IOP Concise Physics | A Morgan & Claypool Publication

Theories of Matter, Space and Time

Quantum theories

N Evans
S F King

VOLUME TWO

IOP ebooks

Theories of Matter, Space and Time, Volume 2

Quantum theories

N Evans and S F King
University of Southampton

Morgan & Claypool Publishers

ISBN 978-1-6817-4983-9 (ebook)
ISBN 978-1-6817-4980-8 (print)
ISBN 978-1-6817-4981-5 (mobi)

DOI 10.1088/978-1-6817-4983-9

Version: 20180601

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 prequel (*Theories of Matter Space and Time: Classical theories*) grew out of courses that we have both taught as part of the third and fourth year of the undergraduate degree program 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 the first book we introduced several key areas of study. Firstly the principle of least action, an alternative treatment of Newtonian dynamics, that provides new understanding of conservation laws. Secondly special relativity including four-vector notation. Thirdly we discussed 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 were transparent.

This second book of the pair will move the ideas to the arena of quantum mechanics. We first quickly review the basics of quantum mechanics which should be familiar to the reader from a first course. Then we will link the Schrödinger equation to the principle of least action introducing Feynman's path integral methods. Next we present the relativistic wave equations of Klein, Gordon and Dirac. Finally we convert Maxwell's equations of electromagnetism to a wave equation for photons and make

contact with quantum electrodynamics (QED) at a first quantized level. Between the two 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.

Acknowledgements

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. Chapters 3 and 4 evolved from courses we gave at the STFC School for Young High Energy Physicists—so thanks to the lecturers who came before us including David Miller, Rob Thorne, Adrian Signer, Jeff Forshaw and again Tim Morris for influencing our thinking.

Author biographies

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/

Steve King

Steve completed his PhD in QCD perturbation theory in 1980 at Manchester University. He was a postdoctoral fellow at Oxford University, there 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, Volume 2

Quantum theories

N Evans and S F King

Chapter 1

Non-relativistic quantum mechanics

To set the scene for the work to come we begin here by reviewing the basics of non-relativistic quantum mechanics. We will mostly work in one dimension. We will motivate the form of the Schrödinger equation, discuss the information content and interpretation of the wave function, and finally work through the simple example of the square well, including introducing the ideas of orthogonality and completeness.

1.1 One dimensional, time dependent Schrödinger equation

In quantum mechanics the behaviour of a particle is controlled by a wave equation. A free particle is associated with a wave

$$\psi = e^{i(kx-\omega t)} \tag{1.1}$$

where the wave number k and angular frequency ω are related to the momentum and energy of the particle

$$p = \frac{h}{\lambda} \quad \rightarrow \quad k = \frac{p}{\hbar} \tag{1.2}$$

$$E = h\nu \quad \rightarrow \quad \omega = \frac{E}{\hbar} \tag{1.3}$$

here h is Planck's constant and $\hbar = h/2\pi$.

The properties of the particle can therefore be obtained from the wave by acting on it with *operators* (which we mark by a hat over the symbol)

$$\hat{E}\psi = i\hbar\frac{\partial}{\partial t}\psi \tag{1.4}$$

$$\hat{p}\psi = -i\hbar\frac{\partial}{\partial x}\psi \tag{1.5}$$

The free wave function (1.1) is an *eigenfunction* of these operators with the values of E and p being the *eigenvalues*.

For a classical particle in a potential V we require that energy is conserved so

$$E = \frac{p^2}{2m} + V \tag{1.6}$$

which, using the operators, we can rewrite as a wave equation

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi + V\psi \equiv \hat{H}\psi \tag{1.7}$$

where $\hat{H}$ is the Hamiltonian operator. This is the *time dependent Schrödinger equation* which is central to quantum mechanics.

1.2 Time independent Schrödinger equation

In problems where *V* is independent of time there are always solutions to the Schrödinger equation of the form

$$\psi(x,t) = u(x)e^{-iEt/\hbar} \tag{1.8}$$

where *u*(*x*) satisfies (simply substitute this solution into the full Schrödinger equation) the *time independent Schrödinger equation*

$$\hat{H}u(x) \equiv -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}u(x) + V(x)u(x) = Eu(x) \tag{1.9}$$

1.3 Interpretation

The amplitude of the wave function *ψ**(*x*, *t*) *ψ* (*x*, *t*) (which in the time independent case is just $u^*(x)u(x)$) is associated with the probability of finding a particle at *x*. Remembering that *x* is continuous, the precise statement is

$$u^*(x)u(x)\,dx = \text{probability of finding particle between } x \text{ and } x + dx \tag{1.10}$$

Graphically this is shown in figure 1.1 which shows that the probability of finding the particle in the *dx* spatial slice is just the area under the curve *u***u* in that slice.

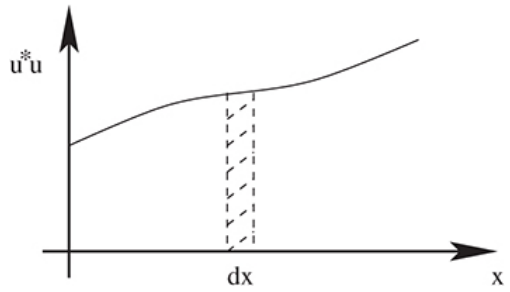


Figure 1.1. *u**(*x*) *u*(*x*) *dx*—the area under the curve shown—gives the probability to find the particle in that region of *x*.

Since the particle must be somewhere with probability 1 we must have[1]

$$\int_{-\infty}^{\infty} u^{*}(x) u(x)\, dx = 1 \tag{1.11}$$

Formally we find observable properties of the particles using the operators

$$\langle x \rangle = \int_{-\infty}^{\infty} u^{*}(x)\, \hat{x}\, u(x)\, dx = \int_{-\infty}^{\infty} u^{*}(x)\, x\, u(x)\, dx \tag{1.12}$$

$$\langle p \rangle = \int_{-\infty}^{\infty} u^{*}(x)\, \hat{p}\, u(x)\, dx = \int_{-\infty}^{\infty} u^{*}(x) \left(-i\hbar \frac{\partial}{\partial x} \right) u(x)\, dx \tag{1.13}$$

1.4 Proof that probability is conserved

To back up this interpretation of the wave function we can show that probability is conserved in the theory. This means that if the probability of the particle being in some area decreases then the probability that it lies outside must increase. In other words there is a flow of probability current density (see figure 1.2) satisfying the usual conservation equation (cf electric charge)

$$\int_{S} \vec{J} . d\vec{A} = - \int \frac{\partial \rho}{\partial t}\, dV \tag{1.14}$$

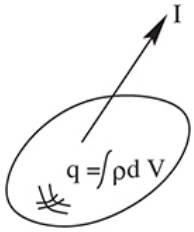


Figure 1.2. The change in a conserved quantity, *q*, in a volume matches to a current leaving the volume.

Using Gauss' theorem ($\int \vec{A} . d\vec{S} = \int \vec{\nabla} . \vec{A}\, dV$) we have

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} . \vec{J} = 0 \tag{1.15}$$

or in one dimension

$$\frac{\partial \rho}{\partial t} + \frac{\partial J^x}{\partial x} = 0 \tag{1.16}$$

Now we can show using the Schrödinger equation that $\rho = \psi^*\psi$ satisfies such a relation. We add two copies of the Schrödinger equation as follows

$$-i\psi^* (\mathrm{SE}) + (\mathrm{SE})^* i\psi \tag{1.17}$$

This gives

$$\begin{aligned} \hbar\psi^* \tfrac{\partial \psi}{\partial t} + \hbar\psi \tfrac{\partial \psi^*}{\partial t} \; = \; & \tfrac{i\hbar^2}{2m}\psi^* \tfrac{\partial^2}{\partial x^2}\psi - i\psi^* V\psi \\ & - \tfrac{i\hbar^2}{2m}\psi \tfrac{\partial^2}{\partial x^2}\psi^* + i\psi^* V\psi \end{aligned} \tag{1.18}$$

and hence

$$\frac{\partial}{\partial t}\left(\psi^*\psi\right) = \frac{i\hbar}{2m}\frac{\partial}{\partial x}\left(\psi^*\frac{\partial}{\partial x}\psi - \psi\frac{\partial}{\partial x}\psi^*\right) \tag{1.19}$$

which indeed has the form of a conservation equation with $\rho = \psi^*\psi$.

1.5 Momentum space wave functions

In the above discussion we have described the particle by its wave function at a particular point in space and then shown how to calculate its momentum with an operator. Alternatively we could write a wave function that describes the probability of the particle having momentum in some *dp* interval directly and then calculating the position becomes more complicated.

In fact it is possible to set up this momentum space wave function such that

$$\phi^*(p)\,\phi(p)\,dp = \text{probability of particle having momentum } p \text{ to } p+dp \tag{1.20}$$

$$\int_{-\infty}^{\infty} \phi^*(p)\,\phi(p)\,dp = 1 \tag{1.21}$$

with the properties of the particle being given by the operator relations

$$\int_{-\infty}^{\infty} \phi^*(p)\,p\,\phi(p)\,dp = \langle p \rangle \tag{1.22}$$

$$\int_{-\infty}^{\infty} \phi^*(p)\left(i\hbar\frac{\partial}{\partial p}\right)\phi(p)\,dp = \langle x \rangle \tag{1.23}$$

Note the difference in sign on $\hat{x}$ relative to the position space operator $\hat{p}$. The relationship between $\psi(x)$ and $\phi(p)$ is given by a Fourier transform

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{-ipx/\hbar}\, dx \tag{1.24}$$

or inversely

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \phi(p) e^{ipx/\hbar}\, dp \tag{1.25}$$

We can demonstrate that the Fourier transform indeed has the correct properties by checking the consistency of the three operator equations above. Firstly consider

$$\begin{aligned} \int \phi^*(p)\, \phi(p)\, dp &= \tfrac{1}{2\pi\hbar} \int dp \left[\int dx\prime e^{\frac{ipx\prime}{\hbar}} \psi^*(x\prime) \right] \left[\int dx\prime\prime e^{\frac{-ipx\prime\prime}{\hbar}} \psi(x\prime\prime) \right] \\ &= \int dx\prime \int dx\prime\prime \tfrac{1}{2\pi\hbar} \psi^*(x\prime) \psi(x\prime\prime) \int dp e^{\frac{-ip(x\prime\prime - x\prime)}{\hbar}} \end{aligned} \tag{1.26}$$

We recognise the *dp* integral as the Fourier expansion of a Dirac delta function

$$\delta(x - x_0) = \frac{1}{2\pi} \int e^{-ik(x - x_0)}\, dk \tag{1.27}$$

where the basic property of the Dirac delta function is

$$\int f(x) \delta(x - x_0)\, dx = f(x_0) \tag{1.28}$$

It is a height one ‘box’ across just the *dx* slice at $x = x_0$.

So with $k = p/\hbar$ and $dk = dp/\hbar$

$$\begin{aligned} \int \phi^*(p)\, \phi(p)\, dp &= \int dx\prime \int dx\prime\prime \delta(x\prime\prime - x\prime) \psi^*(x\prime) \psi(x\prime\prime) \\ &= \int dx\prime \psi^*(x\prime) \psi(x\prime) \\ &= 1 \end{aligned} \tag{1.29}$$

The equations are consistent.

Secondly we can check the relation for the expectation value of the particle’s position

$$\begin{aligned}\int \phi^{*}(p) & \left(i\hbar\tfrac{\partial}{\partial p}\right)\ \phi(p)\,dp \\ &= \tfrac{1}{2\pi\hbar}\int dp\left[\int dx\prime e^{\frac{ipx\prime}{\hbar}}\psi^{*}(x\prime)\right]\left(i\hbar\left(\tfrac{-ix\prime\prime}{\hbar}\right)\right)\left[\int dx\prime\prime e^{\frac{-ipx\prime\prime}{\hbar}}\psi(x\prime\prime)\right] \\ &= \int dx\prime\int dx\prime\prime\tfrac{1}{2\pi\hbar}\ \psi^{*}(x\prime)\ x\prime\prime\,\psi(x\prime\prime)\int dp\,e^{\frac{-ip(x\prime\prime-x\prime)}{\hbar}} \\ &= \int dx\prime\int dx\prime\prime\delta(x\prime\prime-x\prime)\psi^{*}(x\prime)\ x\prime\prime\,\psi(x\prime\prime) \\ &= \int dx\prime\psi^{*}(x\prime)\ x\prime\,\psi(x\prime) \\ &= \langle x\rangle\end{aligned} \tag{1.30}$$

Finally we check the expectation value for momentum

$$\begin{aligned}\int \phi^{*}(p)\,p\,\phi(p)\,dp = \tfrac{1}{2\pi\hbar}\int dp\left[\int dx\prime e^{\frac{ipx\prime}{\hbar}}\psi^{*}(x\prime)\right] \\ \times\left[\int dx\prime\prime\left(i\hbar\tfrac{\partial}{\partial x\prime\prime}e^{\frac{-ipx\prime\prime}{\hbar}}\right)\psi(x\prime\prime)\right]\end{aligned} \tag{1.31}$$

The differential has been inserted adhocly to simply bring down a factor of p. Now we integrate by parts throwing away surface terms at infinity

$$\begin{aligned}\int \phi^{*}(p)\,p\,\phi(p)\,dp &= \tfrac{1}{2\pi\hbar}\int dp\left[\int dx\prime e^{\frac{ipx\prime}{\hbar}}\psi^{*}(x\prime)\right] \\ &\quad\times\left[\int dx\prime\prime e^{\frac{-ipx\prime\prime}{\hbar}}\left(-i\hbar\tfrac{\partial}{\partial x\prime\prime}\psi(x\prime\prime)\right)\right] \\ &= \int dx\prime\int dx\prime\prime\delta(x\prime\prime-x\prime)\psi^{*}(x\prime)\left(-i\hbar\tfrac{\partial}{\partial x\prime\prime}\right)\psi(x\prime\prime) \\ &= \int dx\prime\psi^{*}(x\prime)\left(-i\hbar\tfrac{\partial}{\partial x\prime}\right)\psi(x\prime) \\ &= \langle p\rangle\end{aligned} \tag{1.32}$$

Everything is nicely consistent.

1.6 Heisenberg uncertainty principle

In general the wave function of a particle $\psi(x)$ will correspond to some localised wave packet whose Fourier transform is the momentum space wave function $\phi(p)$, as in equations (1.24) and (1.25). From the theory of Fourier transforms, it is seen that any wave packet that is more strongly peaked in position space will be less strongly peaked in momentum space, and vice versa. For example, a wave function which is a plane wave in position space (and hence its position is completely undetermined) will have a sharp value of momentum with no uncertainty. It is possible to derive a relation between the spread or width of the wave packet in position space Δx and in momentum space Δp, namely,

$$\Delta x\Delta p \geqslant \hbar/2 \tag{1.33}$$

The equality follows directly from the theory of Fourier transforms for the idealised wavepackets. The inequality expresses the fact that, in real experiments which measure the position and momentum of a particle simultaneously, the product of uncertainties in the respective measurements must always exceed the above bound.

There is also a similar uncertainty relation for energy and time of a quantum state,

$$\Delta E \Delta t \geqslant \hbar/2 \tag{1.34}$$

For example, for an atomic transition, the shorter the transition time Δt the greater the width of the associated spectral line ΔE, and vice versa.

The above relations in equations (1.33) and (1.34) are collectively known as the Heisenberg uncertainty principle. They highlight the fact that the quantum world represents a major departure from classical physics, since, even in the most accurate idealised experiment, two quantities such as position and momentum cannot ever be known simultaneously to arbitrary precision. Even great physicists such as Albert Einstein never accepted this, and this led to a series of high profile debates with Niels Bohr. It is now generally accepted that Bohr was correct and Einstein was wrong. Quantum mechanics, though completely counter to our intuition, has been thoroughly vindicated in all experiments to date involving atoms and subatomic particles.

1.7 Square well example

A simple, interesting example of a quantum mechanics system is the square potential well, as shown in figure 1.3. We assume that the particle cannot penetrate the infinite barriers

$$\psi = 0, \ \text{ for } \ x \leqslant 0, \ \ x \geqslant a \tag{1.35}$$

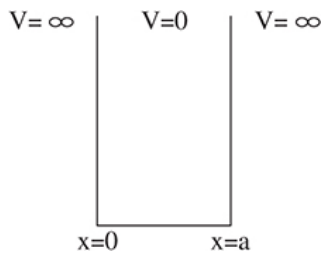


Figure 1.3. The potential of an infinite square well.

Since the potential is time independent the solution takes the form

$$\psi(x,t) = u(x) e^{-iEt/\hbar} \tag{1.36}$$

and we must solve the time independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}u\left(x\right)+V\left(x\right)u\left(x\right)=EU\left(x\right) \tag{1.37}$$

Of course in the region of interest the potential is just $V = 0$.

The solutions to this equation take the form

$$u\left(x\right)=A\ \sin\ kx+B\ \cos\ kx \tag{1.38}$$

The integration constants are fixed by the boundary conditions of ψ vanishing at $x = 0, a$ so

$$u_n\left(x\right)=A\ \sin\ \frac{n\pi x}{a} \tag{1.39}$$

with n integers 1,2,3, ...

Substituting this solution into the Schrödinger equation we find

$$E_n=\frac{\hbar^2}{2m}\left(\frac{n\pi}{a}\right)^2 \tag{1.40}$$

Finally, to find the constant A we can require $\psi(x, t)$ is correctly normalized

$$\begin{aligned}\int_{-\infty}^{\infty}\psi^*\psi\,dx&=1\\&=\int_0^a A^2\sin^2\tfrac{n\pi x}{a}\,dx\\&=A^2\tfrac{a}{2}\end{aligned} \tag{1.41}$$

The full solution is therefore

$$\psi_n\left(x,t\right)=\sqrt{\frac{2}{a}}\ \sin\ \frac{n\pi x}{a}e^{-iE_nt/\hbar} \tag{1.42}$$

1.8 Completeness

The consideration of how a particular initial condition for the wave function in a square well evolves with time provides interesting insight into the uniqueness of the solutions we have found. In particular since the solutions are sine waves of period 2a there is a strong connection to problems one encounters when studying Fourier analysis such as wave forms on a string.

For example, if we take an initial wave function, at $t = 0$, of the triangular form show in figure 1.4 then we can write

$$\psi\left(x,t=0\right)=\sum_{n=1}^{\infty}c_nu_n\left(x\right)$$

where the c_n are the Fourier-like coefficients (we will explain how to derive them in the next section) which are given by

$$c_n = \frac{8k}{n^2\pi^2}\sqrt{\frac{a}{2}}\ \sin\ \frac{n\pi}{2} \tag{1.43}$$

$$\tag{1.44}$$

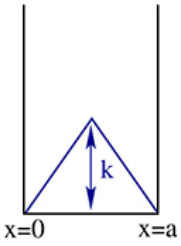


Figure 1.4. The initial conditions for the square well problem considered in section 1.8.

We now know the time evolution since we know that each individual term evolves as

$$u_n(x,0) \to e^{-iE_nt/\hbar}u_n(x,0) \tag{1.45}$$

Resuming the series at time t gives the evolution of the initial condition (to a precision determined by how many terms you resum).

This is an example of a general rule in QM called *completeness*: any wave function may be expanded as a series of the eigenfunction solutions of the Schrödinger equation relevant to that problem. In other words in any problem we may write

$$\phi(x) = \sum_n c_n u_n(x) \tag{1.46}$$

for any function $\phi(x)$, where

$$Hu_n = E_n u_n \tag{1.47}$$

We will not prove this here but if it were not true it would be quite surprising! Imagine we had found all the solutions of the Schrödinger equation and then wrote down an initial condition that could not be rewritten in terms of those solutions ... we would have missed the evolution of that initial condition and hence we cannot have had all the solutions! Completeness is usually the case for a theory to make sense and it allows us to evolve all initial states with time.

1.9 Orthogonality

It is also important in these initial condition problems that there is a unique way of writing

$$\psi(x,0)=\sum_n c_n u_n(x) \tag{1.48}$$

If it were not unique then a given initial condition would have more than one expansion which would evolve differently. Again the theory would not make sense.

Each $u_n(x)$ therefore contains unique information. Orthogonality is a mathematical statement of this fact

$$\int_{-\infty}^{\infty} u_n^*(x) u_m(x)\,dx=\delta_{nm} \tag{1.49}$$

where $\delta_{nm}=1$ if $m=n$ and $\delta_{nm}=0$ if $m\neq n$.

You can think of this expression as similar to a dot product between the coordinate axes vectors ($\hat{\underline{i}},\hat{\underline{j}},\hat{\underline{k}}$)—the axes contain the separate information about the three directions in the space and the dot product is zero between any two orthogonal directions.

Proof: The u_n are eigenfunctions of the Hamiltonian H satisfying $Hu_n=E_nu_n$ so consider

$$\int u_i^* H u_j\,dx \tag{1.50}$$

We can act with H to either the left or right in which case we will find

$$E_j\int u_i^* u_j\,dx=E_i\int u_i^* u_j\,dx \tag{1.51}$$

which can only be true for $i\neq j$ if the wave functions are orthogonal and both sides are zero. When $i=j$ the integral over the wave function squared is just the usual probability of finding the particle in all space and is set equal to one.

Now we know enough to derive the coefficients in (1.44). Given

$$u(x,t=0)=\sum_n c_n u_n(x) \tag{1.52}$$

we multiply by some u_m^* and integrate over all space

$$\int u_m^* u(x,t=0)\,dx=\sum_n\int u_m^* c_n u_n(x)\,dx \tag{1.53}$$

and using orthogonality we find only one term of the sum on the right survives and hence

$$c_m = \int u_m^* u\left(x, t=0\right)\, dx \tag{1.54}$$

using the initial conditions show in figure 1.4 and performing the integrals leads to (1.44).

1.10 The 3D Schrödinger equation

We have concentrated on one dimensional problems but the analysis is easily extended to three dimensions. The momentum operator is

$$\vec{\hat{p}} = -i\hbar\vec{\nabla} \tag{1.55}$$

The Schrödinger equation becomes

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi \tag{1.56}$$

The probability to find a particle in some infinitesimal box of volume δV is

$$\text{probability} = \psi^*\psi\delta V \tag{1.57}$$

where, for example, in spherical coordinates $\delta V = r^2\ \sin\ \theta\ d\theta\ d\phi\ dr$.

1.11 Wave function collapse and all that

The most mysterious feature of QM is that a particle is described by a probability wave which ‘collapses’ during a ‘measurement’ to leave the particle at just one point. In some sense one should think of a quantum of the particle’s energy as being smeared through the wave. If we probe the wave at a point and it releases a quantum then it will look like the particle was at that point. This idea has to allow the wave at a point to ‘know’ what is going on in the rest of the wave instantaneously and this is a rather uncomfortable fact. A number of unresolved ideas to understand things better are:

- Copenhagen interpretation—do not philosophise about it, use it!
- Hidden variables—secretly there is a deterministic description of QM which the wave function is an ‘average’ over.
- Many worlds—all outcomes happen in parallel Universes (this does not explain why a measurement splits the Universes though).

None of these are really satisfactory—not least because it is not precisely clear what constitutes a measurement. Nevertheless QM is the most successful theory physics has and so is clearly correct. The real impact of these issues is that it is hard to have an intuitive feel for the subject. In the next chapter we will investigate an alternative formalism for QM in which the idea of a trajectory for the particle is central, rather than a wave function, and it allows some classical intuition to be used.

Exercise 1.1. Make an odd continuation of the solutions to the infinite square well problem and calculate the momentum space wave functions $\phi(p)$. What is the physical

significance of your result?

Appendix A. Time independent perturbation theory

Perturbation theory is a technique for solving problems where a system we understand is ‘tweaked’ by a small change. For example, we know how to find the solutions of the Schrödinger equation for a particle in a square well potential, or a simple harmonic potential—perturbation theory can tell us the solutions (approximately) if these potentials are modified a little. Let us begin by studying the case where the modification is *time independent* (we will return to the time dependent case in the next chapter).

Consider a time-independent problem in QM we can already solve. This means we have found solutions to the Schrödinger equation

$$H_0\, u\,(x) = E_0\, u\,(x) \tag{A.1}$$

where

$$H_0 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\,(x) \tag{A.2}$$

We assume that the different solutions ϕ_i have different energies E_{0i}.

