
connected, 434
differentiable, 434
differential, 434
flat, 435
metric, 434
metrically connected, 435
non-flat, 435
non-metrically connected, 435
non-symmetric, 435
Riemannian, 436
symmetric, 435
Mass, 1
Mathematical induction, 88, 265
Matrix, 16, 90, 384
algebra, 17, 54, 55, 81, 120, 132
change-of-basis, 68, 69
defective, 159, 162
determinant of a, 19, 55, 60
diagonal, 143, 158
diagonalizable, 143

diagonalization, 143
exponential, 268
form, 55, 60, 65, 120, 140, 339
identity, 16, 17, 56
invertible, 68, 145
Jacobian, 441
multiplication, 16, 54
non-defective, 162
nonsingular, 161
non-symmetric, 69
notation, 16, 29, 49, 53, 54, 58, 81, 117, 185, 191, 447
orthogonal, 65, 68, 69, 148
proper orthogonal, 70
reflection, 166
representation, 70, 72, 179
rotation, 166
single-column, 17, 117, 118, 185
square, 28
symmetric, 118, 443
transformation, 68, 179
upper triangular, 162
Maxwell 2-form, 423, 430
Maxwell's equations, 422, 424–427, 429
Mean curvature, 433, 458, 459, 492, 514, 515, 535, 544
Meridian, 533, 534, 547
Metric coefficients, 177, 184, 318, 324, 330, 335, 434, 435
contravariant, 182, 199, 200, 318, 331, 385, 443
covariant, 182, 199, 204, 386, 442
mixed, 180
Metric compatibility property, 466, 468
Metric tensor, 188
co-contravariant, 189
contra-covariant, 189
contravariant, 183, 189, 228
covariant, 182, 189, 228
Metrinilic, 325, 593–595, 598–601
Meusnier, 519
Midsurface, 604, 605, 607, 611, 612
Minkowski spacetime, 375, 386, 426
Minor
antisymmetric, 483
radius, 438
skew-symmetric, 116
symmetric, 115
symmetries, 115, 119, 256, 257, 296, 307
Möbius strip, 374
Monge patch, 620
Monkey saddle, 640
Monomials, 44, 111, 265
Multibody dynamics, 67
Multiplicative plasticity, 230, 308
Multiplicity
algebraic, 135, 136, 159, 160
geometric, 135, 159, 160
Musical, 393
Mutually, 7, 11, 52, 69, 148, 181, 199, 445, 504

N

Natural parameter, 499, 501
Necessary and sufficient, 60, 61, 121, 248, 250, 252, 254, 545
Negative, 242
definite, 60, 66
semi-definite, 60
Newton's
identity, 137
law for gravitation, 360
method, 345
Newton-Raphson method, 295
Non
-Euclidean, 434
-collinear, 289
-defective, 131, 143, 144, 162, 163
-flat, 435
-homogeneous, 427
-metricity, 434, 435
-orientable, 375
-repeated, 14
-symmetric, 435
-umbilic, 539
-umbilical, 539–541, 543–545
Noncommutativity, 330, 332, 478, 481, 557, 569
Noncylindrical, 509, 510, 513
Nonlinear solid mechanics, 29, 56, 98, 230, 231, 290, 293, 564
Nonnegative, 44, 62, 111, 501
Nonsingular, 79
Norm, 7, 20, 50, 111
Normal, 436, 442, 456
curvature, 433, 494, 495, 515–519
curvature vector, 516
section, 519
section curve, 519, 520, 524
variation, 628
Normalization condition, 293
Notation, 12
absolute, 12
direct, 12
index, 13, 77, 127, 203, 215

indicial, 13, 42, 112, 220
matrix, 49, 53
subscript, 13
suffix, 13
symbolic, 12
tensorial, 12, 113
Voigt, 118, 119, 128, 129, 332
Numerical differentiation, 230, 310

