LECTURE NOTES ADVANCED QUANTUM MECHANICS
AP3051

Jos Thijssen

Autumn 2013

An electronic version of these notes is available at
`http://blackboard.tudelft.nl/`.

PREFACE

These notes are intended to be used for the course “Advanced Quantum Mechanics”, which is part of the master programme in Applied Physics at Delft University of Technology. They are been the result of teaching activities in the field of quantum mechanics over a range of years and at different levels.

Some relatively elementary topics are covered in the first few chapters of these lecture notes. After some review of complex analysis and bachelor-level quantum mechanics, we shall dig into the formalism of quantum mechanism a bit more deeply, thereby uncovering interesting relations with classical mechanics. Further topics such as variational calculus and the WKB approximation are covered, before entering into the exciting world of Green’s functions, which has applications in scattering theory, open quantum systems and interacting many-body systems. After applying the Born approximation in scattering theory, we enter the main topics of this lecture course: The quantization of harmonic fields such as lattice vibrations and the electromagnetic field, the formalism of second quantization, open quantum systems and relativistic quantum mechanics. Part of the topics are taught in an optional add-on course.

Elaborate exercises are an integral part of the process of learning quantum mechanics, and the material can simply not be mastered without going through these exercises.

In the course of the years, I have learnt a lot from students and from co-teachers. I want to thank them all, and mention here in particular Leo Di Carlo whose many insightful remarks and comments have contributed significantly to this version.

Quantum mechanics at this level has two faces; it is hard material and requires a lot of effort. But it is also great fun; after spending long hours trying to find the solution to the problems, the satisfaction lets you forget all the misery you’ve gone through. I hope that the fun part will persist in the memory of the students after following this course.

Delft, August 2013

CONTENTS

1

A COMPLEX FUNCTION THEORY SURVIVAL GUIDE

Not every student may have had enough complex analysis to appreciate the manipulations we must carry out when – for example – calculating Green's functions in this course; others may have forgotten most of what they've learnt about this subject. Hence, I list the most important results of complex function theory without proof.

An *analytic* function is a complex function which can be differentiated an infinite number of times. It turns out that a complex function which is differentiable, satisfies the *Cauchy-Riemann* equations. In order to formulate these equations, we first introduce some notation. A point in the complex plane is given as

$$z = x + \mathrm{i}y.$$

A complex function $f(z)$ may then be written as

$$f(z) = u(x, y) + \mathrm{i}v(x, y).$$

u is the real, and v the imaginary part of the complex function. The Cauchy-Riemann equations are then

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}; \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

It turns out that this condition is sufficiently strong to ensure infinite differentiability: a complex function which can be differentiated once, can be differentiated an infinite number of times. Functions that satisfy this requirement are called *analytic*.

We often deal with integrations over closed curves ('contours'). It can be shown that the integral of an analytic function taken over such a contour gives zero:

$$\oint_\Gamma f(z)dz = 0,$$

where Γ denotes the contour. We adopt the convention that in complex integration, the contour is always traversed in the anti-clockwise direction. Reversing the direction reverses the sign of the result (which in the case of an analytic function has no effect, as the result of the integration is 0).

We often deal with functions having singularities. Point-like singularities are called *poles*. We say that a function f has a pole of order n in the point a on the complex plane if $(z-a)^n f(z)$ is analytic in a, but $(z-a)^{n-1} f(z)$ is not. Now suppose we expand f around a as a series expansion in $z-a$, including negative powers:

$$f(z) = \sum_{n=-\infty}^{\infty} (z-a)^n f_n(a).$$

The *residue* of f in a, denoted as $\mathrm{res}_a f$ is defined as the coefficient of $(z-a)^{-1}$ in this expansion. For a pole of order one, also a called a simple pole, the residue of f is given as $\lim_{z\to a}(z-a)f(a)$. In general, for an isolated pole, the residue is defined as

$$\mathrm{res}_a f(z) = \frac{1}{(n-1)!}\lim_{z\to a}\frac{d^{n-1}}{dz^{n-1}}\left[(z-a)^n f(z)\right].$$

The most important result of complex analysis that we shall be using frequently, is about functions with a set of isolated poles $a_1, a_2, \dots, a_k$ *within* the closed contour Γ. We then have

$$\oint_\Gamma f(z)dz = 2\pi \mathrm{i}\sum_{j=1}^{k}\mathrm{res}_{a_j} f.$$

This is the so-called *residue theorem.*

Let us consider an example. We calculate the integral

$$\oint \frac{1}{1+z^2}dz$$

over a circle of radius 2 around the origin. The contour contains the points $\pm\mathrm{i}$. We note that

$$\frac{1}{1+z^2} = \frac{\mathrm{i}}{2}\frac{1}{z+\mathrm{i}} - \frac{\mathrm{i}}{2}\frac{1}{z-\mathrm{i}}.$$

Both terms yield a standard integral with a simple pole, and working them out using the residue theorem yields the value

$$\oint \frac{1}{1+z^2}dz = \frac{\mathrm{i}}{2}(2\pi\mathrm{i}) - \frac{\mathrm{i}}{2}(2\pi\mathrm{i}) = 0.$$

Another result is important for the cases we will be dealing with. Consider a semi-circle with radius R in the upper complex plane, and centred around 0. We call this circle Γ^+. Then, *Jordan's lemma* says that if the function $f(z)$ is bounded on Γ^+, the integral

$$\int_{\Gamma^+} e^{\mathrm{i}kz} f(z)dz$$

is finite, and for $R\to\infty$ it approaches zero.

As an example, we calculate

$$\int_{-\infty}^{\infty}\frac{e^{\mathrm{i}kx}}{x-a}dx,$$

where a is a complex number with a positive imaginary part and k is real. We now evaluate the integral over the contour Γ shown in figure 1.1. Because of Jordan's Lemma, this integral is equal to the integral over the real axis only. Using the residue theorem, we immediately have

$$\int_{-\infty}^{\infty}\frac{e^{\mathrm{i}kx}}{x-a}dx = \oint\frac{e^{\mathrm{i}kx}}{x-a}dx - \int_{\Gamma^+}\frac{e^{\mathrm{i}kx}}{x-a}dx.$$

Now we use Jordan's lemma which says that the second term vanishes, and the integral over the closed contour can be evaluated using the residue theorem:

$$\oint\frac{e^{\mathrm{i}kx}}{x-a}dx = 2\pi\mathrm{i}e^{ika}.$$

Finally, we consider the integral

$$\int\frac{f(x)}{x}dx$$

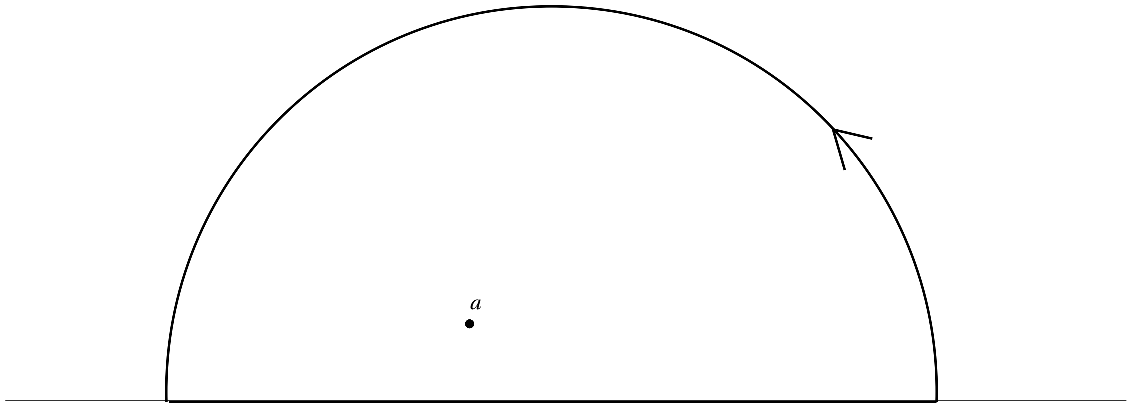


FIGURE 1.1: The contour for evaluating an integral along the real axis.

over the real axis. This integral is not well defined at $x = 0$ unless f vanishes there (and is continuously differentiable at $x = 0$). Now let us consider the same integral but running just above the real axis:

$$\int \frac{f(x+\mathrm{i}\epsilon)}{x+\mathrm{i}\epsilon} dx,$$

where ϵ is small. Supposing again that f is regular, the small ϵ does not change its value significantly, but we have avoided the singularity at $x = 0$. For x not very close to zero, the integral is approximately equal to the one running over the real axis, and we should focus on what happens near $x = 0$. Let us work out the imaginary part of $1/(x+\mathrm{i}\epsilon)$:

$$\mathrm{Im}\left(\frac{1}{x+\mathrm{i}\epsilon}\right) = \frac{1}{2\mathrm{i}}\left(\frac{1}{x+\mathrm{i}\epsilon} - \frac{1}{x-\mathrm{i}\epsilon}\right) = -\frac{\epsilon}{x^2+\epsilon^2}.$$

The right hand side of this function is a narrow peak centred around $x = 0$. Its integral is $-\pi$. Therefore, for small ϵ, the integral will be $-\pi\delta(x)$. All in all, we see that we can write

$$\frac{1}{x+\mathrm{i}\epsilon} = 1/x \text{ (for } x \text{ away from zero) } - \mathrm{i}\pi\delta(x).$$

The first part of this expression is called the *principal value* and denoted as $\mathscr{P}$:

$$\frac{1}{x+\mathrm{i}\epsilon} = \mathscr{P}\left(\frac{1}{x}\right) - \mathrm{i}\pi\delta(x).$$

More precisely

$$\mathscr{P}\left(\frac{1}{x}\right) = 1/x \text{ for } |x| > \epsilon, \epsilon \to 0.$$

For completeness, we write down the principal value integral over a function $f(x)$ having a singularity in a is:

$$\mathscr{P}\int f(x)dx = \lim_{\epsilon\downarrow 0}\int_{-\infty}^{a-\epsilon} f(x)dx + \int_{a+\epsilon}^{\infty} f(x)dx.$$

2

A QUANTUM MECHANICS SURVIVAL GUIDE

In this chapter we review standard quantum mechanics, highlighting important results. Quantum mechanics describes the evolution of the state of a mechanical system. The state is a vector in a special kind of complex vector space, the *Hilbert space*. This is a vector space in which an inner product is defined.