Now imagine perturbing the problem by changing the potential by a small amount. Thus

$$H = H_0 + H_p \tag{A.3}$$

Since it is a small change most likely the wave function solutions have not changed much and we can write the new solutions as

$$u_i = \phi_i + \delta\phi_i \tag{A.4}$$

The Schrödinger equation now becomes

$$(H_0 + H_p)(\phi_i + \delta\phi_i) = (E_{0i} + E_p)(\phi_i + \delta\,\phi_i) \tag{A.5}$$

H_p, $\delta\phi_i$ and E_p are all small so we can expand this equation

$$\begin{aligned} \text{zeroth order:} \quad & H_0\,\phi_i = E_{0i}\,\phi_i \\ \text{first order:} \quad & H_0\,\delta\,\phi_i + H_p\,\phi_i = E_{0i}\,\delta\,\phi_i + E_p\,\phi_i \end{aligned}$$

We have dropped terms that are the *square of a small quantity*.

Now we use the *completeness* of the set of states ϕ_i to write

$$\delta\,\phi_i = \sum_{n\neq i} c_n\,\phi_n \tag{A.6}$$

(Note: the $\delta\phi_i$ is the amount that ϕ_i shifts away from being ϕ_i so we do not include ϕ_i in the sum.)

The *first order* expression is

$$H_0 \sum_{n \neq i} c_n \phi_n + H_p \phi_i = E_{0i} \sum_{n \neq i} c_n \phi_n + E_p \phi_i \tag{A.7}$$

In the first term we can act with H_0 on ϕ_n and get a factor of E_{0n}. Now multiply on the left by ϕ_j^* and integrate over all space

$$\int \phi_j^* \textstyle\sum_{n \neq i} E_{0n} c_n \phi_n \, dx + \int \phi_j^* H_p \phi_i \, dx = E_{0i} \int \phi_j^* \textstyle\sum_{n \neq i} c_n \phi_n \, dx + E_p \int \phi_j^* \phi_i \, dx \tag{A.8}$$

Using the orthogonality of the wave functions ($\int \phi_j^* \phi_i \, dx = \delta_{ij}$) we find

$$\sum_{n \neq i} E_{0n} c_n \delta_{nj} + \int \phi_j^* H_p \phi_i \, dx = E_{0i} \sum_{n \neq i} c_n \delta_{jn} + E_p \delta_{ij} \tag{A.9}$$

and performing the sums

$$c_j E_{0j} + \int \phi_j^* H_p \phi_i \, dx = E_{0i} c_j + E_p \delta_{ij} \tag{A.10}$$

Now set $i = j$ so the first term on each side cancels and

$$E_p = \int \phi_i^* H_p \phi_i \, dx \tag{A.11}$$

or if $j \neq i$ so the δ_{ij} are zero

$$c_j = \frac{\int \phi_j^* H_p \phi_i \, dx}{(E_{0i} - E_{0j})} \tag{A.12}$$

We have obtained the lowest order perturbation theory results

$$\boxed{\begin{aligned} E_i &= E_{0i} + \textstyle\int \phi_i^* H_p \phi_i \, dx \\ u_i &= \phi_i + \textstyle\sum_{n \neq i} \frac{\int \phi_n^* H_p \phi_i \, dx}{(E_{0i} - E_{0n})} \phi_n \end{aligned}} \tag{A.13}$$

Of course these are not exact—we threw away some small terms. We can get a better approximation by taking the above answers and allowing small corrections on top

$$(H_0 + H_p)(\psi_i + \delta\psi_i) = \left(E_i + E_p'\right)(\psi_i + \delta\psi_i) \tag{A.14}$$

We then repeat everything above so for example

$$E = E_i + \int u_i^* H_p u_i \, dx \tag{A.15}$$

or substituting in terms of the unperturbed results

$$E = E_{0i} + \int \phi_i^* H_p \phi_i \, dx + \sum_{n \neq i} \frac{\int \phi_n^* H_p \phi_i \, dx}{(E_i - E_n)}^2 \tag{A.16}$$

By repeatedly doing this we can make the result arbitrarily good.

The first order shift in the energy is easy to calculate since it only requires knowledge of the unperturbed ϕ_i in question. The first order shift in the wave function, and hence second order shift in the energy, requires a knowledge of all the ϕ_n to calculate the shifts for one ϕ_i.

A.1 Example: perturbed square well

Consider the square well problem we solved before. Now imagine that the potential is perturbed by a term (shown in figure A1)

$$\Delta V = -\alpha x \qquad 0 < x < a/2 \tag{A.17}$$

We can calculate the shift in the energy of the solutions u_n as follows

$$\begin{aligned} \Delta E_n &= \int_{-\infty}^{\infty} u_n^*(x) \Delta V u_n(x) \, dx \\ &= -\frac{2\alpha}{a} \int_0^{a/2} x \sin^2 \frac{n\pi x}{a} \, dx \\ &= -\frac{\alpha a}{8} - \frac{\alpha a}{4n^2\pi^2} + \frac{\alpha a}{4n^2\pi^2} \cos n\pi \end{aligned} \tag{A.18}$$

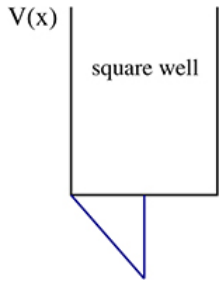


Figure A1. The perturbation to the square well discussed in (1.74).

Appendix B. Orbital and spin angular momentum

Here we provide a brief introduction to orbital and intrinsic spin in quantum mechanics.

In three-dimensional problems with rotational symmetry (for example an electron moving in the central potential $V(r)$ from the proton in a hydrogen atom) we can write the Laplacian, ∇^2, in spherical polar coordinates

$$\nabla^2 = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\,\theta}\frac{\partial}{\partial\theta}\left(\sin\,\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \tag{B.1}$$

The Schrödinger equation can then be solved by an ansatz of the form $\Psi(\vec{x},t) = R(r)Y(\theta,\phi)e^{-iEt}$. If we substitute this into the Schrödinger equation (1.56) we find the top and bottom expressions in

$$\begin{aligned} l\left(l+1\right) &= \tfrac{1}{R(r)}\left[\tfrac{\partial}{\partial r}\left(r^2\tfrac{\partial R(r)}{\partial r}\right) + \tfrac{2\,mr^2}{\hbar^2}\left(E - V\left(r\right)\right)R\left(r\right)\right] \\ &= -\tfrac{1}{Y(\theta,\phi)}\left[\tfrac{1}{\sin\,\theta}\tfrac{\partial}{\partial\theta}\left(\sin\,\theta\tfrac{\partial}{\partial\theta}\right) + \tfrac{1}{\sin^2\theta}\tfrac{\partial^2}{\partial\phi^2}\right]Y\left(\theta,\phi\right) \end{aligned} \tag{B.2}$$

Here we have separated the variables so each of the two sides of the equation can be written as equal to a constant which we have written as $l(l+1)$ where l is some constant for the moment.

The top equation depends on $V(r)$ and depends on the particular problem—so we will leave this unsolved but instead concentrate on the bottom equation for the angular dependence which is common to all such problems. We will just quote the form of the smooth solutions (which are called spherical harmonics) here

$$Y\left(\theta,\phi\right) = P_{l,m}\left(\cos\,\theta\right)e^{im\phi} \tag{B.3}$$

m here can only take integer values and $l = \mid m \mid, \mid m \mid +1, \mid m \mid +2, \ldots$. The $P_{l,m}$ are polynomials that depend on the choice of l and m. For example some early cases are

$$\begin{aligned} l &= 0, \quad m = 0, \quad P_{0,0} = 1 \\ l &= 1, \quad m = 0, \quad P_{1,0} = \cos\,\theta \\ l &= 1, \quad \mid m \mid = 1, \quad P_{1,1} = \sin\,\theta \\ l &= 2, \quad m = 0, \quad P_{2,0} = \tfrac{3}{2}\cos^2\theta - \tfrac{1}{2} \\ l &= 2, \quad \mid m \mid = 1, \quad P_{2,1} = \sin\,\theta\,\cos\,\theta \\ l &= 2, \quad \mid m \mid = 2, \quad P_{2,2} = \sin^2\theta \end{aligned} \tag{B.4}$$

You may recognise the relation between the degeneracy of these states for a given l and the degeneracy of atomic orbitals. For example there is one s-state, three p-states and five d-states.

Let us now make the connection to angular momentum. The angular momentum operators are

$$\vec{\hat{L}} = \vec{\hat{r}} \times \vec{\hat{p}} = -i\hbar(\vec{r} \times \vec{\nabla}) \tag{B.5}$$

In spherical polar coordinates these are explicitly

$$\begin{aligned}\vec{\hat{L}} = i\hbar\Big[&\Big(\sin\,\phi\tfrac{\partial}{\partial\theta} + \cot\,\theta\,\cos\,\phi\tfrac{\partial}{\partial\phi}\Big)\hat{x} \\ &+ \Big(-\cos\,\phi\tfrac{\partial}{\partial\theta} + \cot\,\theta\,\sin\,\phi\tfrac{\partial}{\partial\phi}\Big)\hat{y} - \tfrac{\partial}{\partial\phi}\hat{z}\Big]\end{aligned} \tag{B.6}$$

We can immediately see that the $Y(\theta, \phi)$ are eigenvalues of L_z with eigenvector $m\hbar$.

Further if we compute $L^2 = L_x^2 + L_y^2 + L_z^2$ we find

$$\hat{L}^2 = -\left[\frac{1}{\sin\,\theta}\frac{\partial}{\partial\theta}\left(\sin\,\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \tag{B.7}$$

which is precisely the quantity that occurs in ∇^2. Now we can see from (1.77) that the $Y(\theta, \phi)$ are eigenvalues of $\hat{L}^2$ with eigenvector $l(l + 1)\,\hbar^2$. Spherical harmonics with the same l have the same angular momentum magnitude. m then chooses the state with a particular projection of the angular momentum onto the z-axis.

This, in brief, is the story of orbital angular momentum. To understand intrinsic spin we should abstract a little. The angular momentum operators satisfy an interesting *algebra*. To find this take the definition of L in (1.80) in Cartesian coordinates and one can simply show that

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y \tag{B.8}$$

Now consider for example the $l = 1$ spherical harmonics of which there are three. Let us associate them with the basis vectors (1,0,0), (0,1,0) and (0,0,1). These are eigenvectors of $\hat{L}_z$ with eigenvalues $m\hbar$ so we can write $\hat{L}_z$ as the matrix

$$\hat{L}_z = \hbar\begin{matrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{matrix} \tag{B.9}$$

$\hat{L}_x$ and $\hat{L}_y$ can also be written as 3 × 3 matrices—one acts on the appropriate spherical harmonic with the operator from (B.6) then rewrites the answer in terms of the three $l = 1$ spherical harmonics. The coefficients in that expansion are the matrix elements and one finds

$$\hat{L}_x = \hbar \begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{matrix}, \quad \hat{L}_y = \hbar \begin{matrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{matrix} \tag{B.10}$$

These three matrices satisfy the algebra (B.8). The spherical harmonics are a particular representation of the basis vectors of this algebra. You can play the same trick with higher *l* values producing larger dimension matrices that satisfy the algebra.

The interesting thing is that there are additional representations of the algebra (B.8) that the orbital angular momentum solutions do not create. For example there is a two-dimensional representation where the matrices are the Pauli matrices $\hat{L}_i = \hbar\sigma_i/2$. Explicitly here

$$\sigma_1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B.11}$$

The $\hat{L}_i$ act on a two vector and the L_z eigenvalues are $m = \pm \hbar/2$. It turns out that nature uses these representations as an abstraction of angular momentum. All fundamental fermions (e.g. electrons, or quarks) have intrinsic spin 1/2 and need to be written as a two vector of wave functions, the top element of which describes the spin up particle and the bottom component the spin down particle. In non-relativistic quantum mechanics this seems like an add-on but we will see in chapter 3 that in relativistic quantum mechanics spin naturally emerges in the context of the Dirac equation.

[1]Note that for a free particle wave function the normalization of the wave function is interpreted as the flux of particles per unit volume or within a finite box.

IOP Concise Physics

Theories of Matter, Space and Time, Volume 2

Quantum theories

N Evans and S F King

Chapter 2

Path integral approach to quantum mechanics

New insights into classical mechanics can be obtained from Hamilton's principle in which a classical particle is viewed as following the path which minimizes an action (we introduce these ideas in detail in the partner book *Theories of Matter Space and Time: Classical Theories*). Feynman developed a quantum mechanics version of this idea which we will study here. We are going to start with his prescription and see that it is indeed the same theory as the Schrödinger equation. Although it returns some classical intuition to the quantum world, it is still a very strange place!

2.1 Proposal for the quantum mechanical amplitude

To motivate the form of the theory consider the usual double slit type experiment shown in figure 2.1. A classical description in which the particle goes through a single slit will clearly not do. We will adopt a much more radical idea that the particle travels by ALL possible paths!

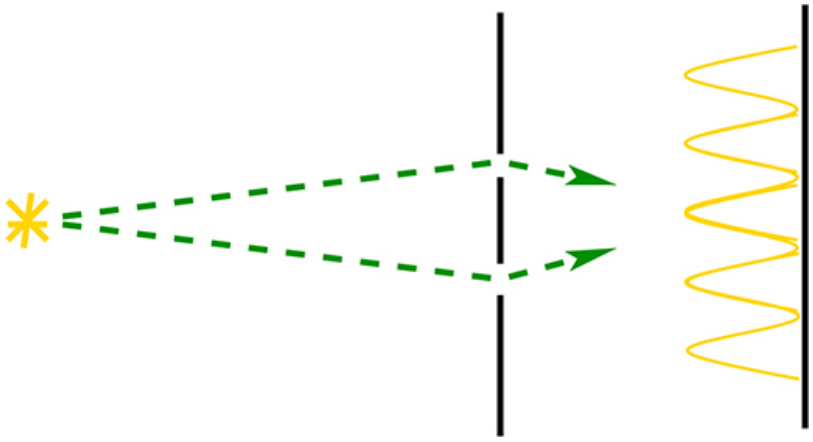


Figure 2.1. The classic double slit experiment showing the wave nature of particles.

The interference pattern suggests that there should be cancelling and reinforcing phases in the description. We are therefore led to the proposal of Feynman that the probability amplitude for a particle to travel from point *A* to point *B* is given schematically by

$$K\left(B,A\right)=\text{constant}\ \sum_{\text{all paths}} e^{iS[\text{path}]/\hbar} \tag{2.1}$$

where *S* is the classical action of each particular path, and every possible path contributes in the sum.

The probability for a particle to travel from point *A* to point *B* is then given by

$$P\left(B,A\right)=\mid K\left(B,A\right)\mid^{2} \tag{2.2}$$

where *K*(*B*, *A*) in equation (2.1) is called the quantum mechanical *kernel*.

Our proposal looks nutty (!)—every possible path is contributing the same constant amount up to a phase. Can this ever reproduce Hamilton's principle as the classical limit of the theory?

2.2 The classical limit

If we consider a particle (with momentum *p*) incident on a hole (of radius *r*) then we will see large quantum effects only when the wavelength of the wave function associated with the particle is

$$\lambda \gtrsim r \tag{2.3}$$

Of course $\lambda=\frac{h}{p}$ so it is because *h* is small in nature that we do not see quantum effects when we throw cricket balls through doors (of course there might well be some serious classical effects, so do not try this at home!).

From this discussion we can see that if we take

$$h\to 0 \tag{2.4}$$

then all wavelengths become very small and the theory becomes classical at all length scales.

Note that also in this limit the uncertainty principle ($\Delta p\Delta x\geqslant\hbar$) allows both *p* and *x* to be measured together which again corresponds to classical physics.

So what does our prescription give in this classical limit $h\to 0$? In general for a set of paths close to each other (as shown in figure 2.2), in this limit, we will find the difference in the classical action between neighbouring paths

$$\Delta S\gg\hbar$$

just because $\hbar$ is so small. This means that these paths have very different phases in the kernel above. The phase just points out a direction in the complex plane. The sum over these paths will just average the phase ... but if the phases are essentially random as in this case we will get precisely zero. (2.5)

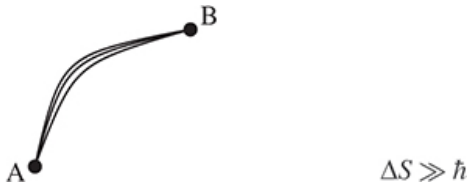


Figure 2.2. A collection of paths away from the minimum of the action have rapidly varying phase in the kernel and cancel.

The only time this will not be true is if we find a cluster of paths for which $\Delta S < \hbar$. This will only be true around a minimum of S where there is little change in S. A little cluster of paths here will all have roughly the same phase and add in such a way as to dominate the kernel. Thus in the classical limit our prescription does reproduce Hamilton's principle.

Incidentally, this tells us that in a quantum theory a classical trajectory gets smeared since it is equally likely to travel on a neighbouring path provided $\Delta S \leqslant \hbar$.

2.3 Wave functions

We will not really believe that this new prescription is quantum mechanical until we have seen that it gives identical physics to the Schrödinger equation. To move towards that proof let us see how to relate the kernel to wave functions. We had that for motion from a point $A = (x_a, t_a)$ to a point $B = (x_b, t_b)$

$$\text{probability}\,(A \to B) = |\, K\,(B, A)\,|^2 \tag{2.6}$$

If we imagine that the particle began at A at t_a, because its wave function was such that

$$|\, \psi(t_a)\,|^2 = \delta(x - x_a), \tag{2.7}$$

then we can identify the wave function at a later time $t_b > t_a$ with the kernel

$$\psi(x_b, t_b) = K(B, A) \tag{2.8}$$

where we allow x_b to be any general point at time t_b.

Using this result it is possible to derive an expression for the evolution of any wave function at some time into the wave function at some later time in terms of an integral over the product of the initial wave function and the kernel. In order to do this, consider the set of paths shown in figure 2.3. For a path going through C the action divides

$$S_{\text{path}} = S_{AC} + S_{CB} = \int_{t_0}^{t_1} L dt + \int_{t_1}^{t_2} L dt \tag{2.9}$$

So the contribution to the kernel from all possible paths from A to B through C is given by

$$K(B, A, \text{ via } C) = \sum_{A \to C} e^{iS_{AC}/\hbar} \cdot \sum_{C \to B} e^{iS_{CB}/\hbar} \tag{2.10}$$

Note that the cross terms in the multiplication of the sums gives all combinations of route A to C with all routes C to B. We therefore have

$$K(B, A, \text{ via } C) = \text{constant} \;\; K(C, A) \;\; K(B, C) \tag{2.11}$$

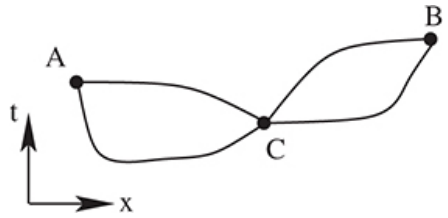


Figure 2.3. Paths from A to B via C.

These are not all the paths from A to B though because they all go through the special point C. To get all paths from A to B we must let C vary over all possible positions so that

$$K(B, A) = \text{constant} \int_{-\infty}^{\infty} K(C, A) \;\; K(B, C) \, dx_c \tag{2.12}$$

We previously, in (2.8), identified $K(B, A)$ as the wave function at time t_b and similarly we can identify here $K(C, A) = \psi(x_c, t_c)$, the wave function at time t_c. In both cases the wave functions have evolved from the delta function

form at time t_a in (2.7) but they can be arbitrarily complicated depending on the evolution, for example, through some potential. Thus this expression tells us how one wave function evolves into another

$$\psi(x_b, t_b) = \text{constant} \int_{-\infty}^{\infty} \psi(x_c, t_c) \; K(B, C) \; dx_c \tag{2.13}$$

The evolution is controlled by the kernel.

2.4 Deriving the Schrödinger equation

We want to show that the path integral expression for the evolution of a wave function is the same as the Schrödinger equation. The analysis below makes use of Gaussian integrals which are reviewed in Appendix C.

To derive the standard Schrödinger equation we must look at a particle with the Lagrangian

$$L = \tfrac{1}{2} m \dot{x}^2 - V(x) \tag{2.14}$$

The path integral expression for how the wave function evolves is

$$\psi(x\prime, t\prime) = A \int_{-\infty}^{\infty} K(x\prime, t\prime; x, t) \psi(x, t) dx \tag{2.15}$$

We need a way to keep track of all possible paths in order to work out the kernel. One way to do this is to divide time up into infinitesimal time slices and assume that the particle travels in a straight line at constant speed in any such time slice as shown in figure 2.4.

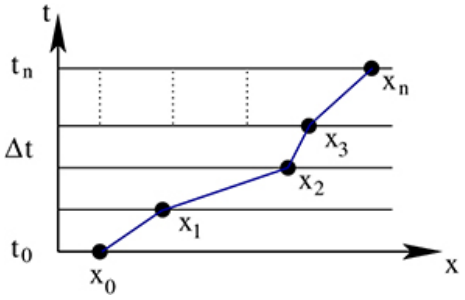


Figure 2.4. Paths a particle might take from the point x at time t to x' at time t' divided into many very short straight segments.

Now we can consider the time evolution of the wave function just across one Δt time slice. We will assume that the particle does not travel too far in any time slice (so we will write $x = x' + \Delta x$) and we will assume that its velocity is constant along the way

$$\psi\left(x\prime, t+\Delta t\right) = A\int_{-\infty}^{\infty} K\left(x\prime, t+\Delta t; x, t\right)\psi\left(x,t\right)dx \tag{2.16}$$

We know the kernel here because the paths are always straight lines (it is just $\exp(i\,S_{\text{path}}/\hbar)$)

$$\begin{aligned} S_{x\to x\prime} &= \int_t^{t+\Delta t} L\left(x,\dot{x}\right)dt \\ &= L\left(\tfrac{x+x\prime}{2}, \tfrac{x\prime-x}{\Delta t}\right)\Delta t \\ &= \left[\tfrac{1}{2}m\left(\tfrac{x\prime-x}{\Delta t}\right)^2 - V\left(\tfrac{x+x\prime}{2}\right)\right]\Delta t \end{aligned} \tag{2.17}$$

Thus our wave function evolves as

$$\psi\left(x\prime, t+\Delta t\right) = A\int_{-\infty}^{\infty} e^{i\frac{\Delta t}{\hbar}\left[\frac{1}{2}m\left(\frac{x\prime-x}{\Delta t}\right)^2 - V\left(\frac{x+x\prime}{2}\right)\right]}\psi\left(x,t\right)dx \tag{2.18}$$

There are lots of small terms in this expression so we can perform an expansion in them

$$\begin{aligned} x - x\prime &= \Delta x \\ \psi\left(x\prime, t+\Delta t\right) &= \psi\left(x\prime,t\right) + \Delta t\frac{\partial\psi(x\prime,t)}{\partial t} + \cdots \\ \psi\left(x,t\right) &= \psi\left(x\prime,t\right) + \Delta x\frac{\partial\psi(x\prime,t)}{\partial x\prime} + \frac{\left(\Delta x\right)^2}{2}\frac{\partial^2\psi(x\prime,t)}{\partial x\prime^2}\cdots \\ e^{\frac{-i\Delta t}{\hbar}V\left(\frac{x+x\prime}{2}\right)} &= 1 - \frac{i\Delta t}{\hbar}V\left(x\prime\right) + \cdots \end{aligned}$$

- To *zeroth order* our expression is, keeping x' constant in the integral

$$\psi\left(x\prime,t\right) = A\int_{-\infty}^{\infty} e^{i\frac{m\Delta x^2}{2\hbar\Delta t}}\psi\left(x\prime,t\right)d\left(\Delta x\right) \tag{2.19}$$

Note we have changed from summing over all x to summing over all Δx but these are equivalent! Of course we had assumed that Δx was small whilst in the integral we are letting it take all possible values—this is an

approximation. We can get away with this because the integrand is peaked around $\Delta x = 0$ so the large Δx pieces are not very important.

The integral is just a Gaussian integral and so

$$\psi\left(x\prime,t\right) = A\left(\tfrac{2\pi i\hbar\Delta t}{m}\right)^{1/2}\psi\left(x\prime,t\right) \tag{2.20}$$

which can only be true if

$$\boxed{A = \left(\tfrac{2\pi i\hbar\Delta t}{m}\right)^{-1/2}} \tag{2.21}$$

We have derived an expression for the constant in the wave function evolution equation.

