IOP Concise Physics | A Morgan & Claypool Publication

Theories of Matter, Space and Time

Classical theories

N Evans

S F King

IOP | ebooks

Theories of Matter, Space and Time

Classical theories

N Evans
University of Southampton

S F King
University of Southampton

Morgan & Claypool Publishers

ISBN 978-1-6817-4696-8 (ebook)
ISBN 978-1-6817-4697-5 (print)
ISBN 978-1-6817-4698-2 (mobi)

DOI 10.1088/978-1-6817-4696-8

Version: 20171201

IOP Concise Physics
ISSN 2053-2571 (online)
ISSN 2054-7307 (print)

A Morgan & Claypool publication as part of IOP Concise Physics
Published by Morgan & Claypool Publishers, 1210 Fifth Avenue, Suite 250, San Rafael, CA, 94901, USA

IOP Publishing, Temple Circus, Temple Way, Bristol BS1 6HG, UK

*Steve King dedicates this to his wife Margaret and to the memory of his Mother,
who always wanted him to write a book.*

*Nick Evans would like to take this opportunity to thank the teachers that inspired him
on his path: Barry Evans, Nigel Wood, Nick Tumber, John Johnstone, Alan Martin
and Chris Maxwell.*

Contents

Preface

This book and its sequel (*Theories of Matter, Space and Time: Quantum Theories*) grew out of courses that we have both taught as part of the third and fourth year of the undergraduate degree programme in Physics at Southampton University, UK. Our goal was to guide the full MPhys undergraduate cohort through some of the trickier areas of theoretical physics that we expect our undergraduates to master. In particular the aim is to move beyond the initial courses in classical mechanics, special relativity, electromagnetism and quantum theory to more sophisticated views of these subjects and their interdependence. Our approach is to keep the analysis as concise and physical as possible whilst revealing the key elegance in each subject we discuss.

In this book we first introduce the key areas of the principle of least action, an alternative treatment of Newtonian dynamics, that provides new understanding of conservation laws. In particular here we show how the formalism evolved from Fermat's principle of least time in optics. In the second chapter we introduce special relativity leading quickly to the need and form of four-vectors. We develop four-vectors for all kinematic variables and generalize Newton's second law to the relativistic environment. We return to the principle of least action for a free relativistic particle. The third chapter presents a review of the integral and differential forms of Maxwell's equations before massaging them to four-vector form so that the Lorentz boost properties of electric and magnetic fields are transparent. Again we return to the action principle to formulate minimal substitution for an electrically charged particle.

The second book of this pair will move the ideas developed here to the arena of quantum mechanics. We will present the relativistic wave equations of Klein, Gordon and Dirac, generalize the least action principle to Feynman's path integral methods, and include electromagnetism to make contact with quantum electrodynamics at a first quantized level. Between these volumes we hope to move a student's understanding from their first courses to a place where they are ready, beyond, to embark on graduate level courses on quantum field theory.

Acknowledgments

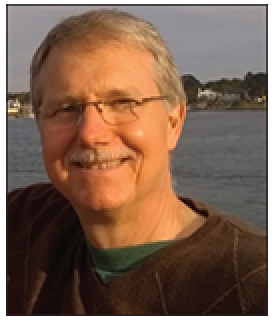
The authors are grateful to Professor Tim Morris and Dr Beatriz de Carlos in Southampton who have both taught this material and whose helpful suggestions along the way have hopefully improved this volume.

Author biographies

Professor Nick Evans

Nick completed his PhD in collider phenomenology in 1993 at Southampton University. He performed his early research work at Yale and Boston Universities in the US before returning to Southampton in 1999 on a UK government 5 year fellowship. His work centred on strongly interacting particle systems, including composite Higgs models, and he played a large role in applying string theory to study the strong nuclear force and the mechanism of mass generation. Much of his work centres on the structure of the vacuum so in a sense he works on nothing. He is now a Professor at Southampton University and the Director of the Faculty of Physical Science and Engineering Graduate School. Nick's outreach work includes the on-line physics with murder, mystery, thriller ‘The Newtonian Legacy’ which you can read for free online at: http://www.southampton.ac.uk/~evans/NL/.

Professor Steve King



Steve completed his PhD in QCD perturbation theory in 1980 at Manchester University. He was a postdoctoral fellow at Oxford University, where he worked on composite models, before moving to Harvard and Boston Universities in the US, where he worked on technicolour and collider phenomenology. Returning to Southampton in 1987, he won a 5 year fellowship to work on lattice QCD and top quark condensates. Soon after becoming a Lecturer, he turned his attention to supersymmetry, cosmology, strings, unification, flavour symmetry models and neutrinos. He is now Professor and First Year Director of Studies in Physics and Astronomy at Southampton. For more details see the recent on-line interview: https://jphysplus.iop.org/2015/12/01/an-interview-with-stephen-king-physicist/

IOP Concise Physics

Theories of Matter, Space and Time

Classical theories

N Evans and S F King

Chapter 1

Least action

Newton, through his three laws of dynamics, developed an extremely successful description of the motion of objects. These laws can, for example, describe the elliptical orbits of planets to remarkable precision. There is though an alternative presentation of these successes, the principle of least action, which we will explore here. It is a formalism that grew out of optics and will allow us to study an area of mathematics called 'calculus of variation'. Of course it must turn out to be the same as Newton's laws. This alternative formalism makes some dynamics problems easier to solve but, more importantly, it will give us new insights into conservation laws. It is important to master these methods since as one moves to the forefront of modern quantum theories the least action principle becomes the only way to define theories such as that of the strong nuclear force.

1.1 Optics

Our starting point will be to think about the path that light travels by. In these enlightened times we might start from Maxwell's equations and derive a wave equation with light waves as solutions to determine how the light propagates. Before this technology though Fermat proposed:

Fermat's principle of least time: *light propagates between two points so as to minimize its travel time.*

Thus for example in a uniform medium where the speed of light c is a constant the minimum time of travel

$$t = \frac{d}{c} \tag{1.1}$$

is given by the path of shortest distance d, i.e. a straight line. This is still a perfectly good (if limited) description of light.

We can obtain more interesting results by thinking about media where the speed of light changes.

1.1.1 Snell's law

Consider two neighbouring regions of space in which light travels at different speeds v_1, v_2 —for example a glass–air interface. We will be interested in the light that travels from the point (x_1, y_1) in the first medium to the point (x_2, y_2) in the second (figure 1.1).

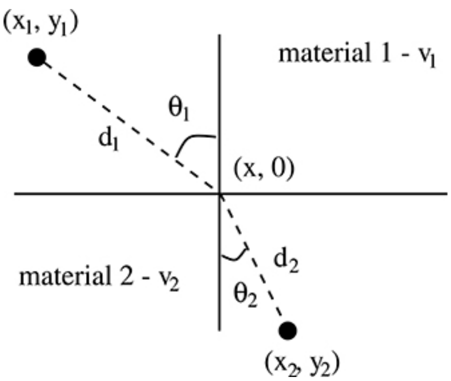


Figure 1.1. Possible paths that light might follow transiting across the interface between two materials.

In any one medium light travels in a straight line but in this case we have some choice in where the light crosses between the media. Let's consider the arbitrary crossing point $(x, y = 0)$. The time of travel is

$$\begin{aligned} T[x] &= \tfrac{d_1}{v_1} + \tfrac{d_2}{v_2} \\ &= \tfrac{\sqrt{(x-x_1)^2 + y_1^2}}{v_1} + \tfrac{\sqrt{(x-x_2)^2 + y_2^2}}{v_2} \end{aligned} \tag{1.2}$$

We now want to find the path (i.e. the value of *x* through which it passes) which minimizes the time taken. Thus

$$\frac{dT}{dx} = \frac{(x - x_1)}{v_1\sqrt{(x-x_1)^2 + y_1^2}} - \frac{(x_2 - x)}{v_2\sqrt{(x_2-x)^2 + y_2^2}} = 0 \tag{1.3}$$

This equation though is just

$$\boxed{\frac{\sin\,\theta_1}{v_1} = \frac{\sin\,\theta_2}{v_2}} \tag{1.4}$$

which is *Snell's law*.

In terms of index of refraction which is defined, relative to the vacuum, as

$$n_1 = \frac{c}{v_1} \tag{1.5}$$

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \tag{1.6}$$

1.1.2 Complicated problems

We can imagine more complicated problems than that above where the index of refraction is an arbitrary function of position. For example consider light moving in a plane where the speed of the light is $v(x, y)$ (figure 1.2).

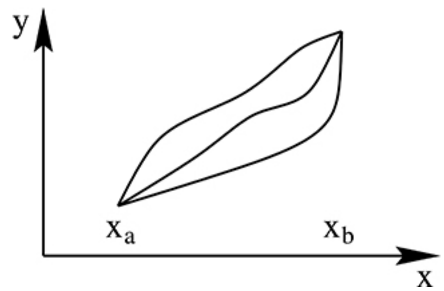


Figure 1.2. Possible paths that light might travel in a plane with varying speed of light $v(x, y)$.

Different paths are described by different functions $y(x)$. The time to travel along an arbitrary little piece of path is

$$dT = \frac{\text{distance}}{\text{velocity}} = \frac{\sqrt{dx^2 + dy^2}}{v(x, y)} \tag{1.7}$$

Summing such contributions up along a path gives the total time of travel

$$T[y(x)] = \int_{x_a}^{x_b} \frac{1}{v(x, y)} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \; dx \tag{1.8}$$

To rewrite this in a more standard form we have found that the time taken to traverse a path is

$$T\left[y\right] = \int_{x_a}^{x_b} L\left(y, \dot{y}, x\right) \; dx \tag{1.9}$$

where

$$L\left(y, \dot{y}, x\right) = \frac{1}{v(x, y)} \sqrt{1 + \dot{y}^2}$$

Now we want to find the path $y(x)$ that gives the minimum time. (1.10)

Exercise 1.1: To remind yourself about partial differentiation, define a function by

$$T = a(t)\ \ b\,(t)^3 \dot{b}\,(t) t^{10}$$

where the dot indicates a derivative with respect to *t*. Give expressions for

$$\frac{\partial T}{\partial t}, \quad \frac{\partial T}{\partial b}, \quad \frac{\partial T}{\partial \dot{b}}, \quad \frac{dT}{dt}$$

This is the sort of problem that calculus of variation is designed to address, as discussed in appendix A.

As we show in appendix A, the problem of finding the path that minimizes the time, is equivalent to solving a differential equation called the Euler–Lagrange equation,

$$\frac{d}{dx}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \tag{1.11}$$

which corresponds to equation (A.9) with *q* identified as *y* and *s* identified as *x*.

Let's look at a couple of examples.

1.1.3 Light in vacuum

In vacuum the speed of light is a constant so $v(x,y) = c$.

The Euler–Lagrange equation is

$$\frac{d}{dx}\left(\frac{\dot{y}}{\sqrt{1+\dot{y}^2}}\right) = 0 \tag{1.12}$$

Integrating this gives

$$\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = \text{constant} \tag{1.13}$$

The only solution of this is

$$\dot{y} = \text{constant}, m \tag{1.14}$$

or integrating

$$y = mx + c \tag{1.15}$$

i.e. a straight line. This is our first example of the solution of the Euler–Lagrange equation giving the path that minimizes *T*. *m* and *c* are determined by the initial and final position of the light.

1.1.4 Light in the atmosphere

In the atmosphere the air temperature and density change with height resulting in the speed of light depending on height—$v(h)$. Equivalently we can write the refractive index $n(h)$ with

$$v\left(h\right) = \frac{c}{n\left(h\right)} \tag{1.16}$$

Our result for the length of time light takes to travel some path $h(x)$ can be written as an *optical path length*

$$cT\left[h\right] = \int_{x_1}^{x_2} dx L, \quad L = n\left(h\right)\sqrt{1+\dot{h}^2} \tag{1.17}$$

We can use the fact that *L* is independent of *x* to simplify the Euler–Lagrange equation as follows. Note that

$$\frac{dL}{dx} = \frac{\partial L}{\partial x} + \frac{\partial L}{\partial h}\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h} \tag{1.18}$$

The first term on the right is zero. Now replace $\frac{\partial L}{\partial h}$ using the Euler–Lagrange equation

$$\frac{\partial L}{\partial h} = \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right) \tag{1.19}$$

and we find

$$\frac{dL}{dx} = \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right)\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h} \tag{1.20}$$

which is just

$$\frac{d}{dx}\left[L - \dot{h}\frac{\partial L}{\partial \dot{h}}\right] = 0 \tag{1.21}$$

which gives us

$$L - \dot{h}\frac{\partial L}{\partial \dot{h}} = \text{constant}, D \tag{1.22}$$

Note that this is only a first order equation rather than the second order Euler–Lagrange equation so is simpler to solve.

In our problem, using the explicit form for *L* above we have

$$n\sqrt{1+\dot{h}^2} - \frac{\dot{h}^2 n}{\sqrt{1+\dot{h}^2}} = D \tag{1.23}$$

which simplifies to

$$\frac{n}{\sqrt{1+\dot{h}^2}} = D \tag{1.24}$$

Note that the physical meaning of *D* is the value of the index of refraction at the point where the light ray becomes horizontal so that $\dot{h} = 0$.

Squaring and rearranging we find

$$\frac{dh}{dx} = \sqrt{\frac{n^2}{D^2} - 1} \tag{1.25}$$

Thus

$$\boxed{x - x_0 = \int_{h_0}^{h} \frac{dh}{\sqrt{\frac{n^2}{D^2} - 1}}} \tag{1.26}$$

Explicit example: Consider a ray of light that begins moving horizontally ($\dot{h} = 0$) at *h* = 0 in an atmosphere where

$$n\left(h\right) = n_0 - \lambda h \tag{1.27}$$

where *λ* is some constant. We must solve the integral

$$x = \int \frac{dh}{\sqrt{\frac{\left(n_0 - \lambda h\right)^2}{D^2} - 1}}$$

This can be done by changing variables to (1.28)

$$n_0 - \lambda h = D \cosh \phi$$

The integral becomes (1.29)

$$x = -\int \frac{D}{\lambda} d\phi = -\frac{D}{\lambda}\phi + c$$

Returning to the original coordinates and requiring the boundary conditions $\dot{h}(x=0)=0$ and $h(x=0)=0$ gives the result (1.30)

$$\boxed{h = \frac{n_0}{\lambda}\left(1 - \cosh \frac{\lambda x}{n_0}\right)}$$ (1.31)

- When λ is positive $n(h)$ decreases with altitude—this is what normally happens in the atmosphere. The plot of the form of the solution is given in figure 1.3.
  Thus if we look up at the Empire State building it will appear taller than it actually is.
- If there is a temperature inversion then λ is negative so $n(h)$ increases with altitude. When we plot the form of the solution we get figure 1.4.
  We see 'the sky on the ground'—a mirage.

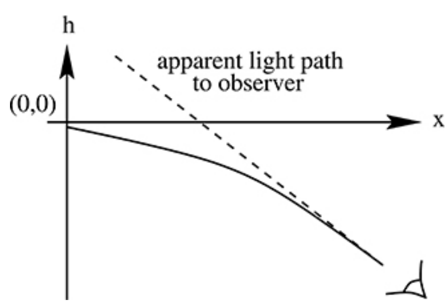


Figure 1.3. The solution for the light path $h(x)$ when $n(h)$ decreases with height.

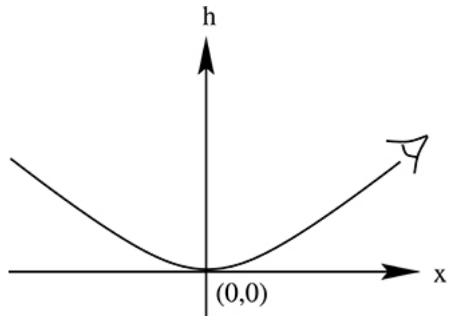


Figure 1.4. The solution for the light path $h(x)$ when $n(h)$ increases with height.

Exercise 1.2:

(a) Consider a fibre optic cable lying in the *z* direction. The cable is made of glass with index of refraction $n(r)$, where *r* is the radial distance from the centre of the cable. Working in cylindrical coordinates (r, θ, z) show that Fermat's principle implies that light travels on the path minimizing the quantity

$$\int_{z_1}^{z_2} f(r(z), \theta(z), r\prime(z), \theta\prime(z))dz = \int_{z_1}^{z_2} n(r)\sqrt{r'^2 + r^2\theta'^2 + 1}dz.$$

where a prime indicates differentiation with respect to *z*. z_1 and z_2 are the *z*-coordinates of the end points of the path.

(b) If a light ray initially has $\theta\prime = 0$ show, from the appropriate Euler–Lagrange equation, that the θ independence of *f* implies the path followed by the light is described by a constant value of θ.

(c) Use the *z* independence of *f* to deduce that the first order differential equation for rays travelling paths with constant θ is

$$f - \frac{\partial f}{\partial r\prime}\ r\prime = \text{constant}$$

1.2 Newtonian dynamics

We have seen that the motion of light can be described by a ‘principle of least time’. Is there an equivalent rule that would describe the motion of a particle in Newtonian dynamics? There is and it is enshrined as

Hamilton's principle: *A particle travels by the path between two points that minimizes the action*.

We need to know what the ‘action’ is. Let's write it first for one dimensional motion. The action is

$$\boxed{S[\text{path}] = \int_{t_a}^{t_b} L(x, \dot{x}, t)dt}$$

(1.32)

where the dot indicates differentiation with respect to the time, t. L is known as the *Lagrangian* and is given by

$$\boxed{L = \text{kinetic energy} - \text{potential energy} = T - V} \tag{1.33}$$

From appendix A, we know that the path that minimizes the action satisfies the Euler–Lagrange equation, analogous to the case of optics equation (1.11),

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0} \tag{1.34}$$

which corresponds to equation (A.11) with q identified as x and s identified as t.

We can now check to see if any of this makes sense (!). For a non-relativistic particle in a one dimensional potential we have

$$L = T - V = \frac{1}{2}m\dot{x}^2 - V(x) \tag{1.35}$$

The Euler–Lagrange equation is therefore

$$\frac{d}{dt}(m\dot{x}) + \frac{\partial V}{\partial x} = 0 \tag{1.36}$$

which is Newton's second law since

$$F = -\frac{\partial V}{\partial x} \tag{1.37}$$

Note that the momentum of the particle is given by

$$p = m\dot{x} = \frac{\partial L}{\partial \dot{x}} \tag{1.38}$$

1.2.1 Multiple coordinates

Suppose that we now have several coordinates

$$q_i \quad i = 1, \ldots, n \tag{1.39}$$

For example for one particle moving in three dimensions we might call $x = q_1$, $y = q_2$, $z = q_3$.