We represent the state of the system by a ket-vector, such as $\left|\psi\right\rangle$. The inner product between vectors $\left|\psi\right\rangle$ and $\left|\phi\right\rangle$ is denoted $\langle\phi|\psi\rangle$. For a quantum mechanical system for which we want to calculate the time evolution, the vector describing the system becomes itself time-dependent: $\left|\psi(t)\right\rangle$.

The equation determining the time evolution of a known state vector $\left|\psi(0)\right\rangle$ at $t = 0$ is known as the *time-dependent Schrödinger equation*:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = \hat{H}\left|\psi(t)\right\rangle. \tag{2.1}$$

Here, $\hat{H}$ is an *operator* acting on vectors in the Hilbert space. $\hat{H}$ is *Hermitian*, which means it is equal to its Hermitian conjugate. The Hermitian conjugate $\hat{A}^\dagger$ of an operator $\hat{A}$ is defined as follows: $\hat{A}^\dagger$ must be such that for any two vectors ψ and ϕ it must hold that

$$\langle\phi|\hat{A}|\psi\rangle = \left(\left\langle\psi|\hat{A}^\dagger|\phi\right\rangle\right)^*,$$

We summarise:

$\hat{A}^\dagger = \hat{A}$ means that $\hat{A}$ is Hermitian.

In particular, the Hamiltonian $\hat{H}$ is Hermitian: $\hat{H}^\dagger = \hat{H}$.

Hermitian operators have important properties:

- The eigenvalues λ of a Hermitian operator are all real: $\lambda = \lambda^*$.
- Eigenvectors $\left|\phi_\lambda\right\rangle$ and $\left|\phi_\mu\right\rangle$ belonging *different* eigenvalues λ and μ are always mutually orthogonal:
$$\langle\phi_\lambda|\phi_\mu\rangle = 0; \qquad \lambda \neq \mu.$$
- *Degenerate* eigenvectors can be chosen orthogonal.
- All eigenvectors span the Hilbert space.

The fact that the eigenvectors form a basis of the Hilbert space leads to the often-used *resolution of the identity*, in which the unit operator $\mathbb{1}$ is written as:

$$\mathbb{1} = \sum_j |\phi_\lambda\rangle\langle\phi_\lambda|$$

where the sum is over all the eigenvectors.

The solution to the time-dependent Schrödinger equation (2.1) is easy to find: it is just

$$|\psi(t)\rangle = e^{-\mathrm{i}\hat{H}t/\hbar}|\psi(0)\rangle.$$

This solution is verified by substituting it back into that equation. However, working out this solution is very difficult, as it contains the exponential of an operator. The easiest way to handle this exponent is by diagonalising the operator. Suppose we have a complete set of eigenvectors $|\phi_n\rangle$ and eigenvalues E_n for $\hat{H}$:

$$\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle, \tag{2.2}$$

and that we know how to expand the initial state $|\psi(0)\rangle$ into these eigenstates. Given that the exponent of an operator can be written as the diagonal operator with the exponents of the eigenvalues on its diagonal (I write this here for a finite-dimensional matrix):

$$e^{-\mathrm{i}t\hat{H}/\hbar} = \begin{pmatrix} e^{-\mathrm{i}E_1t/\hbar} & 0 & \cdots & 0 & 0 \\ 0 & e^{-\mathrm{i}E_2t/\hbar} & \cdots & 0 & 0 \\ & & & & \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 0 & e^{-\mathrm{i}E_Nt/\hbar} \end{pmatrix},$$

we obtain

$$|\psi(t)\rangle = \sum_{j=1}^{N} c_j e^{-\mathrm{i}E_jt/\hbar}|\phi_j\rangle,$$

where

$$c_j = \langle\phi_j|\psi(0)\rangle.$$

Eq. (2.2) is called the *stationary Schrödinger equation.*

In case you have forgotten how to diagonalise a matrix, I recall that, since for an eigenvector $|\phi\rangle$ of an operator $\hat{A}$,

$$\hat{A}|\phi\rangle = \lambda|\phi\rangle,$$

$|\phi\rangle$ should be a non-zero vector for which

$$(\hat{A} - \lambda\mathbb{1})|\phi\rangle = 0.$$

From your first lecture on linear algebra, you should know that this can only be true if the determinant of the matrix $\hat{A} - \lambda\mathbb{1}$ vanishes. This leads to an algebraic equation for λ. As an example, consider

$$\hat{H} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The determinant condition is

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0,$$

which leads to

$$\lambda^2 = 1,$$

so $\lambda = \pm 1$. This could be anticipated as you may have recognised $\hat{H}$ as the Pauli matrix σ_x and know that the Pauli matrices all have eigenvalues ± 1. Obviously, the larger the matrix, the higher the order of the equation for λ and the more work it takes to find the eigenvalues.

Once you have the eigenvalue, you may find the corresponding eigenvector by solving linear equations. Calling the eigenvector (a, b) we have,for the $+1$ eigenvalue,

$$\begin{aligned} b &= a \\ a &= b \end{aligned}$$

and the normalised eigenvector becomes $(1,1)/\sqrt{2}$. Similarly we find $(1,-1)/\sqrt{2}$ for the eigenvector with eigenvalue -1. These eigenvectors could also have been guessed if you let yourself be guided by the symmetric structure of the matrix you are diagonalising. Another helpful fact is that, for a Hermitian matrix (operator), the eigenvectors belonging to different eigenvalues are orthogonal, and that for equal eigenvalues (degeneracy), all eigenvectors can always be chosen orthogonal. The matrix eigenvalue problem can be solved analytically only if the matrix size is modest (typically smaller than or equal to 3) or if the matrix has a simple and/or very regular structure. In all other cases we use numerical routines for solving the eigenvalue problem.

We see that the time evolution of a wave function is determined by the *time evolution operator* $\hat{U}(t) = e^{-\mathrm{i}t\hat{H}/\hbar}$. From the hermiticity of $\hat{H}$, it is easy to see that $\hat{U}(t)$ satisfies

$$\hat{U}(t)\hat{U}^\dagger(t) = \hat{U}^\dagger(t)\hat{U}(t) = \mathbb{1},$$

where $\mathbb{1}$ is the unit operator. An operator satisfying this equation is called *unitary*. We see that unitarity of the time evolution operator directly follows from the hermiticity of the Hamiltonian. Interestingly, this unitarity also guarantees that the norm of the wave function is conserved. To see this, we use

$$\begin{aligned} \left|\psi(t)\right\rangle &= \hat{U}(t)\left|\psi(0)\right\rangle; \\ \left\langle\psi(t)\right| &= \left\langle\psi(0)\right|\hat{U}^\dagger(t), \end{aligned}$$

to evaluate what happens to the norm as time evolves:

$$\left\langle\psi(t)|\psi(t)\right\rangle = \left\langle\psi(0)|\hat{U}^\dagger(t)\hat{U}(t)|\psi(0)\right\rangle = \left\langle\psi(0)|\psi(0)\right\rangle,$$

which shows that the norm is indeed preserved.

The Schrödinger equation is quite a general equation and does not specify the structure of the Hilbert space, nor the specific form of the Hamiltonian. Physicists have guessed both in the first decades of the twentieth century and good guesses have turned out to yield results for physical measurements in excellent agreement with experiment. Here we list a few.

- Spinless point particle in one dimension. Hilbert space: class of square integrable functions (L_2) on the real axis. Hamiltonian:

  $$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x).$$

- Spinless point particle in three dimensions. Hilbert space: class of square integrable functions (L_2) in $\mathbb{R}_3$. Hamiltonian:

  $$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}).$$

- Particles with spin 1/2, neglecting their motion. Hilbert space: two-dimensional vector space. Hamiltonian:

  $$\hat{H} = \frac{eB}{m}\sigma_z,$$

  where B is a magnetic field along the z-axis and σ_z is the Pauli matrix

  $$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to extend this list with numerous other cases.

For any physical quantity A, appropriate for the system at hand, there exists a Hermitian operator $\hat{A}$ whose eigenvalues λ_n are the possible values of A found in a measurement. These values occur with probability $\left|\langle\phi_n|\psi\rangle\right|^2$, where ϕ_n is the eigenvector corresponding to λ_n and $\left|\psi\right\rangle$ is the state of the system. The expectation value of A in a system in quantum state $\left|\psi\right\rangle$ is given by $\langle\psi|\hat{A}|\psi\rangle$.