O
Objective, 293
Operator(s), 53
bilinear, 21
Del, 332
derivative, 433, 583, 591
deviatoric, 103, 119
double-star, 383
flat, 393
Hessian, 344
Hodge star, 383, 386, 387, 419, 423, 424
Laplace, 342
linear, 232
Nabla, 332, 338, 353, 484
polarization, 280
sharp, 393
skew-symmetrizer, 103
spherical, 103, 119
star, 383
substitution, 14
super-symmetric deviatoric, 104
symmetrizer, 103
Orientation, 20, 64, 286, 374, 375, 403, 404, 441, 480, 492, 501, 505
-preserving, 401–403
-reversing, 404
Origin, 7, 11, 21, 64, 69, 169, 243, 346, 468, 541
Orthogonal, 52, 62, 81, 141, 523
directions, 544
group, 64, 235, 237
improper, 64, 72
parametrization, 523, 530, 535, 536, 554, 622, 639
projection(s), 65, 464
proper, 64, 69, 72, 507
subspaces, 66
transformation, 63, 64, 151, 236, 243, 293

P
Parabolic point, 494, 540
Parallel, 533, 534, 547
postulate, 434
surface, 642, 644
transport, 434, 462, 463, 467, 468, 478, 479, 632, 634
Parallelepiped, 22, 23, 55, 176
curvilinear, 366, 367
rectangular, 368
Parallelogram, 17, 64, 379, 410, 480, 481, 552, 571
Parametric, 436, 488–490
Parametrization, 499
arc length, 623
Clairaut, 530, 533
curve, 499
isothermal, 622
isothermic, 622
natural, 499, 500
orthogonal, 523, 622
surface, 436, 499
unit-speed, 500
Partial
derivative, 212, 215
differentiation, 331, 556, 656
time derivative, 585, 587, 594, 596, 602
Perturbation parameter, 230
Piola transformation, 296
Plane(s), 17, 25, 65, 83, 150, 175
curve, 495, 497, 505, 549
Frenet, 504
inclined, 409
normal, 504
osculating, 501, 504
rectifying, 504
sheet, 435
stress, 612
tangent, 456, 493
tangential, 441, 444, 453, 494
Poincaré's Lemma, 390, 392, 429
Point, 440, 442
elliptic, 494, 540
flat, 495, 540
function, 170, 436, 497
hyperbolic, 495, 540
parabolic, 494
planar, 495, 540
regular, 442
saddle, 495, 540
singular, 442, 498, 500
space, 10, 497
Poisson equation, 344
Poisson's ratio, 128, 612
Polar decomposition, 291, 293
Polyconvexity, 293

Polynomial representations, 250
Position, 315, 319
  current, 565
  increment vector, 196
  initial, 529
  original, 565
  vector, 197, 486, 492
Positive
  definite, 60–62, 66, 291, 386, 443
  -definiteness, 6, 62, 122
  semi-definite, 60, 62, 296
Positively oriented, 11, 20, 64, 370, 371, 414, 444, 451
Potential, 374
  electric, 424, 425, 429
  magnetic 1-form, 429
  magnetic vector, 424, 429
Powers, 44, 111, 140, 163, 264
Principal, 228
  axes, 131
  curvatures, 540, 541, 543, 545, 549, 640, 647
  directions, 131, 287, 288, 537, 539, 541–544, 647, 648
  normal surface, 510, 514, 645
  normal vector, 501, 503, 504, 509
  scalar invariants, 133, 137, 163, 216, 254, 255, 272, 294, 458
  stretches, 292, 301
  values, 131, 134, 137, 153, 243, 274, 287, 288, 305, 433
Principle of material objectivity, 293
Product, 344, 367, 376, 471
  box, 22
  cross, 17, 20, 21, 27, 83, 96, 184, 202, 206, 377, 393, 410, 447, 451
  direct, 38
  dot, 5, 42, 43, 45, 178, 180, 342, 376, 393, 434, 442, 447, 591
  dyadic, 38, 42, 96
  factor, 281
  inner, 5–7, 10, 36, 44, 62, 142, 149, 184, 384, 517
  mixed, 22
  rule, 223, 226, 321, 322, 413, 414, 462, 465–467, 470, 471, 571, 572, 575
  scalar, 5, 15, 44, 50, 184, 199, 202, 206, 478, 591
  scalar triple, 22, 394, 626
  tensor, 38, 40, 188, 281
  triadic, 94
  triple vector, 22, 25, 74
  vector, 17, 18, 20
  wedge, 377, 379, 380
Pull-back, 572
Push-forward, 567