As discussed in appendix A, equation (A.11), for the n dimensional case we have to solve a set of n Euler–Lagrange equations—one associated with each coordinate,

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0} \tag{1.40}$$

i.e. we need to write down *n* copies of the Euler–Lagrange equation, for $i = 1, 2, \ldots, n$ and try to solve them simultaneously.

Generalized coordinates

The reason that we have written the coordinates so generally as q_i rather than for example using x, y, z is that in some problems these are not the appropriate coordinates because of a *constraint*. A simple example to illustrate this is a ball on a wire hoop (figure 1.5).

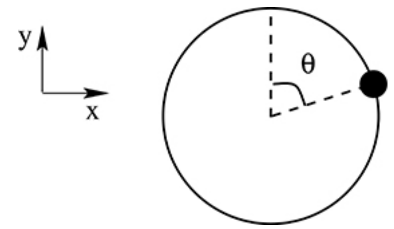


Figure 1.5. The coordinates describing a ball constrained to run around a hoop.

The hoop stops the ball moving in the radial direction so the ball cannot be at any arbitrary (x, y). The sensible coordinate to use is the angle θ.

Such a reduced set of coordinates is called *generalized coordinates*.

Generalized momentum

A generalization of the idea of momentum can be defined in the spirit of equation (1.38). The *generalized momentum* associated with a generalized coordinate is given by

$$\boxed{p_i = \frac{\partial L}{\partial \dot{q}_i}} \tag{1.41}$$

1.2.2 Example: projectile motion

Consider the familiar problem of a projectile in a uniform gravitational field (figure 1.6).

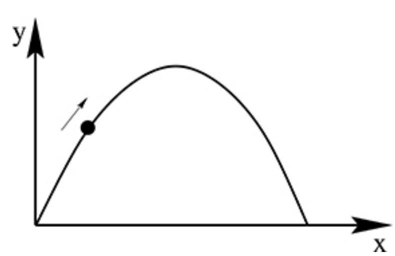


Figure 1.6. The motion of a projectile in the x, y plane subject to constant gravity in the vertical direction *y*.

We can obtain the normal Newtonian equations of motion from the Euler–Lagrange equations. We need expressions for the kinetic and potential energy of the system so we can build the Lagrangian. The kinetic energy is just

$$T = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 \tag{1.42}$$

and the potential energy

$$V = mgy \tag{1.43}$$

So the Lagrangian is just

$$L = T - V = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 - mgy \tag{1.44}$$

Now we find the two Euler–Lagrange equations. The first associated with the *x* coordinate is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \tag{1.45}$$

which gives

$$\boxed{m\ddot{x} = 0} \tag{1.46}$$

The second equation associated with the *y* coordinate is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \tag{1.47}$$

which gives

$$\boxed{\ddot{y} = -g} \tag{1.48}$$

The two boxed equations are the standard Newtonian equations of motion.

Hopefully you’re starting to see the power of this technique now—the kinetic and potential energies of a system are fairly easy to work out and then we just do some maths. There's not all that resolving forces business! The next problem is an example that would be very hard by the standard methodology.

1.2.3 Example 2: double pendulum

Consider a double pendulum as shown in figure 1.7.

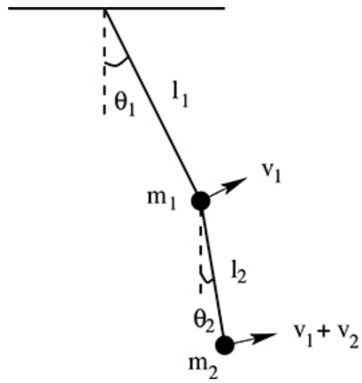


Figure 1.7. The coordinates relevant for a double pendulum.

It would be pretty hard work to determine all the forces in play here. However, the Lagrangian technique means we only have to calculate the energies of the two masses to get to the equations of motion.

The first mass has a velocity $\vec{v}_1$ with magnitude $l_1\dot{\theta}_1$ ($v = \omega r$). The second mass has both this motion plus a second contribution from the swing of the second pendulum $\bar{v}_2$ with magnitude $l_2\dot{\theta}_2$. The total velocity of the second mass is therefore

$$\overrightarrow{v}_{\mathrm{tot}} = \overrightarrow{v}_1 + \overrightarrow{v}_2 \tag{1.49}$$

so

$$\begin{aligned} v_{\mathrm{tot}}^2 &= (\vec{v}_1 + \vec{v}_2) \,.\, (\vec{v}_1 + \vec{v}_2) \\ &= (l_1\,\dot{\theta}_1)^2 + (l_2\,\dot{\theta}_2)^2 + 2l_1\,\dot{\theta}_1\,l_2\,\dot{\theta}_2\,\cos(\theta_2 - \theta_1) \end{aligned} \tag{1.50}$$

where $\theta_2 - \theta_1$ is the angle between $\overrightarrow{v}_1$ and $\overrightarrow{v}_2$.

Thus the total kinetic energy of the system is

$$T = \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 + \frac{1}{2}m_2\Big[(l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + 2l_1\dot{\theta}_1 l_2\dot{\theta}_2\,\cos(\theta_2 - \theta_1)\Big] \tag{1.51}$$

The potential energy is determined by the heights of the masses

$$V = -m_1 g l_1\,\cos\,\theta_1 - m_2 g(l_1\,\cos\,\theta_1 + l_2\,\cos\,\theta_2) \tag{1.52}$$

and the Lagrangian is

$$L = T - V \tag{1.53}$$

There are two Euler lagrange equations—one associated with θ_1

$$\frac{d}{dt}\left[m_1 l_1^2\dot{\theta}_1 + m_2 l_1^2\dot{\theta}_1 + m_2 l_1 l_2 \dot{\theta}_2\,\cos(\theta_2-\theta_1)\right] - m_2 l_1 l_2 \dot{\theta}_1\dot{\theta}_2\,\sin(\theta_2-\theta_1) + (m_1+m_2) g l_1\,\sin\,\theta_1 = 0 \tag{1.54}$$

and one with θ_2

$$\frac{d}{dt}\left[m_2 l_2^2\dot{\theta}_2 + m_2 l_1 l_2 \dot{\theta}_1\,\cos(\theta_2-\theta_1)\right] + m_2 l_1 l_2 \dot{\theta}_1\dot{\theta}_2\,\sin(\theta_2-\theta_1) + m_2 g l_2\,\sin\,\theta_2 = 0 \tag{1.55}$$

These are pretty messy (but that was the point!). Things simplify a bit if we assume that both θ_1 and θ_2 are small and expand to linear order. We then get

$$\begin{aligned}\left(m_1+m_2\right) l_1^2\ddot{\theta}_1 + m_2 l_1 l_2\ddot{\theta}_2 &= -(m_1+m_2) g l_1\theta_1\\ m_2 l_2^2\ddot{\theta}_2 + m_2 l_1 l_2\ddot{\theta}_1 &= -m_2 g l_2\theta_2\end{aligned} \tag{1.56}$$

These coupled equations in fact have normal mode solutions of the form

$$\begin{aligned}\ddot{\theta}_1 &= -\omega^2\theta_1\\ \ddot{\theta}_2 &= -\omega^2\theta_2\end{aligned} \tag{1.57}$$

i.e. the two pendulums oscillate with the same frequency.

To find ω you can try substituting in the form of the solution in equations (1.57) into (1.56). You'll find two simultaneous equations for θ_1 and θ_2 with two solutions. You'll find in one case θ_1/θ_2 is positive and in the other it is negative. So in one case the pendulums swing together and in the other case in opposite directions.

Exercise 1.3: If a system with generalized coordinate q has the Lagrangian

$$L = \frac{1}{2}\dot{q}^2 - q^3$$

what is the Euler–Lagrange equation describing the system?

Exercise 1.4: If a system with generalized coordinates ξ and ψ has the Lagrangian

$$L = \frac{1}{2}\dot{\xi}^2 + \cos\,\xi\dot{\psi} - \xi e^{\psi}$$

what are the Euler–Lagrange equations describing the system?

Exercise 1.5: Two blocks of equal mass *M* are connected by a flexible string of length ℓ. One block is placed on a smooth horizontal table and the other block hangs over the edge. Using the length *z* of string hanging over the edge as a generalized coordinate, write down the Lagrangian and use the Euler–Lagrange equation to find the acceleration of the hanging mass in the following cases:

(i) The mass of the string is negligible.

(ii) The string is heavy with mass *m* distributed uniformly along it.

Exercise 1.6:

(a) Show that for a non-relativistic, free particle of mass *m* travelling with constant velocity *v* the action *S* describing its motion reduces to

$$S = mvd/2$$

where *d* is the distance travelled. This was a form for the action proposed by Maupertuis who believed it reflected the simplicity and economy of the Creator-God ....

(b) Consider such a particle rolling on a table in the x, y plane with speed v_1. Along the *y*-axis there is a height discontinuity in the table which the particle can move over at the cost of potential energy which reduces its velocity to v_2. If the particle starts at (x_1, y_1) to the left of the *y*-axis and ends to the right at (x_2, y_2) show that the action for it passing across the *y*-axis at arbitrary *y* (assuming it travels in a straight line except when it crosses the *y*-axis) is given by

$$S = mv_1\sqrt{x_1^2 + (y - y_1)^2} + mv_2\sqrt{x_2^2 + (y - y_2)^2}$$

By minimizing the action deduce the relation

$$v_1 \sin\theta_1 = v_2 \sin\theta_2$$

where the angles are the angles between the particle's direction of motion and the *x*-axis before and after it crosses the *y*-axis. Contrast this result with Snell's law for light.

1.3 Conservation laws

Finally let's look at one of the most surprising pieces of insight to come out of the Lagrangian formalism—that is a deeper understanding of conservation laws. The mathematics of this is discussed in more detail in appendix B.

1.3.1 Ignorable coordinates

If the Lagrangian does not depend on some coordinate q_i it is called an *ignorable coordinate*. Then $\partial L/\partial q_i = 0$ and its associated generalized momentum is conserved as we can see from the Euler–Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = \frac{dp_i}{dt} = 0 \tag{1.58}$$

so

$$p_i = \text{constant} \tag{1.59}$$

This is clearly a mathematical fact but there is a deeper interpretation. If L only depends on $\dot{q}_i$ not q_i itself then we can shift

$$q_i \to q_i + \text{const} \tag{1.60}$$

and leave the Lagrangian, L, (and hence the physics) invariant. This is a *symmetry*—translation invariance in the q_i direction.

Thus we learn that the true relation is

$$\boxed{\text{symmetry} \to \text{conserved momentum}}$$

This is a new insight we have not seen before in Newtonian mechanics.

1.3.2 Energy conservation

Consider the case that L does not depend explicitly on t. This implies that a quantity known as the *Hamiltonian* is conserved. The *Hamiltonian* is defined as

$$H = \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L \tag{1.61}$$

To prove that it is conserved we explicitly calculate

$$\frac{dH}{dt} = \sum_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i - \frac{\partial L}{\partial t} - \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i - \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i = 0 \tag{1.62}$$

using the Euler–Lagrange equations and $\frac{\partial L}{\partial t} = 0$.

In simple systems the Hamiltonian is just the total energy of the system as we can see for example in one dimension where

$$L = \frac{1}{2}m\dot{x}^2 - V(x) \tag{1.63}$$

so using the definition above

$$H = \frac{1}{2}m\dot{x}^2 + V(x)$$ (1.64)

In conclusion here we have learnt that time translation invariance implies energy conservation.

1.3.3 Example: central forces

Consider a particle moving subject to a central force i.e. in a potential $V(r)$ (figure 1.8).

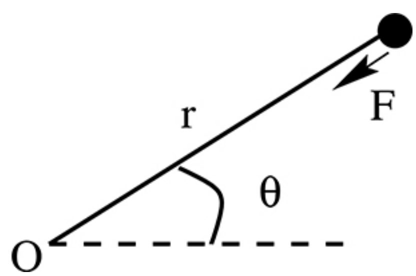


Figure 1.8. The coordinates relevant for a particle moving in a central potential.

The kinetic energy of the particle is

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2$$ (1.65)

thus

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - V(r)$$ (1.66)

There is an Euler–Lagrange equation associated with the *r* coordinate

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0$$ (1.67)

giving

$$m\ddot{r} = mr\dot{\theta}^2 - \frac{\partial V}{\partial r}$$ (1.68)

Plus a second equation for θ, which since L is independent of θ, is just

$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) = 0 \qquad (1.69)$$

which tells us that angular momentum is conserved.

The Hamiltonian is also conserved and is given here by

$$H = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + V \qquad (1.70)$$

which is the total energy.

Exercise 1.7: If a system with generalized coordinates *x* and *y* has the action

$$S = \int\left(\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 + \cos\ \dot{y} - x\right)dt$$

what quantities are conserved?

1.3.4 Hamiltonian and energy

Finally it is worth stressing that the Hamiltonian is not always the energy of the system. As an example consider a bead on a hoop that is being rotated at a fixed angular velocity ω, as shown in figure 1.9.

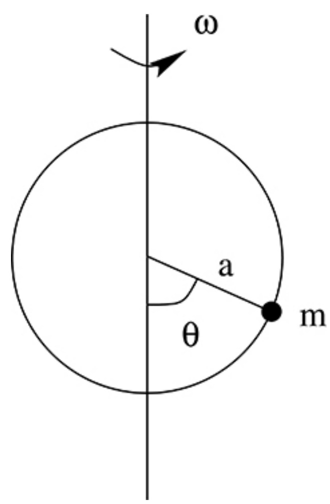


Figure 1.9. A bead on a hoop which is rotating at angular speed ω.

To be explicit, the hoop is in a vertical plane near the surface of the Earth, where that vertical plane is subject to a steady rotation about a fixed axis passing through the centre of the hoop, due to an external turning force or torque. The fact that the energy is not conserved is due to the fact that the turning force, which is required to maintain the steady rate of rotation, is external to the system. However, we shall show that, even in this case, the Hamiltonian is conserved, even though the Hamiltonian cannot be identified with the energy.

The single coordinate θ (which is a function of time t) as shown in the diagram is sufficient to describe the position of the bead so this is a good generalized coordinate. The kinetic energy is given by

$$T = \frac{1}{2}m(a^2\dot{\theta}^2 + a^2\sin^2\theta\omega^2) \tag{1.71}$$

and the potential energy by

$$V = -mga\ \cos\ \theta \tag{1.72}$$

Thus

$$L = \frac{1}{2}m(a^2\dot{\theta}^2 + a^2\sin^2\theta\omega^2) + mga\ \cos\ \theta \tag{1.73}$$

Since L does not depend on t the Hamiltonian is conserved. In particular

$$H = \frac{1}{2}ma^2\dot{\theta}^2 - \frac{1}{2}ma^2\sin^2\theta\omega^2 - mga\ \cos\ \theta \tag{1.74}$$

Although H is conserved the total energy of the system is not since to keep the hoop rotating a constant external torque must be applied, thereby doing work on the system.

Appendix A. Calculus of variation

In this appendix we derive the Euler–Lagrange equation from the calculus of variation, using a general notation which is applicable both to optics and dynamics.

Consider a set of curves $q(s)$ between two points (q_1, s_1) and (q_2, s_2) in the s, q plane (we will only consider curves where the trajectory is single valued at each value of s)[1] (see figure A.1).

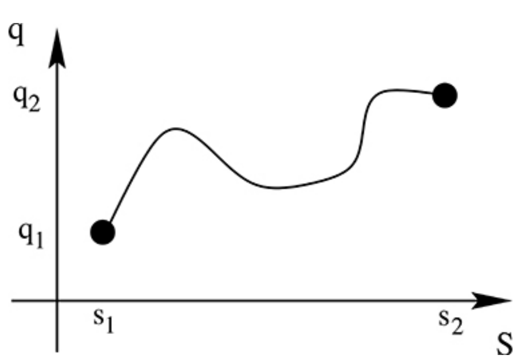


Figure A.1. An arbitrary path in the $q - s$ plane between fixed end points.

Imagine we are interested in one curve that minimizes the quantity

$$S\left[q\left(s\right)\right]=\int_{s_1}^{s_2}L\left(q,\dot{q},s\right)ds \tag{A.1}$$

L is just a number at each point on a given curve determined by the values of *q* and *s* at that point and the gradient $\dot{q}=dq/ds$. The integral sums these numbers along the line.

If the curve that minimizes *S* is $\bar{q}\left(s\right)$ we can write the other curves as deviations from it

$$q\left(s\right)=\bar{q}\left(s\right)+\delta q\left(s\right) \tag{A.2}$$

subject to the boundary conditions

$$\delta q\left(s_1\right)=\delta q\left(s_2\right)=0 \tag{A.3}$$

The value of *S* for these curves varies from the value for $\bar{q}\left(s\right)$ by

$$\delta S=S\left[\bar{q}+\delta q\right]-S\left[\bar{q}\right] \tag{A.4}$$

Since $\bar{q}\left(s\right)$ is the *minimum* though $\delta S=0$ to lowest order in δq.

Let's calculate $S[\bar{q}+\delta q]$ to order δq:

$$\begin{aligned} S\left[\bar{q}+\delta q\right] &= \int_{s_1}^{s_2}L\left(\bar{q}+\delta q,\dot{\bar{q}}+\delta\dot{q},s\right)ds \\ &\simeq \int_{s_1}^{s_2}\left(L\left(\bar{q},\dot{\bar{q}},s\right)+\delta\dot{q}\,\tfrac{\partial L}{\partial\dot{q}}+\delta q\tfrac{\partial L}{\partial q}+\cdots\right)ds \\ &\simeq S\left[\bar{q}\right]+\int_{s_1}^{s_2}\left(\delta\dot{q}\,\tfrac{\partial L}{\partial\dot{q}}+\delta q\tfrac{\partial L}{\partial q}\right)ds+\mathrm{O}\left(\delta q^2\right)\end{aligned} \tag{A.5}$$

Integrating the second term by parts ($u=\partial L/\partial\dot{q}$, $dv/ds=\delta\dot{q}$ etc)

$$\int_{s_1}^{s_2}\delta\dot{q}\,\frac{\partial L}{\partial\dot{q}}\,ds=\left[\delta q\frac{\partial L}{\partial\dot{q}}\right]_{s_1}^{s_2}-\int_{s_1}^{s_2}\delta q\frac{d}{ds}\left(\frac{\partial L}{\partial\dot{q}}\right)ds \tag{A.6}$$

The first term vanishes since δq vanishes at the ends of the path.

Thus

$$S\left[\bar{q}+\delta q\right]-S\left[\bar{q}\right]=-\int_{s_1}^{s_2}\delta q\left(\frac{d}{ds}\left(\frac{\partial L}{\partial\dot{q}}\right)-\frac{\partial L}{\partial q}\right)ds+\cdots \tag{A.7}$$

This is only zero (at order δq) if

$$\boxed{\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0} \tag{A.8}$$

This is the *Euler–Lagrange equation*.