We now concentrate on electrons in 3D, moving in the field of a radial potential depending only on the distance r to the origin: $V \equiv V(r)$, $r = |\mathbf{r}|$. This is a special example of a system exhibiting a symmetry. If there is a symmetry, there usually is *degeneracy*, meaning that two or more eigenvalues of the Hamiltonian (i.e., the energies) have the same values – we shall return to this point in chapter 4. It turns out that if there is symmetry, there is one or more operators that commute with the Hamiltonian. The expectation values of these operators (which are assumed to have no explicit time dependence) then remain unchanged in time as can easily be checked from the time evolution:

$$\langle A\rangle_t = \left\langle\psi_0\left|e^{\mathrm{i}tH/\hbar}Ae^{-\mathrm{i}tH/\hbar}\right|\psi_0\right\rangle.$$

Taking the time derivative of this expression yields the commutator $[H, A]$ which vanishes by assumption. It is possible to find a set of vectors that are eigenvectors of all independent operators which commute with H. To be specific, states whose *energy* eigenvalues are degenerate may have different eigenvalues for an operator $\hat{A}$ other than $\hat{H}$. In order to identify all the states uniquely, we need in addition to $\hat{H}$ a set of operators $\hat{A}$, $\hat{B}$, $\hat{C}$,... such that each (simultaneous) eigenvector of this set of operators has a *unique* set of eigenvalues E_n, a_j, b_k, c_l, ... The set of all independent operators which commutes with H, including H is called *observation maximum*:

An *observation maximum* is the set of all independent operators, including H, that commute amongst themselves and with H. The eigenvalues of all these operators label the simultaneous eigenvectors of all the operators of the observation maximum. They form a basis of the Hilbert space.

Degeneracy is related to symmetry which is present in the Hamiltonian (this is formally substantiated by the quantum mechanical version of Noether's theorem, which we shall not go into here). This relation is the reason why the energies of a 3D system which is spherically symmetric (that is, a system with a radial potential), are degenerate. The operators commuting with the Hamiltonian for a spinless particle in a radial potential are the angular momentum operators L^2 and L_z. These have the eigenvalues $\hbar^2 l(l+1)$ and $\hbar m$ respectively, where l is $0, 1, 2, \ldots$ and m is an integer running from $-l$ to l. The energy eigenvalues depend on l and an additional quantum number, n – they are written as E_{nl}. For each l, m runs from $-l$ to l in integer steps. So there are $2l+1$ m-values for each l. As the energy eigenvalues do *not* depend on m, they are (at least) $2l+1$-fold degenerate. If the particles have spin-1/2, there are additional quantum numbers: s which always takes on the value 1/2 (as we are dealing with spin-1/2 particles) and m_s which takes the value $\pm 1/2$. All in all, the states of an electron in a radial potential are denoted $|n, l, m, s, m_s\rangle$. The quantum number s being always 1/2 for an electron, is often left out. If we include the spin, each level with quantum number l is $2(2l+1)$-fold degenerate. Figure 2.1 shows a schematic representation of a spectrum of a particle in a radial potential.

The Coulomb potential is a special case: this potential has some hidden symmetry [1] which causes several of the E_{nl} to coincide. Whereas we normally label the energy eigenvalues for each l by $n = 1, 2, \ldots$, this degeneracy allows us to label the states as in Figure 2.2. The quantum number along the vertical axis is called the *principal quantum number*. We see that for each principal quantum number n, we have states with l-values between 0 and $n-1$.

---

[1] This symmetry is related to the four-dimensional rotational group $O(4)$.

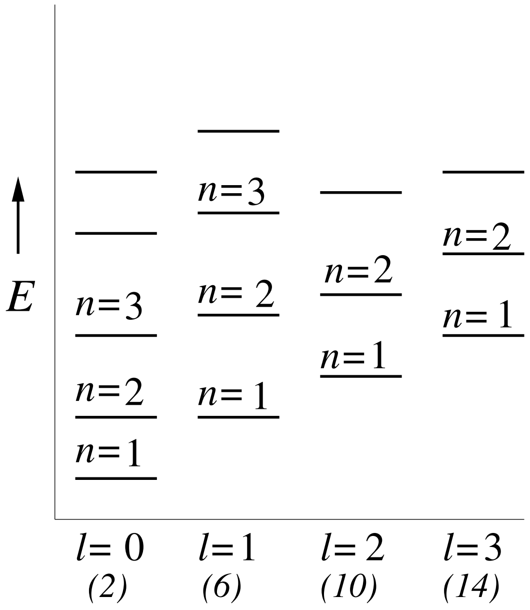


FIGURE 2.1: The spectrum of an electron in a radial potential, grouped according to the l-value. In parenthesis, the degeneracy of the levels (including the two-fold spin-degeneracy) is given.

This is a special degeneracy of the hydrogen atom – the degeneracies of the energy levels can be found by adding the degeneracies for each $l = 0, \ldots, n-1$. These degeneracies correspond to the so-called 'noble gas' atoms (for higher n, deviations from this series occur due to effects not taken into account here).

2.1 SPIN-1/2 AND THE BLOCH SPHERE

In this section, we elaborate a bit on the states of a particle whose Hilbert space is two-dimensional; this is the simplest non-trivial example of a Hilbert space. The standard example of such a system is a spin-1/2 particle, but many other realisations of systems with a two-dimensional Hilbert space are possible. Systems with such a Hilbert space are denoted as 'two-level systems' (TLS).

Any Hermitian operator in this space can be represented as a 2×2 hermitian matrix. This means that such a matrix has in principle 4 degrees of freedom (the diagonal elements must be real, and the off-diagonal elements must be each-other's complex conjugate). This means that any hermitian operator can be written as a linear combination of four basis operators. These are taken to be the unit matrix and the three Pauli matrices:

$$\begin{aligned}
\mathbb{1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\sigma_y &= \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \\
\sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned}$$

The Pauli matrices satisfy the properties:

$$\{\sigma_j, \sigma_k\} = \sigma_j \sigma_k + \sigma_k \sigma_j = 2\delta_{jk}$$

where the indices j and k stand for x, y or z. The braces $\{,\}$ are generally used for the *anti-commutator* in these notes. From this it follows in particular that $\sigma_j^2 = \mathbb{1}$. If we apply

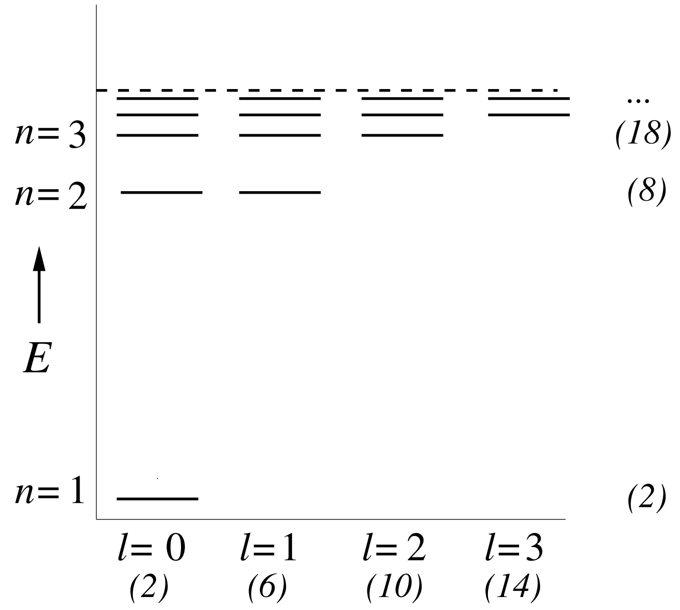


FIGURE 2.2: The spectrum of an electron in a Coulomb potential, grouped according to their l-value. The numbers along the vertical axis is the principal quantum number. The degeneracies per l-value are the same as for Figure 2.1. Adding the degeneracies for each principal quantum number n gives the degeneracies in parentheses on the right hand side.

a space rotation, the spin also changes. This change is expressed by the rotation operator $\exp(\mathrm{i}\boldsymbol{\alpha}\cdot\boldsymbol{\sigma}/2)$, which represents a rotation over an angle $|\boldsymbol{\alpha}|$ about an axis with direction $\boldsymbol{\alpha}$. Although this looks like a complicated expression to work out, the property $\sigma_j^2 = \mathbb{1}$ enables us to turn it into a simple expression. For a rotation over an angle α about the z-axis, we have

$$\exp(\mathrm{i}\alpha\sigma_z/2) = \cos(\alpha/2)\mathbb{1} + \mathrm{i}\sin(\alpha/2)\sigma_z,$$

where the equality can be verified from the Taylor expansions of the exponential function and of the sine and cosine.