• The *Schrödinger equation* emerges at the next leading order

$$\begin{aligned}\Delta t\tfrac{\partial\psi(x\prime,t)}{\partial t} = A\int_{-\infty}^{\infty} e^{i\frac{m\Delta x^2}{2\hbar\Delta t}}\Big[-i\tfrac{\Delta t}{\hbar}V\left(x\prime\right)\psi\left(x\prime,t\right) \\ + \Delta x\tfrac{\partial\psi(x\prime,t)}{\partial x} + \tfrac{\left(\Delta x\right)^2}{2}\tfrac{\partial^2\psi(x\prime,t)}{\partial x\prime^2}\Big]d\left(\Delta x\right)\end{aligned} \tag{2.22}$$

Each term on the right-hand side is a Gaussian style integral again. The middle term has a single power of Δx so is an odd integral and zero. The remaining terms give

$$\Delta t\tfrac{\partial\psi(x\prime,t)}{\partial t} = -i\tfrac{\Delta t}{\hbar}V\left(x\prime\right)\psi\left(x\prime,t\right) + \tfrac{i\hbar\Delta t}{2m}\tfrac{\partial^2\psi(x\prime,t)}{\partial x\prime^2} \tag{2.23}$$

or in other words the Schrödinger equation.

2.5 Path integral for a free particle

The path integral provides a nice way to think about quantum mechanics but in truth the Schrödinger equation is usually easier to solve. Let us look at a very simple problem—a free particle—using the path integral approach though.

We will split the free particle's trajectory up into Δt time slices again (see figure 2.4 but now with $V = 0$). We have already determined that the kernel for motion over one time slice is

$$K(x_1,x_0) = \sqrt{\tfrac{m}{2\pi i\hbar\Delta t}}\,e^{i\frac{m}{2\hbar}\left[\frac{(x_1-x_0)^2}{\Delta t}\right]} \tag{2.24}$$

To combine two time slices we multiply the kernels for the two separate motions and integrate over the position of the central point as in (2.12)

$$K(x_2, x_0) = \left(\frac{m}{2\pi i\hbar\Delta t}\right) \int e^{i\frac{m}{2\hbar}\left[\frac{(x_1-x_0)^2}{\Delta t}+\frac{(x_2-x_1)^2}{\Delta t}\right]} dx_1 \tag{2.25}$$

which we can do using the final Gaussian integral result from Appendix C.

$$K\left(x_2, x_0\right) = \sqrt{\frac{m}{2\pi i\hbar 2\Delta t}} e^{i\frac{m}{2\hbar}\left[\frac{\left(x_2-x_0\right)^2}{2\Delta t}\right]} \tag{2.26}$$

Note that all that has happened is that we have recovered the result for one time slice but with the time doubled and the distance travelled lengthened. One can keep repeating the above calculation adding time slices and the final result for the whole motion is then just

$$\begin{aligned} K(x_n, x_0) &= \sqrt{\frac{m}{2\pi i\hbar n\Delta t}} e^{i\frac{m}{2\hbar}\left[\frac{(xn-x_0)^2}{n\Delta t}\right]} \\ &= \sqrt{\frac{m}{2\pi i\hbar(t_n-t_0)}} e^{i\frac{m}{2\hbar}\left[\frac{(xn-x_0)^2}{(t_n-t_0)}\right]} \end{aligned} \tag{2.27}$$

Note the form of the exponential is easy to remember because it is just exp($i\Delta t$ KE/ $\hbar$) with KE the classical kinetic energy assuming constant velocity.

2.6 Interpreting the free particle kernel

We can see that this answer encodes a number of QM results we already know. First set $x_0 = 0$ and $t_0 = 0$ for simplicity so

$$K\left(x, t\right) = \sqrt{\frac{m}{2\pi i\hbar t}} e^{i\frac{mx^2}{2\hbar t}} \tag{2.28}$$

From (2.8) we know that $K(x, t) = \psi(x, t)$ is a free particle wave function if the particle started from a Dirac delta function at the origin. Now if we plot the real part of $K(x, t)$ at some later t it looks like figure 2.5.

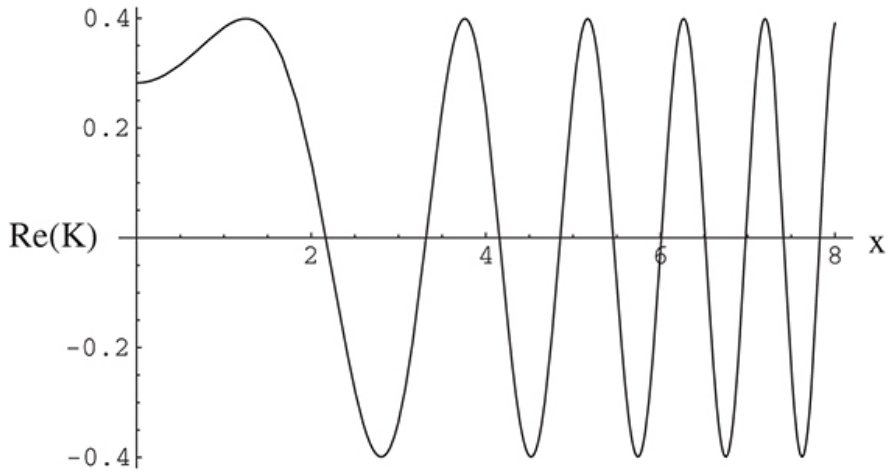


Figure 2.5. The real part of the kernel for a free particle plotted against position at some fixed time (it takes the form $\cos\ x^2$).

It is a wave whose wavelength shortens as we go to larger *x*. Classically for a particle to have got to some *x* in time *t* it must have

$$p = m\frac{x}{t} \tag{2.29}$$

The quantum mechanical version of this result is that the approximate wavelength of the kernel at some *x* is given by

$$\begin{aligned} \Delta\,\text{phase} &= 2\pi \\ &= \tfrac{m(x+\lambda)^2}{2\hbar t} - \tfrac{mx^2}{2\hbar t} \\ &\simeq \tfrac{mx\lambda}{\hbar t} \end{aligned} \tag{2.30}$$

where we have expanded in λ/x. We find

$$\lambda = \frac{2\pi\hbar}{mx/t} = \frac{h}{p} \tag{2.31}$$

a familiar result. The interpretation is that the higher momentum (smaller wavelength) components of the wavepacket travel further out in a given time.

Similarly we can fix *x* in $K(x, t)$ and plot the real part against *t* as shown in figure 2.6. We can work out the period of the wave at some *t* as we did the

wavelength above

$$\begin{aligned} 2\pi &= \tfrac{mx^2}{2\hbar t} - \tfrac{mx^2}{2\hbar(t+T)} \\ &= \tfrac{mx^2}{2\hbar t}\left(1-(1+T/t)^{-1}\right) \\ &\simeq \tfrac{mx^2}{2\hbar t^2}T \end{aligned} \tag{2.32}$$

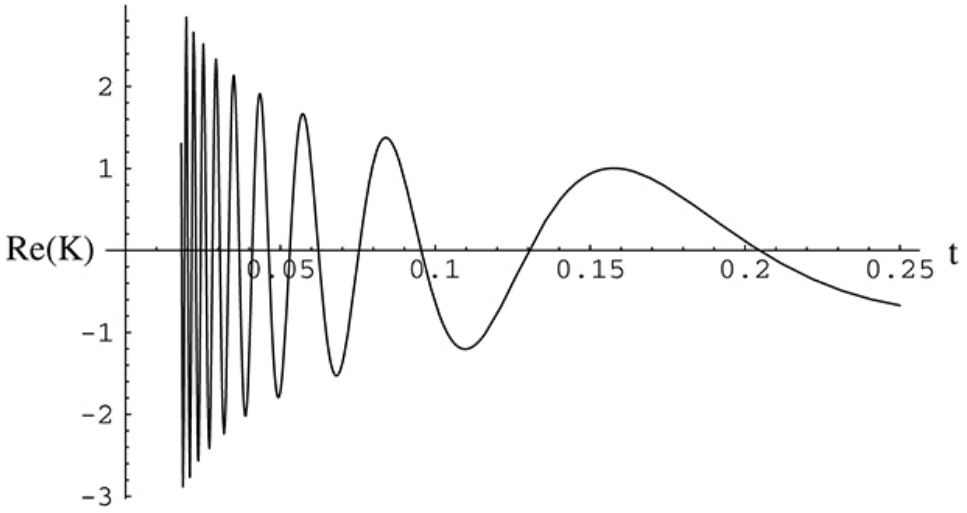


Figure 2.6. The real part of the kernel for a free particle plotted against time at some fixed position (it takes the form $\cos(1/t)$).

The angular frequency is

$$\omega = 2\pi/T = \tfrac{1}{\hbar}\tfrac{mx^2}{2t^2} \tag{2.33}$$

which, up to the factor of $\hbar$ is just the kinetic energy of the particle and hence

$$E = \hbar\omega \tag{2.34}$$

The interpretation is that the higher energy (higher frequency) components of the wavepacket pass by a fixed point earlier in time.

2.7 Barrier problems

Knowing the kernel for a free particle we can solve a number of problems involving particles starting from a point source, passing through a barrier and

eventually ending up on a screen.

To find the kernel associated with the particles' motion from the source to the screen we must sum $e^{iS/\hbar}$ for all the paths not blocked by the barrier. On these paths the particles are free, so

$$K = C(t)e^{\frac{imx^2}{2\hbar t}} \tag{2.35}$$

If we assume the source is at infinity then the distance from the source to any point on the barrier is the same. We can therefore treat each point on the barrier as an equal emitter of particles and just sum $e^{iS/\hbar}$ for the paths from the barrier to the screen. We find

$$K(\text{screen}) = A(t)\int_{\text{barrier}} e^{\frac{imx_{\text{path}}^2}{2\hbar t}} \; f(s) \; ds \tag{2.36}$$

Here $A(t)$ is a constant depending only on time, the exponential is the contribution from the action of each path, $f(s)$ is either 1 or 0 depending upon whether that point on the barrier is a hole or blocking the particle and finally ds sums over all points on the barrier. Compare this to (2.12).

Example: single slit

Let us look at a simple barrier with a single slit opening of width d as shown in figure 2.7. We will work in the narrow width approximation where $d \ll L_0$. The distance from a point P on the screen to each element of the slit is

$$L_0 + x\,\sin\,\theta, \quad -\frac{d}{2} < x < \frac{d}{2} \tag{2.37}$$

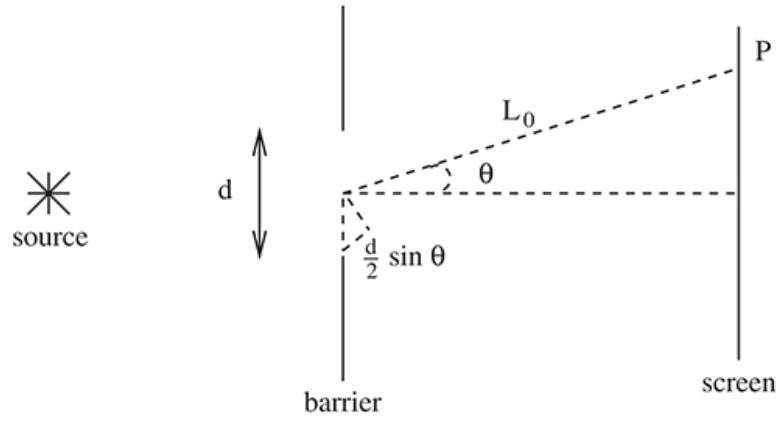


Figure 2.7. A single slit barrier problem.

Our expression for the kernel is therefore

$$K(P,t) = A(t)\int_{-\frac{d}{2}}^{\frac{d}{2}} e^{im(L_0 + x\,\sin\theta)^2/2\hbar t}\,dx \tag{2.38}$$

Since $L_0 \gg d$ then

$$\begin{aligned} K(P,t) &\simeq A\int_{-\frac{d}{2}}^{\frac{d}{2}} e^{imL_0^2/2\hbar t} e^{i2mL_0 x\,\sin\theta/2\hbar t}\,dx \\ &\simeq Ae^{imL_0^2/2\hbar t}\frac{\hbar t}{imL_0\,\sin\theta}\left[e^{imL_0 x\,\sin\theta/\hbar t}\right]_{-\frac{d}{2}}^{\frac{d}{2}} \\ &\simeq \frac{-iA(t)\hbar t e^{imL_0^2/2\hbar t}}{mL_0\,\sin\theta} 2\,\sin\left(\frac{mL_0\,d\,\sin\theta}{2\hbar t}\right) \end{aligned} \tag{2.39}$$

The probability of finding a particle at P is

$$\begin{aligned} |\,K(P,t)\,|^2 &= \frac{|A|^2\hbar^2 t^2}{m^2 L_0^2\sin^2\theta} 4\sin^2\left(\frac{mL_0\,d\,\sin\theta}{2\hbar t}\right) \\ &\simeq \text{constant}\ \frac{\sin^2(\alpha\,\sin\theta)}{\beta\sin^2\theta} \end{aligned} \tag{2.40}$$

where α and β are just constants. We can plot the rough form of this solution and find the form in figure 2.8.

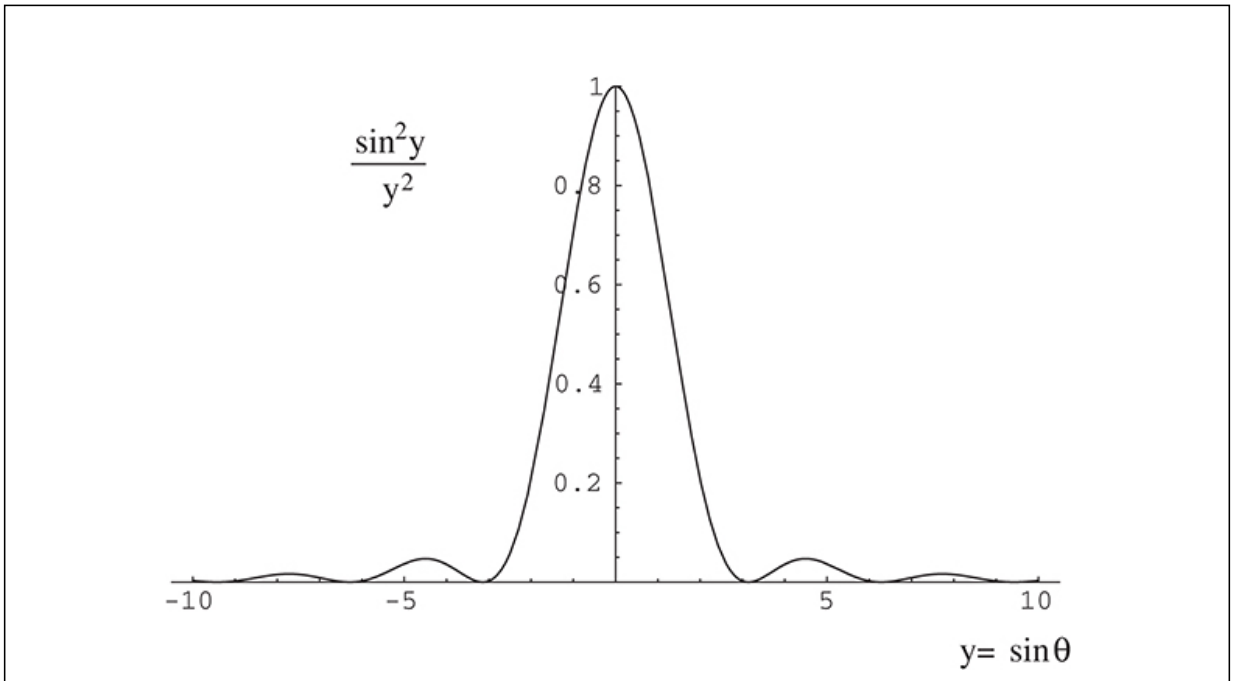


Figure 2.8. The probability function for the end-point of a particle passing through a single slit.

Note that the minima are when

$$\frac{mL_0\, d}{2\hbar t}\,\sin\,\theta = n\pi \tag{2.41}$$

i.e. when

$$d\,\sin\,\theta = \frac{2\hbar t}{mL_0} n\pi = n\frac{h}{p} = n\lambda \tag{2.42}$$

The usual result for destructive interference.

2.8 The kernel in terms of wave functions

In order to switch between the Schrödinger equation formalism and the path integral formalism it is helpful to have an expression for the kernel in terms of wave functions.

To find this form remember that

$$\psi(x,t_2) = \int_{-\infty}^{\infty} K(x,t_2;y,t_1)\psi(y,t_1)dy \tag{2.43}$$

Let us try now to get an equivalent statement starting from the time independent Schrödinger equation

$$H\psi = E\psi \tag{2.44}$$

If we start with some wavepacket at time t_1 we can use completeness to write it as

$$\psi(x,t_1) = \sum_{n=1}^{\infty} c_n\,\phi_n\,(x) \tag{2.45}$$

Equally we can invert this expression to give

$$c_n = \int_{-\infty}^{\infty} \phi_n^{*}\,(y)\psi\,(y,t_1)dy \tag{2.46}$$

We have used the orthogonality of the wave functions here to pick out the coefficient of a particular ϕ_n by multiplying by ϕ_n^* and integrating over all space. Again we have switched $x \to y$ to remind that the answer does not depend on the integration variable.

Furthermore, we know how $\psi(x, t_1)$ evolves in time to time t_2

$$\psi(x, t_2) = \sum_{n=1}^{\infty} c_n \phi_n(x) e^{iE_n(t_2 - t_1)/\hbar} \tag{2.47}$$

Substituting in our expression for the c_n we find

$$\psi(x, t_2) = \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \phi_n(x) e^{iE_n(t_2 - t_1)/\hbar} \phi_n^*(y) \psi(y, t_1) dy \tag{2.48}$$

Comparing back to the path integral result (2.43) we see that

$$\boxed{K(x, t_2; y, t_1) = \sum_{n=1}^{\infty} \phi_n(x) \phi_n^*(y) e^{iE_n(t_2 - t_1)/\hbar}} \tag{2.49}$$

Exercise 2.1: Show that for a free particle travelling from x_a at t_a to x_b at t_b the classical action is given by

$$S_{\text{classical}} = \frac{1}{2} m \frac{(x_b - x_a)^2}{(t_b - t_a)}$$

Exercise 2.2: Perform the Gaussian integral

$$\int_{-\infty}^{\infty} e^{-\alpha x^2 - \beta x} dx$$

Hint: Complete the square!

Exercise 2.3: Consider a non-relativistic, free particle of mass m travelling in two dimensions between two points A and B on the x-axis equally spaced about the y-axis. Consider paths where the particle travels in a straight line at constant speed to an arbitrary point on the y-axis and then in a straight line at the same speed to B, taking total time T. Calculate the action for these paths. Argue that classically the particle will travel in a straight line. Quantum mechanically the path is smeared. Estimate the width of the path when the particle crosses the y-axis.

Exercise 2.4: A massive, non-relativistic particle emitted by a source at infinity encounters a sheet of absorbing material with a circular hole of side a in it. Derive an expression for the quantum probability for finding the particle at a distance d along the axis of the hole on the far side at a time T.

Appendix C. Gaussian integrals

We will need to know the results of the following integrals (but we will need a few tricks in order to calculate them)

$$\boxed{I_n(\alpha) = \int_{-\infty}^{\infty} x^n e^{-\alpha x^2} dx} \tag{C.1}$$

- Firstly consider when $n = 0$. The trick is to calculate I_0^2

$$I_0^2(\alpha) = \int_{-\infty}^{\infty} e^{-\alpha x^2} dx \int_{-\infty}^{\infty} e^{-\alpha y^2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\alpha(x^2+y^2)} dx dy \tag{C.2}$$

This is a two dimensional integral in the x, y-plane and we can switch to polar coordinates r, θ

$$I_0^2(\alpha) = \int_0^{\infty} \int_0^{2\pi} e^{-\alpha r^2} (r dr d\theta) \tag{C.3}$$

since $rdr = dr^2/2$

$$I_0^2(\alpha) = \tfrac{1}{2} \tfrac{1}{(-\alpha)} [\theta]_0^{2\pi} \left[e^{-\alpha r^2} \right]_0^{\infty} = \tfrac{\pi}{\alpha} \tag{C.4}$$

and thus

$$\boxed{I_0(\alpha) = \sqrt{\tfrac{\pi}{\alpha}}} \tag{C.5}$$

- When n is an ODD number the integral is ODD and therefore zero.
- To obtain the result for EVEN n note that

$$\boxed{I_{2n}(\alpha) = (-1)^n \tfrac{d^n}{d\alpha^n} I_0(\alpha)} \tag{C.6}$$

Thus, for example

$$I_2 = \int_{-\infty}^{\infty} x^2 e^{-\alpha x^2} dx = -\tfrac{d}{d\alpha} I_0(\alpha) = \tfrac{1}{2\alpha} \sqrt{\tfrac{\pi}{\alpha}} \tag{C.7}$$

Finally, we shall also need the related integral

$$J = \int_{-\infty}^{\infty} e^{-\alpha \left[(x_1 - x_0)^2 + (x_2 - x_1)^2 \right]} dx_1 \tag{C.8}$$

which is simplified by noting that

$$(x_1 - x_0)^2 + (x_2 - x_1)^2 = 2\left[x_1 - \frac{(x_2+x_0)}{2}\right]^2 + \frac{(x_2-x_0)^2}{2} \tag{C.9}$$

now if we change the integration variable to $w = x_1 - (x_2 + x_0)/2$ ($dw = dx_1$) we find

$$J = \int_{-\infty}^{\infty} e^{-2\alpha w^2} e^{-\frac{\alpha}{2}(x_2-x_0)^2} dw \tag{C.9}$$

$$J = \sqrt{\frac{\pi}{2\alpha}} e^{-\frac{\alpha}{2}(x_2-x_0)^2} \tag{C.10}$$

Appendix D. Scattering theory

Consider an experiment such as that in figure D1 in which a particle scatters off some weak potential. We have a trajectory based intuition in such problems and so the path integral formalism is a natural place to start. The kernel governing the motion in the potential is

$$K_V(b,a) = \sum_{\text{paths}} e^{\frac{i}{\hbar}\int_{t_a}^{t_b}\left(\frac{m}{2}\dot{x}^2 - V\right)dt} \tag{D.1}$$

If the potential is weak we can expand the exponential in V

$$\begin{aligned} K_V(b,a) &= \sum_{\text{paths}} e^{\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m}{2}\dot{x}^2 dt}\left(1 - \frac{i}{\hbar}\int_{t_a}^{t_b} V(x,t)dt + \cdots\right) \\ &= K_0(b,a) - \frac{i}{\hbar}\int_{t_a}^{t_b} F(s)ds + \cdots \end{aligned} \tag{D.2}$$

where

$$F(s) = \sum_{\text{paths}} e^{\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m}{2}\dot{x}^2 dt} V(x(s), s) \tag{D.3}$$

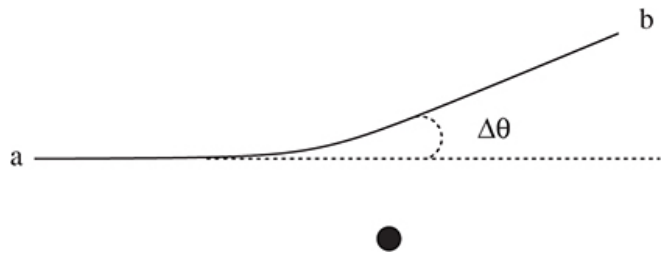


Figure D1. A moving particle scattering off a stationary target.

F(*s*) is the free kernel but with each path weighted by the value of the potential where the path is at time *s*. If that point is *c* we can draw the motion as in figure D2.

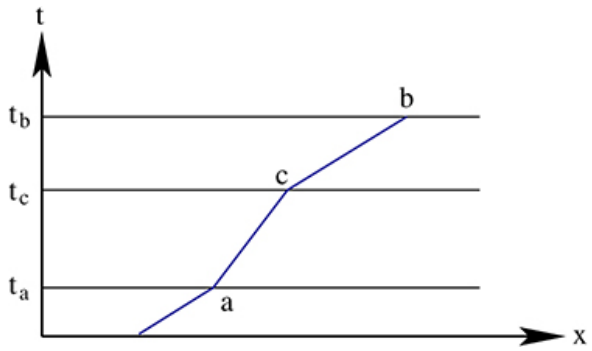


Figure D2. Splitting a trajectory from *a* to *b* at point *c*

As we saw before, because the action can be split into two parts (the motion $a \to c$ then $c \to b$) we can rewrite the kernel as two propagators, provided we remember to let the mid-point take all possible values, so

$$F\left(t_c\right) = \int_{-\infty}^{\infty} K_0\left(b,c\right) V\left(x_c,t_c\right) K_0\left(c,a\right) dx_c \tag{D.4}$$

Thus

$$K_V\left(b,a\right) = K_0\left(b,a\right) - \tfrac{i}{\hbar} \int_{t_a}^{t_b} \int_{-\infty}^{\infty} K_0\left(b,c\right) V\left(x_c,t_c\right) K_0\left(c,a\right) dx_c dt_c + \cdots \tag{D.5}$$

This expression is easily interpreted as a set of diagrams where the particle travels freely except at points where it scatters one or more times with the potential as shown in figure D3.