In general, we will want to solve problems in more than one dimension. For example, there may be several such generalized coordinates, q_i corresponding to three dimensions, x, y, z or multiple angles θ_i. The above formalism is easily adapted for such cases. The definition of the action above in terms of the Lagrangian ($L = T - V$) remains the same, however we now have several coordinates

$$q_i \qquad i = 1, \ldots, \ n \tag{A.9}$$

In the derivation above of the Euler–Lagrange equation, it is straightforward to take into account deviations in the path in all of these coordinates. We would find that the change in the action of a path close to the minimizing path would have the form

$$\Delta S = -\int_{s_1}^{s_2} \sum_i \delta q_i \left(\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i}\right) ds \tag{A.10}$$

At the minimum the coefficients of each δq_i must vanish independently so we get a set of Euler–Lagrange equations—one associated with each coordinate

$$\boxed{\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0} \tag{A.11}$$

Exercise 1.8: Work through the above derivation in the case where *L* depends on two coordinates *q* and *p*. What two equations must then be satisfied by the minimizing curve?

Appendix B. Mathematics of conservation laws

Under certain circumstances the Euler–Lagrange equation simplifies from a second order equation to a first order equation. This has important applications in Newtonian dynamics, where the physical interpretation is the connection between symmetry and conservation laws, although here we just focus on the mathematics.

There are two particularly interesting special cases:

(1) If $L(q, \dot{q}, s)$ is independent of the coordinate *q*

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0 \tag{B.1}$$

So

$$\boxed{\frac{\partial L}{\partial \dot{q}} = \text{constant}}$$

(B.2)

(2) If $L(q, \dot{q}, s)$ is independent of the coordinate *s*

$$\frac{dL}{ds} = \frac{\partial L}{\partial q}\dot{q} + \frac{\partial L}{\partial \dot{q}}\ddot{q}$$

(B.3)

using the Euler–Lagrange equation gives

$$\frac{dL}{ds} = \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) + \frac{\partial L}{\partial \dot{q}}\ddot{q}$$

(B.4)

which is just

$$\frac{d}{ds}\left[L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right] = 0$$

(B.5)

so that

$$\boxed{L - \dot{q}\,\frac{\partial L}{\partial \dot{q}} = \text{constant}}$$

(B.6)

Exercise 1.9: This is an exercise in using calculus of variation outside of optics or dynamics. A smooth curved wire connects the origin to the lower point (x_1, y_1). A bead on the wire slides without friction from rest at the upper to the lower point under the influence of gravity. Its mechanical energy is conserved as it moves along the wire. Choose down to be the positive *y* direction.

(a) Show that the time, *T*, required for the bead's journey is

$$T = \frac{1}{\sqrt{2g}}\int_0^{x_2}\sqrt{\frac{(1+y\prime^2)}{y}}dx$$

(b) Given that the integrand of the above integral is independent of *x* show that the curve *y*(*x*) making *T* stationary satisfies the differential equation

$$\frac{dy}{dx} = \sqrt{\frac{(b-y)}{y}}$$

(c) Change the dependent variable from *y* to *ϕ* where $y = b\sin^2\phi/2$ and show that the above can be integrated to give the *brachistochrone*

$$x = b/2\,(\phi - \sin\,\phi)$$

[1]For example, in the case of light in two dimensions we identify $q(\vec{s})y(x)$, while for a particle in one dimension we identify $q(\vec{s})x(t)$, or for a simple pendulum we identify $q(s)\vec{}\theta(t)$.

IOP Concise Physics

Theories of Matter, Space and Time

Classical theories

N Evans and S F King

Chapter 2

Special relativity

Light travels at the very high speed, $c \simeq 3 \times 10^8$ m s^{-1}. In the late 1800s and early 1900s physicists realized that the familiar Newtonian laws of motion break down when particles travel near this speed, which turns out to be a maximum speed in our Universe. Einstein reconciled these discoveries in his special theory of relativity which he wrote down in 1905. Originally these ideas emerged in Maxwell's theory of electromagnetism but it is now standard to present the laws of dynamics first then move to the more complicated case of electromagnetism. This is the ordering we will take in this and the next chapter. The special theory of relativity deals with observations of dynamics by an observer moving at a constant speed. Here we will learn how to write the laws of dynamics in a form consistent with special relativity's postulates. These laws are needed to explain essentially any event in a particle accelerator, many observations in astronomy, but also are crucial to our everyday lives. For example, the GPS satellite system our mobile phones use continually is very sensitive to relativistic corrections from the satellites' motions.

2.1 The postulates

The two fundamental postulates of special relativity are:

- *The speed of light*, c, is the same when measured in any inertial frame. This was the crucial result from the Michelson–Morley experiment.
- *The laws of physics are the same for an observer in any inertial frame.* This is the statement that there is no observer (for example stationary relative to some 'aether') for whom the laws are especially simple.

An observer moving at constant speed is said to be in an *inertial frame* that can be thought of as a combination of:

- A rigid, stationary (relative to the observer) lattice grid by which position coordinates are specified.
- A set of synchronized clocks at each lattice point so time can be recorded (figure 2.1).

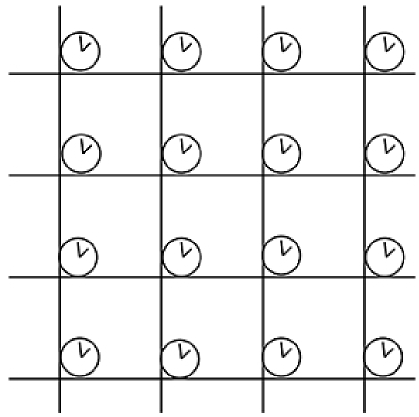


Figure 2.1. A spatial grid with a clock at each point that can be used to specify the coordinates of an event.

With these observational tools the observer can specify any *event* by the set of coordinates

$$(x, y, z), \quad \text{and} \quad t \tag{2.1}$$

Note that moving from one inertial frame to another is often described as performing a *boost*.

2.2 Lorentz transformations

To see the bizarre implications of the first postulate, consider a light wave front emitted from a stationary source at the origin (figure 2.2).

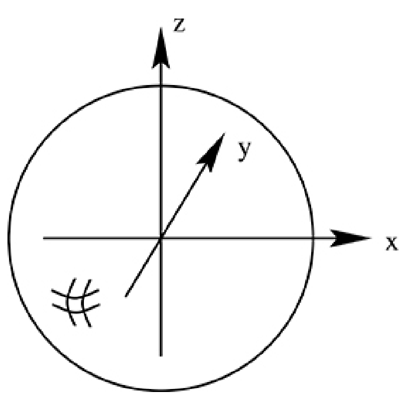


Figure 2.2. A spherical light front emitted from the origin.

Let us call this inertial frame (stationary relative to the light source) frame *S*. The light wave moves away from the source at speed *c* as a spherical shell described by

$$x^2 + y^2 + z^2 = (ct)^2 \tag{2.2}$$

Now consider an inertial frame S' moving with speed *v* in the positive *x* direction. For convenience let's set the origin of both sets of coordinates at time *t* = 0 at the same place (figure 2.3).

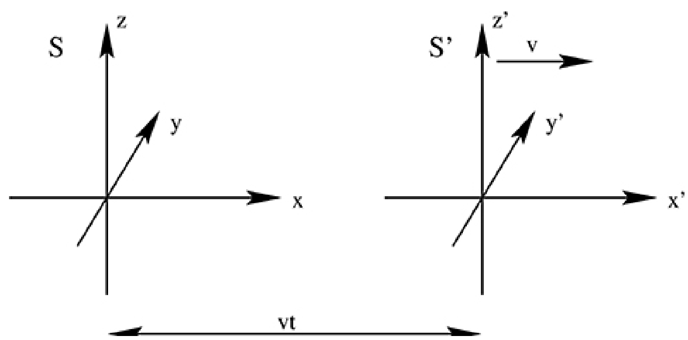


Figure 2.3. Two sets of coordinate axes separating at speed *v* in the *x* direction.

The origins of the two sets of coordinates separate by a distance *vt* in time *t*.

The first postulate says that the observer in S' sees light travel at speed *c* too. Thus in this frame too the light forms a spherical shell centred on the origin in S' described now by

$$x\prime^2 + y\prime^2 + z\prime^2 = (ct')^2 \tag{2.3}$$

This is very surprising—you would have guessed that the observer moving relative to the light source would not be in the centre of the spherical light shell.

The only way to reconcile the two viewpoints is if the two observers disagree on the values of times and positions. The two equations for the position of the shell, equations (2.2) and (2.3), in the two frames moving at relative speed *v* are reconciled by the *Lorentz transformations*

$$\boxed{\begin{aligned} t' &= \gamma\left(t - \tfrac{v}{c^2}x\right) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned}} \tag{2.4}$$

where

$$\gamma = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}} \tag{2.5}$$

Exercise 2.1: Explicitly check that substituting the Lorentz transformations into equation (2.3) one obtains equation (2.2).

Exercise 2.2: How would the Lorentz transformations differ if the boost was in the *z* direction rather than the *x* direction?

An immediate check we should make on these transformations is that they make sense in the slow moving world we live in. When $v \ll c$, $\gamma \simeq 1$ and

$$t' \simeq t, \quad x' \simeq x - vt \tag{2.6}$$

which is indeed what we would expect.

The Lorentz transformations imply that observers moving relative to each other will not agree on the *simultaneity* of events. For example, if a stationary observer sees an event happen at *t* = 0 a distance of 10 m away

$$(t = 0, \ \ x = 10) \tag{2.7}$$

then an observer moving in the *x* direction at speed *v* will record the event as occurring at a time

$$t' = -\gamma \frac{v}{c^2}(10) \tag{2.8}$$

i.e. earlier than when the two observers passed each other ($t = t' = 0$). The implications of this are that the observers do not agree on measurements of periods and lengths.

2.2.1 Time dilation

Imagine an observer at the origin in the frame *S* marks a second by letting off two flashes of light separated by 1 second on his watch. The flashes of light are two events with coordinates

$$(x = 0, \ \ t = 0) \quad (x = 0, \ \ t = 1) \tag{2.9}$$

The moving observer in the S' frame sees the events as

$$(x\prime = 0, \ \ t\prime = 0) \quad (x\prime = -\gamma vt, \ \ t\prime = \gamma)$$

The S' observer has recorded a time (2.10)

$$\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \geqslant 1 \tag{2.11}$$

longer than one second. The S' observer therefore declares that the *S* observer's watch (which is moving relative to S') is running slow:

$$a\ moving\ clock\ runs\ slow$$

2.2.2 Lorentz contraction

Consider a ruler of length *L* at rest in the frame *S*. An observer in *S* might make measurements of the position of the two ends to deduce its length. Those measurements can be represented by the events

$$(t = 0,\ \ x = 0), \quad (t = 0,\ \ x = L) \tag{2.12}$$

A moving observer in the frame S' watches this process and is somewhat bemused. He sees the measurement events as

$$(t' = 0,\ \ x' = 0) \quad \left(t' = -\gamma \frac{v}{c^2} L,\ \ x' = \gamma L\right) \tag{2.13}$$

The measurements were taken according to S' at *different times*. Remember that S' sees the ruler moving, so if you measure the end points at different times you'll not correctly measure the length.

S' wants *S* to make the second measurement at $t' = 0$. In *S* the position of the ruler doesn't change but when should *S* make the measurement so that S' says $t' = 0$?

$$t' = \gamma\left(t - \frac{v}{c^2} L\right) = 0 \tag{2.14}$$

Thus

$$t = \frac{v}{c^2} L \tag{2.15}$$

(The *S* observer doesn't see what is special about this time of course!)

Now where is the second end in the S' coordinates when this new *S* measurement is made?

$$x' = \gamma (x - vt) = \gamma \left(L - \frac{v^2}{c^2} L \right) = \frac{L}{\gamma}$$

(2.16)

Thus S' says the two correct simultaneous measurements of the end points are

$$(t' = 0, \;\; x' = 0) \qquad \left(t' = 0, \;\; x' = \frac{L}{\gamma} \right)$$

(2.17)

S' therefore sees the moving ruler to be shorter by a factor of γ relative to *S*:

$$moving\ objects\ contract\ in\ the\ direction\ of\ motion.$$

Exercise 2.3: Repeat the computation of the length of the ruler in the frame S' but assuming that the ends of the ruler are at the points $x = 1m, x = 2m$ in the *S* frame. Show that the contracted length is again L/γ.

2.3 An analogy to rotations

It's helpful to think of the Lorentz transformations as a generalization of the idea of rotations in the following sense.

Consider first rotations in two dimensions. We can set up two observers who are using coordinates rotated by an angle θ relative to each other (figure 2.4).

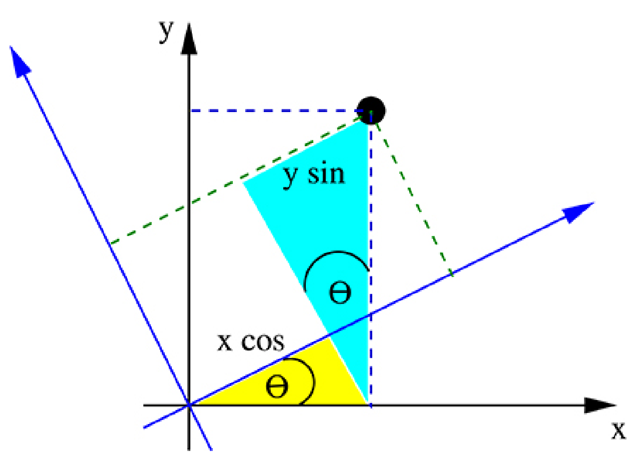


Figure 2.4. The motion of coordinate axes under a rotation.

The coordinates transform between the two coordinate systems as

$$x' = x\ \cos\ \theta + y\ \sin\ \theta$$
$$y' = y\ \cos\ \theta - x\ \sin\ \theta$$

(2.18)

The different coordinate choices are, in a sense, a distraction from the physics involved (of say a moving particle) which is really the same for the observer using either coordinates. The elegant way to express this is to use *vectors*. The vector (e.g. from the origin to a particle) is the same for both observers although its components may be different for the different observers. We write

$$\vec{x} \quad \text{or} \quad \mathbf{x} = (x, y)$$

(2.19)

The coordinate transformation can then be written as a matrix multiplication on the vector

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\ \theta & \sin\ \theta \\ -\sin\ \theta & \cos\ \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

(2.20)

There is something invariant about the position of a particle under rotations—it's distance from the origin, i.e.

$$L^2 = x^2 + y^2 = x'^2 + y'^2$$

(2.21)

We can extract this from the vector by the dot product of the vector with itself

$$L^2 = \vec{x} \cdot \vec{x}$$

(2.22)

Now consider Lorentz transformations in the *x* and *t* directions where the coordinates are mixed up by a boost. The Lorentz transformations, although not exactly like the mixing of spatial coordinates under rotations, do have a similar form. Let's try to draw a diagram with the coordinate axes of two different inertial frame observers both shown (figure 2.5).

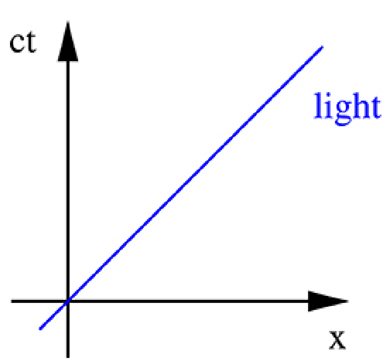


Figure 2.5. The path light follows in the $x - ct$ plane.

We begin with one stationary observer's coordinates in the $x - (ct)$ plane. We use ct rather than just t because it has the same dimensions as x.

Note that light travels on the line at 45° to the axes since it reaches a distance $x = ct$ in ct time.

We can use the Lorentz transformations to plot the position of the equivalent axes in a frame moving relative to this frame. The coordinate axes are given when $ct' = 0$ and $x' = 0$ so

$$ct' = \gamma ct - \frac{v}{c}\gamma x$$

$$ct' = 0 \rightarrow ct = \frac{v}{c}x \tag{2.23}$$

$$x' = \gamma x - \frac{v}{c}\gamma ct$$

$$x' = 0 \rightarrow ct = \frac{c}{v}x \tag{2.24}$$

Now we can plot these axes in the original frame's coordinates (figure 2.6).

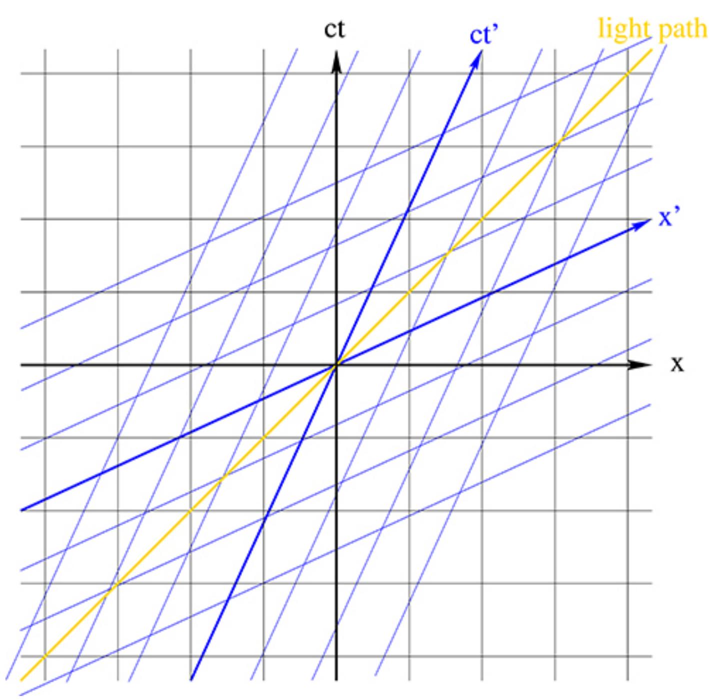


Figure 2.6. Coordinate axes before and after a boost superimposed.

The marked lines are the S' coordinate axes—they agree with the original coordinates as to the point (0, 0). The plot also shows the grid $x' = 0, 1, 2..$ $ct' = 0, 1, 2 .. \text{etc.}$ Note that in the new coordinate system the path light takes is given by the same line—it goes through the points (0, 0) (1, 1) (2, 2) etc.

This is an equivalent plot to the one we drew for rotations. We can place an event on the plot and then read off its coordinates in either the original frame using the square grid or in the boosted frame using the skewed grid.

The grid can be used to see time dilation and length contraction (figure 2.7).