A vector in the spin-1/2 Hilbert space is represented by a vector

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \text{ with } |a|^2 + |b|^2 = 1.$$

This vector is characterised by three real numbers (two complex numbers contain four real numbers, minus one because of the normalisation condition). Furthermore, when we multiply this vector by a phase factor $\exp(\mathrm{i}\gamma)$, the state does not change. Therefore by choosing, say, a to be real, there are only two numbers left. These can be represented by a point on a 3D sphere (which itself is a two-dimensional manifold) – see figure 2.3. The point is defined by the two polar angles ϑ and φ. The point on the sphere is simply given by the expectation value of the three Pauli matrices:

$$\langle\psi|\sigma_x|\psi\rangle = a^* b + ab^*$$

and similar for y and z. The point with these coordinates is the *polarisation.*

Up[to an overall phase factor, the relation between the components a and b and the polar angles ϑ and φ is given by

$$a = \exp(-\mathrm{i}\varphi/2)\cos(\vartheta/2);$$
$$b = \exp(\mathrm{i}\varphi/2)\sin(\vartheta/2),$$

as can be verified (see problem 9). The sphere of polarisation points is called *Bloch sphere.*

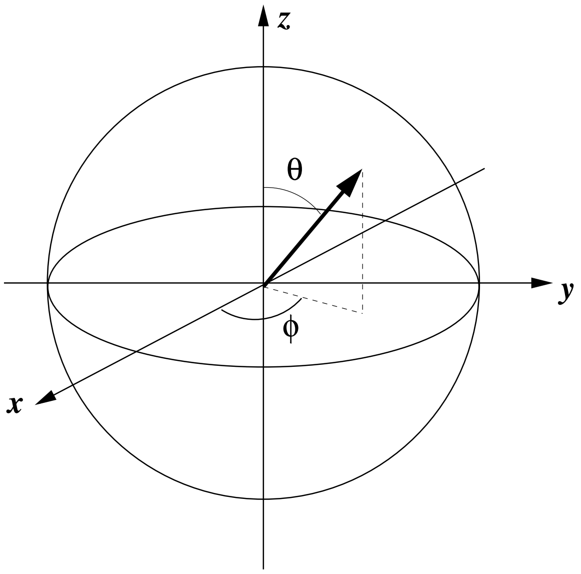


FIGURE 2.3: The Bloch sphere.

2.2 SCHRÖDINGER AND HEISENBERG PICTURES

In quantum mechanics, experimental results are expressed in terms of matrix elements of operators. In general, for an operator $\hat{A}$ (which we assume to be time-independent), such a matrix element is

$$\langle \phi | \hat{A} | \psi \rangle .$$

These matrix elements change in time as the wave functions are time-dependent:

$$|\psi(t)\rangle = e^{-\mathrm{i}Ht/\hbar} |\psi\rangle$$

where $|\psi\rangle$ on the right hand side is the wave function at $t = 0$.

Therefore, we can write

$$\langle \phi(t) | \hat{A} | \psi(t) \rangle = \left\langle \phi \left| e^{\mathrm{i}Ht/\hbar} \hat{A} e^{-\mathrm{i}Ht/\hbar} \right| \psi \right\rangle .$$

From this formulation, it is immediately clear that we can take two viewpoints:

- We take $\hat{A}$ independent of time and let $|\psi\rangle$ and $|\phi\rangle$ evolve in time according to the time-dependent Schrödinger equation, or
- We take the wave functions $|\psi\rangle$ and $|\phi\rangle$ fixed and introduce a time-dependent operator $\hat{A}(t)$:
$$\hat{A}(t) = e^{\mathrm{i}Ht/\hbar} \hat{A} e^{-\mathrm{i}Ht/\hbar}.$$
In that case, the matrix element is written as
$$\langle \phi | \hat{A}(t) | \psi \rangle .$$

The first viewpoint is called the *Schrödinger picture* and the second the *Heisenberg picture*. For systems where the Hamiltonian can be split into an ‘easy’ and a ‘difficult’ part, it makes sense to use a third picture, called *interaction picture*. This will be covered later in these notes.

2.3 PROBLEMS

1. An electron in a hydrogen atom finds itself in a state

$$\left|\psi\right\rangle = |1,0,0\rangle + \frac{1}{2}|2,1,1\rangle + \frac{1}{\sqrt{2}}|2,1,0\rangle + \frac{i}{\sqrt{2}}|2,1,-1\rangle .$$

We neglect spin and the states are labeled as $|n,l,m\rangle$.

(a) Normalise this state.

(b) Calculate the probability of finding the electron with energy E_2.

(c) Calculate the probability of finding the electron with angular momentum component $L_z = 0$.

(d) Calculate the probability of finding $L_x = \hbar$ in a measurement.

(e) The hydrogen atom is subject to a field which adds a term αL_x to the Hamiltonian. Give $\left|\psi(t)\right\rangle$, if the state given above is the state at $t = 0$.

2. Consider a spin-1/2 particle in a spherically symmetric potential. The state of the particle is denoted $\left|\psi\right\rangle$. **L** is the orbital angular momentum and **S** the spin operator. The functions ψ_+ and ψ_- are defined by

$$\psi_\pm(\mathbf{r}) = \left\langle \mathbf{r},\pm|\psi\right\rangle, \tag{2.3}$$

where the second argument in the bra-vector on the right-hand side denotes the spin and where

$$\psi_+(\mathbf{r}) = R(r)\left[Y_0^0(\theta,\phi) + \frac{1}{\sqrt{3}}Y_0^1(\theta,\phi)\right]; \tag{2.4a}$$

$$\psi_-(\mathbf{r}) = \frac{R(r)}{\sqrt{3}}\left[Y_1^1(\theta,\phi) - Y_0^1(\theta,\phi)\right]; \tag{2.4b}$$

with R some given function of r and $Y_m^l(\theta,\phi)$ the eigenfunctions of the angular momentum operators L^2, L_z.

(a) Which condition must be satisfied by $R(r)$ in order for $\left|\psi\right\rangle$ to be normalised?

(b) A measurement of the spin z-component S_z is performed. What are the possible results with respective probabilities? Same questions for L_z and S_x.

(c) A value 0 is measured for the quantity L^2. What is the state of the particle immediately after the measurement?

3. An electron subject to a magnetic field in the z direction evolves under the Hamiltonian

$$H = -\frac{1}{2}\Delta_1 S_z.$$

At $t = 0$, the electron spin points along the positive x direction.

(a) Calculate the time-dependent wavefunction.

(b) Formulate the equations of motion of the expectation values of the spin components, $\langle S_x\rangle$, $\left\langle S_y\right\rangle$ and $\langle S_z\rangle$.

From now, we consider the stationary behaviour rather than the dynamics.

(c) Consider now a second electron. This second electron experiences a field in the z direction, but with different magnitude. The two electrons interact via a weak transverse coupling. The total Hamiltonian is thus given by:

$$H = -\frac{1}{2}\Delta_1 S_{z1} - \frac{1}{2}\Delta_2 S_{z2} + J\left(S_{x1}S_{x2} + S_{y1}S_{y2}\right),$$

where $J << \Delta_1, \Delta_2, |\Delta_1 - \Delta_2|$.

Use first-order perturbation theory to estimate the energies and corresponding eigenstates for this Hamiltonian.

(d) Use second-order perturbation theory to estimate the energies and corresponding eigenstates for this Hamiltonian.

(e) Find the exact solution for the energies and eigenstates.

4. Consider two spin-1/2 particles subject to a Heisenberg-type interaction

$$H_0 = -\alpha \mathbf{S}_1 \cdot \mathbf{S}_2,$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the spin operators for particle 1 and particle 2, respectively.

(a) Find the energies and expand the corresponding eigenstates in the basis $|s, m_s\rangle$, where s and m_s denote the quantum numbers for the total momentum operators $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ and $S_z = S_{z1} + S_{z2}$, respectively.

(b) Expand the eigenstates in the basis $|m_{s1}, m_{s2}\rangle$, where m_{s1} en m_{s2} are the quantum numbers of the operators S_{z1} en S_{z2}.

(c) Consider now the modified Hamiltonian

$$H = H_0 + \beta S_z^2,$$

Give the exact energies and corresponding eigenstates.

(d) Determine whether the ground state of the system is entangled. Note: an entangled state cannot be written as a product of states for particle 1 and particle 2.

5. Consider two quantum dots in close proximity. A quantum dot is a small structure that can be occupied by an electron. We consider the case where only one state is available in each dot. In this problem, we assume that exactly one electron is present in the system. Denote the energies of the states as ϵ_i, $i = 1, 2$ where i labels the dot. To these levels correspond the states $|\phi_i\rangle$, $i = 1, 2$. As the dots are placed close together, they are coupled. The coupling constant is (the complex number) τ:

$$\tau = \langle \phi_1 | H | \phi_2 \rangle,$$

where H is the Hamiltonian.

(a) Write the Hamiltonian of the system in the form of a 2×2 matrix.

(b) Find the spectrum of this Hamiltonian. Plot the spectrum as a function of $\epsilon_1 - \epsilon_2$ for fixed τ. Also find the eigenfunctions.

(c) Suppose we place an electron in dot 1 at $t = 0$. Give the time evolution of the wave function and show that the probability to find the electron in dot 2 as a function of time has the form

$$P(t) = C\left[1 - \cos(\omega t)\right].$$

Determine ω. Hint: write the quantum state $|\psi\rangle$ as

$$|\psi\rangle = a|\phi_1\rangle + b|\phi_2\rangle$$

and use the spectrum found in (b) together with the time-dependent solution of the Schrödinger equation.

6. A particle is located at the origin of a line where the potential V is zero. At $t = 0$, the particle is released. Find the wave function in the x-representation at time $t > 0$.