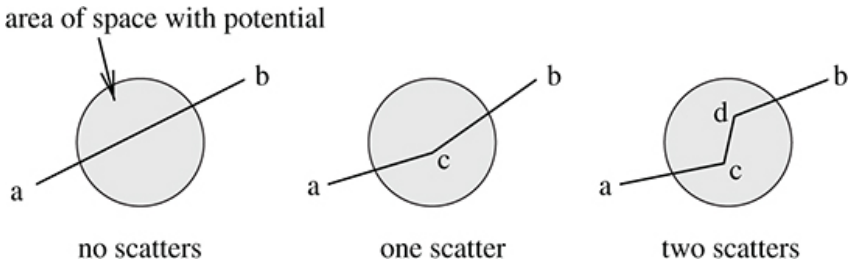
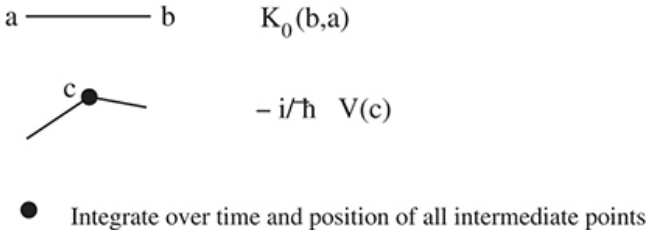


Figure D3. A diagrammatic representation of the perturbative expansion of the kernel in (2.65).

Feynman constructed a set of rules that relate each diagram to the mathematical formula:

Feynman rules



Thus for example we can immediately write down the form of the third diagram above

$$K_2 = \left(-\frac{i}{\hbar}\right)^2 \int\int K_0\left(b,d\right)V\left(d\right)K_0\left(d,c\right)V\left(c\right)K_0\left(c,a\right)\,dt_c\,dx_c\,dt_d\,dx_d \tag{D.6}$$

This is clearly a useful mnemonic.

D.1 Traditional time dependent perturbation theory

We can rewrite the path integral perturbation theory results in terms of wave functions. Remember we showed the kernel could be written in terms of eigenfunctions of the Hamiltonian in problems where the potential *U* is time independent

$$K_U\left(b,a\right) = \sum_n \phi_n\left(x_b\right)\phi_n^*\left(x_a\right)e^{-iE_n(t_b-t_a)/\hbar} \tag{D.7}$$

Now imagine adding a small perturbation $V_p(x,\ t)$. We can again expand the exponent of the path integral and obtain

$$K_V\left(b,a\right)\ =\ K_U\left(b,a\right)-\frac{i}{\hbar}\int\int K_U\left(b,c\right)V_p\left(c\right)K_U\left(c,a\right)dx_c\,dt_c+\cdots. \tag{D.8}$$

Substituting in we get

$$K_V\left(b,a\right)=\sum_n\sum_m\lambda_{mn}(t_b,t_a)\phi_m(x_b)\phi_n^*(x_a) \tag{D.9}$$

where

$$\begin{aligned}\lambda_{mn}\ &=\delta_{mn}e^{-iE_n(t_b-t_a)/\hbar}\\&\quad-\frac{i}{\hbar}\int\int\phi_m^*(x_c)V(x_c,t_c)\phi_n(x_c)dx_c\ \ e^{-i[E_m(t_c-t_b)-E_n(t_c-t_a)]/\hbar}dt_c\end{aligned} \tag{D.10}$$

What is the physical meaning of λ_{mn}? Remember that if we begin in the state ϕ_n at t_a then the wave function evolves so at t_b

$$\begin{aligned}\psi\left(x_b,t_b\right)&=\int K_V\left(b,a\right)\phi\left(x_a\right)dx_a\\&=\sum_k\sum_l\lambda_{kl}\phi_k\left(x_b\right)\int\phi_l^*\left(x_a\right)\phi_n\left(x_a\right)dx_a\\&=\sum_k\lambda_{kn}\phi_k\left(x_b\right)\end{aligned} \tag{D.11}$$

The λ_{mn} are the expansion coefficients of the wave form at t_2. Since we started in the single state ϕ_n their square is the probability that the state will be in a different state at time t_b. They are therefore called the *transition amplitudes*.

D.2 Initial response to a perturbation

One example we can make progress with is studying the initial response of a system to the imposition of a perturbation at, say, the time $t_a = 0$. Thus

$$V_p=0\qquad\text{for }\ t<0\qquad\qquad V_p=V_p\left(x\right)\qquad\text{for }\ t\geqslant0 \tag{D.12}$$

Suppose the system begins in the state Φ_{0n} what is the transition amplitude to a *different* state Φ_{0m} at time $t_b = T$?

$$\begin{aligned}e^{i(E_mt_a-E_nt_b)/\hbar}\lambda_{mn}\ &\simeq-\frac{i}{\hbar}\int\phi_m^*(x_c)V_p(x_c)\phi_n(x_c)dx_c\int_0^T e^{-i(E_n-E_m)t_c/\hbar}dt_c\\&\simeq\frac{(V_p)_{mn}}{E_m-E_n}\left[e^{-i(E_m-E_n)T/\hbar}-1\right]\end{aligned} \tag{D.13}$$

where for convenience we have defined $(V_p)_{mn} = \int \overset{*}{\phi}_m V_p \phi_n dx$

The probability of a transition from the initial state *n* to a final state *m* is given by

$$P_{mn}(T) \simeq \overset{*}{\lambda}_{mn}(T)\lambda_{mn}(T) = \mid \lambda_{mn}(T) \mid^2 \tag{D.14}$$

which is

$$\boxed{P_{mn}(t) \simeq \frac{\frac{\overset{*}{(V_p)}_{mn}(V_p)_{mn}}{\hbar^2}\sin^2\left[\frac{(E_m - E_n)t}{2\hbar}\right]}{\left[\frac{(E_m - E_n)}{2\hbar}\right]^2}} \tag{D.15}$$

Plotting the factor $\sin^2 x/x^2$ at fixed *t* gives a curve of the form in figure D4.

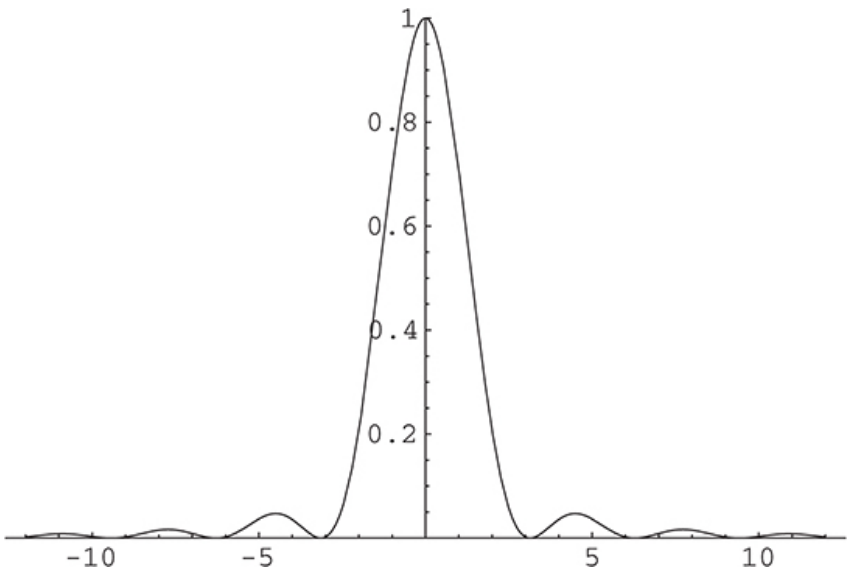


Figure D4. A plot of $\sin^2 x/x^2$ which gives the form of the probability for a system to scatter by energy Δ*E* after a perturbation.

Its maximum is at $E_m = E_n$ and it falls to zero when $\frac{E_m - E_n}{2\hbar} = \frac{\pi}{T}$. Essentially only transitions between states in this energy range occur.

D.3 Example: perturbed square well II

Consider the perturbed square well from A.1. The probability of a transition from the ground state to an excited state as a result of imposing the perturbation is

$$P_{1n}(t) \simeq \frac{|(V_p)_{1n}|^2}{\hbar^2} \frac{\sin^2\left[\frac{(E_n-E_1)}{2\hbar}t\right]}{(E_n-E_1)^2/4\hbar^2} \tag{D.16}$$

The energies of the level are already known so the only thing to calculate is

$$\begin{aligned}(V_p)_{1n} &= \int_{-\infty}^{\infty} \psi_1^*(x) V_p \psi_n(x) dx \\ &= -\frac{2\alpha}{a}\int_0^{a/2} x \sin\frac{\pi x}{a}\sin\frac{n\pi x}{a}dx \\ &= \frac{\alpha}{a}\int_0^{a/2} x\left[\cos\frac{(n+1)\pi x}{a} - \cos\frac{(n-1)\pi x}{a}\right]\end{aligned} \tag{D.17}$$

Integrating by parts one finds, for example

$$(V_p)_{13} = \frac{a\alpha}{2\pi^2} \tag{D.18}$$

D.4 Fermi's golden rule

In most systems there are many states with the same energy. If there are dN_j states with energy E_{0j} to $E_{0j} + dE_{0j}$ then the total transition probability from a state i is

$$P_i(t) = \frac{1}{\hbar^2}\int_{-\infty}^{\infty} \lambda_j^*(t)\lambda_j(t) dN_j = \int_{-\infty}^{\infty} \lambda_j^*(t)\lambda_j(t)\frac{dN_j}{dE_j}dE_j \tag{D.19}$$

dN_j/dE_{0j} is just the 'density of states', ρ_j. Substituting in for the $a_j(t)$ we have

$$P_i(t) \simeq \int_{-\infty}^{\infty} (V_p)_{ji}^* (V_p)_{ji} \rho_j \frac{\sin^2\left[\frac{(E_{0j}-E_{0i})t}{2\hbar}\right]}{\left[\frac{(E_{0j}-E_{0i})}{2\hbar}\right]^2} dE_{0j} \tag{D.20}$$

The $\sin^2 x/x^2$ bit of the integrand is so peaked though we can assume the integral only gets contributions around the central peak. Since it is so narrow $(V_p)_{ji}$ and ρ_j will not vary much in this range and we can write

$$P_i(t) \simeq \frac{1}{\hbar^2}\left(V_p\right)^*_{ji}\left(V_p\right)_{ji}\rho_j \int_{-\infty}^{\infty} \frac{\sin^2\left[\frac{(E_{0j}-E_{0i})t}{2\hbar}\right]}{\left[\frac{(E_{0j}-E_{0i})}{2\hbar}\right]^2} dE_{0j} \tag{D.21}$$

In fact the integral can now be calculated

$$\int_{-\infty}^{\infty} \frac{\sin^2 z}{z^2} dz = \pi \tag{D.22}$$

and thus

$$P_i(t) \simeq \frac{2\pi}{\hbar}\left(V_p\right)^*_{ji}\left(V_p\right)_{ji}\rho_j t \tag{D.23}$$

Note it is proportional to the elapsed time. The *transition rate* is the time derivative of this probability

$$\boxed{R_i \equiv \frac{dP_i(t)}{dt} = \frac{2\pi}{\hbar}\left|\left(V_p\right)_{ji}\right|^2 \rho_j} \tag{D.24}$$

This is 'Fermi's golden rule' for the transition rate from a state *i* to all nearby states as a result of a perturbation.

Exercise D1: A hydrogen atom consists of an electron subject to the potential

$$V(x) = -\frac{e^2}{4\pi\epsilon_0 r}$$

the ground state wave function for the electron is given by

$$\psi(x,t) = \frac{1}{\sqrt{\pi}} a_0^{-3/2} e^{-r/a_0} e^{-iEt/\hbar}$$

where a_0 is the Bohr radius ($a_0 = 4\pi\epsilon_0\hbar^2/m_e\, e^2$), and the ground state energy is

$$E = \frac{m_e e^4}{\left(4\pi\epsilon_0\right)^2 2\hbar^2}$$

- If the atom is placed in a uniform weak electric field pointing in the *z*-direction find the first order shift in energy of this state.

Exercise D2: For a particle in a one-dimensional, simple harmonic potential well

$$V(x) = \tfrac{1}{2}kx^2$$

the wave functions for the system are given by

$$\psi_n(x,t) = \chi_n(x)e^{-iE_n t/\hbar}$$

where

$$\chi_n(x) = \left(\sqrt{\tfrac{\alpha}{\pi}}\left(\tfrac{1}{2^n n!}\right)\right)^{1/2}\mathrm{H}_n\left(\sqrt{\alpha}x\right)e^{-\frac{1}{2}\alpha x^2}$$

for n = 0, 1, 2, ... and where $\alpha = m\omega/\hbar(\omega = \sqrt{k/m})$. The H_n are the Hermite polynomials of order n.

$$\mathrm{H}_0(y) = 1, \qquad \mathrm{H}_1(y) = 2y, \qquad \mathrm{H}_0(y) = 2 - 4y^2, \cdots.$$

The states have energy $\left(\frac{1}{2} + n\right)\hbar\omega$.

- Show that for n = 0 and n = 1 these are the solutions of the time independent Schrödinger equation subject to the problem's boundary conditions.
- If the potential is perturbed by a term

$$\Delta V(x) = -\kappa x^4$$

calculate the shift in the energy of the ground state a long time after the perturbation is imposed. You will need to make use of the Gaussian integral results in Appendix C.

- What are the probabilities, for small times, of a transition from the ground state to the first and second excited states if this perturbation is applied at t = 0?

IOP Concise Physics

Theories of Matter, Space and Time, Volume 2

Quantum theories

N Evans and S F King

Chapter 3

Relativistic quantum mechanics

In this chapter we study the quantum theory of relativistic particles. In a particle accelerator we are interested in, for example, the interactions of highly energetic electrons so the need to combine relativity and quantum mechanics is pressing. Some remarkable results will come out of this synthesis. In particular we will theoretically predict the existence of anti-particles and also fermion spin. Sit back and enjoy!

3.1 Relativity review

A more thorough introduction to special relativity is provided in the prequel book *Theories of Matter, Space and Time: Classical Theories* but here is a quick recap.

In relativity an event is described by the four coordinates of a four-vector

$$x^\mu = \left(ct, \vec{x}\right) \tag{3.1}$$

Under Lorentz transformations (LT) it transforms—a familiar example of a LT is a boost along the z-axis, for which

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}$$

with, as usual, $\beta = v/c$ and $\gamma = \left(1-\beta^2\right)^{-1/2}$. LTs can be thought of as generalized rotations.

x^μ is then a 4-vector since it transforms as

$$x^\mu \to x\prime^\mu = \Lambda^\mu_\nu x^\nu$$

The Greek labels μ, ν ... $\in\{0, 1, 2, 3\}$ denote Lorentz indices and the summation convention is used. (3.2)

The 'length' of the 4-vector $(c^2t^2 - \mid \vec{x} \mid^2)$ is invariant to LTs. In general we define the Minkowski scalar product of two 4-vectors x and y as

$$x \cdot y = x^\mu y^\nu g_{\mu\nu} = x^\mu y_\mu \tag{3.3}$$

where the metric

$$g^{\mu\nu} = g_{\mu\nu} = \mathrm{diag}\,(1, -1, -1, -1), \qquad g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu = \begin{cases} 1 \text{ if } \mu = \nu \\ 0 \text{ if } \mu \neq \nu \end{cases} \tag{3.4}$$

has been introduced. The last step in (3.3) is nothing but the definition of a covariant 4-vector (sometimes referred to as a co-vector)

$$x_\mu \equiv g_{\mu\nu} x^\nu \tag{3.5}$$

To formulate a coherent relativistic theory of dynamics we define kinematic variables that are also 4-vectors (i.e. transform as described above). For example, we define a 4-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} \tag{3.6}$$

where τ is the *proper time* measured by a clock moving with the particle. Everyone will agree as to what the clock says at a particular event so this measure of time is Lorentz invariant and u^μ transforms as x^μ. Note that since $t = \gamma\tau$ (this is the usual time dilation result)

$$u^\mu = \frac{dt}{d\tau}\frac{dx^\mu}{dt} = \gamma(c, \vec{v}) \tag{3.7}$$

and has invariant length

$$u^\mu u_\mu = \gamma^2(c^2 - \mid \vec{v} \mid^2) = c^2 \tag{3.8}$$

Similarly 4-momentum provides a relativistic definition of energy and momentum

$$p^\mu = mu^\mu \equiv \left(E/c, \vec{p}\right) \tag{3.9}$$

The invariant length gives us the crucial relation

$$p^\mu p_\mu = E^2/c^2 - \mid \vec{p} \mid^2 = m^2 c^2 \tag{3.10}$$

Note that ∂_μ is defined by,

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu}, \quad \partial_\mu x^\nu = \delta^\nu_\mu, \tag{3.11}$$

so $\nabla^i = -\partial^i$ and $\partial^\mu = (\partial^0, -\vec{\nabla})$.

We will use natural units henceforth. This firstly means redefining the unit of distance so that $c = 1$. Secondly, we will redefine the unit of energy so that $E = h\nu = 2\pi\nu$, i.e. set $\hbar = 1$. So mass, energy, inverse length and inverse time all have the same dimensions. Generally, think of energy E as the basic unit, e.g. mass m has units of GeV and distance x has unit GeV^{-1}.

3.2 The Klein–Gordon equation

For a free relativistic particle the total energy E is no longer given by the equation we used to derive the Schrödinger equation in chapter 1. Instead it is given by the Einstein equation

$$E^2 = \vec{p}^{\,2} + m^2. \tag{3.12}$$

In position space we write the energy–momentum operator as

$$\hat{p}^\mu \to i\partial^\mu \quad (\hat{E}, \hat{\vec{p}}) = \left(i\frac{\partial}{\partial t}, -i\vec{\nabla}\right) \tag{3.13}$$

Note that the minus sign in the spatial parts of ∂^μ match and explain the sign in the standard operator relations (1.4) and (1.5).

Substituting these operators into (3.10) acting on a wave function $\phi(x)$ generates the Klein–Gordon equation

$$(\Box + m^2)\phi(x) = 0$$

where we have introduced the box notation,

$$\Box = \partial_\mu \partial^\mu = \partial^2 / \partial t^2 - \nabla^2 \tag{3.14}$$

and x is the 4-vector $(t, \vec{x})$. (3.15)

The Klein–Gordon equation has plane wave solutions:

$$\phi(x) = N e^{-i(Et - \vec{p}.\vec{x})} \tag{3.16}$$

where N is a normalization constant and if we substitute the solution into the equation we recover

$$E = \pm \sqrt{|\ \vec{p}\ |^2 + m^2} \tag{3.17}$$

3.2.1 Problems in the Klein–Gordon equation

There are two problems with this equation though. Indeed historically Schrödinger originally began by writing down this relativistic equation but then retreated to his non-relativistic equation because of the issues we will discuss here.

Firstly, there are both positive and negative energy solutions because of the square root in (3.17). The negative energy solutions pose a severe problem if you try to interpret ϕ as a wave function as we are trying to do. The spectrum is no longer bounded from below, and you can extract arbitrarily large amounts of energy from the system by driving it into ever more negative energy states. The system is completely unstable! Any external perturbation capable of pushing a particle across the energy gap of $2m$ between the positive and negative energy continuum of states can uncover this difficulty. Furthermore, we cannot just throw away these solutions as unphysical since they appear as part of the complete set of states (as discussed in section 1.8) for the Klein–Gordon equation and so emerge in almost any problem.

A second problem with the wave function interpretation arises when trying to find a probability density. In relativity a density transforms under boosts, since lengths contract, and forms part of a 4-vector with the current density. Here since ϕ is Lorentz invariant, $|\ \phi\ |^2$ does not transform like a density so we will not have a Lorentz covariant continuity equation

$$\partial_t \rho + \vec{\nabla}.\vec{J} = 0 \quad \text{or} \quad \partial_\mu J^\mu = 0$$

(3.18)

We can derive a candidate for the probability density/current by finding something which does satisfy such a continuity equation as we did in section 1.4 for the Schrödinger equation. As there, one starts with the Klein–Gordon equation multiplied by ϕ^* and subtracts the complex conjugate of the KG equation multiplied by ϕ. (3.18) emerges with $J^\mu = (\rho, \vec{J})$ and

$$\rho \equiv i\left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right), \tag{3.19}$$

$$\vec{J} \equiv -i(\phi^* \vec{\nabla}\phi - \phi \vec{\nabla}\phi^*) \tag{3.20}$$

It is thus natural to interpret ρ as a probability density and $\vec{J}$ as a probability current.

However, for a plane wave solution (3.16), $\rho = 2 \mid N \mid^2 E$, so ρ is not positive definite since we have already found E can be negative. This clearly makes no sense!

We should note that the equation is a candidate to describe spinless relativistic particles only since there is just a single probability density describing a particle state (as in the Schrödinger equation).

Exercise 3.1: Derive (3.19) and (3.20).

3.2.2 Feynman–Stückelberg interpretation

The Klein–Gordon equation appears to have unacceptable negative energy states and negative probabilities for those states if ϕ is interpreted as the single particle wave function. Many years later Feynman and Stückelberg came to the rescue and proposed a way forwards to make sense of the equation. It is linked to Pauli's idea that one does not directly measure the number of particles. You can only detect them via their charges through an interaction. This means you cannot observe the probability density but only the charge density/current (qJ^μ) and that can be negative!

The Klein–Gordon equation has a time reversal symmetry so in addition to states propagating forwards in time that look like e^{-iEt} there are solutions that travel backwards in time like e^{+iEt}. Normally we would throw away these backwards propagating solutions for causality's sake (you do not want to be able to kill your Grandfather!). However, if E can

be negative these two sets of states become confused. Does $e^{-i(-E)t}$ propagate forwards in time with negative energy or backwards in time with positive energy?

Feynman and Stückelberg proposed that it is possible to consistently keep just half of the solutions to the Klein–Gordon equations but not the ones you would immediately guess. They suggested to keep positive energy states propagating forwards in time, but only negative energy states that propagate backwards in time! We interpret these states as positive energy states moving forwards in time ($e^{+i(-E)t}$). However, in the solutions the charge density/current is opposite sign. These particles look like opposite charge versions of the normal particle states propagating forwards in time. This is a prediction of anti-particles!

Now we find a theory that is consistent with the requirements of causality and that has none of the aforementioned problems. In fact, the negative energy states cause us problems only so long as we think of them as real physical states propagating forwards in time. Therefore, we should interpret the emission (absorption) of a negative energy particle with momentum p^μ as the absorption (emission) of a positive energy antiparticle with momentum $-p^\mu$.

In order to get more familiar with this picture, consider a process with a π^+ and a photon in the initial state and final state. In figure 3.1(a) the π^+ starts from the point A and at a later time t_1 emits a photon at the point $\vec{x}_1$. If the energy of the π^+ is still positive, it travels on forwards in time and eventually will absorb the initial state photon at t_2 at the point $\vec{x}_2$. The final state is then again a photon and a (positive energy) π^+.

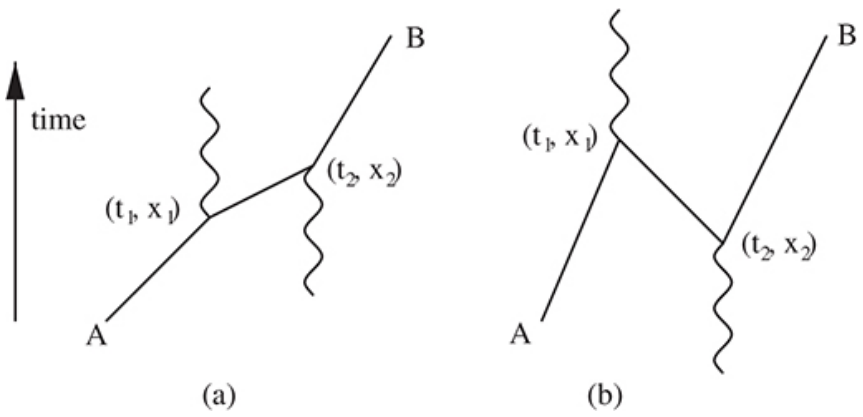


Figure 3.1. Pion–photon scatterings in which the intermediate pion has (a) positive energy and travels forwards in time and (b) has negative energy and travels backwards in time.