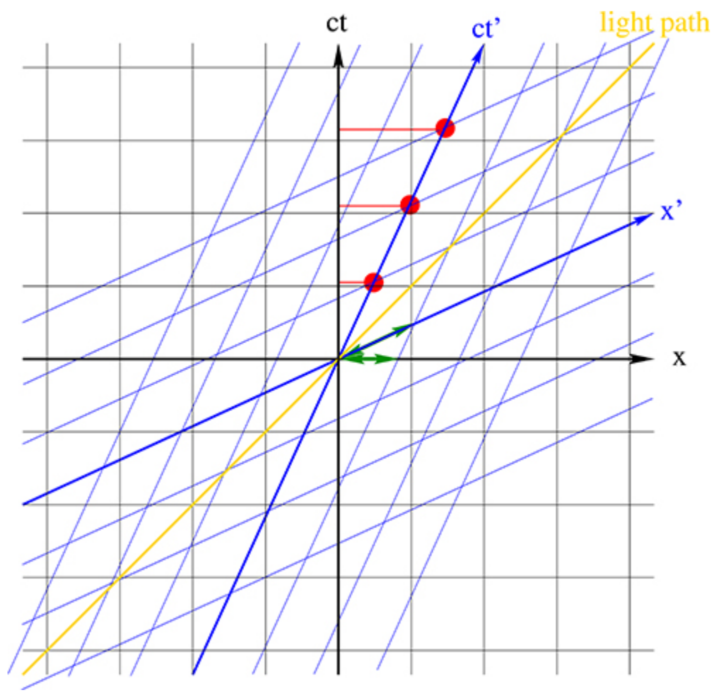


Figure 2.7. Coordinate axes before and after a boost with events marked relevant to measuring a time and a length.

The circles are events positioned at $x' = 0$ every second in S'. Reading the time of the event on the original axes though shows that *S* sees more than 1 s has passed between events—a clock in a moving inertial frame measures time more slowly—time dilation.

The solid line represents a rod in S'. In *S* if we measure distance at the same time for each end we get a smaller length—lengths appear contracted in a moving inertial frame.

Although *x* and *t* change between *S* and S' for this picture, like the coordinates for the rotations, we want a frame invariant way to discuss events. This will lead us to introduce vectors in this plane which have space- and time-like components. In the rotation case the vector had an invariant length that was the same for all observers. For Lorentz transformations we have shown in section 2.2 that the quantity

$$ct^2 - \mid \mathbf{x} \mid^2 = \text{constant} \tag{2.25}$$

is left invariant. This will be the ‘length’ of our new ‘4-vectors’.

Exercise 2.4: Sketch a space–time diagram showing a stationary and a moving coordinate frame with relative speed *v*. Explain from the diagram how a ruler lying on the *x*-axis of the *moving* frame between $x' = 1$ and $x' = 2$ is seen contracted in the stationary frame.

2.4 Four-vectors

Our previous discussion leads us to consider a four component vector with *ct* as the time-like component, and *x*, *y* and *z* position components describing an event or object. We will write this four-vector as

$$x^\mu = (ct, x, y, z) = \left(x^0, x^1, x^2, x^3\right) \tag{2.26}$$

In this notation the index μ on x^μ takes the values $0, 1, 2, 3$ corresponding to the components as shown.

We have identified two properties of the four-vector already. Firstly, under Lorentz boosts in the positive *x* direction by speed *v* it transforms as

$$x^\mu \to x\prime^\mu = \begin{pmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{2.27}$$

Secondly, we know that it has a Lorentz invariant length

$$(x_0)^2 - (x_1)^2 - (x_2)^2 - (x_3)^2 = (ct)^2 - \mid \vec{x} \mid^2 \tag{2.28}$$

2.4.1 Index convention

At this point we are going to adopt a rather compact notation for multiplying four-vectors. It will take a little getting used to but is not intrinsically deep! There are two rules that will apply to the ‘μ’ index on a four-vector:

- A given label for an index may occur at most twice in any term in an expression.

- A repeated index is said to be ‘contracted’. Typically people write a repeated index once up and once down. Such a repeated index is ‘summed over’.

The best way to explain this is with an example. We can write the Lorentz transformation of x^μ in the following form

$$x^\mu \rightarrow x\prime^\mu = \Lambda^\mu{}_\nu x^\nu \tag{2.29}$$

The new object $\Lambda^\mu{}_\nu$ has two indices each of which can take the values $0, 1, 2, 3$ and so there are 4 × 4 = 16 components. These 16 components are just the 16 components of the Lorentz transformation matrix we’ve written above (for example let μ count the row and ν the column).

In the expression the ν index occurs twice and this implies we must let ν take all possible values and add up the answers we get in each case.

For example, consider the case where we set $\mu = 0$ then

$$\begin{aligned} x\prime^0 &= \Lambda^0{}_\nu x^\nu \\ &= \Lambda^0{}_0 x^0 + \Lambda^0{}_1 x^1 + \Lambda^0{}_2 x^2 + \Lambda^0{}_3 x^3 \\ &= \gamma x^0 - \gamma \tfrac{v}{c} x^1 \end{aligned} \tag{2.30}$$

This has reproduced the Lorentz transformation for $x^0 = ct$.

Exercise 2.5: Convince yourself that equations (2.27) and (2.29) both reproduce the four equations (2.4).

We can also write the Lorentz invariant length in this way. Formally we do this as follows. We define a two index object called the *metric* with the 16 components

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.31}$$

Now we can write

$$x_\mu = g_{\mu\nu} x^\nu \tag{2.32}$$

This four-vector with a lowered index has components

$$x_\mu = (ct, -x, -y, -z)$$

So finally we can define the length of the four-vector as (2.33)

$$x^\mu x_\mu = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3$$

$$= (ct)^2 - x^2 - y^2 - z^2$$ (2.34)

This notation, which is common, is a little sloppy because the lowered index on a four-vector secretly contains the metric and its minus signs. In practice you may just want to remember to insert the minus signs as they appear in the above expression when you contract the indices on four-vectors, as here, rather than always write the metric factors! BEWARE though that there are not these minus signs in the Lorentz transformation expression (2.29) where $\Lambda^\mu{}_\nu$ is not a four-vector-like object!

Exercise 2.6: Calculate $x^\mu x_\mu$ and $x^\mu y_\mu$ for the four-vectors

$$x^\mu = (3, 1, 0, 2) \quad y^\mu = (4, 5, 3, 0)$$

Show explicitly by performing a Lorentz boost by speed *v* in the *x*-direction that these products are Lorentz invariant.

2.5 The laws of dynamics

We have seen the consequences of relativity for observations of lengths and periods. Now we will turn to thinking about how to formulate the laws of dynamics. Simple Newtonian formulae such as $f = ma$ do not work because they contain time dependence and different observers don't agree on lengths of time.

Our guiding principle should be the second postulate which says that physical laws should be the same for an observer in any inertial frame. We will cast the laws in a way where this is manifestly true. Four-vectors will be the tool that allows this since they are a frame invariant way of describing the properties of a particle. Our laws will only:

- Contain Lorentz invariant quantities such as $x^\mu x_\mu$
- Or take the form $X^\mu = Y^\mu$.

This latter form is explicitly Lorentz invariant because the two sides of the equation transform in the same way under Lorentz transformations.

So far we only have a four-vector describing position. We will now construct four-vectors describing the kinematic properties of a particle.

2.5.1 Four-velocity

It is not sensible to use

$$v = \frac{dx^\mu}{dt}$$

as our definition of velocity because both x^μ and t transform under Lorentz boosts. (2.35)
The resulting transformation is very messy.

Ideally we would like a measure of time that is Lorentz invariant so that v would transform only through the transformation of x^μ. It would then be a four-vector itself. Such a Lorentz invariant measure of time is:

Proper time: *the time elapsed on a clock in the rest frame of a moving object.* Essentially we imagine that everything has a watch and we time an event for the object by the time on its watch not the observer's. Observers in any reference frame will then get the same answer.

Finally, we can make a sensible choice for our variable four-velocity

$$\boxed{u^\mu = \tfrac{dx^\mu}{d\tau}} \tag{2.36}$$

Let's stress again that this four-vector transforms just like x^μ under boosts, i.e.

$$u\prime^\mu = \Lambda^\mu{}_\nu u^\nu \tag{2.37}$$

It is useful to know how four-velocity relates to the more standard velocity measured by an observer using his own watch (we can call this *coordinate velocity*)

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau} \tag{2.38}$$

We can work out $dt/d\tau$ from the Lozentz transformations. τ is the time in the rest frame, where the particle is sat at the origin, so in a moving frame

$$t = \gamma\tau \quad \rightarrow \quad \frac{dt}{d\tau} = \gamma \tag{2.39}$$

Thus the components of four-velocity are

$$u^\mu = \gamma\left(c, v_x, v_y, v_z\right) \tag{2.40}$$

From this expression we can finally work out the invariant ‘length’ of this four-vector from the product

$$\boxed{u^\mu u_\mu = \gamma^2\left(c^2 - \mid \vec{v} \mid^2\right) = c^2} \tag{2.41}$$

2.5.2 Four-acceleration

The definition of acceleration is now straightforward

$$a^\mu = \frac{du^\mu}{d\tau} \tag{2.42}$$

Again it's worth stressing that this object is a four-vector which transforms in the same way as x^μ.

2.5.3 Four-momentum

The natural generalization of momentum is given by

$$p^\mu = mu^\mu = m\frac{dx^\mu}{d\tau} \tag{2.43}$$

Here we have introduced the mass of the particle, *m*—it is a constant, intrinsic property of the particle.

p^μ is again a four-vector that transforms as

$$p\prime^\mu = \Lambda^\mu{}_\nu\, p^\nu \tag{2.44}$$

Interestingly though we have been led to introduce a time-like version of momentum. What does this correspond to? To find out we should take the classical limit of the theory ($v \ll c$) and see what it corresponds to in Newtonian dynamics. Remember that the time-like component of four-velocity was $u_0 = \gamma c$ so

$$\begin{aligned} p^0 &= mc\gamma \\ &= mc\big(1 - v^2/c^2\big)^{-1/2} \\ &\simeq mc\left(1 + \tfrac{1}{2}\tfrac{v^2}{c^2} + \cdots\right) \end{aligned} \tag{2.45}$$

The first term is a constant. The second term though is recognisable since $\frac{1}{2}mv^2$ is kinetic energy in the low *v* limit. This suggests we should interpret p_0 as the relativistic version of energy (divided by *c*). Then we have a surprising interpretation of the first, constant, term—a particle at rest has energy

$$E_{\rm rest} = mc^2 \tag{2.46}$$

We can write the components of p_μ in a number of ways now

$$p^\mu = \left(\frac{E}{c}, \vec{p}\right) = mu^\mu = m\gamma\left(c, \vec{v}\right) \tag{2.47}$$

The relativistic expression for energy is therefore

$$E = \gamma mc^2 \tag{2.48}$$

and the relativistic version of kinetic energy (the energy when moving minus the energy at rest)

$$T = (\gamma - 1)mc^2 \tag{2.49}$$

The invariant length of the four-vector follows from $u^\mu u_\mu = c^2$ so

$$p^\mu p_\mu = \boxed{\frac{E^2}{c^2} - \left|\,\vec{p}\,\right|^2 = m^2c^2} \tag{2.50}$$

Exercise 2.7: Calculate, by explicitly performing a boost, the relativistic energy and momentum of a proton moving at speed $v = 0.5c$. The rest mass of a proton is approximately 1 GeV c^{-2}.

2.5.4 Hypothesis for dynamical law

Armed with these four-vector variables we can now have a guess as to the form of the relativistic version of Newton's second law. The obvious equation to try is

$$f^\mu = \frac{dp^\mu}{d\tau} \tag{2.51}$$

This is manifestly Lorentz invariant and has the correct non-relativistic limit if f^μ is a relativistic extension of force. As yet though we haven't mentioned forces and we won't until we discuss electromagnetism! In fact this guess is the correct law.

The law tells us something interesting even when $f^\mu = 0$

$$\frac{dp^\mu}{d\tau} = 0 \quad \rightarrow \quad p^\mu = \text{constant} \tag{2.52}$$

In other words, if no external force acts on a system four-momentum is conserved. This is the relativistic analogue of conservation of energy (p_0) and conservation of the usual three-component momentum (p_1, p_2, p_3).

2.6 Physics with four-momentum

To gain experience with four-vectors we will now look at four physics problems where using four-momentum makes the solutions much easier than without.

To make our life easier we will use a trick that is common. Instead of using the usual units system we will work in a new system where

$$c = 1 \tag{2.53}$$

In other words we redefine the unit of length so that it is the distance light travels in 1 second! This would not be sensible for everyday life but in problems where everything is travelling at the speed of light a metre is an absurdly small distance. In practice we will be able to drop all the factors of *c* from computations. It's pretty easy to put them back into the final answer using dimensional analysis as we will see.

2.6.1 The Doppler effect

What frequency will an observer see a light wave have if he is moving relative to it (figure 2.8)?

Figure 2.8. A photon approaching an observer.

Consider first a static observer in the frame of the light source. The photons of light carry four-momentum

$$p^{\mu} = \left(E, \vec{p}\right) = \left(hf, -\frac{h}{\lambda}\widehat{\mathbf{x}}\right) = (hf, -hf, 0, 0) \tag{2.54}$$

Note that the photon is moving in the negative *x*-direction towards the observer. We have used the quantum mechanical relations between the energy and frequency of the photon and between its momentum and wavelength. We have also used $f\lambda = c = 1$.

We can now ask what would happen to the frequency of the light if the observer was moving in the positive *x*-direction at speed *v*. We just perform a boost on the four-vector

$$p\prime^{\mu} = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} hf \\ -hf \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma(1+v)hf \\ -\gamma(1+v)hf \\ 0 \\ 0 \end{pmatrix} \tag{2.55}$$

Now if we just concentrate on the time-like component we have

$$p\prime^{0} = E\prime = hf\prime = \sqrt{\frac{1}{1-v^2}}(1+v)hf = \sqrt{\frac{(1+v)^2}{(1+v)(1-v)}}hf \tag{2.56}$$

or

$$\boxed{f' = \sqrt{\frac{(1+v)}{(1-v)}}f} \tag{2.57}$$

Finally, we can reintroduce the factors of *c* since the factors of $(1+v)$ are not dimensionally correct. We should have

$$\boxed{f' = \sqrt{\frac{(1+v/c)}{(1-v/c)}}f} \tag{2.58}$$

2.6.2 The Compton effect

The Compton effect relates the angle of scattering of a photon off a static electron to its final wavelength. The classic experiment is schematically shown in figure 2.9.

Figure 2.9. A schematic of the Compton experiment.

You've probably calculated this relationship for the change in the wavelength of the photon as a function of its scattering angle, $\Delta\lambda(\theta)$, previously. Using four-momentum will get us to the answer much quicker.

Set up the four-momentum of the particles to be:

$$\begin{gathered}\text{initial photon}:\ p^{\mu}_{\gamma i}=\left(\frac{h}{\lambda},\frac{h}{\lambda}\widehat{x}\right)\\ \text{initial electron}:\ p^{\mu}_{ei}=(m_e,0)\\ \text{final photon}:\ p^{\mu}_{\gamma f}=\left(\frac{h}{\lambda'},\frac{h}{\lambda'}\widehat{f}\right)\\ \text{final electron}:\ p^{\mu}_{ef}\end{gathered}$$

Here $\hat{\mathbf{f}}$ is a unit vector in the direction of the motion of the final photon, which is at an angle θ to the x-axis.

Since no external force acts, four-momentum is conserved in the collision so

$$p^{\mu}_{\gamma i}+p^{\mu}_{ei}=p^{\mu}_{\gamma f}+p^{\mu}_{ef} \tag{2.59}$$

It turns out to be helpful to rearrange this equation so that p^{μ}_{ef} is isolated—we know least about p^{μ}_{ef} so will want to eliminate it

$$p^{\mu}_{\gamma i}+p^{\mu}_{ei}-p^{\mu}_{\gamma f}=p^{\mu}_{ef} \tag{2.60}$$

Now we consider the Lorentz invariant product

$$\begin{aligned}p^{\mu}_{ef}p_{ef\mu}&=m_e^2\\ &=\left(p^{\mu}_{\gamma i}+p^{\mu}_{ei}-p^{\mu}_{\gamma f}\right)\left(p_{\gamma i\mu}+p_{ei\mu}-p_{\gamma f\mu}\right)\\ &=p^{\mu}_{\gamma i}p_{\gamma i\mu}+p^{\mu}_{ei}p_{ei\mu}+p^{\mu}_{\gamma f}p_{\gamma f\mu}+2\left(p^{\mu}_{\gamma i}p_{ei\mu}-p^{\mu}_{\gamma i}p_{\gamma f\mu}-p^{\mu}_{\gamma f}p_{ei\mu}\right)\\ &=0+m_e^2+0+2\left(\frac{h}{\lambda_i}m_e-0\right)-2\frac{h}{\lambda_i}\frac{h}{\lambda_f}(1-\cos\,\theta)-2\left(\frac{h}{\lambda_f}m_e-0\right)\end{aligned} \tag{2.61}$$

We have used two crucial facts here. Firstly, when the four-momentum of a particle is contracted with itself we simply obtain the invariant m_e^2. Secondly, we have used the contraction law $p^{\mu}_1 p_{2\mu}=\left(p^0_1 p^0_2-\vec{p}_1\cdot\vec{p}_2\right)$.

Rearranging we find

$$\frac{h}{\lambda_i}m_e-\frac{h}{\lambda_f}m_e=\frac{h}{\lambda_i}\frac{h}{\lambda_f}(1-\cos\,\theta) \tag{2.62}$$

Multiplying through by $\lambda_i\lambda_f/\left(hm_e\right)$ gives

$$\lambda_f - \lambda_i = \frac{h}{m_e}(1 - \cos\,\theta)$$

(2.63)

which is the answer we want. Again we can insert *c* on dimensional grounds

$$\boxed{\lambda_f - \lambda_i = \frac{h}{m_e c}(1 - \cos\,\theta)}$$

(2.64)

2.6.3 Fixed target experiments

A simple way in which to create fundamental particles is by colliding a high energy proton or electron into a fixed target of, for example, lead (figure 2.10).

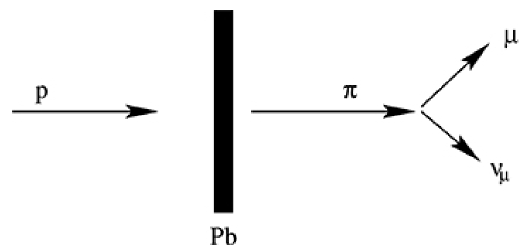


Figure 2.10. A proton collision with a fixed lead target generating a pion that then decays.