7. (a) Consider the three spin triplet states $|1m\rangle$ (in the standard notation $|s, m_s\rangle$) and the singlet state $|00\rangle$, which are constructed out of two spin-$\frac{1}{2}$ particles, A and B. The vector $\boldsymbol{\sigma}_{\mathrm{A}} = (\sigma_{\mathrm{A}x}, \sigma_{\mathrm{A}y}, \sigma_{\mathrm{A}z})$ has the three Pauli spin matrices as its components, and $\boldsymbol{\sigma}_{\mathrm{B}}$ is defined in a similar way. Show that

   $$(\boldsymbol{\sigma}_{\mathrm{A}} \cdot \boldsymbol{\sigma}_{\mathrm{B}})\,|1m\rangle = +\,|1m\rangle$$

   and

   $$(\boldsymbol{\sigma}_{\mathrm{A}} \cdot \boldsymbol{\sigma}_{\mathrm{B}})\,|00\rangle = -3\,|00\rangle\,.$$

   Obtain the eigenvalues of $(\boldsymbol{\sigma}_{\mathrm{A}} \cdot \boldsymbol{\sigma}_{\mathrm{B}})^n$.

   (b) A system consisting of two spin $\frac{1}{2}$ particles is described by the Hamiltonian

   $$H = \begin{cases} \lambda\,(\boldsymbol{\sigma}_{\mathrm{A}} \cdot \boldsymbol{\sigma}_{\mathrm{B}})\,, & t \geq 0, \\ 0 & t < 0. \end{cases}$$

   Assume $\left|\psi(t=0)\right\rangle = |{\downarrow\uparrow}\rangle$, i.e. particle A has spin along the $-z$ axis, and particle B has its spin oriented along the $+z$ axis. Express H in terms of the vector operator $\boldsymbol{\sigma} = \boldsymbol{\sigma}_{\mathrm{A}} + \boldsymbol{\sigma}_{\mathrm{B}}$. Obtain $\left|\psi(t)\right\rangle$ for $t > 0$ and determine the probability for finding the system in the states $|{\uparrow\uparrow}\rangle$, $|{\downarrow\downarrow}\rangle$ and $|{\uparrow\downarrow}\rangle$.

8. We consider some properties of the Pauli matrices σ_x, σ_y and σ_z.

   (a) Show that $\sigma_i^2 = 1$, for $i = x, y, z$.

   (b) Show that $[\sigma_i, \sigma_j] = 2\mathrm{i}\epsilon_{ijk}\sigma_k$ where $\epsilon_{ijk} = 1$ whenever i, j, k is an *even* permutation of x, y, z and $\epsilon_{ijk} = -1$ when i, j, k is an *odd* permutation of i, j, k.

   (c) For angular momentum quantum number $l = 1$, we have three possible states $|1, m\rangle$, where $m = 1, 0, -1$.
   Write down the matrix form of L_z in the basis of these states.

   (d) Show that the matrices

   $$L_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } L_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix},$$

   together with the matrix found in (c), satisfy the angular momentum commutation relations

   $$\left[L_i, L_j\right] = \mathrm{i}\hbar\epsilon_{i,j,k}L_k.$$

   (e) Calculate the eigenvalues and eigenvectors of L_x and L_y.

9. *It may be useful to refer to the previous problem to refresh your knowledge concerning the Pauli matrices!*
   The state of a quantum two-level system (also called quantum bit, qubit or Q-bit) can be written as

   $$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle,$$

   where the coefficients α and β are complex, and properly normalised: $|\alpha|^2 + |\beta|^2 = 1$.

   (a) Explain why, without loss of generality, we can also write the state as

   $$|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{\mathrm{i}\phi}\sin(\theta/2)\,|1\rangle,$$

   with θ ranging from 0 to π and ϕ ranging from 0 to 2π.

(b) There is a one-to-one correspondence between the state $|\psi\rangle$ and a unit vector in 3-D, whose orientation is defined by polar and azimuthal angles θ and ϕ, respectively (see figure 2.3). Such a vector is called the *Bloch vector* associated with $|\psi\rangle$. Draw the Bloch vectors associated with the states $|0\rangle$, $|1\rangle$, $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$, and $\frac{1}{2}|0\rangle - \frac{i\sqrt{3}}{2}|1\rangle$.

(c) Find $\langle\psi|\sigma_x|\psi\rangle$, $\langle\psi|\sigma_y|\psi\rangle$, and $\langle\psi|\sigma_z|\psi\rangle$ as a function of θ and ϕ. Here, σ_x, σ_y and σ_z are the Pauli matrices.

(d) Show that $\langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 = 1$.

(e) How are $\langle\sigma_x\rangle$, $\langle\sigma_y\rangle$ and $\langle\sigma_z\rangle$ related to the cartesian coordinates of the Bloch vector?

(f) Unitary operations on one qubit correspond to rotations in the *Bloch sphere*. What kind of rotations correspond to the operators σ_x, σ_y and σ_z? Specify the axis of rotation, and the rotation angle for each.

(g) What does the operator $\hat{R}_x(\alpha) \equiv \cos(\alpha/2)\mathbb{1} - \mathrm{i}\sin(\alpha/2)\sigma_x$ do?

(h) What does the operator $\hat{R}_{\vec{n}}(\alpha) \equiv \cos(\alpha/2)\mathbb{1} - \mathrm{i}\sin(\alpha/2)\vec{n}\cdot\{\hat{\sigma}_x,\hat{\sigma}_y,\hat{\sigma}_z\}$ do? Here, $\vec{n}$ is a 3-D unit vector.
(You only need to show that $\hat{R}_{\vec{n}}$ leaves the state with Bloch vector $\vec{n}$ invariant.)

10. The non-relativistic Hamiltonian that describes the interaction of a charged particle with the electromagnetic field is

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{r})\right)^2 + q\phi(\mathbf{r})$$

(a) Assume that the wave function is changed by a constant phase

$$\psi(\mathbf{r},t) \to \psi'(\mathbf{r},t) = e^{\mathrm{i}\alpha}\psi(\mathbf{r},t).$$

Show that $\psi'(\mathbf{r},t)$ satisfies the original Schrödinger equation with the original vector potential $\mathbf{A}(\mathbf{r},t)$.

(b) Assume that the wave function is changed by a *non-constant* phase

$$\psi(\mathbf{r},t) \to \psi'(\mathbf{r},t) = e^{\mathrm{i}\alpha(\mathbf{r},t)}\psi(\mathbf{r},t).$$

Show that ψ' *does not* satisfy the original Schrödinger equation with the original vector potential $\mathbf{A}(\mathbf{r},t)$.

(c) Show that $\psi'(\mathbf{r},t)$ *does* satisfy the original Schrödinger equation but with a new vector potential

$$\mathbf{A}'(\mathbf{r},t) = \mathbf{A}(\mathbf{r},t) + \boldsymbol{\chi}(\mathbf{r},t),$$

where $\boldsymbol{\chi}$ is a vector field. How is the gauge term $\boldsymbol{\chi}(\mathbf{r},t)$ related to the phase term $\alpha(\mathbf{r},t)$?
How does the scalar potential change:

$$\phi(\mathbf{r},t) \to \phi'(\mathbf{r},t) = \phi(\mathbf{r},t) + ?$$

3

FORMAL QUANTUM MECHANICS AND THE PATH INTEGRAL

When we consider classical mechanics, we start from Newton's laws and derive the behaviour of moving bodies subject to forces from these laws. This is a nice approach as we always like to see a structured presentation of the world surrounding us. We should however not forget that people have thought for thousands of years about motion and forces before Newton's compact formulation of the underlying principles was found. It is not justified to pretend that physics only consists of understanding and predicting phenomena from a limited set of laws. The 'dirty' process of walking in the dark and trying to find a comprehensive formulation for the phenomena under consideration is an essential part of physics.

This also holds for quantum mechanics, although it was developed in a substantially shorter amount of time than classical mechanics. In fact, quantum mechanics started at the beginning of the twentieth century, and its formulation was more or less complete around 1930.

The previous chapter contained a brief review of quantum mechanics at a level where you should feel comfortable. Now we make a step partly into new material by considering quantum mechanics from a formal viewpoint. Part of this material can be found in Griffiths' book (chapter 3), in particular for section 3.1, where we introduce quantum mechanics by formulating the postulates on which the quantum theory is based. In sections 3.2, 3.3 and 3.4, we establish the link between the classical mechanics and quantum mechanics via Poisson brackets and via the path integral.

3.1 THE POSTULATES OF QUANTUM MECHANICS

Quantum theory can be formulated in terms of a set of *postulates* which however do not have a canonised form similar to Newton's laws: most books have their own version of these postulates and even their number varies.

We now present a particular formulation of these postulates.

1. The state of a physical system at any time t is given by the *wave function* of the system at that time. This wave function is an element of the Hilbert space of the system. The evolution of the system in time is determined by the Schrödinger equation:

   $$\mathrm{i}\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = \hat{H}\left|\psi(t)\right\rangle.$$

   Here $\hat{H}$ is an Hermitian operator, called the *Hamiltonian*.

2. Any physical quantity Q is represented by an Hermitian operator $\hat{Q}$.
   When we perform a measurement of the quantity Q, we will always find one of the

eigenvalues of the operator $\hat{Q}$. For a system in the state $|\psi(t)\rangle$, the probability of finding a particular eigenvalue λ_i, with an associated eigenvector $|\phi_i\rangle$ of $\hat{Q}$ is given by

$$P_i = \frac{|\langle\phi_i|\psi(t)\rangle|^2}{\langle\psi(t)|\psi(t)\rangle\langle\phi_i|\phi_i\rangle}.$$

Immediately after the measurement, the system will find itself in the state $|\phi_i\rangle$ corresponding to the value λ_i which was found in the measurement of λ_i.

Several remarks can be made.