There is another process however, with the same initial and final state, shown in figure 3.1(b). Again, the π^+ starts from the point A and at a later time t_2 emits a photon at the point $\vec{x}_1$. But this time, the energy of the photon emitted is bigger than the energy of the initial π^+. Thus, the energy of the π^+ becomes negative and it is forced to travel backwards in time. Then at an *earlier* time t_1 it absorbs the initial state photon at the point $\vec{x}_2$, thereby rendering its energy positive again. From there, it travels forward in time and the final state is the same as in figure 2.1(a), namely a photon and a (positive energy) π^+.

In today's language, the process in figure 3.1(b) would be described as follows: in the initial state we have a π^+ and a photon. At time t_1 and at the point $\vec{x}_2$ the photon creates a π^+-π^- pair. Both propagate forwards in time. The π^+ ends up in the final state, whereas the π^- is annihilated at (a later) time t_2 at the point $\vec{x}_1$ by the initial state π^+, thereby producing the final state photon. To someone observing in real time, the negative energy state moving backwards in time looks to all intents and purposes like a negatively charged pion with positive energy moving forwards in time.

We have discovered anti-matter! The Feynman–Stückelberg interpretation revives the Klein–Gordon equation as a perfectly sensible theory of spinless particles and their anti-particles. Note the pions we have already mentioned are an example of spinless particles although they are fundamentally made from quarks. In 2012 the Higgs boson was discovered at the Large Hadron Collider and, to date, it seems to be the first example of a fundamental spinless particle.

3.3 Dirac equation

Historically the Klein–Gordon equation was believed to be sick although now we understand it is telling us about anti-particles. Dirac tried an alternative route to obtain a relativistic wave equation. To try to solve the problem of negative energy solutions Dirac wanted an equation which was first order in time derivatives. The free particle solution would then only generate one power of E and the pesky square root that gave negative energy solutions in the Klein–Gordon equation might be avoided.

His starting point was to assume a Hamiltonian of the form,

$$H_D = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m$$

where p_i are the three components of the momentum operator $\vec{p}$, and α_i and β are some unknown quantities, which, as will be seen below, cannot simply be commuting numbers. (3.21)

We should write the momentum operators explicitly in terms of their differential operators, using equation (3.13). Then the Dirac equation becomes, using the Dirac Hamiltonian in equation (3.21),

$$i\,\frac{\partial\psi}{\partial t} = \left(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\psi \tag{3.22}$$

which is the position space Dirac equation.

If ψ is to describe a free particle it must though satisfy the Klein–Gordon equation so that it has the correct energy–momentum relation. This requirement imposes relationships among α_1, α_2, α_3 and β. To see these, apply the operator on each side of equation (3.22) twice, i.e. iterate the equation,

$$-\frac{\partial^2\psi}{\partial t^2} = \left[-\alpha^i\alpha^j\nabla^i\nabla^j - i(\beta\alpha^i + \alpha^i\beta)m\nabla^i + \beta^2 m^2\right]\psi$$

with an implicit sum over i and j from 1 to 3. The Klein–Gordon equation by comparison is

$$-\frac{\partial^2\psi}{\partial t^2} = \left[-\nabla^i\nabla^i + m^2\right]\psi \tag{3.23}$$

If we do not assume that the α^i and β commute then the KG will be satisfied if

$$\begin{aligned}\alpha_i\alpha_j + \alpha_j\alpha_i &= 2\delta_{ij}\\ \beta\alpha_i + \alpha_i\beta &= 0\\ \beta^2 &= 1\end{aligned} \tag{3.24}$$

for $i, j = 1, 2, 3$.

The α_i and β cannot be ordinary numbers since they do not commute, but it is possible to give them a realization as matrices. In this case, ψ must become a column vector that the matrices act on. These multi-component objects are called *spinors*. Since these have more than one component our wave equation will describe probabilities for more than one particle—what these extra particles are we will investigate shortly.

In two dimensions a natural set of matrices for the $\vec{\alpha}$ would be the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.25}$$

However, there is no other independent 2 × 2 matrix with the right properties for β, and it turns out that the smallest number of dimensions for which the Dirac matrices can be realized is four. One choice is the *Dirac representation*:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.26}$$

Note that each entry above denotes a two-by-two block and that the 1 denotes the 2 × 2 identity matrix.

Exercise 3.2: Consider an eigenvector of α_i or β. Show, that since these matrices square to 1, their eigenvalues must be ±1.

Next, by considering the trace of the relations in (3.24), show that α_i and β are traceless.

Hence argue that the α_i and β matrices must be even dimensional.

There is a theorem due to Pauli that states that all sets of matrices obeying the relations in (3.24) are equivalent.

Exercise 3.3: Check explicitly that the matrices in (3.26) satisfy the Dirac algebra in (3.24).

3.3.1 Continuity equation

Again we need a candidate for $J^\mu = (\rho, \vec{J})$, the probability density/current for the particles being described. One proceeds as before by simply adding $\psi DE^\dagger + \psi^\dagger DE$ and rearranging.

Note here the 'dagger' symbol means complex conjugation and transposing. If one wants to make a single real number from a complex vector such as

$$A = \begin{pmatrix} a + ib \\ c + id \end{pmatrix} \tag{3.27}$$

then

$$A^\dagger A = (a - ib, c - id)\begin{pmatrix} a + ib \\ c + id \end{pmatrix} = a^2 + b^2 + c^2 + d^2 \tag{3.28}$$

does the trick!

Back to the continuity equation one finds

$$\rho = J^0 = \psi^\dagger \psi \qquad \vec{J} = \psi^\dagger \vec{\alpha} \psi \tag{3.29}$$

satisfy the continuity equation $\partial_\mu J^\mu = 0$.

Note that ρ is now positive definite, unlike in the Klein–Gordon equation—this seemed initially like a major achievement to Dirac.

Exercise 3.4: Derive (3.29).

3.3.2 Solutions to the Dirac equation

The wave function in the Dirac equation is a four component vector. To shed light on what this means let us look at free particle solutions.

We look for plane wave solutions of the form

$$\psi = \begin{pmatrix} \chi\left(\vec{p}\right) \\ \phi\left(\vec{p}\right) \end{pmatrix} e^{-i\left(Et - \vec{p}\cdot\vec{x}\right)} \tag{3.30}$$

Here $\phi\left(\vec{p}\right)$ and $\chi\left(\vec{p}\right)$ are two-component spinors that depend on momentum $\vec{p}$ but are independent of $\vec{x}$. If they depended on $\vec{x}$ then we would mess up the action of the energy and momentum operators which give the correct eigenvalues just by differentiation of the exponential term. We want to keep that property.

Using the Dirac representation of the matrices, and inserting the trial solution into the Dirac equation gives the pair of simultaneous equations

$$E\begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} m & \vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -m \end{pmatrix}\begin{pmatrix} \chi \\ \phi \end{pmatrix} \tag{3.31}$$

Particle at rest

The simplest example we can solve the Dirac equation for is when $\vec{p} = 0$, $m \neq 0$, which might represent an electron in its rest frame. The equations (3.31) decouple and become simply,

$$E\chi = m\chi \qquad E\phi = -m\phi \tag{3.32}$$

So, in this case, we see that χ corresponds to solutions with $E = m$, while ϕ corresponds to solutions with $E = -m$. Dirac had therefore failed to remove these negative energy solutions! In light of our earlier discussion of the Feynman–Stückelberg interpretation, we no longer need to recoil in horror at the appearance of these negative energy states. The same interpretation works here although Dirac came up with his own interpretation which is still a useful way to think about the problem.

Dirac's interpretation of negative energy

Dirac's interpretation depends on the assumption that we are describing fermions that obey the Pauli exclusion principle. He postulated the existence of a 'sea' of negative energy states (see figure 3.2). The vacuum or ground state has all the negative energy states full. An additional electron must now occupy a positive energy state since the Pauli exclusion principle forbids it from falling into one of the filled negative energy states. On promoting one of these negative energy states to a positive energy one, by supplying energy, an electron–hole pair is created, i.e. a positive energy electron and a hole in the negative energy sea. The hole is seen in nature as a positive energy positron. This was a radical new idea, and brought pair creation and antiparticles into physics.

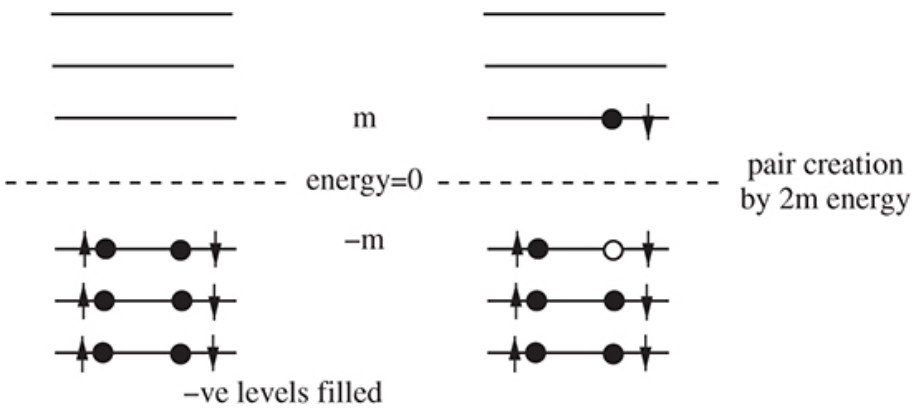


Figure 3.2. Dirac's filled negative energy states.

The problem with Dirac's hole theory is that it does not work for bosons. Such particles have no exclusion principle to stop them falling

into the negative energy states, releasing their energy.

General solutions

The negative energy solutions persist for an electron with $\vec{p} \neq 0$ for which the solutions to equation (3.31) are

$$\phi = \frac{\vec{\sigma} \cdot \vec{p}}{E + m} \chi, \qquad \chi = \frac{\vec{\sigma} \cdot \vec{p}}{E - m} \phi \tag{3.33}$$

Now we can substitute one of these equations into the other and use $\left(\vec{\sigma} \cdot \vec{p}\right)^2 = \left|\vec{p}\right|^2$. Explicitly

$$\left(\vec{\sigma}.\vec{p}\right)^2 = \begin{pmatrix} p_3 & p_1 - ip_2 \\ p_1 + ip_2 & -p_3 \end{pmatrix}^2$$

$$= \begin{pmatrix} \left(p_1\right)^2 + \left(p_2\right)^2 + \left(p_3\right)^2 & 0 \\ 0 & \left(p_1\right)^2 + \left(p_2\right)^2 + \left(p_3\right)^2 \end{pmatrix} = \left|\vec{p}\right|^2 I \tag{3.34}$$

We find that

$$\phi = \frac{\left(\vec{\sigma} \cdot \vec{p}\right)^2}{E^2 - m^2} \phi = \frac{\left|\vec{p}\right|^2}{E^2 - m^2} \phi \tag{3.35}$$

from which we deduce that the familiar $E = \pm \mid \sqrt{\left|\vec{p}\right|^2 + m^2} \mid$ is indeed satisfied by the solutions.

We write the positive energy solutions with $E = + \mid \sqrt{\left|\vec{p}\right|^2 + m^2} \mid$ as

$$\psi(x) = \begin{pmatrix} \chi \\ \frac{\vec{\sigma} \cdot \vec{p}}{E + m} \chi \end{pmatrix} e^{-i\left(Et - \vec{p} \cdot \vec{x}\right)} \tag{3.36}$$

while the general negative energy solutions with $E = - \mid \sqrt{\left|\vec{p}\right|^2 + m^2} \mid$ are

$$\psi(x) = \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E-m}\phi \\ \phi \end{pmatrix} e^{-i\left(Et-\vec{p}\cdot\vec{x}\right)} \tag{3.37}$$

for arbitrary constant ϕ and χ. Clearly when $\vec{p} = 0$ these solutions reduce to the positive and negative energy solutions discussed previously.

Let us rewrite the solutions, (3.36) and (3.37), in a standard form introducing the spinors $u_\alpha(s,\vec{p})$ and $v_\alpha(s,\vec{p})$. The label α = 1, 2, 3, 4 is a spinor index that often will be suppressed. Take the positive energy solution equation (3.36) and define

$$\sqrt{E+m}\begin{pmatrix} \chi_r \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi_r \end{pmatrix} e^{-ip\cdot x} \equiv u(s,p)e^{-ip\cdot x}. \tag{3.38}$$

Here we have used four vector notation to write $p\cdot x = Et - \vec{p}\cdot\vec{x}$.

For the negative energy solution of equation (3.37), we change the sign of the energy, $E \to -E$ (note that since E is already negative this makes it positive), and the three-momentum, $\vec{p} \to -\vec{p}$, to obtain,

$$\sqrt{E+m}\begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\phi_r \\ \phi_r \end{pmatrix} e^{ip\cdot x} \equiv v(s,p)e^{ip\cdot x} \tag{3.39}$$

In these two solutions E is now always positive and given by $E = \left(\left|\vec{p}\right|^2 + m^2\right)^{1/2}$. The argument s takes the values 1, 2 with

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.40}$$

The u-spinor solutions will correspond to particles and the v-spinor solutions to antiparticles. The role of the two χ's will become clear in a following section, where it will be shown that the two choices of s are spin labels.

Orthogonality and completeness

Our solutions to the Dirac equation take the form

$$\psi = N u_s e^{-ip\cdot x} \qquad \psi = N v_r e^{ip\cdot x} \qquad r,s = 1,2 \tag{3.41}$$

The N is a normalization factor. We have already included a normalization factor $\sqrt{E+m}$ in our spinors. With this factor,

$$u^\dagger(r,p)u(s,p) = v^\dagger(r,p)v(s,p) = 2E\delta^{rs}. \tag{3.42}$$

This corresponds to the standard relativistic normalization of $2E$ particles per unit volume—this makes $u^\dagger u$ and hence $\psi^* \psi$ transforms like the time component of a 4-vector under Lorentz transformations, as it must to be the zeroth component of J^μ. Note that the spinors are *orthogonal*.

We must further normalize the spatial wave functions. In fact a plane wave is not normalizable in an infinite space so we will work in a large box of volume V

$$\int \psi_a^\dagger \psi_b d^3x = 2E\, N^2\, V\, \delta_{ab} \tag{3.43}$$

where a, b run over the possible values of r, s and the value of p. Note again the orthogonality of the states. To normalize to $2E$ particles per unit volume we must set $N = 1/\sqrt{V}$. Sometimes it is helpful to normalize so that $\int \psi_a^\dagger \psi_b d^3x = \delta_{ab}$ (so that there is one particle per unit volume) in which case $N = 1/\sqrt{2EV}$—this is not a Lorentz invariant normalization so must be done in a particular frame.

Remember that the solutions to the wave equation form a complete set of states meaning that we can expand (like a Fourier expansion) an arbitrary function $\chi(x)$ in terms of them

$$\chi(x) = \sum_n a_n \psi_n(x) \tag{3.44}$$

The a_n are the equivalent of Fourier coefficients and if χ is a wave function in some quantum mixed state then $\mid a_n \mid^2$ is the probability of being in the state ψ_n.

3.3.3 Spin

Now it is time to justify the statements we have been making that the Dirac equation describes spin. We will see that the two components of each of the positive and negative energy solutions describe spin up and spin down states of a spin 1/2 fermion.

Conserved quantities: A conserved quantity in quantum mechanics is described by a time independent operator that commutes with the Hamiltonian. To prove this we evaluate the time derivative of the expectation value of some operator, $\hat{F}$,

$$\frac{d\langle f\rangle}{dt} = \frac{d}{dt}\int \psi^\dagger \hat{F}\psi dx$$

$$= \int \frac{\partial\psi^\dagger}{\partial t}\hat{F}\psi dx + \int \psi^\dagger \hat{F}\frac{\partial\psi}{\partial t}dx \tag{3.45}$$

Note that $\hat{F}$ is time independent here. Now we use the wave equation

$$\hat{H}\psi = i\hbar\frac{\partial\psi}{\partial t} \tag{3.46}$$

to find

$$\frac{d\langle f\rangle}{dt} = \frac{i}{\hbar}\int \psi^*\left(\hat{H}\hat{F} - \hat{F}\hat{H}\right)\psi dx \tag{3.47}$$

So if the commutator $[\hat{F},\hat{H}]$ vanishes the expectation value is conserved.

Now the Dirac Hamiltonian in momentum space is given in equation (3.21) as

$$H_D = \vec{\alpha}\cdot\vec{p} + \beta m \tag{3.48}$$

and the orbital angular momentum operator is

$$\vec{L} = \vec{R}\times\vec{p}$$

$\vec{L}$ and H_D may not commute because they contain x and p which do not commute ($[x_i, p_j] = i\delta_{ij}$). Evaluating the commutator of $\vec{L}$ with H_D,

$$\begin{aligned}\left[\vec{L}, H_D\right] &= \left[\vec{R}\times\vec{p}, \vec{\alpha}\cdot\vec{p}\right] \\ &= \left[\vec{R}, \vec{\alpha}\cdot\vec{p}\right]\times\vec{p} \\ &= i\vec{\alpha}\times\vec{p}\end{aligned} \tag{3.49}$$

we see that the orbital angular momentum is not conserved (otherwise the commutator would be zero).

We would like to find a *total* angular momentum $\vec{J}$ that *is* conserved, by adding an additional operator $\vec{S}$ to $\vec{L}$,

$$\vec{J} = \vec{L} + \vec{S}, \quad \left[\vec{J}, H_D\right] = 0 \tag{3.50}$$

To this end, consider the three matrices,

$$\vec{\Sigma} \equiv \begin{pmatrix}\vec{\sigma} & 0 \\ 0 & \vec{\sigma}\end{pmatrix} = -i\alpha_1\alpha_2\alpha_3\vec{\alpha} \tag{3.51}$$

where the first equivalence is merely a definition of $\vec{\Sigma}$ and the last equality can be verified by an explicit calculation. The $\vec{\Sigma}/2$ have the correct commutation relations to represent angular momentum, since the Pauli matrices do, and their commutators with $\vec{\alpha}$ and β are,

$$[\vec{\Sigma}, \beta] = 0 \qquad [\Sigma_i, \alpha_j] = 2i\varepsilon_{ijk}\alpha_k \tag{3.52}$$

Here ε_{ijk} is a totally anti-symmetric tensor which is zero if any of the three indices are the same: ε_{123} is +1, and we get a minus sign if we interchange any two indices so $\varepsilon_{213} = -1$.

From the relations in (3.52) we find that

$$\left[\vec{\Sigma}, H_D\right] = -2i\vec{\alpha}\times\vec{p} \tag{3.53}$$

Exercise 3.5: It is instructive to explicitly work out the relations (3.51), (3.52) and (3.53) in the Dirac representation.

Comparing equation (3.53) with the commutator of $\vec{L}$ with H_D in equation (3.49), you see that

$$\left[\vec{L}+\frac{1}{2}\vec{\Sigma},H_D\right]=0 \tag{3.54}$$

and we can identify

$$\vec{S}=\frac{1}{2}\vec{\Sigma} \tag{3.55}$$

as the additional quantity that, when added to $\vec{L}$ in equation (3.50), yields a conserved total angular momentum $\vec{J}$. We interpret $\vec{S}$ as an angular momentum *intrinsic* to the particle. It is hopefully clear from the form of $\vec{S}$ that we are describing the spin of a spin 1/2 fermion. More formally

$$\vec{S}^2=\frac{1}{4}\begin{pmatrix}\vec{\sigma}\cdot\vec{\sigma} & 0\\ 0 & \vec{\sigma}\cdot\vec{\sigma}\end{pmatrix}=\frac{3}{4}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{3.56}$$

and, recalling that the eigenvalue of $\vec{J}^2$ for spin j is $j(j+1)$, we conclude that $\vec{S}$ represents spin-1/2 and the solutions of the Dirac equation have spin-1/2 as promised. We worked in the Dirac representation of the matrices for convenience, but the result is necessarily independent of the representation.

Now consider the u-spinor solutions $u(s, p)$ of equation (3.38). Choose $\overrightarrow{p}=(0,0,p_z)$and write

$$u_\uparrow\equiv u\left(1,p\right)=\begin{pmatrix}\sqrt{E+m}\\ 0\\ \sqrt{E-m}\\ 0\end{pmatrix}\qquad u_\downarrow\equiv u\left(2,p\right)=\begin{pmatrix}0\\ \sqrt{E+m}\\ 0\\ -\sqrt{E-m}\end{pmatrix} \tag{3.57}$$

With these definitions, we get

$$S_z u_\uparrow=\frac{1}{2}u_\uparrow,\qquad S_z u_\downarrow=-\frac{1}{2}u_\downarrow$$

So, these two spinors represent spin up and spin down along the *z*-axis respectively. For the *v*-spinors, with the same choice for $\vec{p}$, write,

$$v_{\downarrow} = v\,(1,p) = \begin{pmatrix} \sqrt{E-m} \\ 0 \\ \sqrt{E+m} \\ 0 \end{pmatrix} \qquad v_{\uparrow} = v\,(2,p) = \begin{pmatrix} 0 \\ -\sqrt{E-m} \\ 0 \\ \sqrt{E+m} \end{pmatrix} \tag{3.58}$$

where now,

$$S_z v_{\downarrow} = \frac{1}{2} v_{\downarrow} \qquad S_z v_{\uparrow} = -\frac{1}{2} v_{\uparrow}$$

This apparently perverse choice of up and down for the *v*'s is actually quite sensible when one realizes that a negative energy electron carrying spin +1/2 backwards in time looks just like a positive energy positron carrying spin −1/2 forwards in time.

3.3.4 Lorentz covariant notation

There is a more compact way of writing the Dirac equation, which requires that we get to grips with some more notation. Define the *γ-matrices*,

$$\gamma^0 = \beta \qquad \vec{\gamma} = \beta\vec{\alpha} \tag{3.59}$$

In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \tag{3.60}$$

In terms of these, the relations between the $\vec{\alpha}$ and β in equation (3.24) can be written compactly as, the Clifford algebra,

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu} \tag{3.61}$$

Exercise 3.6: Again it is a good idea to work out the relations (3.61) explicitly in the Dirac representation.

Combinations like $a_{\mu}\gamma^{\mu}$ occur frequently and are conventionally written as,

$$\not{a} = a_\mu \gamma^\mu = a^\mu \gamma_\mu, \tag{3.62}$$

pronounced ‘a slash’.

The Dirac equation we had previously multiplied through by β from the right was

$$\beta \times i\frac{\partial \psi}{\partial t} = \beta \times (-i\vec{\alpha}.\vec{\nabla}\psi) + \beta \times \beta m\psi \tag{3.63}$$

using $\beta^2 = 1$ and the relabelling the remaining matrices as γ^i gives

$$i\gamma^\mu \partial_\mu \psi - m\psi = 0 \tag{3.64}$$

Or more compactly yet

$$\left(i\not{\partial} - m\right)\psi = 0 \tag{3.65}$$

or, in momentum space,

$$\left(\not{p} - m\right)\psi = 0 \tag{3.66}$$

The spinors *u* and *v* satisfy

$$\left(\not{p} - m\right)u\left(s,p\right) = 0 \tag{3.67}$$

$$\left(\not{p} + m\right)v\left(s,p\right) = 0 \tag{3.68}$$

since for $v(s,p)$, $E \to -E$ and $\vec{p} \to -\vec{p}$.

We want the Dirac equation (3.65) to preserve its form under Lorentz transformations (3.2). We have just naively written the matrices in the Dirac equation as γ_μ, however, this does not make them a 4-vector! They are just a set of numbers in four matrices and there is no reason they should change when we do a boost. However, the notation is deliberately suggestive, for when combined with Dirac fields you can construct quantities that transform like vectors and other Lorentz tensors (we will

not show this here). Since ∂^μ does transform, for the equation to be Lorentz covariant we are led to propose that ψ transforms too. As an example of such a transformation let us look at Parity transformations.

Parity

Consider parity (space inversion) transformations, $\hat{P}$: $t, \vec{x} \to t, -\vec{x}$.