It's not immediately obvious how much energy is available to make rest mass energy of the new particle because momentum conservation requires the final state to be moving and have kinetic energy. A sensible thing to do is to move to the centre of mass frame where the particle and target (a particle in the wall) approach each other with equal and opposite momentum (figure 2.11).

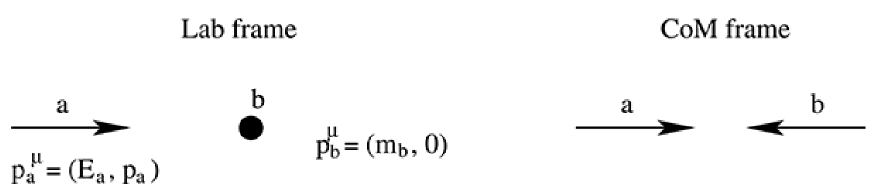


Figure 2.11. The lab frame with one fixed target and the centre of mass frame where both particles have equal momentum.

In this frame the particle produced will be at rest and all the energy of the initial state will become rest mass energy of the product.

We can work out the Lorentz boost needed to move from the original 'lab' frame to the centre of mass frame. We boost the four-momenta in the lab frame by an amount *v*

$$p_a^{\mu'} = \begin{pmatrix} \gamma & -\gamma v \\ -\gamma v & \gamma \end{pmatrix} \begin{pmatrix} E_a \\ p_a \end{pmatrix} = \begin{pmatrix} \gamma\left(E_a - vp_a\right) \\ \gamma\left(p_a - vE_a\right) \end{pmatrix} \tag{2.65}$$

$$p_b^{\mu'} = \begin{pmatrix} \gamma & -\gamma v \\ -\gamma v & \gamma \end{pmatrix} \begin{pmatrix} m_b \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma m_b \\ -\gamma v m_b \end{pmatrix} \tag{2.66}$$

In the centre of mass frame the momenta must be equal and opposite so

$$p_a'^x = -p_b'^x$$

$$\gamma v m_b = \gamma\left(p_a - vE_a\right) \tag{2.67}$$

and the required boost is by

$$\boxed{\frac{v}{c} = \frac{p_a}{m_b c + E_a/c}} \tag{2.68}$$

If after this boost the particles are ultra-relativistic so that $E_a \simeq \mid p_a \mid = \mid p_b \mid \simeq E_b$ then the total available energy is

$$E_{\mathrm{com}} = 2\gamma m_b c = \sqrt{\frac{4m_b^2 c^2}{1 - \left(\frac{p^a}{m_b c + E_a/c}\right)^2}} = \sqrt{\frac{4m_b^2 c^2 \left(m_b c + E_a/c\right)^2}{\left(m_b c + E_a/c\right)^2 - p_a^2}} \tag{2.69}$$

If we now expand in the limit with $E_a \gg m_b c^2, m_a c^2$ we find (remember that in this high energy limit $E_a/c = p_a$)

$$E_{\mathrm{com}} = \sqrt{\frac{4m_b^2 E_a^2}{2m_b E_a}} \tag{2.70}$$

$$\boxed{E_{\mathrm{com}} = \sqrt{2m_b E_a}} \tag{2.71}$$

We could have obtained this result more quickly by calculating the invariant rest mass of the whole system in the original coordinates

$$\begin{aligned} p^{\mu}_{\mathrm{TOT}} p_{\mathrm{TOT}\,\mu} &= m^2_{\mathrm{TOT}} c^2 \\ &= \left(p^{\mu}_a + p^{\mu}_b\right)\left(p_{a\mu} + p_{b\mu}\right) \\ &= p^{\mu}_a p_{a\mu} + p^{\mu}_b p_{b\mu} + 2 p^{\mu}_a p_{b\mu} \end{aligned} \tag{2.72}$$

which in the limit where E_a is large compared to the rest masses gives

$$m^2_{\mathrm{TOT}} c^2 = 2 \frac{E_a}{c} \frac{m_b c^2}{c} = 2 E_a m_b \tag{2.73}$$

2.6.4 The GZK bound

Active galaxies accelerate protons to very high energies but there is a maximum energy we should expect to see (first calculated by Greisen, Zatsepon and Kuzmin). The reason for the maximum is that the Universe is full of photons left over from the Big Bang which higher energy protons can interact with. These photons are responsible for the ambient background temperature of the Universe $T \sim 3$ K ($E_\gamma = k_B T = 8 \times 10^{-4}$ eV). The protons interact as follows

$$p\gamma \to \Delta\left(M_\Delta \sim 1.2\ \mathrm{GeV}/c^2\right) \to \pi^+ n \tag{2.74}$$

The Δ is a short lived particle and the final decay is by far its most dominant decay process. If there is sufficient energy in the collision to create a Δ then the proton is converted to other particles very efficiently. We can calculate the minimum energy the proton must have.

Let's assign the proton and photon initial four-momenta

$$p^{\mu}_p = (E_p, k, 0, 0), \quad p^{\mu}_\gamma = (h\nu, -h\nu, 0, 0) \tag{2.75}$$

Note that we’ve set up the process so the photon and proton will collide head-on. This maximizes the energy available for new particle creation and will therefore give us the minimum proton energy for the process.

Four-momentum will be conserved in the interaction so

$$p^{\mu}_\Delta = p^{\mu}_p + p^{\mu}_\gamma \tag{2.76}$$

Rearranging and squaring gives

$$p^{\mu}_{\Delta} p_{\Delta\mu} = m^2_{\Delta} = \left(p^p_{\mu} + p^{\gamma}_{\mu}\right)\left(p^{p\mu} + p^{\gamma\mu}\right)$$

$$= m^2_p + m^2_{\gamma} + 2E_p h\nu - 2k\left(-h\nu\right) \tag{2.77}$$

For a relativistic proton $E_p \simeq k$ and so

$$E_p = \frac{m^2_{\Delta} - m^2_p}{4h\nu} \simeq 2 \times 10^{20} \ \mathrm{GeV} \tag{2.78}$$

Protons with energy of this or above will undergo this interaction. Factoring in the density of photons it turns out that the mean free path for such protons is about 3 Mpc (our Galaxy group is about 20 Mpc across). We shouldn't expect to see any protons of this energy from active galaxies.

Surprisingly though experimenters have reported observations of cosmic ray protons with higher energy than this bound (although they do see a decrease in the number of events above the bound limit). If these events are real we might be doing something wrong! Could special relativity break down at such high energies? Could there be a source of high energy protons within 3 Mpc (either an astronomical source or very massive particles left over from the Big Bang that decay to these protons)? At the moment this issue is an open question.

Exercise 2.8: A charged pion ($m_{\pi} = 140\ \mathrm{MeV}/c^2$) at rest decays to a charged muon ($m_{\mu} = 105$ MeV/c^2) and a massless muon neutrino. Calculate the energy and momenta of the neutrino and the muon.

Exercise 2.9: In the original (Homestake) solar neutrino detection experiment neutrinos from the Sun interact with Cl^{37} atoms to form Ar^{37} and an electron. Assuming the Cl atoms are at rest what boost is required to move to the centre of mass frame? Determine the minimum energy the neutrino must have for this reaction to proceed.

Exercise 2.10: Speculative models of particle physics predict that at very high energies all matter is unified into a single form. If this were true one would expect, very rarely, that protons would decay to, for example, a positron and a photon. Derive an expression for the wavelength of the emerging photon in the proton's rest frame.

2.7 Tensors

We are now familiar with four-vectors. They are though just one part of a family of objects called *tensors* which can have more than one index. We will need these later when we study electromagnetism. To introduce them think about angular momentum.

Non-relativistically angular momentum is given by

$$\vec{l} = \vec{r} \times \vec{p}$$

with components (2.79)

$$l^1 = yp_z - zp_y$$

$$l^2 = zp_x - xp_z$$

$$l^3 = xp_y - yp_x$$

Relativistically these components are naturally part of the tensor (2.80)

$$L^{\mu\nu} = x^\mu p^\nu - p^\mu x^\nu$$

For example (2.81)

$$L^{12} = xp_y - yp_x = l^3$$

Tensors have a number of properties which in this case we can deduce from their 'composite' nature. Thus (2.82)

- Under Lorentz transformations: $L'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta L^{\alpha\beta}$
- Lorentz invariant: $L^{\mu\nu}L_{\mu\nu} = \left(L^{00}\right)^2 - \left(L^{01}\right)^2 + \left(L^{11}\right)^2 + \cdots$ = constant

There are 16 terms in the final sum here (in fact because of the anti-symmetry of $L^{\mu\nu}$ in equation (2.81) the diagonal terms are zero and e.g. $L^{12} = -L^{21}$ so there are only six independent components). Finally we note that the metric we introduced earlier is itself a tensor.

Exercise 2.11: Show that the metric tensor is invariant to a boost by speed *v*.

2.8 Relativistic action

The action that reproduces the relativistic equation of motion for a free particle

$$\frac{dp^\mu}{d\tau} = 0$$

has an interesting form. It is given by (2.83)

$$S = -m \int \sqrt{\frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau}}\, d\tau$$

(2.84)

Note that formally here τ need not be the proper time because we can parametrize the path by any other $\tau'(\tau)$. Since $\frac{dx^\mu}{d\tau} = \frac{dx^\mu}{d\tau'}\frac{d\tau'}{d\tau}$ the action transforms to the same form as equation (2.84) but with $\tau \to \tau'$.

The Euler–Lagrange equations take the form

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial\left(\frac{dx_\mu}{d\tau}\right)}\right) - \frac{\partial L}{\partial x^\mu} = 0 \tag{2.85}$$

or explicitly

$$\frac{1}{2}\frac{d}{d\tau}\left[m\frac{dx^\mu}{d\tau}\left(\frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau}\right)^{-1/2}\right] = 0 \tag{2.86}$$

from which we learn that

$$\frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau} = u^\mu u_\mu = c^2 \tag{2.87}$$

and

$$\frac{d}{d\tau}\left[m\frac{dx^\mu}{d\tau}\right] = \frac{dp^\mu}{d\tau} = 0 \tag{2.88}$$

the correct equation of motion if we do identify the parameter with the proper time.

If we look closely at equation (2.84) though, we realize that it has an interesting form. The proper time is being used to parameterize the path of the particle but if we move $d\tau$ into the square root we see it cancels and what we are actually doing is calculating the length of the path. This is very elegant in that the length of the path is the only physical characteristic of the motion—it's nice that the action is so simple.

2.9 Lorentz transformations and rotations II

As a final amusement we show here that boosts and rotations are intimately tied together in relativity. Naively you might think Lorentz transformations form a closed set of operations (that is doing two boosts is equivalent to doing one other boost) but in fact things are more complicated as the following procedure shows.

We will do the following Lorentz transformations on a four-vector $x^\mu = (t, \mathbf{x})$:

(1) Boost by δv in the *x*-direction
(2) Boost by δu in the *y*-direction
(3) Boost by $-\delta v$ in the *x*-direction

(4) Boost by $-\delta u$ in the *y*-direction

You might think we'd be back where we started but let's see …

(1) **Boost by *δ*v in the *x*-direction:**

$$x\prime^{\mu} = \Lambda^{\mu}{}_{\nu} X^{\nu} = \begin{pmatrix} \gamma(\delta v) & -\gamma(\delta v)\frac{\delta v}{c} & 0 & 0 \\ -\gamma(\delta v)\frac{\delta v}{c} & \gamma(\delta v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{2.89}$$

Now since δv is small

$$\gamma(\delta v) = \left(1 - \frac{\delta v^2}{c^2}\right)^{-1/2} \simeq 1 + \frac{1}{2}\frac{\delta v^2}{c^2} + \cdots \tag{2.90}$$

$$\gamma(\delta v)\frac{\delta v}{c} \simeq \frac{\delta v}{c} + \cdots \tag{2.91}$$

We've kept all the terms up to order $\frac{\delta v^2}{c^2}$.

$$x\prime^{\mu} = \begin{pmatrix} 1 + \frac{\delta v^2}{2c^2} & -\frac{\delta v}{c} & 0 & 0 \\ -\frac{\delta v}{c} & 1 + \frac{\delta v^2}{2c^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{2.92}$$

(2) **Boost by *δ*u in the *y*-direction:**

$$x''^{\mu} = \Lambda^{\mu}{}_{\nu} x''^{v} = \begin{pmatrix} 1+\frac{\delta u^2}{2c^2} & 0 & -\frac{\delta u}{c} & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{\delta u}{c} & 0 & 1+\frac{\delta u^2}{2c^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{2.93}$$

(3) **Boost by −δv in the** ***x*****-direction:**

$$x'''^{\mu} = \Lambda^{\mu}{}_{\nu} x''^{v} = \begin{pmatrix} 1+\frac{\delta v^2}{2c^2} & \frac{\delta v}{c} & 0 & 0 \\ \frac{\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{2.94}$$

(4) **Boost by −*δu* in the** ***y*****-direction:**

$$x''''^{\mu} = \Lambda^{\mu}{}_{\nu} x'''^{v} = \begin{pmatrix} 1+\frac{\delta u^2}{2c^2} & 0 & \frac{\delta u}{c} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\delta u}{c} & 0 & 1+\frac{\delta u^2}{2c^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{2.95}$$

Note: that (1) and (3) are each others' inverse:

$$\begin{pmatrix} 1+\frac{\delta v^2}{2c^2} & \frac{\delta v}{c} & 0 & 0 \\ \frac{\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1+\frac{\delta v^2}{2c^2} & -\frac{\delta v}{c} & 0 & 0 \\ -\frac{\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.96}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \mathrm{O}\left(\frac{\delta v/u^3}{c^3}\right)$$

(2.97)

The loop of four transformations is more complicated though because we're doing (2) in between (1) and (3). So …

$$X''''^{\mu} = \Lambda^{\mu}{}_{\nu}(-\delta u)\Lambda^{\alpha}{}_{\beta}(-\delta v)\Lambda^{\beta}{}_{\gamma}(\delta u)\Lambda^{\gamma}{}_{\nu}(\delta v)X^{\nu}$$

(2.98)

$$\Lambda^{\mu}{}_{\alpha}(-\delta u)\Lambda^{\mu}{}_{\beta}(-\delta v)\Lambda^{\mu}{}_{\gamma}(\delta u)\Lambda^{\mu}{}_{\nu}(\delta v) =$$

(2.99)

$$\begin{pmatrix} 1+\frac{\delta u^2}{2c^2} & 0 & \frac{\delta u}{c} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\delta u}{c} & 0 & 1+\frac{\delta u^2}{2c^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1+\frac{\delta v^2}{2c^2} & \frac{\delta v}{c} & 0 & 0 \\ \frac{\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times$$
$$\begin{pmatrix} 1+\frac{\delta u^2}{2c^2} & 0 & -\frac{\delta u}{c} & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{\delta u}{c} & 0 & 1+\frac{\delta u^2}{2c^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1+\frac{\delta v^2}{2c^2} & -\frac{\delta v}{c} & 0 & 0 \\ -\frac{\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(2.100)

$$= \begin{pmatrix} 1+\frac{\delta u^2+\delta v^2}{2c^2} & \frac{\delta v}{c} & \frac{\delta u}{c} & 0 \\ \frac{\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ \frac{\delta u}{c} & 0 & 1+\frac{\delta u^2}{2c^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1+\frac{\delta u^2+\delta v^2}{2c^2} & \frac{-\delta v}{c} & \frac{-\delta u}{c} & 0 \\ \frac{-\delta v}{c} & 1+\frac{\delta v^2}{2c^2} & 0 & 0 \\ \frac{-\delta u}{c} & 0 & 1+\frac{\delta u^2}{2c^2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(2.101)

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{\delta u \delta v}{c^2} & 0 \\ 0 & -\frac{\delta u \delta v}{c^2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \mathrm{O}\left(\frac{\delta v / u^3}{c^3} \right) \tag{2.102}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\,\theta & \sin\,\theta & 0 \\ 0 & -\sin\,\theta & \cos\,\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.103}$$

where $\cos\,\theta \simeq 1 + \cdots$ and $\sin\,\theta \simeq \delta u \delta v + \cdots$

The result is a rotation about the *z*-axis!! It is therefore necessary to consider the combination of both rotations and Lorentz transformations as a single set of transformations.

IOP Concise Physics

Theories of Matter, Space and Time

Classical theories

N Evans and S F King

Chapter 3

Relativistic electromagnetism

In this section of the book we will study electromagnetism. We begin by reviewing Maxwell's equations in integral and differential form. Our main task here though will be to understand how these equations already encode relativity. To do this we will need to rewrite them in terms of potentials to find a manifestly Lorentz invariant form. We will then understand how electric and magnetic fields change under a boost.

3.1 Integral form of Maxwell's equations

We begin by reviewing the physics of Maxwell's equations in integral form.

3.1.1 Gauss' law

It is a remarkable fact that the net number of electric field lines exiting a closed surface *S* is proportional to the sum of the electric charges *q* inside the volume enclosed by *S*. Since the electric field $\vec{E}$ is proportional to the number of field lines per unit area, mathematically we have,

$$\boxed{\int_S \vec{E} \cdot d\vec{A} = \frac{q}{\varepsilon_0}} \tag{3.1}$$

- $\vec{E}$ is the (vector) force a unit charge experiences at a point on the closed surface *S*.
- The integral means a sum of $\vec{E} \cdot d\vec{A}$ for the infinitesimal surface elements that make up a whole, closed surface *S*. Remember that a little area element (figure 3.1) is described by a vector normal to its surface.
- *q* is the net charge contained inside the surface.

For example, the electric field around a point charge is given by Gauss' law using a spherical surface *S* of radius *r* around the charge. The integral is then trivially performed and the result is summarized in figure 3.2.

$$4\pi r^2 \mid \vec{E} \mid = \frac{q}{\epsilon_0}$$

$$\mid \vec{E} \mid = \frac{q}{4\pi\epsilon_0 r^2}$$

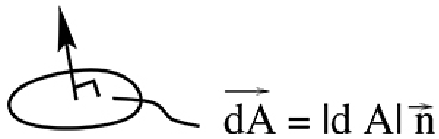


Figure 3.1. An area element vector.

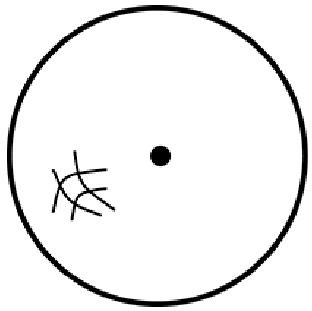


Figure 3.2. A Guassian surface around a point charge.

3.1.2 No magnetic charges

The equivalent of Gauss' law for magnetic fields is just

$$\boxed{\int_S \vec{B} \cdot d\vec{A} = 0} \tag{3.2}$$

since there are no magnetic charges, i.e. no magnetic monopoles.