1. The wave function contains the *maximum amount of information* we can have about the system. In practice, we often do not know the wave function of the system.

2. Note that $\hat{Q}$ being Hermitian implies that the eigenstates $|\phi_i\rangle$ always form a *basis* of the Hilbert space of the system under consideration. Thus the state $|\psi(t)\rangle$ of the system before the measurement can always be written in the form

$$|\psi(t)\rangle = \sum_i c_i(t)\,|\phi_i\rangle.$$

   The probability to find in a measurement the values λ_i is therefore given by

$$P_i = \frac{|c_i|^2}{\sum_j |c_j|^2}$$

   where we have omitted the time-dependence of the c_i. For a normalised state $|\psi(t)\rangle$ it holds that, if the eigenvectors $|\phi_i\rangle$ are orthonormal:

$$\sum_i |c_i|^2 = 1.$$

   In that case

$$P_i = |c_i|^2.$$

3. So far we have suggested in our notation that the eigenvalues and eigenvectors form a *discrete* set. In reality, not only discrete, but also continuous spectra are possible. In those cases, the sums are replaced by integrals.

4. In understanding quantum mechanics, it helps to make a clear distinction between the formalism which describes the evolution of the wave function (the Schrödinger equation, postulate 1) versus the *interpretation scheme*. We see that the wave function contains the information we need to predict the outcome of measurements, using the measurement postulate (number 2).

It now seems that we have arrived at a formulation of quantum mechanics which is similar to that of classical mechanics: a limited set of laws (prescriptions) from which everything can be derived, provided we know the form of the Hamiltonian (this is analogous to the situation in classical mechanics, where Newton's laws do not tell us what the form of the forces is).

However there is an important difference: the classical laws of motion can be understood by using our everyday experience so that we have some intuition for their meaning and content. In quantum mechanics, however, our laws are formulated as mathematical statements concerning objects (vectors and operators) for which we do not have a natural intuition. This is the reason why quantum mechanics is so difficult in the beginning (although its mathematical structure as such is rather simple). You should not despair when quantum mechanics seems difficult: many people find it difficult, and the workings of the measurement process are still the object of intensive debate. Sometimes you must switch your intuition off and use the rules of linear algebra to solve problems.

Above, we have mentioned that quantum mechanics does not prescribe the form of the Hamiltonian. In fact, although the Schrödinger equation, quite unlike the classical equation of motion, is a linear equation, which allows us to make ample use of linear algebra, the structure of quantum mechanics is richer than that of classical mechanics because in principle *any* type of Hilbert space can occur in Nature. In classical mechanics, the space of all possible states of an N-body system is a $6N$ dimensional space (we have $3N$ space and $3N$ momentum coordinates). We may extend this with 2 angles per particle if the particles carry a magnetic or electric dipole moment. In quantum mechanics, wave functions can be part of infinite-dimensional spaces (like the wave functions of a particle moving along a one-dimensional axis) but they can also lie in a finite-dimensional space (for example in the case of *spin*, which has no classical analogue).

3.2 RELATION WITH CLASSICAL MECHANICS

In order to see whether we can guess the structure of the Hamiltonian for systems which have a classical analogue, we consider the time evolution of a physical quantity Q. We assume that Q does not depend on time *explicitly*. However, the expectation value of Q may vary in time due to the change of the wave function in the course of time. For normalised wave functions,

$$\frac{d}{dt}\left(\langle\psi(t)|\hat{Q}|\psi(t)\rangle\right) = \left(\frac{\partial}{\partial t}\langle\psi(t)|\right)\hat{Q}\left|\psi(t)\right\rangle + \langle\psi(t)|\,\hat{Q}\left(\frac{\partial}{\partial t}|\psi(t)\rangle\right).$$

Using the time-dependent Schrödinger equation and its Hermitian conjugate,

$$-\mathrm{i}\hbar\frac{\partial}{\partial t}\langle\psi(t)| = \langle\psi(t)|\,\hat{H},$$

(note the minus sign on the left-hand side which results from the Hermitian conjugate) we obtain

$$\mathrm{i}\hbar\frac{d}{dt}\left(\langle\psi(t)|\hat{Q}|\psi(t)\rangle\right) = \langle\psi(t)|\hat{Q}\hat{H}-\hat{H}\hat{Q}|\psi(t)\rangle = \langle\psi(t)|\left[\hat{Q},\hat{H}\right]|\psi(t)\rangle,$$

where $\left[\hat{Q},\hat{H}\right]$ is the commutator. We see that the time derivative of $\hat{Q}$ is related to the commutator between $\hat{Q}$ and $\hat{H}$. This should wake you up or ring a bell. Perhaps you have seen a derivation of the time derivative of any physical quantity Q in classical mechanics, based on the Hamilton equations of classical mechanics. These equations read

$$\dot{p}_j = -\frac{\partial H}{\partial q_j}; \qquad \dot{q}_j = \frac{\partial H}{\partial p_j}.$$

Note that H is expressed as a function of the generalised coordinates q_j and p_j.

As a side note, we provide a different form of this equation. Introducing the two-dimensional coordinates $z_j = (p_j, q_j)$, we may write the Hamilton equations in the form

$$\frac{\partial z_j}{\partial t} = J\nabla_j H,$$

where $\nabla_j = (\partial/\partial p_j, \partial/\partial q_j)$ and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The time derivative of $Q(p_j, q_j)$ is given by

$$\frac{dQ}{dt} = \sum_j\left(\frac{\partial Q}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial Q}{\partial q_j}\frac{\partial H}{\partial p_j}\right) \equiv \{Q,H\}.$$

We see that this equation is very similar to that obtained above for the time derivative of the expectation value of the operator $\hat{Q}$! The differences consist of replacing the Poisson bracket by the commutator and adding a factor $-\mathrm{i}\hbar$. Note that the commutator and the Poisson bracket both are anti-symmetric under exchange of the two quantities involved:

$$[A,B] = -[B,A]; \qquad \{A,B\} = -\{B,A\}.$$

It seems that classical and quantum mechanics are not that different after all. Could this perhaps be a guide to formulate quantum mechanics for systems for which we have already a classical version? This turns out to be the case. At this stage, we summarise the result by writing down the time-evolution for $\hat{Q}$ and its classical version:

$$\mathrm{i}\hbar\frac{d\langle\hat{Q}\rangle}{dt} = \left\langle\left[\hat{Q},\hat{H}\right]\right\rangle \text{(Quantum)};$$

$$\frac{dQ}{dt} = \left\{Q,H\right\} \qquad \text{(Classical)},$$

which shows the striking similarity between the two formulations. These equations are both called *Liouville equation* – the first is however usually denoted as *quantum Liouville equation*.

As an example, we start by considering a one-dimensional system for which the relevant classical observables are the position x and the momentum p. Classically, we have

$$\left\{x,p\right\} = \frac{\partial x}{\partial x}\frac{\partial p}{\partial p} - \frac{\partial x}{\partial p}\frac{\partial p}{\partial x} = 1.$$

The second term in the expression vanished because x and p are to be considered as *independent* coordinates. From this, and from the factor $\mathrm{i}\hbar$ occurring in the quantum evolution equation above, we may guess the quantum version of this Poisson relation:

$$\left[x,p\right] = \mathrm{i}\hbar$$

which should look familiar (if it does not, please review the second year quantum mechanics course). It seems that our recipe for making quantum mechanics out of classical mechanics makes sense! Therefore we can now state the following rule:

> If the Hamiltonian of some classical system is known, we can use the same form in quantum mechanics, taking into account the fact that the coordinates q_j and p_j become Hermitian *operators* and that their commutator relations are:

$$[q_j,q_k] = 0; \quad [p_j,p_k] = 0; \quad [q_j,p_k] = \mathrm{i}\hbar\delta_{jk}.$$

You can verify these extended commutation relations easily by working out the corresponding classical Poisson brackets.

In the second year, you have learned that

$$\hat{p} = \frac{\hbar}{\mathrm{i}}\frac{d}{dx}.$$

What about this relation? It was not mentioned here so far. The striking message is that this relation *can be derived from the commutation relation*. In order to show this, we must discuss another object you might have missed too in the foregoing discussion: the wave function written in the form $\psi(\mathbf{r})$ (for a particle in 3D). It is important to study the relation between this and the state $|\psi\rangle$. Consider a vector $\mathbf{a}$ in two dimensions. This vector can be represented by two numbers a_1 and a_2, which are the components of the vector $\mathbf{a}$. However, the actual values of the components depend on how we have chosen our basis vectors. The vector $\mathbf{a}$ is an arrow in a two dimensional space. In that space, $\mathbf{a}$ has a particular length and a particular orientation. By changing the basis vectors, we do not change the *object* **a**, but we *do* change the numbers a_1 and a_2.

In the case of the Hilbert space of a one-dimensional particle, we can use as basis vectors the states in which the particle is localised at a particular position x. We call these states $|x\rangle$. They are eigenvectors of the position operator $\hat{x}$ with eigenvalue x:

$$\hat{x}\,|x\rangle = x\,|x\rangle\,.$$

The states $|x\rangle$ are properly normalised:

$$\langle x|x'\rangle = \delta(x-x'),$$

where $\delta(x-x')$ is the Dirac delta function. We now can define $\psi(x)$:

$$\psi(x) = \langle x|\psi\rangle\,,$$

that is, $\psi(x)$ are the 'components' of the 'vector' $|\psi\rangle$ *with respect to the basis* $|x\rangle$. For three dimensions, we have a wave function which is expressed with respect to the basis $|\mathbf{r}\rangle$.