Q
Quadratic approximation, 492, 543
Quantum mechanics, 131

R
Radius, 83, 502, 533
Reciprocal, 56, 121, 174, 200, 502
Redundant, 245–247, 281, 286, 290
Reflection, 64
Regular point, 442
Reissner-Mindlin, 611
Remarkable theorem, 549, 551
Replacement property, 14, 151, 173, 330, 376
Resultant
  force, 608
  moment, 608
  stress, 608
Ricci
  curvature tensor, 433, 557, 559, 561
  scalar, 433, 557, 559, 563, 654
Ricci's theorem, 324, 325
Riemannian, 332
  geometry, 332
  manifold, 436
  metric, 443
Right, 112
  -handed, 11, 12, 20, 21, 64, 176, 370, 499, 503, 505, 648, 649
  mapping, 36, 38, 44, 93, 98, 100, 101, 117, 132, 194, 206
  stretch tensor, 291
Rigid body motion, 507, 549, 550
Rivlin's identities, 233, 278, 280, 281, 286
Rotation, 62, 64, 69, 81, 339, 370, 533, 611
  axis, 82
  axis of, 533, 547
  counterclockwise, 81
  finite, 82
  infinitesimal, 74, 97, 611
  matrix, 166
  Rodrigues, 82
  tensor, 64, 70, 73, 88, 291, 308
Ruled surfaces, 497, 508, 509
Ruling, 508, 510, 512

S

Scalar, 36
  curvature, 557
  field, 485, 587
  multiplication, 4, 5, 9, 36, 37, 93, 150, 377
  projection, 8
  Ricci, 557
Second, 6
  Bianchi identity, 331
  -kind Ricci tensor, 559
  law of thermodynamics, 297
  -order covariant derivatives, 326, 473
  -order tensor, 41
Self-dual, 173, 179
Sesquilinear form, 10
Set(s), 5, 35, 48, 64, 93, 116, 132, 561
  complete, 158
  finite, 286
  generating, 237, 251, 255, 259, 284
  irreducible, 236
  linearly independent, 149
  orthonormal, 149
  spanning, 5
Shearing, 435, 604, 607
Shell, 604, 605, 607, 610–612
  shifter, 606
Signature, 384, 386
Signed area, 379
Similar matrices, 132
Singular, 56
  point, 442, 498, 500, 501, 505
Skew, 48
  -commutative property, 377
  -symmetric, 48, 103, 116, 339, 377, 380, 568
Slant, 531, 532
Small strain
  regime, 611
  tensor, 128
Solenoidal, 337
Space(s), 5, 40
  ambient, 435, 436, 447, 455, 493, 515, 585
  characteristic, 135, 156, 157
  complex vector, 9
  cotangent, 392
  curves, 433, 497, 500, 501, 507, 520
  flat, 328, 332
  inner product, 149
  metrically connected, 472
  null, 135
  point, 10, 497
  tangent, 436, 447
  translation, 10
  vector, 10, 20, 23, 67, 82, 99, 104
Spacetime, 386, 429, 561
Special theory of relativity, 386
Spectral, 254
  components, 142
  decomposition, 131, 140, 143, 144, 149, 162, 241, 253
  form, 140, 147, 149, 239, 242, 308
  representation, 140, 142, 145
Spectrum, 136
Speed of light, 386
Sphere, 332, 434, 435, 437, 494, 524, 534, 614, 631, 654
1-sphere, 561
2-sphere, 561
Stability analysis, 131
Stationary, 315, 585
Steady-state, 315
Stokes' theorem, 198, 365, 370–372, 408, 414, 417
Straightest possible path, 528, 529, 634
Strain measures, 291, 295, 309
Streamlines, 565
Stress tensor, 295, 305
  Biot, 294
  Cauchy, 295, 608
  first Piola-Kirchhoff, 294
  Kirchhoff, 295, 309
  material, 303
  second Piola-Kirchhoff, 158, 295, 305
  true, 158, 295
Stretching, 435, 511, 604, 607
Structural mechanics, 433, 604
Sufficiently smooth, 212, 315, 433, 436
Super-symmetric, 115, 116, 119, 225, 296
Surface, 196, 332
  binormal, 510, 645, 647
  Christoffel symbols, 456, 487, 578
  Christoffel symbols of the first kind, 460
  Christoffel symbols of the second kind, 456
  closed, 198
  contravariant derivative, 471, 472
  contravariant metric coefficients, 443, 487, 594
  contravariant metric tensor, 446
  coordinates, 436, 445, 452, 454, 466, 585, 588
  covariant basis vectors, 440, 446, 543, 557, 648