3.1.3 Faraday's law

Faraday discovered that moving a loop of wire in a magnetic field induces a current in the wire (figure 3.3). The number of magnetic field lines passing through the loop is the magnetic flux given by

$$\Phi = \int_S \vec{B}.d\vec{A} \tag{3.3}$$

where the area *S* which is integrated over in this case is the open surface enclosed by the loop. Now the induced voltage depends on the rate of change of the number of magnetic field lines passing through the loop with respect to time *t*, as given by Faraday's law,

$$e.m.f. = -\frac{\partial \Phi}{\partial t} \tag{3.4}$$

The minus sign reflects Lenz's Law which says the system resists change.

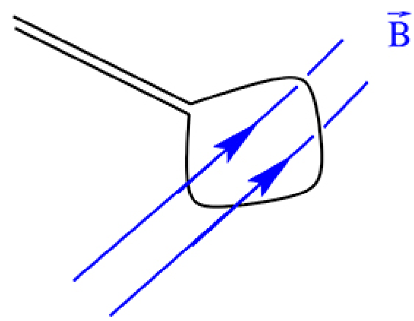


Figure 3.3. Magnetic flux penetrating a wire loop.

The voltage difference around the loop perimeter, *s*, is given by $V = Ed$ but since different bits of the wire point in different directions we must calculate this for each infinitessimal bit of wire and sum the answers,

$$e.m.f = \int_s \vec{E}.\, d\vec{l} \tag{3.5}$$

Finally combining the above equations we have

$$\boxed{\int_s \vec{E}.d\vec{l} = -\int_S \frac{\partial \vec{B}}{\partial t}.\, d\vec{A}} \tag{3.6}$$

3.1.4 Ampere's law

The analogue of Faraday's law for the case of the line integral of the magnetic field around a loop *s* bounding an open surface *S* is given by

$$\boxed{\int_s \vec{B}.\, d\vec{l} = \mu_0 \int_S \vec{J}.\, d\vec{A} + \mu_0 \varepsilon_0 \int_S \frac{\partial \vec{E}}{\partial t} \cdot d\vec{A}} \tag{3.7}$$

Reading just the first two terms in this equation we see the familiar physics that if a current *I* (given by the current density $\vec{J}$ integrated over the area $d\vec{A}$ of the closed surface S) is flowing through some loop then there is a circulating magnetic field (figure 3.4).

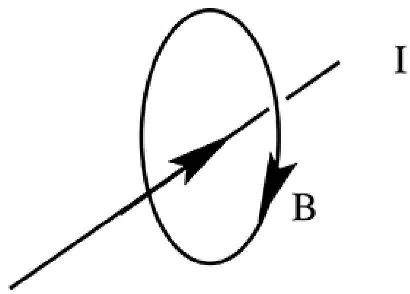


Figure 3.4. The magnetic field induced by a current carrying wire loop.

The final term was added for consistency by Maxwell (we will revisit this shortly) and mirrors the term in Faraday's law.

This integral form of Maxwell's equations are a complete description of electromagnetism. In what follows we shall simply recast the equations in several different ways in order to display their physics content better.

3.2 Differential form of Maxwell's equations

The first rewriting of Maxwell's equations we shall do is to put the equations into a differential equation form. The benefit of this form will be that the equations are true locally at a point. In the integral form one has to pick 'loops' and 'areas' to define the integrals and they are therefore telling you about global properties of a problem. We will need two bits of mathematics (see appendix C for a proof):

Gauss' theorem:

$$\boxed{\int_S \vec{F}.\, d\vec{A} = \int \vec{\nabla} \cdot \vec{F} dV} \tag{3.8}$$

Stokes' theorem:

$$\boxed{\int_s \vec{F}.\, d\vec{l} = (\vec{\nabla} \times \vec{F})\, .\, d\vec{A}} \tag{3.9}$$

We can use these to find the differential form of Maxwell's equations as the following two examples show.

Differential form of Gauss' law

We can now convert the integral form of Gauss' law in equation (3.1)

$$\int_S \vec{E} \cdot d\vec{A} = \frac{q}{\varepsilon_0} \tag{3.10}$$

to the differential form using Gauss' theorem

$$\int_S \vec{E} \cdot d\vec{A} = \int_V \vec{\nabla} \cdot \vec{E}\, dV \tag{3.11}$$

where V is the volume enclosed by the closed surface S. If we also write the charge in terms of a charge density

$$\frac{q}{\varepsilon_0} = \int_V \frac{\rho}{\varepsilon_0} dV \tag{3.12}$$

then comparing the above equations we find

$$\int_V \vec{\nabla} \cdot \vec{E}\, dV = \int_V \frac{\rho}{\varepsilon_0} dV \tag{3.13}$$

Then shrinking the volume V to a point the integrands may be equated to yield Gauss's law in differential form,

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \tag{3.14}$$

Differential form of Faraday's law

We now convert the integral form of Faraday's law in equation (3.6)

$$\int_s \vec{E} \cdot d\vec{l} = -\int_S \frac{\partial \vec{B}}{\partial t} . d\vec{A} \tag{3.15}$$

to the differential form using Stokes' theorem

$$\int_s \vec{E} \cdot d\vec{l} = \int_S (\vec{\nabla} \times \vec{E}) . d\vec{A} \tag{3.16}$$

Equating the right-hand sides of the above equations

$$\int_S (\vec{\nabla} \times \vec{E}).\, d\vec{A} = -\int_S \frac{\partial \vec{B}}{\partial t}.\, d\vec{A}$$

(3.17)

Then, shrinking the surface to a point, we may equate the integrands of the last equation and hence arrive at the differential form of Faraday's law,

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

(3.18)

3.2.1 Maxwell's equations in differential form

Using Gauss' theorem and Stokes' theorem we have rewritten the Maxwell equations for Gauss's law and Faraday's law as the differential equations (3.14) and (3.18). It is straightforward to do the same for the analogous equations for the magnetic field. Then all four Maxwell equations in the vacuum may be expressed in differential form:

$$\boxed{\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \tfrac{\rho}{\varepsilon_0} \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{E} &= -\tfrac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \times \vec{B} &= \mu_0 \vec{J} + \mu_0 \varepsilon_0 \tfrac{\partial \vec{E}}{\partial t} \end{aligned}}$$

(3.19)

Exercise 3.1: How many components do the following nine objects have?

$$\begin{array}{lll} \vec{\nabla} & \nabla^2 \phi & \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) \\ \vec{\nabla}\phi & \nabla^2 \vec{A} & \partial_\mu F^{\mu\nu} \\ \vec{\nabla} \cdot \vec{A} & \nabla \times \vec{A} & \partial^\mu F^{\nu\lambda} \end{array}$$

Exercise 3.2: The vector $\vec{A} = x\hat{\mathbf{i}} + xy\hat{\mathbf{j}} + xz^3\hat{\mathbf{k}}$. Evaluate $\vec{\nabla} \cdot \vec{A}$ and $\vec{\nabla} \times \vec{A}$.

3.2.2 Conservation of charge

Another equation it is useful to put into differential form is that describing charge conservation. Since charge is conserved the current flowing out through the surface of some volume must give the change in charge within the volume (see figure 3.5)

$$\int_S \vec{J} \cdot d\vec{A} = -\int \frac{\partial \rho}{\partial t} dV$$

(3.20)

Applying Gauss' divergence theorem to the left-hand side we have the differential form for charge and current conservation,

$$\boxed{\vec{\nabla} \cdot \vec{J} = -\frac{\partial \rho}{\partial t}}$$

(3.21)

Exercise 3.3: Prove the conservation of charge starting from Maxwell's equations in differential form.

Exercise 3.4: Consider flow within a gas or fluid of density $\rho(\mathbf{r})$. Show that conservation of mass within some volume implies

$$-\frac{d}{dt}\int \rho \, dV = \int \rho\vec{v} \cdot d\vec{A}$$

where $v(\mathbf{r})$ is the flow velocity. By explicitly applying this equation to an infinitesimal volume show that

$$\frac{\partial}{\partial t}\rho + \vec{\nabla} \cdot (\rho\vec{v}) = 0$$

If the velocity of the fluid is subject to the condition $\vec{\nabla} \times \vec{v} = 0$ show, using Stokes' theorem, that the fluid does not support circulation.

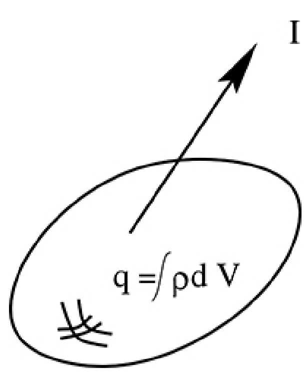


Figure 3.5. Charge leaving a volume as a current.

3.2.3 The displacement current

Prior to Maxwell's involvement the fourth 'Maxwell' equation was just

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} \tag{3.22}$$

However, we can see quite simply in this formalism that this cannot be correct. This is because it is true that for any vector field $\vec{F}$ (The proof is given in appendix D).

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) \equiv 0 \tag{3.23}$$

Let's see if this makes sense for our equation above by taking the divergence

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) \equiv 0 = \mu_0 \vec{\nabla} \cdot \vec{J} \tag{3.24}$$

But this isn't correct since we just saw that $\vec{\nabla} \cdot \vec{J} = -\frac{\partial \rho}{\partial t}$! Maxwell's extra term corrects things as we can see

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) \equiv 0 = \mu_0 \vec{\nabla} \cdot \vec{J} + \mu_0 \varepsilon_0 \frac{\partial}{\partial t} \vec{\nabla} \cdot \vec{E} \tag{3.25}$$

Using the first Maxwell equation ($\vec{\nabla} \cdot \vec{E} = \rho/\varepsilon_0$) we recover the correct formula for the conservation of charge and current in equation (3.21).

Exercise 3.5: Prove the following vector identities

$$\vec{\nabla} \times (\vec{\nabla}\phi) = 0$$

$$\vec{\nabla} \times (\phi\vec{A}) = \phi(\vec{\nabla} \times \vec{A}) + (\vec{\nabla}\phi) \times \vec{A}$$

3.3 Potentials

Potentials are a mathematical trick for making the Maxwell's equations easier to solve. The one you are already familiar with is the electrostatic potential, which we shall discuss first.

3.3.1 Electrostatic potential

In electrostatic problems the Maxwell equations reduce to

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \quad \vec{\nabla} \times \vec{E} = \vec{0} \tag{3.26}$$

If we write

$$\boxed{\vec{E} = -\vec{\nabla}\phi} \tag{3.27}$$

then, because of the identity (see appendix D)

$$\vec{\nabla} \times \vec{\nabla}\phi \equiv 0 \tag{3.28}$$

the second of our two Maxwell equations is automatically satisfied. We are left with only Poisson's equation

$$-\nabla^2\phi = \frac{\rho}{\varepsilon_0} \tag{3.29}$$

This simplifies things since the equation only involves one scalar function rather than the three components of the electric field. The electric field can then readily be obtained from the scalar potential using equation (3.27).

By integrating equation (3.27), we can write

$$\phi = -\int_{\infty}^{\vec{x}} \vec{E} \cdot d\vec{l} \tag{3.30}$$

which shows that ϕ can be interpreted as the 'potential energy' for moving a unit charge from infinity to the point $\vec{x}$. This energy is independent of the path the charge takes to arrive at that point.

Note that ϕ is only defined upto an arbitrary constant (the energy of a charge at infinity) since

$$\vec{E} = -\vec{\nabla}(\phi + C) = -\vec{\nabla}\phi \tag{3.31}$$

Example 1: infinite parallel plate capacitor Consider the capacitor with a potential difference of V across it (figure 3.6).
Between the parallel planes of the plates there is no charge so Poisson's equation reduces to Laplace's equation

$$\nabla^2\phi = 0 \tag{3.32}$$

In this problem, by the symmetry of the assumed infinite capacitor plates, the only variation in ϕ will be in the x direction defined to be perpendicular to the plates,

$$\nabla^2\phi = \frac{d^2}{dx^2}\phi = 0 \tag{3.33}$$

Integrating twice we obtain

$$\phi = Ax + C \tag{3.34}$$

with A, C constants. They can be fixed by imposing the boundary conditions $\phi(x=0)=0,\ \ \phi(x=d)=V$. We obtain

$$\phi = \frac{V}{d}x \tag{3.35}$$

Finally, we can obtain the electric field from the potential

$$\vec{E} = -\vec{\nabla}\phi = \left(-\frac{V}{d}, 0, 0\right) \tag{3.36}$$

Example 2: co-axial cable The case of a co-axial cable is a similar problem but with different symmetry properties (figure 3.7).

Here the potential will only vary radially. In appendix D ∇^2 is calculated in cylindrical polar coordinates (r, θ, z). Only allowing r variation in ϕ we find

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\phi}{dr}\right) = 0 \tag{3.37}$$

Integrating twice we find

$$\phi = A\ \ln\ r + C \tag{3.38}$$

Again we fix the integration constants from the boundary conditions shown in the figure, so

$$\phi(r) = -\frac{V}{\ln(b/a)}\left(\ln\, r - \ln\, b\right) \tag{3.39}$$

Exercise 3.6: Solve Laplace's equation for the potential generated by a charged point particle. The operator ∇^2 in spherical polar coordinates is given by

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\cot\,\theta}{r^2}\frac{\partial}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial \phi^2}$$

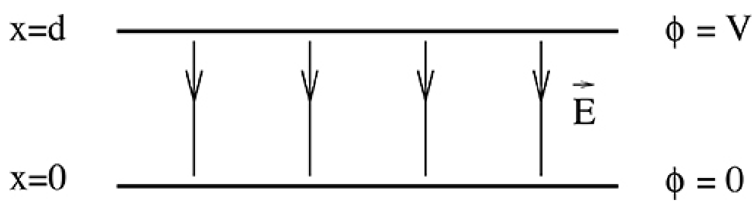


Figure 3.6. A parallel plate capacitor.

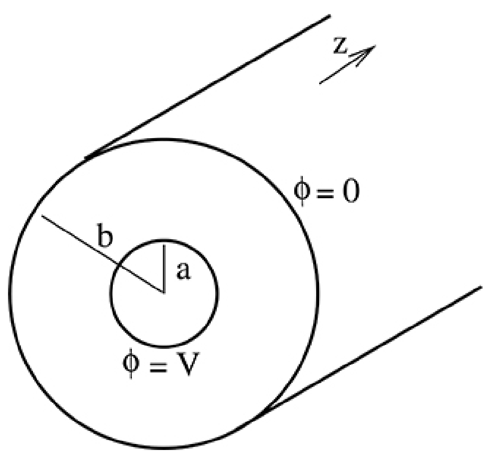


Figure 3.7. A coaxial cable.

3.3.2 The magnetic vector potential

Having introduced an electrostatic scalar potential we might try to introduce a magnetostatic scalar potential in the same way. This does not work though because even in static magnetic problems there must be a current to generate the magnetic field. Thus

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} \tag{3.40}$$

and we cannot use a scalar potential field since $\vec{\nabla} \times \vec{\nabla}\phi \equiv 0$.

On the other hand for all magnetic fields, static or otherwise,

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{3.41}$$

and so we can make use of the alternative identity $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) \equiv 0$.

Thus we can automatically solve the Maxwell equation in equation (3.41) provided we write the magnetic field in terms of a new vector field, the 'vector potential' $\vec{A}$

$$\boxed{\vec{B} = \vec{\nabla} \times \vec{A}} \tag{3.42}$$

Just as there was some freedom in the choice of the electrostatic potential so there is an arbitrariness about the vector potential $\vec{A}$. This is because the magnetic field $\vec{B}$ is left invariant if we transform

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla}\psi\,(x) \tag{3.43}$$

where $\psi(x)$ is an arbitrary scalar field. The invariance of $\vec{B}$ follows from the identity $\vec{\nabla} \times (\vec{\nabla}\psi) \equiv 0$. In other words the same magnetic field $\vec{B}$ results from using either $\vec{A}$ or $\vec{A} + \vec{\nabla}\psi(x)$. Of course trading in one vector field $\vec{B}$ for another $\vec{A}$ does not bring any simplification. However, the winning card for this approach is that, unlike the electrostatic scalar potential ϕ, the magnetic vector potential $\vec{A}$ defined in equation (3.42) is valid for both static and time-varying fields.

3.3.3 A new electric potential

The electrostatic potential we wrote before only worked when there were no time dependent fields, and hence there were no magnetic fields appearing in equations (3.26). Allowing time dependent (non-static) fields means that the electric and magnetic fields are no longer decoupled and in particular the electrostatic potential in equation (3.27) is no longer consistent since it always implies $\vec{\nabla} \times \vec{E} = \vec{0}$. Instead we seek a new potential that implies $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$.

Can we find simultaneous potentials for both $\vec{E}$ and $\vec{B}$ that work in all circumstances? The desired potentials are:

$$\boxed{\begin{aligned} &\vec{E} = -\vec{\nabla}\phi - \tfrac{\partial \vec{A}}{\partial t} \\ &\vec{B} = \vec{\nabla} \times \vec{A} \end{aligned}} \tag{3.44}$$

The second of these equations defines the usual magnetic vector potential which is always valid even in the non-static case since

$$\vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0 \tag{3.45}$$

The first of these equations involves a new, second term on the right-hand side which is designed to yield the correct Maxwell equation, $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$. This is easily seen by taking the curl of the first equation,

$$\begin{aligned} \vec{\nabla} \times \vec{E} = \vec{\nabla} \times \left(-\vec{\nabla}\phi - \tfrac{\partial \vec{A}}{\partial t}\right) \\ &= -\vec{\nabla} \times \left(\vec{\nabla}\phi\right) - \tfrac{\partial(\vec{\nabla} \times \vec{A})}{\partial t} \\ &= -\tfrac{\partial \vec{B}}{\partial t} \end{aligned} \tag{3.46}$$

To summarize, potentials $\vec{A}$ and ϕ may be defined in equation (3.44) which always automatically satisfy the homogeneous Maxwell equations $\vec{\nabla} \cdot \vec{B} = 0$ and $\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$. This should simplify things greatly since now there are only the remaining two inhomogeneous Maxwell equations to solve. Let's write them out in terms of the potentials

$$\vec{\nabla} \cdot \vec{E} = \boxed{-\nabla^2\phi - \tfrac{d(\vec{\nabla} \cdot \vec{A})}{dt} = \tfrac{\rho}{\varepsilon_0}} \tag{3.47}$$

For the $\vec{\nabla} \times \vec{B}$ equation we will again use the identity for this product in appendix D. Thus

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \nabla^2\vec{A} = \mu_0\vec{J} + \mu_0\varepsilon_0\frac{\partial}{\partial t}\left(-\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi\right)$$

or rearranging (3.48)

$$\boxed{-\nabla^2\vec{A} + \mu_0\varepsilon_0\frac{\partial^2\vec{A}}{\partial t^2} = \mu_0\vec{J} - \vec{\nabla}\left(\vec{\nabla}\cdot\vec{A} + \mu_0\varepsilon_0\frac{\partial\phi}{\partial t}\right)}$$ (3.49)

Unfortunately, the two equations we are left with are quite messy! To clean them up we can make use of our ability to redefine the potentials while keeping the $\vec{E}, \vec{B}$ fields the same.