In order to derive the representation of the momentum operator, $\hat{p} = \frac{\hbar}{\mathrm{i}}\frac{d}{dx}$, we first calculate the matrix element of the commutator:

$$\langle x|[\hat{x},\hat{p}]|x'\rangle = \langle x|\hat{x}\hat{p}-\hat{p}\hat{x}|x'\rangle = (x-x')\,\langle x|\hat{p}|x'\rangle\,.$$

The last expression is obtained by having $\hat{x}$ in the first term act on the bra-vector $\langle x|$ on its left, and on the ket $|x'\rangle$ on the right in the second term.

On the other hand, using the commutation relation, we know that

$$\langle x|[\hat{x},\hat{p}]|x'\rangle = \mathrm{i}\hbar\,\langle x|x'\rangle\,.$$

This is an even function of $x-x'$, as interchanging x and x' does not change the matrix element on the right-hand side. Since this function is equal to $(x-x')\,\langle x|\hat{p}|x'\rangle$, we know that $\langle x|\hat{p}|x'\rangle$ must be an odd function of $x-x'$.

Now we evaluate the matrix element $\langle x|\hat{p}|\psi\rangle$. We recall from linear algebra that, since $|x\rangle$ are the eigenstates of an Hermitian operator, they form a complete set, that is:

$$\mathbb{1} = \int |x\rangle\,\langle x|\;dx,$$

where $\mathbb{1}$ is the unit operator. Then we can write

$$\langle x|\hat{p}|\psi\rangle = \int \langle x|\hat{p}|x'\rangle\langle x'|\psi\rangle\;dx'.$$

Now we perform a Taylor expansion around x in order to rewrite $\langle x'|\psi\rangle$:

$$\langle x'|\psi\rangle = \langle x|\psi\rangle + (x'-x)\frac{d}{dx}\langle x|\psi\rangle + \frac{(x'-x)^2}{2!}\frac{d^2}{dx^2}\langle x|\psi\rangle + \cdots$$

We then obtain

$$\langle x|\hat{p}|\psi\rangle = \int \langle x|p|x'\rangle\left[\langle x|\psi\rangle + (x'-x)\frac{d}{dx}\langle x|\psi\rangle + \frac{(x'-x)^2}{2!}\frac{d^2}{dx^2}\langle x|\psi\rangle + \cdots\right]\,dx'.$$

The first term in the square brackets gives a zero after integration, as it is multiplied by $\langle x|\hat{p}|x'\rangle$, which was an odd function of $x-x'$. The second term gives

$$\int \langle x|\hat{p}|x'\rangle\,(x'-x)\frac{d}{dx}\langle x|\psi\rangle\;dx' = -\mathrm{i}\hbar\frac{d}{dx}\langle x|\psi\rangle\,,$$

where we have used the relation

$$\langle x|\hat{p}|x'\rangle (x'-x) = -\mathrm{i}\hbar\delta(x'-x).$$

We use the same relation for the second term. But then we obtain a term of the form

$$(x'-x)\delta(x'-x)$$

in the integral over dx'. This obviously yields a zero. The same holds for all higher order terms, so we are left with

$$\langle x|\hat{p}|\psi\rangle = \frac{\hbar}{i}\frac{d}{dx}\langle x|\psi\rangle,$$

which is the required result.

Having obtained this, we can analyse the form of the eigenstates of the momentum operator:

$$\hat{p}\,|p\rangle = p\,|p\rangle.$$

The states $|p\rangle$ can be represented in the basis $\langle x|$; the components then are $\langle x|p\rangle$. We can find the form of these functions by using the eigenvalue equation and the representation of the momentum operator as a derivative:

$$\begin{aligned} \langle x|\hat{p}|p\rangle &= p\langle x|p\rangle \qquad \text{and} \\ \langle x|\hat{p}|p\rangle &= \frac{\hbar}{\mathrm{i}}\frac{d}{dx}\langle x|p\rangle. \end{aligned}$$

The first of these equation expresses the fact that $|p\rangle$ is an eigenstate of the operator $\hat{p}$, and the second one follows directly from the fact that the momentum operator acts as a derivative in the x-representation. Combining these two we obtain a simple differential equation

$$\frac{\hbar}{\mathrm{i}}\frac{d}{dx}\langle x|p\rangle = p\langle x|p\rangle,$$

with a normalised solution:

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{\mathrm{i}px/\hbar}.$$

This allows us to find any state ψ in the momentum representation, that is, the representation in which we use the states $|p\rangle$ as basis states:

$$\psi(p) = \langle p|\psi\rangle = \int \langle p|x\rangle\langle x|\psi\rangle\, dx = \frac{1}{2\pi\hbar}\int e^{-\mathrm{i}px/\hbar}\psi(x)\, dx.$$

This is nothing but the Fourier transform, the specific form of which we already encountered at the end of the previous chapter. The analysis presented here for a one-dimensional particle can be generalised to three or more dimensions in a natural way.

3.3 THE PATH INTEGRAL: FROM CLASSICAL TO QUANTUM MECHANICS

The path integral is a very powerful concept for connecting classical and quantum mechanics. Moreover, this formulation renders the connection between quantum mechanics and statistical mechanics very explicit. We shall restrict ourselves here to a discussion of the path integral in quantum mechanics. The reader is advised to consult the excellent book by Feynman and Hibbs (*Quantum Mechanics and Path Integrals*, McGraw-Hill, 1965) for more details.

The path integral formulation can be derived from the following heuristics, based on the analogy between particles and waves:

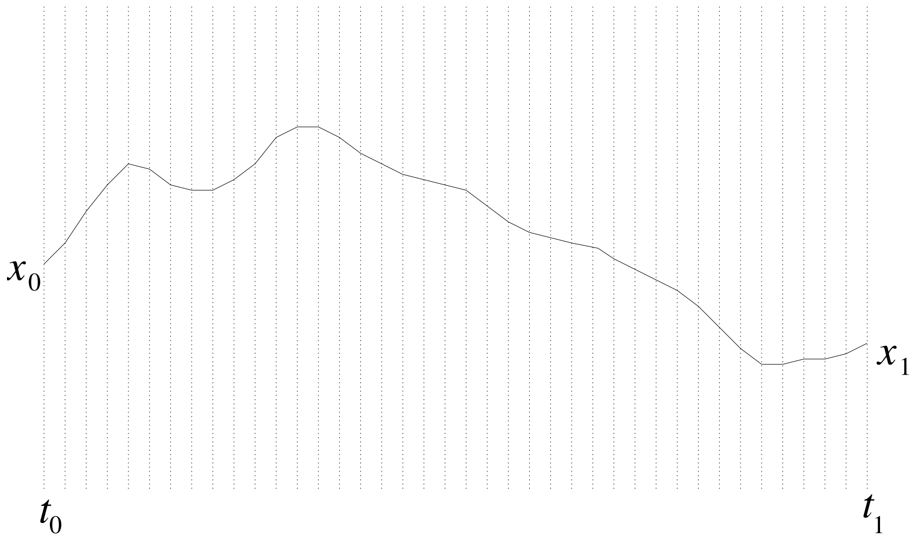


FIGURE 3.1: A possible path running from an initial position x_i at time t_i to a final position x_f at time t_f. The time is divided up into many identical slices.

- A point particle which moves with momentum $\mathbf{p}$ at energy E can also be viewed as a wave with a phase φ given by

  $$\varphi = \mathbf{k}\cdot\mathbf{r} - \omega t$$

  where $\mathbf{p} = \hbar\mathbf{k}$ and $E = \hbar\omega$.

- For a single path, these phases are additive, i.e., to find the phase of the entire path, the phases for different segments of the path should be added.

- The probablity to find a particle which at $t = t_0$ was at $\mathbf{r}_0$, at position $\mathbf{r}_1$ at time $t = t_1$, is given by the absolute square of the sum of the phase factors $\exp(\mathrm{i}\varphi)$ of all possible paths leading from $(\mathbf{r}_0, t_0)$ to $(\mathbf{r}_1, t_1)$:

  $$P(\mathbf{r}_0, t_0; \mathbf{r}_1, t_1) = \left| \sum_{\text{all paths}} e^{\mathrm{i}\varphi_{\text{path}}} \right|^2.$$

  This probability is defined up to a constant which can be fixed by normalisation (i.e. the term within the absolute bars must reduce to a delta-function in $\mathbf{r}_1 - \mathbf{r}_0$).

These heuristics are the analog of the Huygens principle in wave optics.

To analyse the consequences of these heuristics, we chop the time interval between t_0 and t_1 into many identical time slices (see Fig. 3.1) and consider one such slice. Within this slice we take the path to be linear. To simplify the analysis we consider one-dimensional motion. We first consider the contribution of kx to the phase difference. If the particle moves in a time Δt over a distance Δx, we know that its k-vector is given by

$$k = \frac{mv}{\hbar} = \frac{m\Delta x}{\hbar\Delta t}.$$

The phase change resulting from the displacement of the particle can therefore be given as

$$\Delta\varphi = k\Delta x = \frac{m\Delta x^2}{\hbar\Delta t}.$$

We still must add the contribution of $\omega\Delta t$ to the phase. The frequency ω is related to the energy; we have

$$\hbar\omega = \frac{p^2}{2m} + V(x) = \frac{\hbar^2 k^2}{2m} + V(x).$$

Neglecting the potential energy we obtain

$$\Delta\varphi = \frac{m\Delta x^2}{\hbar\Delta t} - \frac{\hbar^2 k^2}{2m\hbar}\Delta t = \frac{m\Delta x^2}{2\hbar\Delta t}.$$

The potential also enters through the $\omega\Delta t$ term, to give the result:

$$\Delta\varphi = \frac{m\Delta x^2}{2\hbar\Delta t} - \frac{V(x)}{\hbar}\Delta t.$$

For x occurring in the potential we may choose any value between the begin and end point – the most accurate result is obtained by substituting the average of the value at the beginning and at the end of the time interval.