covariant curvature tensor, 456, 457, 491, 495
covariant differentiation, 433, 465, 471, 472, 477, 557
covariant metric coefficients, 442, 459, 487, 559
covariant metric tensor, 445
curves, 489, 497, 519, 521
developable, 497, 510, 511
divergence, 485
equiscalar, 333
funnel, 637
gradient, 484
implicit, 522, 529
integral, 367, 368, 370, 372
isothermic, 623
level, 333, 346
Liouville, 535
minimal, 459, 628, 629
mixed curvature tensor, 433, 458, 492, 496, 537, 538, 549, 602
normal vector, 441
open, 198, 370, 371
parallel, 642
permutation symbol, 444
principal normal, 510, 514
regular, 442, 494, 508, 515, 519, 523, 530, 536, 549
of revolution, 437, 438, 533, 547, 637
Riemann-Christoffel curvature tensor, 433, 473, 478, 480
tangent, 510
Sylvester's matrix theorem, 146
Symbol(s), 18, 173, 232, 318
alternating, 18, 28
ambient Christoffel, 456, 461
antisymmetric, 18
Christoffel, 317–321, 335, 352, 455, 459
Landau, 218
Landau order, 213, 230, 316
Levi–Civita, 18
permutation, 18, 19, 25, 96, 151, 173
summation, 14
surface Christoffel, 456, 487, 578
surface permutation, 444
total differential, 232
Symmetric, 47, 103, 115, 139, 149, 237, 435
Symmetry group, 236

T

Tangent, 12, 512
developable, 497, 510–512
line, 494, 502
modulus, 295
plane, 436, 452, 456, 464, 491, 493
space, 392, 433, 436, 447
surface, 510
vector, 479, 499, 501
Tearing, 511
Temperature, 1
Tensor identity
three fields, 278, 283
three variables, 280
two variables, 278
Tensor(s), 5
algebra and calculus, 3, 36, 140
alternating, 96
antisymmetric, 237
basis, 186
Cartesian, 40, 41
coaxial, 158
co-contravariant, 187
co-contravariant unit, 56
cofactor, 60
cofactor of a, 58
conductivity, 35
contra-covariant, 187
contra-covariant unit, 56
contravariant, 56, 187
covariant, 56, 187
covariant curvature, 613
curl of a, 341
curvature, 332, 435
defective, 144, 158
determinant of a, 55, 78, 202
deviatoric, 66
deviatoric projection, 103
diagonalizable, 143
eigenpairs of a, 140
eigenvalues of a, 141
eigenvectors of a, 141
Einstein, 557, 560
equal, 36
field, 315, 316, 320, 322, 327, 455, 471, 476
Finger, 2, 291, 309
first-kind Ricci, 558
first-order, 1, 316
fourth-order, 2, 101, 108
fourth-order unit, 99, 104, 105
fourth-order zero, 99
function, 44, 209, 223, 233, 237, 256
Hessian of a, 345
of higher orders, 326, 328
identity, 36, 55