3.3.4 Gauge transformations

The transformations for these potentials that leave $\vec{E}, \vec{B}$ invariant are the following *gauge transformations*

$$\boxed{\begin{array}{l}\vec{A} \to \vec{A} + \vec{\nabla}\psi \\ \phi \to \phi - \frac{\partial\psi}{\partial t}\end{array}}$$ (3.50)

Exercise 3.7: Show explicitly that the $\vec{E}, \vec{B}$ fields are left invariant by these transformations.

We can make a choice of gauge that transforms $\vec{\nabla}.\vec{A}$ as follows

$$\vec{\nabla}\cdot\vec{A} \to \vec{\nabla}\cdot(\vec{A} + \nabla\psi) = \vec{\nabla}\cdot\vec{A} + \nabla^2\psi$$ (3.51)

Note that $\vec{\nabla}\cdot\vec{A}$ is a number at each point in space. $\nabla^2\psi$ is also a number at each point but here we get to choose it by choosing ψ. The upshot is that we can choose to transform $\vec{\nabla}\cdot\vec{A}$ to anything we want!

$\vec{\nabla}\cdot\vec{A} = 0$ is one sensible choice (known as Coulomb gauge). Another choice, which we shall focus on below, is called Lorenz gauge.

3.3.5 Maxwell's equations in Lorenz gauge

Let's choose to make a gauge transformation designed to cancel the second term on the right-hand side of equation (3.49),

$$\vec{\nabla}\cdot\vec{A} = -\mu_0\varepsilon_0\frac{\partial\phi}{\partial t}$$ (3.52)

In this gauge the two inhomogeneous Maxwell equations in equations (3.47) and (3.49) simplify to

$$\boxed{-\nabla^2\phi + \mu_0\varepsilon_0\frac{\partial^2\phi}{\partial t^2} = \frac{\rho}{\varepsilon_0}} \tag{3.53}$$

$$\boxed{-\nabla^2\vec{A} + \mu_0\varepsilon_0\frac{\partial^2\vec{A}}{\partial t^2} = \mu_0\vec{J}} \tag{3.54}$$

This form of the inhomogeneous Maxwell's equations is much prettier!

Observe the following two points:

Wave equations in free space

In free space $\vec{J} = 0$ and $\rho = 0$ and these equations become wave equations

$$-\nabla^2\phi + \mu_0\varepsilon_0\frac{\partial^2\phi}{\partial t^2} = 0, \quad -\nabla^2\vec{A} + \mu_0\varepsilon_0\frac{\partial^2\vec{A}}{\partial t^2} = 0 \tag{3.55}$$

which have complex wave solutions of the form

$$\vec{A}(\vec{r}, t) = \vec{A}_0 e^{i\left(\omega t - \vec{k}.\vec{r}\right)} \tag{3.56}$$

The physical solutions are obtained by taking the real part of the complex equation. Substituting the complex (or real) solution into the wave equation we find the condition

$$\frac{\omega^2}{k^2} = c^2 = \frac{1}{\mu_0\varepsilon_0} \tag{3.57}$$

In other words these waves move at a speed $c = 1/\sqrt{\mu_0\varepsilon_0}$ which is the speed of light.

Following a similar analysis for the electric and magnetic fields, this is how Maxwell concluded that light is an electromagnetic wave.

Relativistic form

Equations (3.53) and (3.54) also have a very suggestive form for relativity—they are symmetric in time and space. There's also a symmetry between the components of $\vec{A}$ and ϕ—should we promote them to the components of a four-vector? Similarly should the charge density and current become a four-vector?

3.4 Relativistic formulation of electromagnetism

Our goal now is to cast Maxwell’s equations in a manifestly Lorentz invariant form which is compatible with the second postulate of special relativity. The equations in Lorentz gauge in equations (3.53) and (3.54) suggested a four-vector form which we will now explore.

3.4.1 Four-vector current

Consider a uniform distribution of charge in a volume *V* at rest in some frame (figure 3.8).

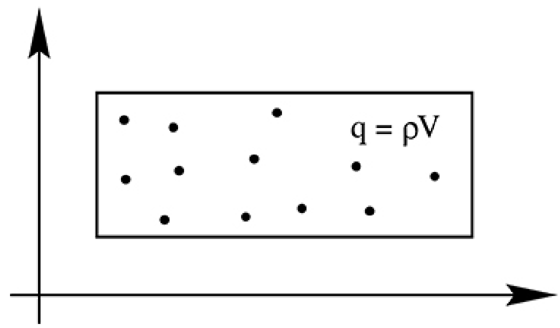


Figure 3.8. Charges distributed in a volume.

If the charge density is ρ_0 then the total charge is $\rho_0 V$.

Now consider boosting to a frame moving with speed *v* relative to the charge. The volume changes because of Lorentz contraction

$$V\prime = \frac{V}{\gamma} \tag{3.58}$$

The total number of charges in the box must be the same for each observer though so the charge density must also change to keep the total charge fixed. Thus

$$\rho\prime = \gamma\rho_0 \tag{3.59}$$

There will also now be a current density since the charges are moving in the new inertial frame. These transformations are all consistent with ρ and $\vec{J}$ being a four-vector.

Thus we define

$$\boxed{J^\mu = (\rho c, \vec{J})} \tag{3.60}$$

Classically the current density is just given in terms of the speed of the particles as $\rho\vec{v}$.

The natural relativistic definition is therefore

$$\boxed{J^\mu = \rho_0 u^\mu} = \rho_0 \frac{dx^\mu}{d\tau} \tag{3.61}$$

The Lorentz invariant ‘length’ of the four-vector then follows from $u^\mu u_\mu = c^2$

$$J^\mu J_\mu = \rho_0^2 c^2 \tag{3.62}$$

Exercise 3.8: Write equations for how each of the four components of the J^μ transform under a Lorentz boost by *v* in the *x*-direction.

3.4.2 Conservation of charge

The conservation of charge equation

$$\vec{\nabla} \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0 \tag{3.63}$$

can now be written in a Lorentz invariant form

$$\partial^\mu J_\mu = 0 \tag{3.64}$$

where

$$\partial^\mu = \left(\frac{\partial}{\partial x^0}, -\vec{\nabla}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right) \tag{3.65}$$

Note the minus sign in the definition of the relativistic derivative four-vector ∂^μ. It looks a bit odd but is needed to get the signs correct here. In fact it is the only prescription compatible with the usual definition of x^μ as we show in the next section.

3.4.3 The four-vector ∂^μ

You might worry that defining

$$\partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right) \tag{3.66}$$

with a minus sign contradicts the fact that

$$x^\mu = (ct, \mathbf{x}) \tag{3.67}$$

For example under a Lorentz boost to a frame moving with speed *v* in the positive *x* direction

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{3.68}$$

i.e.

$$(ct') = \gamma(ct) - \frac{v}{c}\gamma x, \quad x' = \gamma x - \frac{v}{c}\gamma(ct) \tag{3.69}$$

or inverting the relations

$$(ct) = \gamma(ct') + \frac{v}{c}\gamma x', \quad x = \gamma x' + \frac{v}{c}\gamma(ct') \tag{3.70}$$

Similarly the definition in equation (3.66) would imply

$$\partial'^\mu = \Lambda^\mu{}_\nu\, \partial^\nu \tag{3.71}$$

i.e.

$$\frac{1}{c}\frac{\partial}{\partial t'} = \gamma\frac{1}{c}\frac{\partial}{\partial t} + \frac{v}{c}\gamma\frac{\partial}{\partial x}, \;\; -\frac{\partial}{\partial x'} = -\gamma\frac{\partial}{\partial x} - \frac{v}{c}\gamma\frac{1}{c}\frac{\partial}{\partial t} \tag{3.72}$$

Note the signs in the transformations.

To show this is consistent let's work it out from first principles:

$$\frac{\partial}{\partial t'} = \frac{\partial x}{\partial t'}\frac{\partial}{\partial x} + \frac{\partial t}{\partial t'}\frac{\partial}{\partial t} \tag{3.73}$$

$$\frac{\partial}{\partial x\prime} = \frac{\partial x}{\partial x\prime}\frac{\partial}{\partial x} + \frac{\partial t}{\partial x\prime}\frac{\partial}{\partial t} \tag{3.74}$$

from the transformations in equation (3.70) above

$$\frac{\partial x}{\partial t\prime} = v\gamma, \quad \frac{\partial t}{\partial t\prime} = \gamma, \quad \frac{\partial t}{\partial x\prime} = \frac{v}{c^2}\gamma, \quad \frac{\partial x}{\partial x\prime} = \gamma \tag{3.75}$$

Substituting these in equations (3.73) and (3.74) we find equation (3.72)—this shows that there is not an inconsistency (and in fact that the minus sign in equation (3.66) is required).

An alternative quicker statement of this is that one would like

$$\frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \quad \text{or} \quad \partial_\mu x^\mu = 4 \tag{3.76}$$

and again the minus sign in ∂^μ is required.

We can also define a four-vector version of ∇^2 by

$$\partial^\mu \partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{3.77}$$

3.4.4 Four-vector potential

One of the advantages of the relativistic formulation is that we understand how electric and magnetic fields behave under boosts. As an example let's look at the fields around a moving electric charge.

For an electric charge at rest we know that

$$A^\mu = \left(\frac{\phi}{c}, \vec{A}\right) = \left(\frac{q}{4\pi\varepsilon_0 rc}, \vec{0}\right) \tag{3.78}$$

We can make the charge move by boosting to an inertial frame moving at speed *v* in the positive *x* direction

$$A\prime^\mu = \Lambda^\mu{}_\nu A^\nu \tag{3.79}$$

so for example

$$A\prime^0 = \gamma\left(A^0 - \frac{v}{c}A^x\right)$$

which means (3.80)

$$\phi\prime = \frac{\gamma q}{4\pi\varepsilon_0 r\prime}.$$

We must remember also that $r^2 = x^2 + y^2 + z^2$ and *x* transforms too, so (3.81)

$$\phi\prime = \frac{\gamma q}{4\pi\varepsilon_0\left(\gamma^2(x\prime + vt\prime)^2 + y\prime^2 + z\prime^2\right)^{1/2}}$$

Turning to the spatial components we find (3.82)

$$A\prime^x = -\gamma\frac{v}{c}A^0 = -\frac{\gamma v}{c^2}\frac{q}{4\pi\varepsilon_0\left(\gamma^2(x\prime + vt\prime)^2 + y\prime^2 + z\prime^2\right)^{1/2}}$$

and $A'^y = A'^z = 0$. (3.83)

The electric field is then given by

$$\vec{E}\prime = -\vec{\nabla}\prime\phi\prime - \frac{\partial\vec{A}\prime}{\partial t\prime}$$

which works through to (3.84)

$$\begin{aligned} E\prime^x &= \frac{q\gamma}{4\pi\varepsilon_0}\frac{(x\prime + vt\prime)}{\left(\gamma^2(x\prime + vt\prime)^2 + y\prime^2 + z\prime^2\right)^{3/2}} \\ E\prime^y &= \frac{q\gamma}{4\pi\varepsilon_0}\frac{y\prime}{\left(\gamma^2(x\prime + vt\prime)^2 + y\prime^2 + z\prime^2\right)^{3/2}} \\ E\prime^z &= \frac{q\gamma}{4\pi\varepsilon_0}\frac{z\prime}{\left(\gamma^2(x\prime + vt\prime)^2 + y\prime^2 + z\prime^2\right)^{3/2}} \end{aligned}$$

These results are particularly interesting when $v \simeq c$. Look first on the *x*-axis (3.85)

$$E\prime^x \simeq \frac{q}{4\pi\varepsilon_0\gamma^2(x\prime + vt\prime)^2}$$

since *γ* is large this component of the field is reduced relative to that of the stationary charge. On the other hand if we look at the field perpendicular to the (3.86) motion (i.e. at $x\prime = -vt\prime$) $E\prime^y, E\prime^z$ are both enlarged by a factor of γ. Thus the field of a relativistic moving charge is essentially confined to a disc (see figure 3.9).

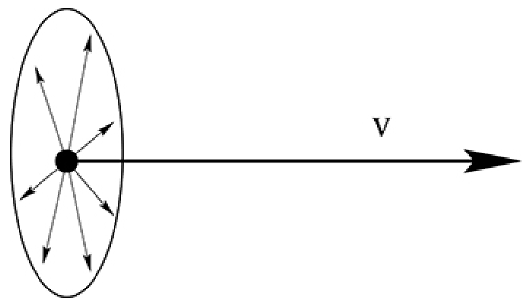


Figure 3.9. The electric field of a highly relativistic charge is confined to a plane transverse to its motion.

3.4.5 The electromagnetic field strength tensor

It is also possible to write Maxwell's equations in a relativistic form involving $\vec{E}$ and $\vec{B}$ fields rather than the potentials. Remember that

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \tag{3.87}$$

so a component is given in terms of A^μ, ∂^ν by

$$\frac{E^i}{c} = \partial^i A^0 - \partial^0 A^i \tag{3.88}$$

Similarly

$$\vec{B} = \vec{\nabla} \times \vec{A} \tag{3.89}$$

so, up to signs, we have the form

$$B^i = \partial^j A^k - \partial^k A^j \tag{3.90}$$

Thus we conclude that the $\vec{E}$ and $\vec{B}$ fields are described by the *electromagnetic field strength tensor*

$$\boxed{F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu} \tag{3.91}$$

Explicitly the components are

$$F^{\mu\nu} = \begin{pmatrix} 0 & -\frac{E^1}{c} & -\frac{E^2}{c} & -\frac{E^3}{c} \\ \frac{E^1}{c} & 0 & -B^3 & B^2 \\ \frac{E^2}{c} & B^3 & 0 & -B^1 \\ \frac{E^3}{c} & -B^2 & B^1 & 0 \end{pmatrix} \tag{3.92}$$

where μ counts the row and ν the column.

Maxwell's equations in terms of $F^{\mu\nu}$ are a little involved and are given by

$$\boxed{\partial_\mu F^{\mu\nu} = \mu_0 J^\nu} \tag{3.93}$$

and

$$\boxed{\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0} \tag{3.94}$$

For example the first equation contains ($\nu = 0$) $\vec{\nabla} \cdot \vec{E} = \rho/\varepsilon_0$ and ($\nu = 1, 2, 3$) $\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$.

The second equation is actually 64 equations so contains many repeats of the remaining two Maxwell equations. For example if we set $\lambda = 1, \mu = 3, \nu = 2$ we obtain $\vec{\nabla} \cdot \vec{B} = 0$ and so forth.

Exercise 3.9: Explicitly extract the differential form of Maxwell's equations from equations (3.93) and (3.94).

Exercise 3.10: Evaluate $F^{\mu\nu} F_{\mu\nu}$ in terms of $\vec{E}$ and $\vec{B}$ fields.

Exercise 3.11: There is a four-component tensor $\varepsilon^{\mu\nu\rho\sigma}$ which is zero if any two indices take the same value. $\varepsilon^{1234} = 1$. Other non-zero components are obtained by interchanging indices of ε^{1234}—interchanging any two indices changes the value by a minus sign. Thus $\varepsilon^{3214} = -1$ while $\varepsilon^{2314} = 1$. Show explicitly that ε^{1234} and ε^{1134} are left invariant by a Lorentz boost.
Show that $\varepsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$ takes the same form as $F^{\mu\nu}$ but with the elctric and magnetic field components interchanged.

Hence, evaluate in terms of $\vec{E}$ and $\vec{B}$ fields the Lorentz invariant quantity $\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$.

3.4.6 Lorentz transformations of electric and magnetic fields

We can calculate the Lorentz transformation properties of the $\vec{E}$ and $\vec{B}$ fields using the fact that $F^{\mu\nu}$ transforms as

$$F'^{\mu\nu} = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}F^{\alpha\beta} \tag{3.95}$$

For example for a boost by speed *v* in the positive *z* direction

$$\begin{aligned}
\frac{E'^1}{c} &= F'^{10} \\
&= \Lambda^{1}{}_{\alpha}\Lambda^{0}{}_{\beta}F^{\alpha\beta} \\
&= \Lambda^{0}{}_{\beta}\left(\Lambda^{1}{}_{0}F^{0\beta} + \Lambda^{1}{}_{1}F^{1\beta} + \Lambda^{1}{}_{2}F^{2\beta} + \Lambda^{1}{}_{3}F^{3\beta}\right) \\
&= \Lambda^{0}{}_{\beta}F^{1\beta} \\
&= \Lambda^{0}{}_{0}F^{10} + \Lambda^{0}{}_{1}F^{11} + \Lambda^{0}{}_{2}F^{12} + \Lambda^{0}{}_{3}F^{13} \\
&= \gamma\left(\frac{E^1}{c} - \frac{v}{c}B^2\right)
\end{aligned} \tag{3.96}$$

The full set of transformations are given by

$$\boxed{\begin{aligned}
\frac{E'^1}{c} &= \gamma\left(\frac{E^1}{c} - \frac{v}{c}B^2\right) \\
\frac{E'^2}{c} &= \gamma\left(\frac{E^2}{c} + \frac{v}{c}B^1\right) \\
\frac{E'^3}{c} &= \frac{E^3}{c} \\
B'^1 &= \gamma\left(B^1 + \frac{v}{c}\frac{E^2}{c}\right) \\
B'^2 &= \gamma\left(B^2 - \frac{v}{c}\frac{E^1}{c}\right) \\
B'^3 &= B^3
\end{aligned}} \tag{3.97}$$

3.4.7 The relativistic force law

When we were studying relativity in chapter 2 we promised to return to the idea of relativistic force when we had studied electromagnetism.

Classically the electromagnetic force is given by

$$\vec{F} = q\left(\vec{E} + \vec{v} \times \vec{B}\right) \tag{3.98}$$

Thus for example the *x* component is given by

$$F^1 = q(E^1 + v^2 B^3 - v^3 B^2) = q(cF^{10} - v^2 F^{12} - v^3 F^{13}) \tag{3.99}$$

to make this more symmetric we can add $-v^1 F^{11}$ since this is just zero!