We would not expect physics to change because of such a redefinition of our axes labelling. For the Dirac equation to remain the same though we must also transform ψ, in the Dirac representation as

$$\psi \to \psi\prime = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \psi \tag{3.69}$$

To see that this works note that under parity

$$\not{\partial} = \begin{pmatrix} \frac{\partial}{\partial t} & \vec{\sigma} \,.\, \vec{\nabla} \\ -\vec{\sigma} \,.\, \vec{\nabla} & -\frac{\partial}{\partial t} \end{pmatrix} \to \begin{pmatrix} \frac{\partial}{\partial t} & -\vec{\sigma} \,.\, \vec{\nabla} \\ \vec{\sigma} \,.\, \vec{\nabla} & -\frac{\partial}{\partial t} \end{pmatrix} \tag{3.70}$$

So

$$\not{\partial} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \to \begin{pmatrix} \frac{\partial}{\partial t} & \vec{\sigma} \,.\, \vec{\nabla} \\ \vec{\sigma} \,.\, \vec{\nabla} & \frac{\partial}{\partial t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \not{\partial} \tag{3.71}$$

This means we can write the parity transformed Dirac equation

$$\left(i \not{\partial} - m \right) \psi\prime = 0 \tag{3.72}$$

as

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \left(i \not{\partial} - m \right) \psi = 0 \tag{3.73}$$

which has the same solutions as the Dirac equation before the transformation we require.

The upshot is that we have discovered that particles and anti-particles have opposite intrinsic parity as can be seen from (3.69).

3.3.5 Massless (ultra-relativistic) fermions

At very high energies we may neglect the masses of particles ($E^2 \simeq \mid \vec{p} \mid^2$). Let us look, therefore, at solutions of the Dirac equation with $m = 0$, on the basis that this will be an extremely good approximation for many situations.

From equation (3.31) we have in this case

$$E\phi = \vec{\sigma} \cdot \vec{p}\chi, \qquad E\chi = \vec{\sigma} \cdot \vec{p}\ \phi \tag{3.74}$$

These equations can easily be decoupled by taking linear combinations and defining the two component spinors N_L and N_R,

$$N_R \equiv \chi + \phi, \quad N_L \equiv \chi - \phi \tag{3.75}$$

which leads to

$$EN_R = \vec{\sigma} \cdot \vec{p} N_R \qquad EN_L = -\vec{\sigma} \cdot \vec{p} N_L \tag{3.76}$$

The system is in fact described by two entirely separated two component spinors. If we take them to be moving in the z-direction, and noting that $\sigma_3 = \mathrm{diag}(1, -1)$, we see that there is one positive and one negative energy solution in each.

Further, since $E = \mid \vec{p} \mid$ for massless particles, these equations may be written

$$\frac{\vec{\sigma} \cdot \vec{p}}{\mid \vec{p} \mid} N_L = -N_L, \qquad \frac{\vec{\sigma} \cdot \vec{p}}{\mid \vec{p} \mid} N_R = N_R \tag{3.77}$$

Now, $\frac{1}{2}\frac{\vec{\sigma}\cdot\vec{p}}{|\vec{p}|}$ is known as the *helicity* operator (i.e. it is the spin operator projected in the direction of motion of the momentum of the particle). We see that the N_L corresponds to solutions with negative helicity, while N_R corresponds to solutions with positive helicity. In other words N_L describes a left-handed particle while N_R describes a right-handed particle, and each type is described by a two-component spinor. For example, N_l can describe massless left-handed neutrinos.

Note that under parity transformations $\vec{\sigma} \to \vec{\sigma}$ (like $\vec{R} \times \vec{p}$), $\vec{p} \to -\vec{p}$, therefore $\vec{\sigma} \cdot \vec{p} \to -\vec{\sigma} \cdot \vec{p}$, i.e. the spinors transform into each other:

$$N_L \leftrightarrow N_R \tag{3.78}$$

So a theory in which N_L has different interactions to N_R (such as the standard model of particle physics in which the weak force only acts on left-handed particles) manifestly violates parity.

Although massless particles can be described very simply using two component spinors as above, they may also be incorporated into the four-component formalism as follows. We use a new matrix

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.79}$$

here written in the Dirac representation again. Now acting with $(1 + \gamma^5)/2$

$$\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \chi \\ \phi \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \chi + \phi \\ \chi + \phi \end{pmatrix} \tag{3.80}$$

leaving a four component spinor that has only the N_R degree of freedom with helicity 1/2 (right handed). Equally $(1 - \gamma^5)/2$ projects out N_L, the particle with helicity $-1/2$ (left handed):

$$\frac{(1+\gamma^5)}{2}\psi \equiv \psi_R, \ \frac{(1-\gamma^5)}{2}\psi \equiv \psi_L \tag{3.81}$$

define the four-component spinors ψ_R and ψ_L.

IOP Concise Physics

Theories of Matter, Space and Time, Volume 2

Quantum theories

N Evans and S F King

Chapter 4

Quantum electrodynamics

We have developed a quantum description of free relativistic particles in Chapter 3. To add interactions we need a theory of force. Here, we will include the electromagnetic interactions of particles within the context of the Dirac equation, including developing a wave equation for photons. This will enable us to consider real collider process such as electron positron annihilation.

4.1 Photon wave equation

To see how to make a relativistic wave equation that describes photons let us begin back at Maxwell's equations in differential form (a more in-depth analysis of classical relativistic electrodynamics is provided in the prequel *Theories of Matter, Space and Time: Classical Theories*)

$$\begin{aligned} \vec{\nabla}\cdot\vec{E} &= \rho & \vec{\nabla}\cdot\vec{B} &= 0 \\ \vec{\nabla}\times\vec{E} &= -\tfrac{\partial\vec{B}}{\partial t} & \vec{\nabla}\times\vec{B} &= \vec{J} + \tfrac{\partial\vec{E}}{\partial t} \end{aligned} \tag{4.1}$$

We have used units here where $\mu_0 = \varepsilon_0 = 1$ (this can be achieved by picking the appropriate units for $\vec{E}, \vec{B}$ and charge). We can solve the Maxwell equations with the following potentials

$$\begin{aligned} \vec{E} &= -\tfrac{\partial\vec{A}}{\partial t} - \vec{\nabla}\phi \\ \vec{B} &= \vec{\nabla}\times\vec{A} \end{aligned} \tag{4.2}$$

which are automatically solutions of the Maxwell equations

$$\vec{\nabla}\cdot\vec{B} = \vec{\nabla}\cdot(\vec{\nabla}\times\vec{A}) \equiv 0 \tag{4.3}$$

and also

$$\begin{aligned} \vec{\nabla}\times\vec{E} &= \vec{\nabla}\times\left(-\tfrac{\partial\vec{A}}{\partial t} - \vec{\nabla}\phi\right) \\ &= -\tfrac{\partial(\vec{\nabla}\times\vec{A})}{\partial t} - \vec{\nabla}\times(\vec{\nabla}\phi) \\ &= -\tfrac{\partial\vec{B}}{\partial t} - 0 \end{aligned} \tag{4.4}$$

This simplifies things greatly since now there are only the remaining two Maxwell equations to solve. Let us write them out in terms of the potentials

$$\vec{\nabla}\cdot\vec{E} = -\nabla^2\phi - \frac{\partial\left(\vec{\nabla}\cdot\vec{A}\right)}{\partial t} = \rho \tag{4.5}$$

and (since $\vec{\nabla}\times\vec{\nabla}\times\vec{A} \equiv -\nabla^2\vec{A} + \vec{\nabla}\cdot(\vec{\nabla}\cdot\vec{A})$)

$$\vec{\nabla}(\vec{\nabla}\cdot\vec{A}) - \nabla^2\vec{A} = \vec{J} + \frac{\partial}{\partial t}\left(-\frac{\partial\vec{A}}{\partial t} - \vec{\nabla}\phi\right) \tag{4.6}$$

or rearranging

$$-\nabla^2\vec{A} + \frac{\partial^2\vec{A}}{\partial t^2} = \vec{J} - \vec{\nabla}\left(\vec{\nabla}\cdot\vec{A} + \frac{\partial\phi}{\partial t}\right) \tag{4.7}$$

Unfortunately, the two equations we are left with, (4.5) and (4.7), are quite messy! To clean them up we can make use of our ability to redefine the potentials whilst keeping the $\vec{E},\vec{B}$ fields the same.

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

$$\begin{aligned}\vec{A} &\to \vec{A} + \vec{\nabla}\alpha \\ \phi &\to \phi - \tfrac{\partial\alpha}{\partial t}\end{aligned} \tag{4.8}$$

where α is an arbitrary scalar function of position and time.
Let us choose to make a gauge transformation such that

$$\vec{\nabla}\cdot\vec{A} = -\frac{\partial\phi}{\partial t} \tag{4.9}$$

In this gauge (Lorenz gauge) Maxwell's equations simplify to

$$-\nabla^2\phi + \frac{\partial^2\phi}{\partial t^2} = \rho \tag{4.10}$$

$$-\nabla^2\vec{A} + \frac{\partial^2\vec{A}}{\partial t^2} = \vec{J} \tag{4.11}$$

This form of our remaining Maxwell's equations is much prettier! They also have a very suggestive form for relativity. They suggest we should define

$$J^\mu = (\rho,\vec{J}), \quad A^\mu = (\phi,\vec{A}) \tag{4.12}$$

so the Maxwell equations can be written as ($\Box = \partial^\mu\partial_\mu$)

$$\Box A^\mu = J^\mu$$

The $\mu = 0$ equation is the ϕ equation (4.10) and the $\mu = 1, 2, 3$ equations give the components of the equation (4.11) for $\vec{A}$.

The Maxwell equations in Lorenz gauge also required the gauge condition (4.9) which becomes

$$\partial^\mu A_\mu = 0 \tag{4.14}$$

To move to a quantum theory we will now treat A^μ as a wave function for photons. In the limit of a large number of photons the wave function can be interpreted as number density. For an observer who is not counting individual photons but just the energy density they provide, A^μ will then look like the classical wave theory. In free space we have

$$\Box A^\mu = 0 \tag{4.15}$$

with solutions

$$A^\mu = \varepsilon^\mu e^{-iq.x} \tag{4.16}$$

where ε^μ is the polarization tensor and $q^2 = 0$ as required for a photon.

The Lorenz condition enforces

$$q^\mu \varepsilon_\mu = 0 \tag{4.17}$$

and we can choose to set the component of ε^μ in the direction of motion to zero.

Further, within Lorenz gauge there are still gauge transformations

$$A^\mu \to A^\mu + \partial^\mu \chi \quad \text{where} \quad \Box\ \chi = 0 \tag{4.18}$$

This can be used to remove one extra degree of freedom from ε^μ for example in Coulomb gauge setting

$$A^0 = 0 \tag{4.19}$$

A^μ only has two degrees of freedom which describe the probability distribution of the two polarizations of the photon. The photon's two polarizations are transverse to the direction of motion.

4.2 Minimal substitution

We now want to return to thinking about coupling the photons to our Dirac field electrons

$$\left(i \not{\partial} - m\right)\psi = 0 \tag{4.20}$$

The obvious thing to do is to just be led by Lorentz invariance

$$\partial^\mu \to \partial^\mu + iqA^\mu \quad \equiv D^\mu \tag{4.21}$$

where the factor of iq is a free constant which as our notation suggests will enter as the electric coupling. This is called *minimal substitution* and in fact matches what nature does. We write

$$\left(i\not{D} - m\right)\psi = 0 \tag{4.22}$$

We must also include the fermion current in the Maxwell equations which we know are

$$\Box A^\mu = J^\mu \tag{4.23}$$

We have seen that the probability current for Dirac equation solutions is given by

$$\begin{aligned}\rho &= \psi^\dagger\psi = \psi^\dagger\beta^2\psi = \bar{\psi}\gamma^0\psi \\ \vec{J} &= \psi^\dagger\vec{\alpha}\psi = \psi^\dagger\beta^2\vec{\alpha}\psi = \bar{\psi}\vec{\gamma}\psi\end{aligned} \tag{4.24}$$

Note here we use the standard notation

$$\bar{\psi} = \psi^\dagger\gamma^0 \tag{4.25}$$

When there are many fermions present (4.24) becomes the number density current for those particles and so clearly the charge density current should be

$$J^\mu = q\bar{\psi}\gamma^\mu\psi \tag{4.26}$$

4.3 Gauge invariance

Minimal substitution in fact works but looks a little ad hoc. It hides a much more fundamental and beautiful symmetry.

Remember that Maxwell's equations are invariant to gauge transformations

$$A^\mu \to A^\mu - \partial^\mu\alpha\left(x\right) \tag{4.27}$$

However, the modified Dirac equation (4.22) we have written above with A^μ in it is not invariant to such a transformation. There is, though, a bigger symmetry which all the equations respect that incorporates the gauge invariance. That larger symmetry is

$$\begin{aligned}\psi &\to e^{iq\alpha(x)}\psi \\ A^\mu &\to A^\mu - \partial^\mu\alpha\left(x\right)\end{aligned} \tag{4.28}$$

Proof: We begin with the Dirac equation

$$[i\gamma_\mu(\partial^\mu + iqA^\mu) - m]\psi = 0 \tag{4.29}$$

When we make the transformations we arrive at

$$[i\gamma_\mu(\partial^\mu + iqA^\mu - iq(\partial^\mu\alpha)) - m]e^{iq\alpha(x)}\psi \tag{4.30}$$

Now for it to be a symmetry we require the solutions of the first equation (4.29) to also be solutions of the second equation (4.30).

The way to show this is to try to move the exp($iq\alpha(x)$) term to the far left. The only term we cannot commute it past is the derivative which will act on $\alpha(x)$. In particular

$$\partial^\mu e^{iq\alpha}\psi = e^{iq\alpha}\big(\partial^\mu\psi\big) + e^{iq\alpha} iq\big(\partial^\mu\alpha\big)\psi \tag{4.31}$$

However, if you look at the term induced in the Dirac equation by the shift in A^μ you will see it precisely cancels this extra term with $\partial^\mu\alpha$. Thus we arrive at

$$e^{iq\alpha}[i\gamma_\mu\left(\partial^\mu + iqA^\mu\right) - m]\psi = 0 \tag{4.32}$$

which clearly has the same solutions as the Dirac equation we started with.

The Maxwell equations we already know are invariant to gauge transformations but we must check that $J^\mu = q\bar{\psi}\gamma^\mu\psi$ which we added is too. The exponentials cancel between $\bar{\psi}$ and ψ and all is well.

The beauty herein: We can look at the gauge transformations from the point of view of the Dirac equation. The free Dirac equation has a symmetry where we shift the solution ψ by a phase $\psi \to e^{i\alpha}\psi$ but where α does not depend on x

$$\Big(i\not{\partial} - m\Big)e^{i\alpha}\psi = e^{i\alpha}\Big(i\not{\partial} - m\Big)\psi = 0 \tag{4.33}$$

This is called a *global* transformation. It is telling us that we are free to place our coordinate axes where we like in the complex plane for ψ.

Now in a relativistic theory you might wonder whether two areas of space that are not causally connected should be forced to have the same choice of coordinate axes. You might choose to impose that α can have dependence on spacetime position x^μ. If you tried to impose this you would find it is not a symmetry of the Dirac equation unless you introduced a field A^μ with the specific transformations we observed nature to have above. In other words you would have had to invent electromagnetism in order to have this symmetry. This is apparently what nature does.

Note that in current thinking we view the symmetry as the fundamental guiding theoretical concept of the theory and consider the existence of A^μ to be derived.

Massless photon: We get one more fact for free too. The Klein–Gordon equation for the photon is

$$\Box A^\mu = 0 \qquad \text{not} \qquad \left(\Box + m^2\right)A^\mu = 0 \tag{4.34}$$

The second term would not be gauge invariant so we must set $m^2 = 0$. The symmetry correctly predicts that the photon is massless!

4.4 QED interactions in perturbation theory

The main technique for computations of particle scatterings is perturbation theory—in other words we assume that the coupling $q \ll 1$. We will be interested in processes such as that shown in figure 4.1. Since we will concentrate on scatterings involving electrons and muons we will set $q = -e$ henceforth. Outside the shaded interaction region we assume the particles are free.

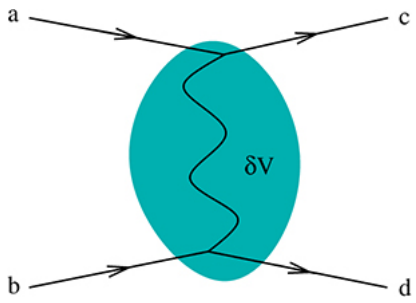


Figure 4.1. The scattering of two particles off each other. We consider the interaction to be non-zero only in the shaded region. Outside that region the solutions are of the free Dirac equation.

Let us write the Dirac equation in a way that displays the smallness of the interaction

$$i\gamma^0 \frac{\partial \psi}{\partial t} + i\gamma^i \partial_i \psi - m\psi - \gamma^0 \, \delta V \, \psi = 0 \tag{4.35}$$

so for the electromagnetic interaction

$$\delta V = -e\gamma^0 \gamma^\mu A_\mu \tag{4.36}$$

Note that $(\gamma^0)^2 = 1$ so the γ^0 have been included simply for notational convenience.

We will assume that the scattering particles begin in a pure $\vec{p}$ state but the interaction then scatters them to another $\vec{p}$ state with some (small) probability. In general we can write

$$\psi = \sum_n \kappa_n \, \phi_n \, (x) e^{-iE_n t} \tag{4.37}$$

The $\phi_n(x)$ are the free Dirac equation solutions with n labelling the spinor state and the $\vec{p}$ state. The κ_n are the probability amplitudes for the given state n. Before the interaction all the κ_n will be zero except one but during the interaction ($-T/2 < t < T/2$) we allow κ_n to change— $\kappa_n(t)$.

If we now substitute the solution into the perturbed Dirac equation above then, at leading order, we obtain zero since we have expanded in solutions of the unperturbed equation. At next order we find

$$i\gamma_0 \sum_n \left(\frac{d\kappa_n}{dt} \right) \phi_n e^{-iE_n t} = \sum_n \gamma_0 \, \delta V \, \kappa_n \, \phi_n \, (x) e^{-iE_n t} \tag{4.38}$$

Now we will make use of the orthogonality of the ϕ_n to extract the final state κ_n. We multiply through by $\int d^3x \, \phi_f^\dagger \gamma_0$

$$\frac{d\kappa_f}{dt} = -i \sum_n \kappa_n \int d^3x \; \phi_f^\dagger \delta V \phi_n e^{-i(E_n - E_f)t} \tag{4.39}$$

For a discussion of normalization of the spinors see section 3.3.2 (we are using $N = 1/\sqrt{2EV}$).

Remembering that at $t = -T/2$ $\kappa_i = 1$ and $\kappa_{i \neq n} = 0$ at leading order we have

$$\frac{d\kappa_f}{dt} = -i \int \psi_f^\dagger \delta V \psi_i \; d^3x \tag{4.40}$$

and integrating with respect to t we find the important result

$$\kappa_f(T/2) = -i \int \psi_f^\dagger \delta V \psi_i \; d^4x \tag{4.41}$$

Note that this equation is a rewritten form of (2.70).

Now let us use our explicit form for δV in QED and concentrate on the scattering of a particle $a \to c$ by a photon A^μ (see figure 4.2).

$$\begin{aligned} \kappa_{ca} &= -i \int \bar{\psi}^c (-e \gamma_\mu A^\mu) \psi^a \; d^4x \\ &= -i \int J_\mu^{ca} A^\mu \; d^4x \end{aligned} \tag{4.42}$$

where

$$J_\mu^{ca} = -e \; \bar{\psi}^c \gamma_\mu \psi^a = e \; N_a N_c \; \bar{u}^c \gamma_\mu u^a \; e^{i(p_c - p_a).x} \tag{4.43}$$

The N's here are the normalizations of the spatial wave functions ψ again from section 3.3.2.

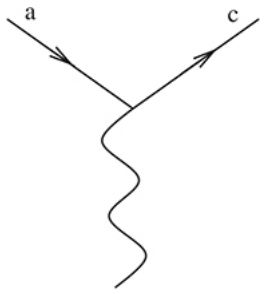


Figure 4.2. The scattering of a particle $a \to c$ by a photon A^μ.

We are really interested in two particles scattering off each other so we would better compute the A^μ field produced when another particle scatters from state $b \to d$ (see figure 4.3)

$$\Box A^\mu = J_{db}^\mu = -e \; N_b N_d \; \bar{u}_d \gamma_\mu u_b \; e^{i(p_d - p_b).x}$$

the solution is (4.44)

$$A^{\mu} = -\frac{1}{q^2} J^{\mu}_{db}, \qquad q = p_d - p_b$$

(4.45)

So finally substituting this back into our expression for κ_{ca} we find

$$\kappa_{fi} = -i\, N_a N_b N_c N_d\, \bar{u}^c\,(-e\gamma_{\mu})u^a \left(-\tfrac{1}{q^2}\right) \bar{u}^d\,(-e\gamma^{\mu})u^b \int e^{i(p_c+p_d-p_a-p_b).x} d^4x$$

(4.46)

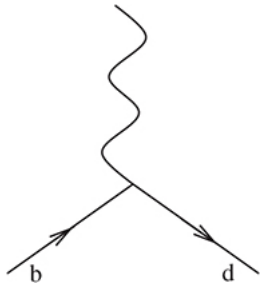


Figure 4.3. The scattering of a particle $b \to d$ that generates the photon that scatters $a \to c$.

Note that the integral is just a delta function that ensures 4-momentum conservation in the interaction.

In order to make this result more memorable, Feynman developed his famous rules that associate different parts of the expression with elements of a diagram of the scattering. For example, consider the Feynman diagram for this scattering shown in figure 4.4 with the appropriate rules shown and where implicitly momentum is conserved at the vertices.

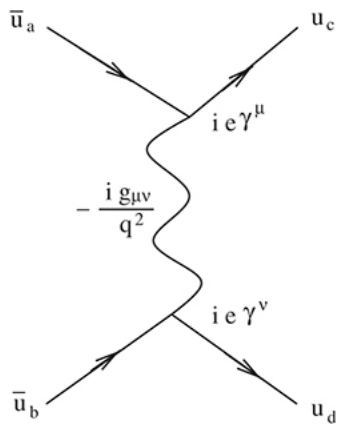


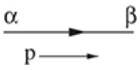
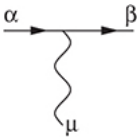
Figure 4.4. The Feynman diagram for a two-two particle scattering showing the appropriate Feynman rules.

Multiplying out the rules shown in figure 4.4 gives us $-i\mathrm{M}_{fi}$ where

$$\kappa_{fi} = -i\, N_a N_b N_c N_d (2\pi)^4 \delta^4 \left(p_f - p_i\right) \mathrm{M}_{fi} \tag{4.47}$$

4.4.1 Summary of Feynman rules of QED

The Feynman rules for computing the amplitude M_{fi} for an arbitrary process in QED are summarized here. They include the rules for internal fermion lines and external photons though we will not derive them directly.

For every …	draw …	write …
Internal photon line		$\frac{-ig^{\mu\nu}}{p^2+i0^+}$
Internal fermion line		$\frac{i(\not{p}+m)_{\alpha\beta}}{p^2-m^2+i0^+}$
Vertex		$-ie\gamma^{\mu}_{\alpha\beta}$
Outgoing electron		$\bar{u}_\alpha(s,p)$
Incoming electron		$u_\alpha(s, p)$
Outgoing positron		$v_\alpha(s, p)$
Incoming positron		$\bar{v}_\alpha(s,p)$
Outgoing photon		$\varepsilon^{*\mu}(\lambda, p)$
Incoming photon		$\varepsilon^{\mu}(\lambda, p)$
• Attach a directed momentum to every internal line		
• Conserve momentum at every vertex, i.e. include $\delta^{(4)}(\sum p_i)$		

Feynman rules for QED. μ, ν are Lorentz indices, α, β are spinor indices and s and λ fix the polarization of the electron and photon respectively.

4.4.2 Electron–muon scattering

So far we have been a little vague about the precise scattering we are studying. Electron–muon scattering is the simplest process at lowest order in the electromagnetic coupling where just the one diagram we have considered so far contributes. It is again shown in figure 4.5. The amplitude obtained by applying the Feynman rules to this diagram is

$$\mathrm{M}_{fi} = (-ie)\,\bar{u}\,(p_c)\gamma^{\mu}u\,(p_a)\left(\frac{-ig_{\mu\nu}}{q^2}\right)(-ie)\,\bar{u}\,(p_d)\gamma^{\nu}u\,(p_b), \tag{4.48}$$

where $q^2 = (p_a - p_c)^2$. Note that, for clarity, we have dropped the spin label on the spinors. We will restore them when we need to. In constructing this amplitude we have followed the fermion lines backwards with respect to fermion flow when working out the order of matrix multiplication (which makes sense if you think of an unbarred spinor as a column vector and a barred spinor as a row vector and remember that the amplitude carries no spinor indices).