inverse of a, 174, 226
invertible, 56
Laplacian of a, 343
left Cauchy-Green strain, 158, 291
Lie derivative of a, 577
metric, 145
mixed curvature, 613
non-defective, 143
non-metricity, 435
nonsingular, 56
of order four, 98
of order three, 94
of order two, 93
orthogonal, 62, 63, 88
permutation, 96
perturbed, 232
polynomial, 44
positive semi-definite, 296
powers, 44
principal invariants of a, 151
product, 38, 40, 188, 281
projection, 104, 124
proper orthogonal, 62
property, 182, 322, 453, 454, 474, 585, 587, 592
rate of deformation, 295
rectangular, 40
referential, 291, 294
Ricci, 557
Ricci curvature, 433, 557
Riemann-Christoffel, 329, 331, 332
Riemann-Christoffel curvature, 329, 332
right Cauchy-Green strain, 158, 291
rotation, 62, 64, 70, 73, 88, 291, 308
second-kind Ricci, 559
second-order, 2, 35, 41, 47, 50, 53, 55
shift, 446, 447, 450, 453, 587
singular, 446
skew, 48, 53, 54
skew-symmetric, 61, 78
skew-symmetric projection, 103
spatial unit, 56
spectral decomposition of a, 143
spectral representation of a, 142
spherical, 66
spherical projection, 103, 126
stress, 127, 612
super-symmetric, 115, 116, 118, 122, 127, 296
super-symmetric deviatoric projection, 104
super-symmetric identity, 119
super-symmetric projection, 103, 116
surface, 453–455, 471, 593, 597
surface covariant curvature, 456
surface mixed curvature, 433, 492, 496, 537, 549
surface Riemann-Christoffel curvature, 433
symmetric, 54, 61, 66, 227
third-order, 2, 93, 95
third-order zero, 94
torsion, 435, 469
trace of a, 52, 56, 164
traceless, 560
transformation law(s), 182, 320, 445, 587
transposed, 53, 81
two-point, 291
unit, 36, 41
zero, 36, 73, 259, 324
zeroth-order, 1
Theorema Egregium, 549, 551
Theory, 290
of algebraic invariants, 233
control, 564
electromagnetic field, 35, 365
of elasticity, 127, 128, 290, 604
infinitesimal strain, 128
of manifolds, 515
of surfaces, 550
Time, 259, 315, 386, 429, 499, 562, 585, 586, 589, 590, 594
-dependent, 587, 588
-independent, 587
-like, 433, 585, 586, 590, 625, 626, 633
rate of change, 297, 305, 593, 594
Torsion-free, 469
Torus, 403, 404, 438, 537, 615
Trace, 206, 266, 497, 644
Traceless, 560
Trajectories, 565
Transformation law(s), 67, 182, 195
tensorial, 72
vectorial, 71
Transport problem of radiation mechanics, 360
Transpose, 16, 46, 47, 54, 88, 113, 189, 206, 211
Transposition, 46, 80, 113
Transverse
normal, 612
shear strains, 612
Trihedron, 550, 649
Darboux, 648, 650
Frenet, 504, 521
moving, 504, 506, 521

Trilinear, 94
Twisted cubic, 627

U
Umbilic, 538
Umbilical, 538–540
Uniform, 315
Upper-left submatrices, 61, 443

V
Vandermonde, 240, 275
Variational principles, 525, 628
Vector(s), 5
  acceleration, 501, 502, 524, 562, 634
  addition, 1, 4, 5, 9, 13, 15
  area, 18, 49, 58
  axial, 48, 78
  binormal, 503, 505, 509
  Cauchy traction, 608
  complex, 9
  contravariant, 183, 350
  covariant, 183, 200, 215, 348
  curvature, 516
  Darboux, 507, 623, 648, 650
  differential, 332
  displacement, 565
  equal, 7
  field, 315, 321, 333, 335, 433, 462
  isotropic, 9
  killing, 574
  length of a, 6
  magnitude of a, 6
  position, 484, 486, 492
  principal normal, 501, 503, 509
  rejection, 8
  space, 3, 5, 6, 9, 35–37, 40, 104, 134, 149, 375, 377, 384, 503
  subtraction, 4
  surface, 198, 452–455, 465, 473, 479, 485, 593
  tangent, 479, 499, 501
  and tensor algebra, 14, 206
  and tensor analysis, 209, 315
  and tensor calculus, 14, 209, 229, 315
  and tensor field, 326
  and tensor identities, 228
  transformation law(s), 68, 567
  unit, 7, 8, 65, 74, 82, 316, 346, 347, 370, 436, 501, 519, 525
  velocity, 499, 524, 561, 586
Velocity, 315
  ambient, 586, 588
Viscoelasticity, 290
Volume form, 198

W
Work conjugate, 295
  pair, 295

Y
Young's modulus, 128, 612