Since (c, v^1, v^2, v^3) are just the non-relativistic limit of u^μ we are led to

$$\boxed{f^\mu = q u_\nu F^{\mu\nu}} \tag{3.100}$$

Now we can ask what the non-relativistic limit of the time-like component of force is?

$$\begin{aligned} f^0 &= q\left(u_0 F^{00} - u_1 F^{10} - u_2 F^{20} - u_3 F^{30}\right) \\ &= -q\gamma \frac{\vec{v} \cdot \vec{E}}{c} \\ &= -\frac{q\gamma}{c} \vec{v} \cdot (\vec{E} + \vec{v} \times \vec{B}) \end{aligned} \tag{3.101}$$

where we have used that $\vec{v} \cdot (\vec{v} \times \vec{B}) \equiv 0.$

Taking $v \ll c$ we obtain $q\vec{v} \cdot \vec{E}$ which is just the work done per second. This indeed should be the rate of change of energy and it makes sense to equate it to $\frac{dp^0}{d\tau}$ in the relativistic generalization of Newton's law.

3.5 The Lagrangian for a charged particle

The equation of motion for a charged, moving particle is given by (for the moment we return to the non-relativistic notation)

$$\frac{d\vec{p}}{dt} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{3.102}$$

The action that reproduces this equation is

$$S = \int L \, dt, \quad L = \frac{1}{2} m \mid \vec{\dot{x}} \mid^2 + q(\vec{\dot{x}} \cdot \vec{A}) - q\phi \tag{3.103}$$

The Euler–Lagrange equation is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \vec{\dot{x}}}\right) - \frac{\partial L}{\partial \vec{x}} = 0 \tag{3.104}$$

or

$$\frac{d}{dt}(m\vec{\dot{x}} + q\vec{A}) - \vec{\nabla}(q\vec{\dot{x}} \cdot \vec{A} - q\phi) = 0 \tag{3.105}$$

To see this is the equation we want we must first be careful about the time dependence of $\vec{A}$. Of course it can explicitly depend on time, but even if it's constant the particle, as it moves, will see a time variation of the field. This is accounted for using the chain rule

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{dx}{dt}\frac{\partial}{\partial x} + \frac{dy}{dt}\frac{\partial}{\partial y} + \frac{dz}{dt}\frac{\partial}{\partial z} = \frac{\partial}{\partial t} + \vec{\dot{x}} \cdot \vec{\nabla} \tag{3.106}$$

So our equation of motion is

$$\frac{d\vec{p}}{dt} + q\frac{\partial \vec{A}}{\partial t} + q\vec{\dot{x}} \cdot \vec{\nabla}\vec{A} - q\nabla(\vec{\dot{x}} \cdot \vec{A}) + q\vec{\nabla}\phi = 0 \tag{3.107}$$

Next we use the identity

$$\vec{\dot{x}} \times \vec{\nabla} \times \vec{A} = \vec{\nabla}(\vec{\dot{x}} \cdot \vec{A}) - (\vec{\dot{x}} \cdot \vec{\nabla})\vec{A} \tag{3.108}$$

We have

$$\frac{d\vec{p}}{dt} = q\left(-\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi\right) + q\vec{\dot{x}} \times \vec{\nabla} \times \vec{A} \tag{3.109}$$

Finally we remember the form for the electric and magnetic field in terms of the potentials equation (3.44) and see that this is precisely the equation of motion

(3.102) we wanted!

Note also the expressions for the generalized momenta

$$\vec{p}_{\text{gen}} = \frac{\partial L}{\partial \dot{\vec{x}}} = m\dot{\vec{x}} + e\vec{A} \tag{3.110}$$

and for the Hamiltonian

$$H = \vec{p}_{\text{gen}} \cdot \dot{\vec{x}} - L = \frac{1}{2}m\left|\,\dot{\vec{x}}\,\right|^2 + e\phi \tag{3.111}$$

These expressions combine to the generalized four-vector momentum

$$p^{\mu}_{\text{gen}} = mu^{\mu} + eA^{\mu} \tag{3.112}$$

Replacing momenta in a problem by this generalized four-momenta is called ‘minimal substitution’.

Appendix C. Gauss’ and Stokes’ theorems

Here are derivations of these two crucial theorems.

C.1 Gauss’ theorem

We want to convert the surface integral

$$\int_S \vec{F} \cdot d\vec{A} \tag{C.1}$$

to a form that is locally true. We do this by calculating the integral for an infinitesimal cubic volume (see figure C.1).

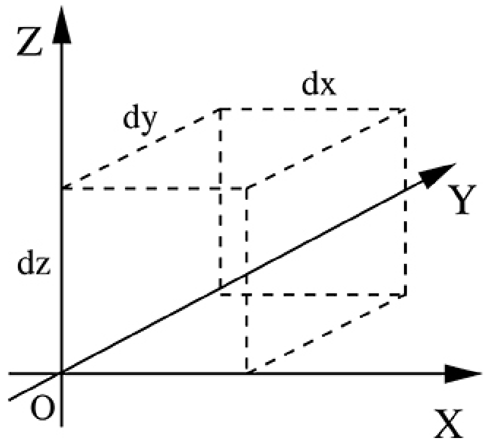


Figure C.1. An infinitesimal cube.

We choose the surface in the integral as the surface of this cube. As an example let's take

$$\vec{F} = F\hat{\mathbf{z}} \tag{C.2}$$

(i.e. the field $\vec{F}$ points in the *z* direction.)

Calculating the surface integral for this $\vec{F}$:

$$\begin{aligned}\text{Flux at bottom surface} &= -F(0)\ \ \delta x\ \delta y \\ \text{Flux at top surface} &= \left(F(0) + \tfrac{\partial F}{\partial z}\delta z\right)\delta x\ \delta y\end{aligned}$$

Here we have Taylor expanded *F* to keep only the leading change in its behaviour as we move in the *z* direction. Note that the top and bottom areas contribute opposite signs because the area vectors point in opposite directions. The other surfaces contribute nothing for this choice of $\vec{F}$. The total integral is therefore

$$\int_S \vec{F} \cdot d\vec{A} = \frac{\partial F}{\partial z}\ \delta x\ \delta y\ \delta z \tag{C.3}$$

This result generalizes, when $\vec{F}$ has *x* and *y* components too, to:

$$\lim \quad \delta \mathrm{V} \to 0 \;\; \frac{\int_S \vec{F} \cdot d\vec{A}}{\delta V} = \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) \tag{C.4}$$

where δV is the volume of the cube.

Alternatively we may write this as

$$\int_S \vec{F} \cdot d\vec{A} = \nabla \cdot \vec{F} \delta V \tag{C.5}$$

where

$$\nabla = \left(\frac{\partial}{\partial x} \hat{\mathbf{x}} + \frac{\partial}{\partial y} \hat{\mathbf{y}} + \frac{\partial}{\partial z} \hat{\mathbf{z}} \right) \tag{C.6}$$

and *dV* is an integral over the whole volume.

Gauss' theorem for extended volumes

It is easy to obtain the equivalent expression for an arbitrary volume—we just build it up out of infinitesimal cubes: e.g. if we put two together (figure C.2).

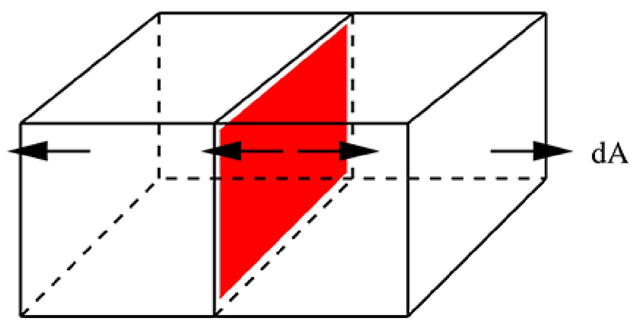


Figure C.2. Two adjacent infinitesimal cubes.

It turns out that

$$\int_{\text{two cubes}} \vec{F} \cdot d\vec{A} = \int_{\text{cube one}} \vec{F} \cdot d\vec{A} + \int_{\text{cube two}} \vec{F} \cdot d\vec{A} \tag{C.7}$$

since the side shared by the two cubes has an area vector with *opposite* sign in the case of the two integrals—the side cancels! We can therefore build any shape in this way and the surface integral is just the sum over the surface integrals of the component cubes so we arrive at Gauss' law

$$\boxed{\int_S \vec{F} \cdot d\vec{A} = \int \vec{\nabla} \cdot \vec{F} dV} \quad \text{(C.8)}$$

C.2 Stokes' theorem

Next we want to convert the line integral

$$\int_S \vec{F} \cdot d\vec{l} \quad \text{(C.9)}$$

to a form that is locally true. We do this by calculating the integral for an infinitesimal rectangular loop (figure C.3).

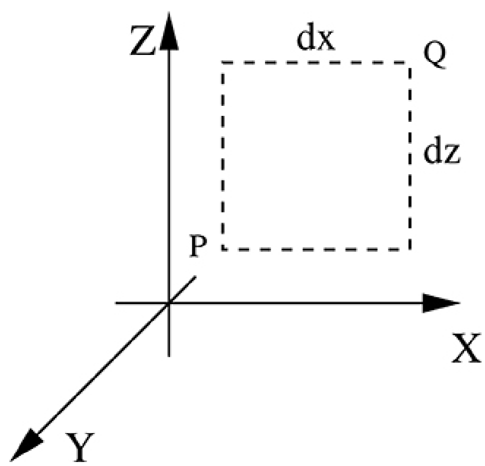


Figure C.3. An infinitesimal square in the *x–z* plane.

We've chosen the loop to lie in the $x - z$ plane.
If at the bottom corner of the rectangle (P)

$$\vec{F} = F_x\hat{\mathbf{x}} + F_y\hat{\mathbf{y}} + F_z\hat{\mathbf{z}} \quad \text{(C.10)}$$

The line integral gets contributions from the top and bottom of the form $F_x\ dx$ and from the sides of the form $F_y\ dy$. However, we must take into account the change in these components across the box. Clockwise round the box we get contributions:

$$\int_S \vec{F}\cdot d\vec{l} = F_z\,\delta z + \left(F_x + \frac{\partial F_x}{\partial z}\delta z\right)\delta x - \left(F_z + \frac{\partial F_z}{\partial x}\delta x\right)\delta z - F_x\,\delta x \qquad \text{(C.11)}$$

$$\int_S \vec{F}\cdot d\vec{l} = \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\delta x\,\delta z \qquad \text{(C.12)}$$

or

$$\int_S \vec{F}\cdot d\vec{l} = c_y\,dA \qquad \text{(C.13)}$$

Note that the area element is in the $\hat{\mathbf{y}}$ direction.

In general c_y is the y-component of a vector called the *curl* of **F**. Its other components are

$$c_x = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) \qquad \text{(C.14)}$$

$$c_z = \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) \qquad \text{(C.15)}$$

We can write the curl as

$$\vec{\nabla}\times\vec{F} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \qquad \text{(C.16)}$$

The calculation above then generalizes, for an area placed at random relative to the axes, to

$$\int_S \vec{F}\cdot d\vec{l} = (\vec{\nabla}\times\vec{F})\,.\,d\vec{A} \qquad \text{(C.17)}$$

Stokes' theorem for extended areas

We can again make larger areas by placing infinitesimal squares next to each other —the common sides cancel from the sum (figure C.4).

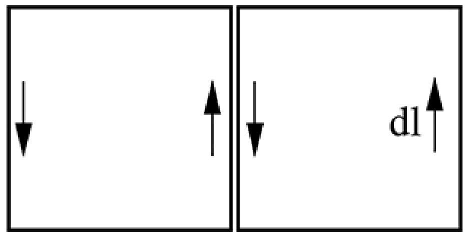


Figure C.4. Two adjacent infinitesimal squares.

Thus

$$\int_{\text{two sq}} \vec{F} \cdot d\vec{l} = \int_{\text{sq one}} \vec{F} \cdot d\vec{l} + \int_{\text{sq two}} \vec{F} \cdot d\vec{l} \tag{C.18}$$

Using our above result we arrive at Stokes' theorem

$$\boxed{\int_S \vec{F} \cdot d\vec{l} = \int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{A}} \tag{C.19}$$

Appendix D. Vector identities

Identity: $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) = 0$
Proof:

$$\begin{aligned} \vec{\nabla} \times \vec{F} &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \\ &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\hat{\mathbf{x}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\hat{\mathbf{y}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\hat{\mathbf{z}} \end{aligned}$$

$$\begin{aligned} \vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) &= \frac{\partial^2 F_z}{\partial x \partial y} - \frac{\partial^2 F_y}{\partial x \partial z} + \frac{\partial^2 F_x}{\partial y \partial z} - \frac{\partial^2 F_z}{\partial y \partial x} + \frac{\partial^2 F_y}{\partial z \partial x} - \frac{\partial^2 F_x}{\partial z \partial y} \\ &= 0 \end{aligned}$$

Identity: $\vec{\nabla} \times (\vec{\nabla}\phi) = 0$
Proof:

$$\begin{aligned}
\vec{\nabla} \times (\vec{\nabla}\phi) &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} & \frac{\partial \phi}{\partial z} \end{vmatrix} \\
&= \left(\frac{\partial^2 \phi}{\partial y \partial z} - \frac{\partial^2 \phi}{\partial z \partial y} \right) \hat{\mathbf{x}} + \left(\frac{\partial^2 \phi}{\partial z \partial x} - \frac{\partial^2 \phi}{\partial x \partial z} \right) \hat{\mathbf{y}} + \left(\frac{\partial^2 \phi}{\partial x \partial y} - \frac{\partial^2 \phi}{\partial y \partial x} \right) \hat{\mathbf{z}} \\
&= 0
\end{aligned}$$

Identity: $\vec{\nabla} \times (\vec{\nabla} \times \vec{F}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{F}) - \nabla^2 \vec{F}$
Proof:

$$\begin{aligned}
\vec{\nabla} \times \vec{F} &= \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \\
&= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{\mathbf{x}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \hat{\mathbf{y}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{\mathbf{z}}
\end{aligned}$$

$$\begin{aligned}
\vec{\nabla} \times (\vec{\nabla} \times \vec{F}) &= \left(\frac{\partial^2 F_y}{\partial y \partial x} - \frac{\partial^2 F_x}{\partial y^2} - \frac{\partial^2 F_x}{\partial z^2} + \frac{\partial^2 F_z}{\partial z \partial x} \right) \hat{\mathbf{x}} \\
&\quad - \left(\frac{\partial^2 F_y}{\partial x^2} - \frac{\partial^2 F_x}{\partial x \partial y} - \frac{\partial^2 F_z}{\partial z \partial y} + \frac{\partial^2 F_y}{\partial z^2} \right) \hat{\mathbf{y}} \\
&\quad + \left(\frac{\partial^2 F_x}{\partial x \partial z} - \frac{\partial^2 F_z}{\partial x^2} - \frac{\partial^2 F_z}{\partial y^2} + \frac{\partial^2 F_y}{\partial y \partial z} \right) \hat{\mathbf{z}} \\
&= \left[\frac{\partial}{\partial x} \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) - \left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) F_x \right] \hat{\mathbf{x}} \\
&\quad + \left[\frac{\partial}{\partial y} \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) - \left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) F_y \right] \hat{\mathbf{y}} \\
&\quad + \left[\frac{\partial}{\partial z} \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) - \left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) F_z \right] \hat{\mathbf{z}} \\
&= \vec{\nabla}(\vec{\nabla} \cdot \vec{F}) - \nabla^2 \vec{F}
\end{aligned}$$

Identity: In cylinderical polar coordinates (r, θ, z)

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial \theta} + \frac{\partial^2}{\partial z^2}$$

Proof:

$$\begin{aligned}
\nabla^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\
&= \frac{\partial}{\partial x}\left(\frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta}\right) + \frac{\partial}{\partial y}\left(\frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta}\right) + \frac{\partial^2}{\partial z^2} \\
&= \left(\frac{\partial r}{\partial x}\right)^2\frac{\partial^2}{\partial r^2} + \frac{\partial^2 r}{\partial x^2}\frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial x}\right)^2\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2 \theta}{\partial x^2}\frac{\partial}{\partial \theta} + \left(\frac{\partial r}{\partial y}\right)^2\frac{\partial^2}{\partial r^2} + \\
&\quad \frac{\partial^2 r}{\partial y^2}\frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial y}\right)^2\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2 \theta}{\partial y^2}\frac{\partial}{\partial \theta} + \frac{\partial^2}{\partial z^2}
\end{aligned}$$

Now we use the relations between x, y and r, θ:

$$\begin{aligned}
r &= \left(x^2+y^2\right)^{1/2} \quad x = r\ \sin\ \theta \\
\tan\ \theta &= x/y \quad y = r\ \cos\ \theta \\
\frac{\partial r}{\partial x} &= \frac{x}{\left(x^2+y^2\right)^{1/2}} = \sin\ \theta \\
\frac{\partial^2 r}{\partial x^2} &= \frac{1}{\left(x^2+y^2\right)^{1/2}} - \frac{x^2}{\left(x^2+y^2\right)^{3/2}} = \frac{1}{r}\left(1-\sin^2\theta\right) = \frac{\cos^2\theta}{r} \\
\frac{\partial r}{\partial y} &= \frac{y}{\left(x^2+y^2\right)^{1/2}} = \cos\ \theta \\
\frac{\partial^2 r}{\partial y^2} &= \frac{1}{\left(x^2+y^2\right)^{1/2}} - \frac{y^2}{\left(x^2+y^2\right)^{3/2}} = \frac{1}{r}\left(1-\cos^2\theta\right) = \frac{\sin^2\theta}{r} \\
\frac{\partial \theta}{\partial x} &= \frac{\cos^2\theta}{x} = \frac{\cos\ \theta}{r} \\
\frac{\partial^2 \theta}{\partial x^2} &= \frac{2\ \cos\ \theta\ \sin\ \theta}{y}\frac{\partial \theta}{\partial x} = \frac{2\ \cos\ \theta\ \sin\ \theta}{r^2} \\
\frac{\partial \theta}{\partial y} &= -\cos^2\theta\frac{x}{y^2} = -\frac{\sin\ \theta}{r} \\
\frac{\partial^2 \theta}{\partial y^2} &= 2\cos\ \theta\ \sin\ \theta\frac{x}{y^2}\frac{\partial \theta}{\partial y} + 2\cos^2\theta\frac{x}{y^3} = -\frac{2\sin\ \theta}{r^2\ \cos\ \theta}\left(\sin^2\theta-1\right) = -\frac{2\sin\ \theta\ \cos\ \theta}{r^2}
\end{aligned}$$

Substituting in the above we find directly

$$\begin{aligned}
\nabla^2 &= \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial \theta} + \frac{\partial^2}{\partial z^2} \\
&= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \theta} + \frac{\partial^2}{\partial z^2}
\end{aligned}$$

A similar procedure may be used in spherical polar coordinates (r, θ, ϕ) where

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\cot\,\theta}{r^2}\frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2}$$