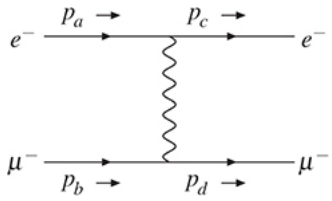


Figure 4.5. Lowest order Feynman diagram for $e^-\,\mu^- \to e^-\,\mu^-$scattering.

The probability for the scattering involves the squared modulus of the amplitude, $\mid \mathrm{M} \mid^2$. Let us see how we obtain a neat form for this—this is a considerable amount of work. Consider

$$(\bar{u}(p_c)\gamma^{\mu}u(p_a))^{*} = (\bar{u}(p_c)\gamma^{\mu}u(p_a))^{\dagger}$$

We could transpose it for free since the whole quantity is just a number. Using rules of matrix algebra we see that this is

$$(u(p_c)^{\dagger}\gamma^0\gamma^{\mu}u(p_a))^{\dagger} = (u(p_a)^{\dagger}\gamma^{\mu\dagger}\gamma^{0\dagger}u(p_c)) \tag{4.49}$$

$$= (u(p_a)^{\dagger}\gamma^{\mu\dagger}\gamma^0 u(p_c)) \tag{4.50}$$

Now it is the case that $\gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^{\mu}$, and so this becomes

$$(u(p_a)^{\dagger}\gamma^0\gamma^{\mu}u(p_c)) = (\bar{u}(p_a)\gamma^{\mu}u(p_c))$$

Using this general result in the expression for $\mid \mathrm{M} \mid^2$ we obtain

$$\begin{aligned} \mid \mathrm{M} \mid^2 &= \tfrac{e^4}{q^4}\,\bar{u}\,(p_c)\gamma^{\mu}u\,(p_a)\bar{u}\,(p_d)\gamma_{\mu}u\,(p_b)\bar{u}\,(p_a)\gamma^{\nu}u\,(p_c)\bar{u}\,(p_b)\gamma_{\nu}u\,(p_d) \\ &= \tfrac{e^4}{q^4}L^{\mu\nu}_{(e)}L_{(\mu)\,\mu\nu} \end{aligned} \tag{4.51}$$

where the subscripts e and μ refer to the electron and muon respectively and

$$L^{\mu\nu}_{(e)} = \bar{u}\,(p_c)\gamma^{\mu}u\,(p_a)\bar{u}\,(p_a)\gamma^{\nu}u\,(p_c)$$

with a similar expression for $L^{\mu\nu}{}_{(\mu)}$.

Exercise 4.1: check explicitly the relations $\gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu$.

Usually a collider experiment has an unpolarized beam and target and they do not measure the polarization of the outgoing particles. Thus we should calculate the squared amplitudes for each possible spin combination, then average over initial spin states and sum over final spin states. Note that we square and then sum since the different spin configurations are in principle distinguishable. In contrast, if several Feynman diagrams contribute to the same process, you have to sum the amplitudes first. We will see examples of this below.

The spin sums are made easy by the following results

$$\begin{aligned} \textstyle\sum_s u\,(s,p)\,\bar{u}\,(s,p) &= \not{p} + m \\ \textstyle\sum_s v\,(s,p)\,\bar{v}\,(s,p) &= \not{p} - m \end{aligned} \tag{4.52}$$

Where, do not forget, we really mean m times the unit 4 × 4 matrix when we write just m.

Exercise 4.2: Prove (4.52) using (3.38) and (3.39).

Using the spin sums we find that

$$\begin{aligned} \tfrac{1}{4}\textstyle\sum_{\text{spins}} |\,\mathrm{M}\,|^2 &= \tfrac{e^4}{4q^4}\left[\gamma^\mu_{\alpha\beta}\left(\not{p}_a + m_e\right)_{\beta\zeta}\gamma^\nu_{\zeta\eta}\left(\not{p}_c + m_e\right)_{\eta\alpha}\right] \\ &\quad\times\left[\gamma_{\mu,\alpha\prime\beta\prime}\left(\not{p}_b + m_\mu\right)_{\beta\prime\zeta\prime}\gamma_{\nu,\zeta\prime\eta\prime}\left(\not{p}_d + m_\mu\right)_{\eta\prime\alpha\prime}\right] \\ &= \tfrac{e^4}{4q^4}\,\mathrm{Tr}\left(\gamma^\mu(\not{p}_a + m_e)\gamma^\nu(\not{p}_c + m_e)\right)\mathrm{Tr}\left(\gamma_\mu(\not{p}_b + m_\mu)\gamma_\nu(\not{p}_d + m \end{aligned} \tag{4.53}$$

where in the first expression, we have made explicit the spinor indices in order that you can see how the trace that appears in the second expression emerges. All calculations of probabilities in QED require the evaluation of traces of products of γ-matrices. Useful theorems can be derived from the fundamental anti-commutation relations of the γ-matrices in equation (3.61) together with the invariance of the trace under a cyclic change of its arguments. For now it suffices to use

$$\begin{aligned} \mathrm{Tr}(\gamma^{\mu_1}\,\ldots\,\gamma^{\mu_n}) &= 0 \quad \text{for } n \text{ odd} \\ \mathrm{Tr}(\gamma^{\mu_1}\,\ldots\,\gamma^{\mu_n}) &= g^{\mu_1\mu_2}\,\mathrm{Tr}(\gamma^{\mu_3}\,\ldots\,\gamma^{\mu_n}) - g^{\mu_1\mu_3}\,\mathrm{Tr}(\gamma^{\mu_2}\gamma^{\mu_4}\,\ldots\,\gamma^{\mu_n}) \\ &\quad + \cdots + g^{\mu_1\mu_n}\,\mathrm{Tr}(\gamma^{\mu_2}\,\ldots\,\gamma^{\mu_{n-1}}) \\ \mathrm{Tr}\left(\not{a}\,\not{b}\right) &= 4\,a.b \\ \mathrm{Tr}\left(\not{a}\,\not{b}\,\not{c}\,\not{d}\right) &= 4(a.b\,c.d - a.c\,b.d + a.d\,b.c) \end{aligned} \tag{4.54}$$

Exercise 4.3: Derive the trace results in equation (4.54). (Hint: for the first one use $(\gamma^5)^2 = 1$.)

Using these results,

$$\begin{aligned}\tfrac{1}{4}\sum_{\text{spins}} \mid \mathrm{M} \mid^2 &= \tfrac{e^4}{4q^4}\Big(\mathrm{Tr}\left(\gamma^\mu(\not{p}_a)\gamma^\nu(\not{p}_c)\right) + \mathrm{Tr}\left(\gamma^\mu\gamma^\nu\right)m_e^2\Big) \\ &\quad \times \Big(\mathrm{Tr}\left(\gamma_\mu(\not{p}_b)\gamma_\nu(\not{p}_d)\right) + \mathrm{Tr}\left(\gamma_\mu\gamma_\nu\right)m_\mu^2\Big) \\ &= \tfrac{e^4}{4q^4}\big(4\big(p_a^\mu p_c^\nu - g^{\mu\nu}p_a\cdot p_c + p_a^\nu p_c^\mu\big) + 4g^{\mu\nu}m_e^2\big) \times (\cdots) \\ &= \tfrac{e^4}{4q^4}\big(4\big(p_a^\mu p_c^\nu + p_a^\nu p_c^\mu + g^{\mu\nu}\big(-p_a\cdot p_c + m_e^2\big)\big)\big) \times (\cdots)\end{aligned}$$

Mandelstam variables: the following kinematic variable are useful and frequently used

$$s = (p_a + p_b)^2 \tag{4.55}$$

$$t = (p_a - p_c)^2 \tag{4.56}$$

$$u = (p_a - p_d)^2 \tag{4.57}$$

Note

$$\begin{aligned}s + t + u &= 3p_a^2 + p_b^2 + p_c^2 + p_d^2 + 2p_a.p_d - 2p_a.p_c - 2p_a.p_d \\ &= m_a^2 + m_b^2 + m_c^2 + m_d^2 + 2p_a^2 + 2p_a.(p_b - p_c - p_d) \\ &= m_a^2 + m_b^2 + m_c^2 + m_d^2\end{aligned} \tag{4.58}$$

Now $t = (p_a - p_c)^2 = -2p_a\cdot p_c + 2m_e^2$ or $t = (p_b - p_d)^2 = -2p_b\cdot p_d + 2m_\mu^2$. So

$$\begin{aligned}\tfrac{1}{4}\sum_{\text{spins}} \mid \mathrm{M} \mid^2 &= \tfrac{4e^4}{t^2}\big(p_a^\mu p_c^\nu + p_a^\nu p_c^\mu + t/2g^{\mu\nu}\big) \\ &\quad \times (p_{b\mu}p_{d\nu} + p_{b\nu}p_{d\mu} + t/2g_{\mu\nu}) \\ &= \tfrac{4e^4}{t^2}(2(p_a\cdot p_b)(p_c\cdot p_d) + 2(p_a\cdot p_d)(p_b\cdot p_c) \\ &\quad + t(p_a\cdot p_c) + t(p_b\cdot p_d) + t^2)\end{aligned}$$

Expressing the answer in terms of the Mandelstam variables, we find

$$\begin{aligned}\tfrac{1}{4}\sum_{\text{spins}} \mid \mathrm{M} \mid^2 &= \tfrac{2e^4}{t^2}\Big[2\Big(\tfrac{s}{2} - \tfrac{\left(m_e^2+m_\mu^2\right)}{2}\Big)^2 + 2\Big(\tfrac{u}{2} - \tfrac{\left(m_e^2+m_\mu^2\right)}{2}\Big)^2 \\ &\quad + t\big(-\tfrac{t}{2} + m_e^2\big) + t\big(-\tfrac{t}{2} + m_\mu^2\big) + t^2\Big] \\ &= \tfrac{2e^4}{t^2}\Big[\tfrac{s^2}{2} - s\left(m_e^2 + m_\mu^2\right) + \tfrac{u^2}{2} - u\left(m_e^2 + m_\mu^2\right) \\ &\quad + \left(m_e^2 + m_\mu^2\right)^2 + t\Big(m_e^2 + m_\mu^2\Big)\Big] \\ &= \tfrac{2e^4}{t^2}\Big(s^2 + u^2 - 4\Big(m_e^2 + m_\mu^2\Big)(s+u) + 6\big(m_e^2 + m_\mu^2\big)^2\Big)\end{aligned} \tag{4.59}$$

In the ultra-relativistic limit where we can neglect the particle masses this is just

$$\frac{1}{4}\sum_{\text{spins}} |\,\mathrm{M}\,|^2 = \frac{2e^4}{t^2}\left(s^2+u^2\right). \tag{4.60}$$

4.5 Cross sections and decay rates

In practice one does not scatter a single particle off another single particle since this is hard to arrange and since most of the time nothing would happen. Instead bunches of particles are collided with each other or a static target or one observes the decay of particles in a sample. Let us introduce the ideas of cross-section and decay rate before we compute them in detail.

Cross-sections

A typical experiment can be schematically represented as in figure 4.6. We have

$$\begin{array}{ccccc} \#\,\text{scatters}\,/\text{s} = & \text{flux} & \times & \#\,\text{target pts} \times \text{cross-section}, \sigma \\ [T]^{-1} & \left[L^2T\right]^{-1} & & [L]^2 \end{array} \tag{4.61}$$

where we have displayed the dimensions of the terms. Algebraically we could write

$$n_s = n_b v\, n_t\, \sigma \tag{4.62}$$

Note that in a classical scattering of say, balls, there is a scattering with probability one if one ball hits any of the area of another. In this case the cross section is precisely the area the ball presents to the scatterer—hence the name.

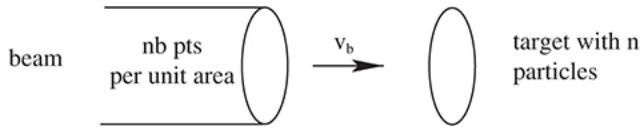


Figure 4.6. A beam on target scattering experiment.

Experimenters measure

$$\sigma = \frac{\#\ \text{scatters per sec}}{\text{flux} \times \#\ \text{target pts}} \tag{4.63}$$

The number of scatters/s depends on our probability $|\,\mathrm{M}\,|^2$ summed over all possible final states. Frequently the number of scatters into some solid angle is measured so people quote the differential cross-section $\frac{d\sigma}{d\Omega}$. We calculate these quantities in detail shortly.

Decay rates

When observing particle decays

$$A \to 1 + 2 + \cdots$$

one measures the number of decays per second per number of A in the sample. This is again just the probability $\mid \mathrm{M} \mid^2$ summed over all possible final states. So we measure (4.64)

$$\Gamma = -\frac{dN_A}{dt}/N_A \tag{4.65}$$

Integrating gives

$$N_A(t) = N_A(0)e^{-\Gamma t} \tag{4.66}$$

So Γ^{-1} is the life-time of the particle since $1/e$ of the particles decay in that time.

Γ is also referred to as the *width* of the decaying particle. To see why, imagine searching for the particle A through

$$e^+e^- \rightarrow A \quad \rightarrow \quad \text{decay products} \tag{4.67}$$

You might expect to find a cross section against energy that is zero until you have enough energy to create A, then a sharp edge at $E = 2m_A$ as shown on the left in figure 4.7. In fact rather than this edge one finds a peak as on the right in figure 4.7.

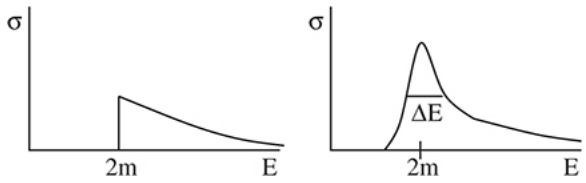


Figure 4.7. The cross-section to pair create particles near the mass threshold of $E = 2m$. On the left is a naive guess. On the right is the structure seen.

Close to $E = 2m_A$ the cross section is much higher than you might naively expect due to a resonance effect. The width of the resonance peak is determined from the uncertainty principle using Γ^{-1} as the uncertainty in time

$$\Delta E \Delta t \sim \hbar \tag{4.68}$$

The width of a resonance therefore counts the number of decay channels a particle has!
Let us now turn to computing these quantities in detail.

4.5.1 Transition rate

Consider an arbitrary scattering process with an initial state i with total 4-momentum P_i and a final state f with total 4-momentum P_f. Let us assume we computed the scattering amplitude for this process, i.e. we know the matrix element

$$-i\prod_{f=1}^{N} N_f \prod_{\text{in}} N_i \, \mathrm{M}(2\pi)^4\delta^4(P_f - P_i) \tag{4.69}$$

Our task is to convert this into a scattering cross section (relevant if there is more than 1 particle in the initial state) or a decay rate (relevant if there is just 1 particle in the initial state), see figure 4.8.

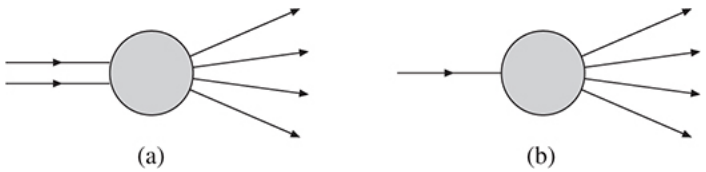


Figure 4.8. Scattering (a) and decay (b) processes.

The probability for the transition to occur is the square of the matrix element, i.e.

$$\text{probability} = \left| -i\prod_{f=1}^{N} N_f \prod_{\text{in}} N_i \, \mathrm{M}(2\pi)^4\delta^4(P_f - P_i) \right|^2 . \tag{4.70}$$

Attempting to take the squared modulus of the amplitude produces a meaningless square of a delta function. This is a technical problem because our amplitude is expressed between plane wave states. These states are states of definite momentum and so extend throughout all of space-time. In a real experiment the incoming and outgoing states are localized (e.g. they might leave tracks in a detector). To deal with this properly we would have to construct normalized wave packet states which do become well separated in the far past and the far future. Instead of doing this we will do a much simpler but rather sloppy derivation. First of all, we will put our system in a box of volume $V = L^3$. We also imagine that the interaction is restricted to act only over a time of order T. The final answers come out independent of V and T, reproducing the ones we would get if we worked with localized wave packets.

Using

$$(2\pi)^4\delta^4(P_f - P_i) = \int e^{i(P_f - P_i)x} d^4x \tag{4.71}$$

we get in our space-time box the result

$$\left| \, (2\pi)^4\delta^4(P_f - P_i) \, \right|^2 \simeq (2\pi)^4\delta^4(P_f - P_i) \int e^{i(P_f - P_i)x} d^4x \simeq VT \, (2\pi)^4\delta^4(P_f - P_i) \tag{4.72}$$

Here we have used the first δ-function to set $p_f = p_i$ in the integral.

We must also use the explicit expressions for the wave function normalizations from section 3.3.2. Above we used the normalization $N = 1/\sqrt{2EV}$. So putting everything together, we find for the transition rate W, i.e. the probability per unit time

$$W = \frac{1}{T} |\ \mathrm{M}\ |^2 V T (2\pi)^4 \delta^4(P_f - P_i) \prod_{f=1}^{N} \left[\frac{1}{2E_f V}\right] \prod_{\text{in}} \left[\frac{1}{2E_i V}\right] \tag{4.73}$$

As expected, the dependence on T cancelled. Usually we are interested in much more detailed information than just the total transition rate. We want to know the differential transition rate dW, i.e. the transition rate into a particular element of the final state phase space. To get dW we have to multiply by the number of available states in the (small) part of phase space under consideration.

For a single particle final state, the number of available states dn in some momentum range $\vec{k}$ to $\vec{k} + d\vec{k}$ is, in the box normalization,

$$dn \ = \ \frac{V}{(2\pi)^3} d^3\vec{k} \tag{4.74}$$

This result is proved by recalling that the allowed momenta in the box have components that can only take on discrete values since $n_x \lambda = L$ so that $k_x = 2\pi\ n_x/L$ where n_x is an integer. Thus $dn = dn_x dn_y dn_z$ and the result follows. For a two particle final state we have

$$dn \ = \ dn_1 dn_2$$

where

$$dn_1 = \frac{V}{(2\pi)^3} d^3\vec{k}_1,\ dn_2 = \frac{V}{(2\pi)^3} d^3\vec{k}_2$$

where dn is the number of final states in some momentum range $\vec{k}_1$ to $\vec{k}_1 + d\vec{k}_1$ for particle 1 and $\vec{k}_2$ to $\vec{k}_2 + d\vec{k}_2$ for particle 2. There is an obvious generalization to an N particle final state,

$$dn \ = \ \prod_{f=1}^{N} \frac{V d^3\vec{k}_f}{(2\pi)^3} \tag{4.75}$$

The transition rate for transitions into a particular element of final state phase space is thus given by, using equations (4.75) and (4.73),

$$\begin{aligned} dW &= |\ \mathrm{M}\ |^2 (2\pi)^4 \delta^4\left(P_f - P_i\right) V \prod_{f=1}^{N} \left[\tfrac{1}{2E_f V}\right] \prod_{\text{in}} \left[\tfrac{1}{2E_i V}\right] \prod_{f=1}^{N} \tfrac{V d^3\vec{k}_f}{(2\pi)^3} \\ &= |\ \mathrm{M}\ |^2 V \prod_{\text{in}} \left[\tfrac{1}{2E_i V}\right] \times \mathrm{LIPS}\left(N\right) \end{aligned} \tag{4.76}$$

where in the second step we defined the Lorentz invariant phase space with N particles in the final state

$$\mathrm{LIPS}\,(N) \equiv (2\pi)^4 \delta^4(P_f - P_i) \prod_{f=1}^{N} \frac{d^3\vec{k}_f}{(2\pi)^3\, 2E_f} \tag{4.77}$$

Observe that everything in the transition rate is Lorentz invariant save for the initial energy factor and the factors of *V*.

Exercise 4.4: Show that $d^3k/2E$ is a Lorentz-invariant element of phase space. (Hint: Think how you would write the phase space in a 4-dimensional integral but with the particle on-shell, i.e. $E = \left(\left|\vec{k}\right|^2 + m^2\right)^{1/2}$).

4.5.2 Decay rates

We turn now to the special case where we have only one particle with mass *m* in the initial state *i*, i.e. we consider the decay of this particle into some final state *f*. In this case, the transition rate is called the partial decay rate and is denoted by Γ_{if}. First of all, we observe that the dependence on *V* cancels. In the rest frame of the particle the partial decay rate is given by

$$\Gamma_{if} = \frac{1}{2m} \int \mid \mathrm{M} \mid^2 \times \mathrm{LIPS} \tag{4.78}$$

The important special case of two particles in the final state deserves further consideration. Consider the partial decay rate for a particle *i* of mass *m* into two particles f_1 and f_2. The Lorentz-invariant phase space is

$$\mathrm{LIPS}\,(N) = (2\pi)^4 \delta^4\left(p_i - p_1 - p_2\right) \frac{d^3\vec{p}_1}{(2\pi)^3\, 2E_1} \frac{d^3\vec{p}_2}{(2\pi)^3\, 2E_2} \tag{4.79}$$

In the rest frame the four-vectors of each particle are

$$p_i = (m, 0) \qquad p_1 = \left(E_1, \vec{p}\right) \qquad p_2 = \left(E_2, -\vec{p}\right) \tag{4.80}$$

Therefore, we can eliminate one three-momentum in the phase space

$$\mathrm{LIPS}\,(N) = \frac{1}{(2\pi)^2} \delta\left(m - E_1 - E_2\right) \frac{d^3\vec{p}_2}{4E_1E_2} \tag{4.81}$$

Hence the partial decay rate becomes

$$\Gamma_{if} = \frac{1}{8m(2\pi)^2} \int \mid \mathrm{M} \mid^2 \delta\left(m - E_1 - E_2\right) \frac{d\mid \vec{p}_f \mid^2 \mid \vec{p}_f \mid d\Omega^*}{E_1E_2} \tag{4.82}$$

where $d\Omega^*$ is the solid angle element for the angle of one of the outgoing particles with respect to some fixed direction, and $\vec{p}_f$ is the momentum of one of the final state particles. But from the on-shell condition $E_1 = (\left|\vec{p}\right|_1^2 + {m_1}^2)^{1/2}$, we have $dE_1 = \mid \vec{p}_f \mid / E_1\, d \mid \vec{p}_f \mid$ and similarly for particle 2 and so

$$d\left(E_1+E_2\right) = \mid \vec{p}_f \mid d \mid \vec{p}_f \mid \frac{E_1+E_2}{E_1 E_2}$$

therefore

$$\mid \vec{p}_f \mid^2 d \mid \vec{p}_f \mid \frac{1}{E_1 E_2} = \frac{\mid \vec{p}_f \mid}{E_1+E_2} d(E_1+E_2) \tag{4.83}$$

Using this in (4.82) and integrating over (E_1 + E_2) we obtain the final result

$$\Gamma_{i \to f_1 f_2} = \frac{1}{32\pi^2 m^2} \int \mid \mathrm{M} \mid^2 \mid \vec{p}_f \mid d\Omega^* \tag{4.84}$$

The total decay rate of particle *i* is obtained by summation of the partial decay rates into all possible final states

$$\Gamma_{\mathrm{tot}} = \sum_f \Gamma_{if} \tag{4.85}$$

The total decay rate is related to the mean life time τ via $(\Gamma_{\mathrm{tot}})^{-1} = \tau$. For completeness we also give the definition of the branching ratio for the decay into a specific final state *f*

$$B_f \equiv \frac{\Gamma_{if}}{\Gamma_{\mathrm{tot}}} \tag{4.86}$$

In an arbitrary frame we find, $W = (m/E)\Gamma_{\mathrm{tot}}$, which has the expected Lorentz dilation factor. In the master formula (equation (4.76)) this is what the product of $1/2E_i$ factors for the initial particles does.

4.5.3 Cross sections

The total cross section for a static target and a beam of incoming particles is defined as the total transition rate for a single target particle and a unit beam flux. The differential cross section is similarly related to the differential transition rate. We have calculated the differential transition rate with a choice of normalization corresponding to a single 'target' particle in the box, and a 'beam' corresponding also to one particle in the box. A beam consisting of one particle per volume *V* with a velocity *v* has a flux N_0 given by

$$N_0 = \frac{v}{V}$$

particles per unit area per unit time. Thus the differential cross section is related to the differential transition rate in equation (4.76) by

$$d\sigma = \frac{dW}{N_0} = dW \times \frac{V}{v} \tag{4.87}$$

Now let us generalize to the case where in the frame in which you make the measurements, the 'beam' has a velocity v_1 but the 'target' particles are also moving with a velocity v_2. In a colliding beam experiment, for example, v_1 and v_2 will point in opposite directions in the laboratory. In this case the definition of the cross section is retained as above, but now the beam flux of particles N_0 is effectively increased by the fact that the target particles are moving towards it. The effective flux in the laboratory in this case is given by

$$N_0 = \frac{\mid \vec{v}_1 - \vec{v}_2 \mid}{V}$$

which is just the total number of particles per unit area which run past each other per unit time. We denote the velocities with arrows to remind you that they are vector velocities, which must be added using the vector law of velocity addition, not the relativistic law. In the general case, then, the differential cross section is given by

$$d\sigma = \frac{dW}{N_0} = \frac{1}{\mid \vec{v}_1 - \vec{v}_2 \mid} \frac{1}{4E_1E_2} \mid \mathrm{M} \mid^2 \times \mathrm{LIPS} \tag{4.88}$$

where we have used equation (4.76) for the transition rate, and the box volume V has again canceled. The amplitude-squared and phase space factors are manifestly Lorentz invariant. What about the initial velocity and energy factors? Observe that (remember that for a particle $p^\mu = (E, \vec{p}) = \gamma(m, m\vec{v})$ so $\vec{v} = \vec{p}/E$)

$$E_1\, E_2(\vec{v}_1 - \vec{v}_2) = E_2\, \vec{p}_1 - E_1\, \vec{p}_2$$

In a frame where $\vec{p}_1$ and $\vec{p}_2$ are collinear (with four momentum (E_1, p_1, 0, 0) and (E_2, $-\mathrm{p}_2$, 0, 0),

$$\mid E_2\, \vec{p}_1 - E_1\, \vec{p}_2 \mid^2 = (p_1 \cdot p_2)^2 - m_1^2 m_2^2$$

as can be shown by explicitly computing both sides and re-writing using $m^2 = E^2 - \mathrm{p}^2$. The last expression is now manifestly Lorentz invariant.

Hence we can define a Lorentz invariant differential cross section. The total cross section is obtained by integrating over the final state phase space:

$$\sigma = \frac{1}{\mid \vec{v}_1 - \vec{v}_2 \mid} \frac{1}{4E_1E_2} \sum_{\text{final states}} \int \mid \mathrm{M} \mid^2 \times \mathrm{LIPS} \tag{4.89}$$

A slight word of caution is needed in deciding on the limits of integration to get the total cross section. If there are identical particles in the final state then the phase space should be integrated so as not to double count.

2-2 Scatters: An important special case is 2 → 2 scattering

$$a(p_a) + b(p_b) \to c(p_c) + d(p_d)$$

Let us work in the centre of mass frame, in the ultra-relativistic limit, where initially the particles have $p_a = -p_b = E$ and finally $p_c = -p_d = E$. The flux factor is given by

$$F = 4E_a\,E_b(\vec{v}_a - \vec{v}_b) = 4(E_a\,p_b + E_b\,p_a) = 8E^2 = 2s \tag{4.90}$$

and the LIPS factor is

$$\text{LIPS} = (2\pi)^4\delta^4(p_a + p_b - p_c - p_d)\frac{d^3p_c}{(2\pi)^3 2E_c}\frac{d^3p_d}{(2\pi)^3 2E_d} \tag{4.91}$$

one three momenta integral just sets conservation of momenta and we have

$$\text{LIPS} = \frac{1}{4\pi^2}\frac{1}{4p_f^2}\delta(2p_i - 2p_f)(p_f^2\,dp_f d\Omega_f) \tag{4.92}$$

where we have also used $E = p_f$. The Ω_f integration is over the angular distribution of the final momenta. The factors of $p_f{}^2$ cancel top and bottom. We change variables to $y = 2p_f$ to make the integration over the final delta function trivial. We have

$$\text{LIPS} = \frac{1}{32\pi^2}d\Omega_f \tag{4.93}$$

Putting the flux and LIPS factors together gives a final answer of

$$\frac{d\sigma}{d\Omega_f} = \frac{1}{64\pi^2 s}\mid \mathrm{M} \mid^2 \tag{4.94}$$

Exercise 4.5: Show that in the centre of mass frame the differential cross section for two particles of mass m scattering to two of mass M is,

$$\frac{d\sigma}{d\Omega^*} = \frac{\sqrt{1 - 4M^2/s}}{64\pi^2 s\sqrt{1 - 4m^2/s}}\mid \mathrm{M} \mid^2. \tag{4.95}$$

Thus for example if we return to the electron–muon scattering problem where we compute M in (4.60). We can now compute the 2 → 2 cross section formula (4.95) to give, in the high energy limit ($s, \mid u \mid \gg m_e^2, m_\mu^2$),

$$\frac{d\sigma}{d\Omega^*} = \frac{e^4}{32\pi^2 s}\frac{s^2 + u^2}{t^2} \tag{4.96}$$

for the differential cross section in the centre of mass frame.

4.6 More scattering processes

The following two examples show a few more of the subtleties of computing scattering processes in QED.

4.6.1 Electron–electron scattering

Since the two scattered particles are now identical fermions, you cannot just replace m_μ by m_e in the calculation we did above. If you look at the diagram of figure 4.5 (with the muons

replaced by electrons) you will see that the outgoing legs can be labeled in two ways. Hence we get the two diagrams of figure 4.9.

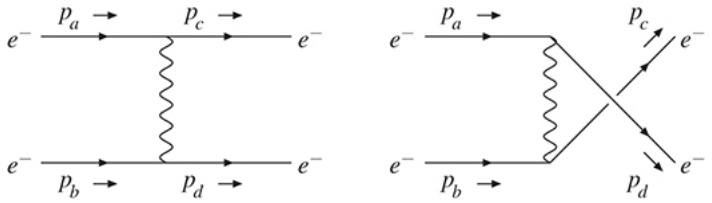


Figure 4.9. Lowest order Feynman diagrams for electron–electron scattering.

The two diagrams give the amplitudes,

$$i\mathrm{M}_1 = \frac{ie^2}{t}\bar{u}(p_c)\gamma^\mu u(p_a)\bar{u}(p_d)\gamma_\mu u(p_b) \tag{4.97}$$

$$i\mathrm{M}_2 = -\frac{ie^2}{u}\bar{u}(p_d)\gamma^\mu u(p_a)\bar{u}(p_c)\gamma_\mu u(p_b) \tag{4.98}$$

Notice the additional minus sign in the second amplitude. This is related to Fermi–Dirac statistics which requires that the two final state fermions cannot be in the same final state. If we set $p_c{}^\mu = p_d{}^\mu$ the result must be zero. The minus sign means in this limit the two diagrams cancel. For a deeper understanding the reader will need to move on to quantum field theory. The minus sign is very important because

$$\begin{aligned} \mid \mathrm{M} \mid^2 &= \mid \mathrm{M}_1 + \mathrm{M}_2 \mid^2 \\ &= \mid \mathrm{M}_1 \mid^2 + \mid \mathrm{M}_2 \mid^2 + 2\ \mathrm{Re}\ \mathrm{M}_1^* \mathrm{M}_2 \end{aligned} \tag{4.99}$$

so the interference term will have the wrong sign if you do not include the extra sign difference between the two diagrams. $\mid \mathrm{M}_1 \mid^2$ and $\mid \mathrm{M}_2 \mid^2$ are very similar to the previous calculation. The interference term is a little more complicated due to a different trace structure.

Performing the calculation explicitly yields (in the limit of negligible fermion masses),

$$\frac{1}{4}\sum_{\text{spins}} \mid \mathrm{M} \mid^2 = 2e^4\left(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} + \frac{2s^2}{tu}\right) \tag{4.100}$$

Exercise 4.6: Prove the result in (4.100). It will be helpful first to prove

$$\begin{aligned} \gamma^\alpha\gamma^\mu\gamma_\alpha &= -2\gamma^\mu \\ \gamma^\alpha\gamma^\mu\gamma^\nu\gamma_\alpha &= 4g^{\mu\nu} \\ \gamma^\alpha\gamma^\mu\gamma^\nu\gamma^\rho\gamma_\alpha &= -2\gamma^\rho\gamma^\nu\gamma^\mu \end{aligned} \tag{4.101}$$

4.6.2 Electron–positron annihilation

The two diagrams for this process are shown in figure 4.10, with the one on the right known as the annihilation diagram. They are just what you get from the diagrams for electron–electron scattering in figure 4.9 if you twist round the fermion lines. The fact that the diagrams are related in this way implies a relation between the amplitudes. The interchange of incoming particles/antiparticles with outgoing antiparticles/particles is called *crossing*. For our particular example, the squared amplitude for $e^+\,e^- \to e^+\,e^-$ is related to that for $e^-\,e^- \to e^-\,e^-$ by performing the interchange $s \leftrightarrow u$. Hence, squaring the amplitude and doing the traces yields (again neglecting fermion mass terms)

$$\frac{1}{4}\sum_{\text{spins}} |\,\mathrm{M}\,|^2 = 2e^4\left(\frac{s^2+u^2}{t^2} + \frac{u^2+t^2}{s^2} + \frac{2u^2}{ts}\right) \tag{4.102}$$

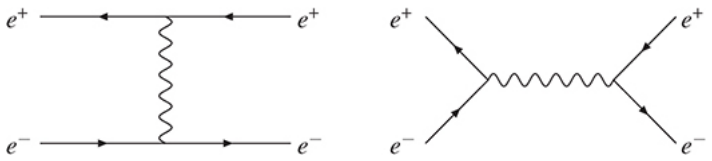


Figure 4.10. Lowest order Feynman diagrams for electron–positron scattering in QED.

If electrons and positrons collide and produce muon–antimuon or quark–antiquark pairs, then the annihilation diagram is the only one that contributes. At sufficiently high energies that the quark masses can be neglected, this immediately gives the lowest order QED prediction for the ratio of the annihilation cross section into hadrons to that into $\mu^+\mu^-$:

$$R \equiv \frac{\sigma(e^+e^- \to \text{ hadrons})}{\sigma(e^+e^- \to\ \mu^+\mu^-)} = 3\sum_f Q_f^2, \tag{4.103}$$

where the sum is over quark flavours f and Q_f is the quark's charge in units of e. The 3 comes from the existence of three colours for each flavour of quark. Historically this was important: you could look for a step in the value of R as your e^+e^- collider's centre of mass energy rose through a threshold for producing a new quark flavour. If you did not know about colour, the height of the step would seem too large.

Finally, we compute the total cross section for $e^+e^- \to \mu^+\mu^-$, neglecting the lepton masses. Here we only have the annihilation diagram, and for the amplitude, we get

$$\begin{aligned}\mathrm{M} &= (-ie)^2\,\bar{u}(p_d)\gamma^\mu v(p_c)\tfrac{-ig_{\mu\nu}}{s}\,\bar{v}(p_a)\gamma^\nu u(p_b)\\ &= \tfrac{ie^2}{s}\,\bar{u}_d\gamma^\mu v_c\bar{v}_a\gamma_\mu u_b\end{aligned} \tag{4.104}$$

Summing over final state spins and averaging over initial spins gives,

$$\frac{1}{4}\sum_{\text{spins}} |\,\mathrm{M}\,|^2 = \frac{e^4}{4s^2}\,\mathrm{Tr}(\gamma^\mu \not{p}_c \gamma^\nu \not{p}_d)\,\mathrm{Tr}(\gamma_\mu \not{p}_b \gamma_\nu \not{p}_a)$$

where we have neglected m_e and m_μ. Using the results in equation (4.54) to evaluate the traces gives,

$$\frac{1}{4}\sum_{\text{spins}} \mid \mathrm{M} \mid^2 = \frac{8e^4}{s^2}\left(p_a \cdot p_d \, p_b \cdot p_c + p_a \cdot p_c \, p_b \cdot p_d\right)$$

Neglecting masses we have,

$$p_a \cdot p_c = p_b \cdot p_d = -t/2 \tag{4.105}$$

$$p_a \cdot p_d = p_b \cdot p_c = -u/2 \tag{4.106}$$

Hence $(1/4)\sum \mid \mathrm{M} \mid^2 u^2)/s^2$, which incidentally is what you get by applying crossing to the electron–muon amplitude of section 4.4.2. We can use this in (4.95) to find the differential cross section in the CM frame,

$$\frac{d\sigma}{d\Omega^*} = \frac{e^4(t^2+u^2)}{32\pi^2 s^3}$$

You could get straight to this point by noting that the appearance of *v* spinors instead of *u* spinors in $\mid \mathrm{M} \mid^2$ does not change the answer since only quadratic terms in m_μ survive the Dirac algebra and we go on to neglect masses anyway. Hence you can use the result of equation (4.96) with appropriate changes.

Neglecting masses, the CM momenta are

$$p_a = \frac{1}{2}\sqrt{s}\,(1,\vec{e}) \qquad p_c = \frac{1}{2}\sqrt{s}\,(1,\vec{e}\prime) \tag{4.107}$$

$$p_b = \frac{1}{2}\sqrt{s}\,(1,\vec{e}) \qquad p_d = \frac{1}{2}\sqrt{s}\,(1,\vec{e}\prime)$$

which gives $t = -s(1-\cos\,\theta)/2$ and $u = -s(1+\cos\,\theta)/2$, where $\cos\,\theta = \vec{e}\cdot\vec{e}\prime$. (4.108)
Hence, finally, the total cross section is,

$$\sigma = \int_{-1}^{1} \frac{d\sigma}{d\Omega^*} 2\pi d\,(\cos\,\theta) = \frac{4\pi\alpha^2}{3s}$$

4.7 Renormalization

Let us very briefly discuss higher order corrections to QED. When we want to calculate scattering amplitudes beyond $\mathrm{O}(e^2)$ we encounter loop diagrams such as those in figure 4.11. Such a loop has a free momentum in it, as shown in figure 4.12.

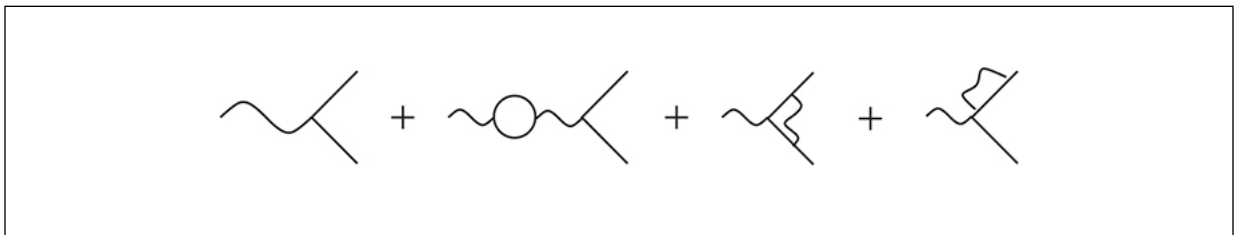


Figure 4.11. Feynman diagrams contributing to an electron photon vertex at one loop in QED.

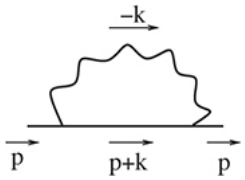


Figure 4.12. The flow of momenta through a loop diagram showing that the momenta *k* is undetermined.

Quantum mechanically we should allow all possible states in the loop, but since there are potentially an infinite number of possible momenta the answer after doing the sum is infinity! The diagram we have drawn contributes to a freely travelling electron and conspires to make the mass infinite and the normalization of the wave function infinite. What is going on?

Actually this is an example of a problem we have in classical physics too. If we treat the electron as a uniform charged ball it has some energy

$$E_{\text{sphere}} = \frac{3}{5}\frac{Q^2}{4\pi\varepsilon_0 R} = mc^2 \tag{4.109}$$

If we believe the electron is a point-like particle we find it has an infinite mass.

All we are learning in these examples is that we are totally ignorant (in both EM and QED) of high energy (ultra-violet) physics that really determines the electron mass. Indeed no one would suggest that QED is a good theory at any energy scale—at the weak scale we must include the weak force and at very high energies gravity.

What we do in both cases is then to 'ignore' this contribution we cannot compute. Formally we can write everywhere in the equations

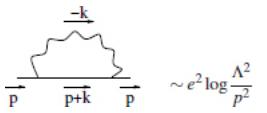
$$m_{\text{physical}} = m_{\text{bare}} + e^2 \log \infty \tag{4.110}$$

where the 'bare' mass is the one we had put into the Dirac equation before we did this computation. It is important that everywhere the physical mass appears there is the same infinite expression but if that happens we can just call the whole lot the observed mass.

If this is going to work we would better be able to absorb all divergences into the four parameters of the theory—the electric charge and mass of the electron and the wave function normalizations of the electron and photon. In fact it is a tougher task to ask of the theory because this must continue to be true at higher loop level where new divergences appear. Remarkably this has been shown to be the case—gauge theories such as QED are unique in having this property as far as we know.

This looks miraculous but in fact physically it is telling us that we have the sensible infra-red end of a correct theory. We call this 'bit' of a theory valid only at low energies an 'effective' theory. We can use this theory in the infra-red and all our ignorance of the ultra-violet is hidden in the parameters of the theory that are handed to us by nature.

All of this though does not mean we just drop the loops in our computations. When the momenta in the loop are of low energy the theory is the correct description. So we might calculate



$$\sim e^2 \log \frac{\Lambda^2}{p^2} \tag{4.111}$$

and shift the mass, say, by this amount. Whilst, when the external particle has momentum q we would get a q rather than p in the answer and want to shift the mass by a different amount. The difference between these two amounts we want to shift by is real and tells us that the particle has a different mass at different energies. The parameters in QED are energy scale dependent.

For example in QED the coupling constant 'runs' with energy according to

$$\frac{e\left(Q^2\right)^2}{4\pi} \equiv \alpha\left(Q^2\right) = \frac{\alpha_0}{1 - \frac{\alpha_0}{3\pi} \ln \frac{Q^2}{M^2}} \tag{4.112}$$

where α_0 is the value of the coupling at the sale M. The form of the solution is sketched in figure 4.13. Note there is a 'Landau pole' at very high energies where the coupling blows up. This is a ridiculously high scale though of 10^{277} GeV. QED will surely be replaced by something else before then!

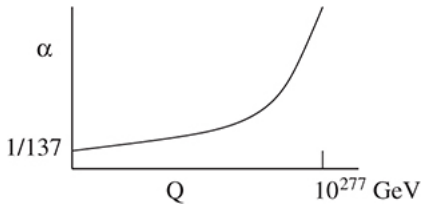


Figure 4.13. A sketch of the running of α with energy scale.

4.8 $g-2$ of the electron

Finally, there is the classic success of renormalization in QED in calculating the gyromagnetic ratio of the electron. Let us sketch how this story goes. The interaction amplitude between an electron and a photon is given by (4.42)

$$\kappa_{fi} = -i \int J^{\mu}_{fi} A_{\mu} \, d^4x \tag{4.113}$$

where

$$J^{\mu}_{fi} = -e\bar{\psi}_f \gamma^{\mu} \psi_i = -e\bar{u}_f \gamma^{\mu} u_i e^{i(p_f - p_i).x} \tag{4.114}$$

There are actually two sorts of interaction present here as we can see using the Gordon decomposition

$$\bar{u}_f \gamma^\mu u_i = \frac{1}{2m} \bar{u}_F [(p_f + p_i)^\mu + i\sigma^{\mu\nu}(p_f - p_i)_\nu] u_i \tag{4.115}$$

where

$$\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \tag{4.116}$$

The Gordon decomposition can be easily derived using the Clifford algebra and the fact that $\not{p}_i u_i = m u_i$.

Inserting the Gordon decomposition into our interaction expression gives us two types of term. The first takes the form

$$J^{fi\,\mu} A_\mu = -e\bar{u}_f \frac{1}{2m}(p_f + p_i)^\mu u_i e^{i(p_f - p_i).x} A_\mu \tag{4.117}$$

which is diagonal in spinor space. This is just the electric coupling of a Klein–Gordon type field.

The second term involves the spin structure and is therefore unique to fermions. We have

$$\kappa_{fi} = -i2\pi\delta(E_i - E_f) \int J^\mu_{fi} A_\mu \, d^3x \tag{4.118}$$

where the time integration has been explicitly carried out to give the energy conserving δ function, and the spatial integral is

$$\int J^\mu_{fi} A_\mu \, d^3x = -\frac{e}{2m} \int \bar{\psi}_f i\sigma^{\mu\nu}(p_f - p_i)^\nu \, \psi_i A^\mu \, d^3x \tag{4.119}$$

To understand this term better we must take the non-relativistic limit. A number of simplifications result:

- the delta function sets $E_f = E_i$ so $(p_f - p_i)^0 = 0$
- the spinors are close to static solutions $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$ so we can drop the bottom two components
- let us also look at the coupling to a time independent magnetic field so $A^\mu = (0, \vec{A})$

These restrictions mean that μ and ν must be spatial indices. Using the explicit form of the γ matrices and restricting to just the top right 2×2 matrix that acts on the top two components of the spinor we have

$$\gamma^0[\gamma^i, \gamma^j] = \begin{pmatrix} -[\sigma^i, \sigma^j] & \cdots \\ \cdots & \cdots \end{pmatrix} \tag{4.120}$$

The upshot of this index structure is that

$$\kappa_{fi} = -i2\pi\delta(E_i - E_f)\int \psi_f^\dagger\Big(\frac{e}{2m}\vec{\sigma}\cdot(\vec{\nabla}\times\vec{A})\Big)\psi_i\, d^3x \tag{4.121}$$

where ψ now has only two components. This is a coupling to the magnetic field $\vec{B} = \vec{\nabla}\times\vec{A}$. It takes the form of a magnetic moment interaction

$$-\vec{\mu}\cdot\vec{B} \tag{4.122}$$

and we see that we are predicting

$$\vec{\mu} = -\frac{e\hbar}{2mc}\vec{\sigma} \tag{4.123}$$

In classical physics the magnetic moment of an orbiting charge e is usually written

$$\vec{\mu}_{\text{orb}} = -\frac{e}{2\,mc}\vec{L} \tag{4.124}$$

and by analogy experimentalists defined the magnetic moment due to intrinsic spin of the charge as

$$\vec{\mu}_{\text{spin}} = -\frac{ge}{2\,mc}\vec{S} = -\frac{ge}{2\,mc}\frac{\vec{\sigma}}{2} \tag{4.125}$$

where g is the gyromagnetic ratio of the particle. The Dirac equation predicts

$$g_{\text{Dirac}} = 2 \tag{4.126}$$

Experimentally one finds for the electron that $g - 2 = 0.00232$ which is pretty good already.

The discrepancy though is due to the next order diagram shown in figure 4.14 which gives a contribution to the vertex Feynman rule of the form

$$\Gamma^\mu = \gamma^\mu F_1\left(q^2\right) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2\left(q^2\right) \tag{4.127}$$

F_1 is the divergent renormalization of the electric charge. F_2 though, which is a contribution to the magnetic moment interaction, is finite. A long calculation gives

$$F_2 = \frac{\alpha}{2\pi} = 0.00232 \tag{4.128}$$

which is even more impressive. At higher order there are many diagrams to consider, and UV divergences enter and must be renormalized. Virtual loops such as that in figure 4.15 probe the physics of quarks and even potentially particles that have not been discovered on-shell yet. To date the computation for the electron has been completed to

order α^4 and matches experiment to 8 significant figures. QED is therefore one of the most stringently tested theories we have ever known.

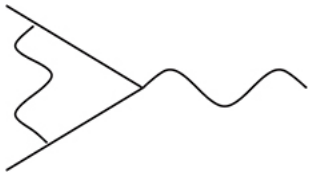


Figure 4.14. The one loop diagram that corrects $g - 2$ of the electron in QED.

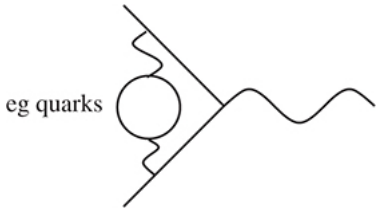


Figure 4.15. A two loop diagram that corrects $g - 2$ of the electron in QED.