If we now use the fact that phases are additive, we see that for the entire path the phases are given by

$$\varphi = \frac{1}{\hbar}\sum_j \left\{ \frac{m}{2}\left[\frac{x(t_{j+1}) - x(t_j)}{\Delta t}\right]^2 - \frac{V[x(t_j)] + V[x(t_{j+1})]}{2} \right\}\Delta t.$$

This is nothing but the discrete form of the *classical action* of the path! Taking the limit $\Delta t \to 0$ we obtain

$$\varphi = \frac{1}{\hbar}\int_{t_0}^{t_1}\left[\frac{m\dot{x}^2}{2} - V(x)\right]dt = \frac{1}{\hbar}\int_{t_0}^{t_1} L(x,\dot{x})\,dt.$$

We therefore conclude that the probability to go from $\mathbf{r}_0$ at time t_0 to $\mathbf{r}_1$ at time t_1 is given by

$$P(\mathbf{r}_0, t_0; \mathbf{r}_1, t_1) = \left| \mathcal{N} \sum_{\text{all paths}} \exp\left[\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t_1} L(x,\dot{x})\,dt\right] \right|^2$$

where $\mathcal{N}$ is the normalisation factor

$$\mathcal{N} = \sqrt{\frac{m}{2\pi\mathrm{i}\Delta t\hbar}}.$$

This now is the path integral formulation of quantum mechanics. Let us spend a moment to study this formulation. First note the large prefactor $1/\hbar$ in front of the exponent. If the phase factor varies when varying the path, this large prefactor will cause the exponential to vary wildly over the unit circle in the complex plane. The joint contribution to the probability will therefore become very small. If on the other hand there is a region in phase space (or 'path space') where the variation of the phase factor with the path is zero or very small, the phase factors will add up to a significant amount. Such regions are those where the action is stationary, that is, we recover the *classical* paths as those giving the major contribution to the phase factor. For $\hbar \to 0$ (the classical case), only the stationary paths remain, whereas for small $\hbar$, small fluctuations around these paths are allowed: these are the quantum fluctuations.

You may not yet recognise how this formulation is related to the Schrödinger equation. On the other hand, we may identify the expression within the absolute signs in the last expression for P with a matrix element of the time evolution operator since both have the same meaning:

$$\langle x_1|\hat{U}(t_1 - t_0)|x_0\rangle = \sum_{\text{all paths}} \mathcal{N}\exp\left[\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t_1} L(x,\dot{x})\,dt\right].$$

This form of the time evolution operator is sometimes called the *propagator*. Let us now evaluate this form of the time evolution operator acting for a small time interval Δt on the wave function $\psi(x,t)$:

$$\psi(x_1, t_1) = \int \langle x_1\left|\hat{U}(t_1 - t_0)\right| x_0\rangle\langle x_0|\psi\rangle\,dx_0 =$$
$$\mathcal{N}\int \mathrm{D}[x(t)]\int_{-\infty}^{\infty}\exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t_1}\left[m\frac{\dot{x}^2(t)}{2} - V[x(t)]\right]dt\right\}\psi(x_0, t_0)\,dx_0.$$

The notation $\int D[x(t)]$ indicates an integral over all possible paths from (x_0, t_0) to (x_1, t_1). We first approximate the integral over time in the same fashion as above, taking t_1 very close to t_0, and assuming a linear variation of $x(t)$ from x_0 to x_1:

$$\psi(x_1, t_1) = \mathcal{N} \int_{-\infty}^{\infty} \exp\left\{\frac{i}{\hbar}\left[m\frac{(x_1 - x_0)^2}{2\Delta t^2} - \frac{V(x_0) + V(x_1)}{2}\right]\Delta t\right\}\psi(x_0, t_0)\, dx_0.$$

A similar argument as used above to single out paths close to stationary ones can be used here to argue that the (imaginary) Gaussian factor will force x_0 to be very close to x_1. The allowed range for x_0 is

$$(x_1 - x_0)^2 \ll \frac{\hbar\Delta t}{m}.$$

As Δt is taken very small, we may expand the exponent with respect to the $V\Delta t$ term:

$$\psi(x_1, t_1) = \mathcal{N} \int_{-\infty}^{\infty} \exp\left[\frac{i}{\hbar} m\frac{(x_1 - x_0)^2}{2\Delta t}\right]\left[1 - \frac{i[V(x_0) + V(x_1)]}{2\hbar}\Delta t\right]\psi(x_0, t_0)\, dx_0.$$

As x_0 is close to x_1 we may approximate $\frac{V(x_0)+V(x_1)}{2\hbar}$ by $V(x_1)/\hbar$. We now change the integration variable from x_0 to $u = x_0 - x_1$:

$$\psi(x_1, t_1) = \mathcal{N} \int_{-\infty}^{\infty} \exp\left(\frac{i}{\hbar} m\frac{u^2}{2\Delta t}\right)[1 - i/\hbar V(x_1)\Delta t]\,\psi(x_1 + u, t_0)\, du.$$

As u must be small, we can expand $\psi(x)$ about x_1 and obtain

$$\psi(x_1, t_1) = \mathcal{N} \int_{-\infty}^{\infty} \exp\left(\frac{im}{\hbar}\frac{u^2}{2\Delta t}\right)\left[1 - \frac{i}{\hbar}V(x_1)\Delta t\right]\left[\psi(x_1, t_0) + u\frac{\partial}{\partial x}\psi(x_1, t_0) + \frac{u^2}{2}\frac{\partial^2}{\partial x^2}\psi(x_1, t_0)\right] du.$$

Note that the second term in the Taylor expansion of ψ leads to a vanishing integral as the integrand is an antisymmetric function of u. All in all, after evaluating the Gaussian integrals, we are left with

$$\psi(x_1, t_1) = \psi(x_1, t_0) - \frac{i\Delta t}{\hbar}V(x_1)\psi(x_1, t_0) + \frac{i\hbar\Delta t}{2m}\frac{\partial^2}{\partial x^2}\psi(x_1, t_0).$$

Using

$$\frac{\psi(x_1, t_1) - \psi(x_1, t_0)}{\Delta t} \approx \frac{\partial}{\partial t}\psi(x_1, t_1),$$

we obtain the time-dependent Schrödinger equation for a particle moving in one dimension:

$$i\hbar\frac{\partial}{\partial t}\psi(x, t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(x, t).$$

You may have found this derivation a bit involved. It certainly is not the easiest way to arrive at the Schrödinger equation, but it has two attractive features;

- Everything was derived from simple heuristics which were based on viewing a particle as a wave and allowing for interference between the waves;
- The formulation shows that the classical path is obtained from quantum mechanics when we let $\hbar \to 0$.

3.4 THE PATH INTEGRAL: FROM QUANTUM MECHANICS TO CLASSICAL MECHANICS

In the previous section we have considered how we can arrive from classical mechanics at the Schrödinger equation. This formalism can be generalised in the sense that for each system for which we can write down a Lagrangian, we have a way to find a quantum formulation in terms of the path integral. Whether a Schrödinger-like equation can be found is not sure: sometimes we run into problems which are beyond the scope of these notes. In this section we assume that we have a system described by some Hamiltonian and show that the time-evolution operator has the form of a path integral as found in the previous section.

The starting point is the time evolution operator, or propagator, which, for a time-independent Hamiltonian, takes the form

$$U(\mathbf{r}_\mathrm{f}, t_\mathrm{f}; \mathbf{r}_\mathrm{i}, t_\mathrm{i}) = \left\langle \mathbf{r}_\mathrm{f} \left| e^{-\frac{\mathrm{i}}{\hbar}(t_\mathrm{f}-t_\mathrm{i})\hat{H}} \right| \mathbf{r}_\mathrm{i} \right\rangle.$$

The matrix element is difficult to evaluate – the reason is that the Hamiltonian which, for a particle in one dimension, takes the form

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$$

is the sum of two *noncommuting* operators. Although it is possible to evaluate the exponents of the separate terms occurring in the Hamiltonian, the exponent of the sum involves an infinite series of increasingly complicated commutators. For any two noncommuting operators $\hat{A}$ and $\hat{B}$ we have

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}} e^{\hat{B}} e^{-1/2[\hat{A},\hat{B}]-1/12([\hat{A},[\hat{A},\hat{B}]]+[\hat{B},[\hat{B},\hat{A}]])+1/24[\hat{A},[\hat{B},[\hat{A},\hat{B}]]]+\ldots}$$

This is the so-called Campbell–Baker–Hausdorff (CBH) formula. The cumbersome commutators occurring on the right can only be neglected if the operators A and B are small in some sense. We can try to arrive at an expression involving small commutators by applying the time slicing procedure of the previous section:

$$e^{-\frac{\mathrm{i}}{\hbar}(t_\mathrm{f}-t_\mathrm{i})\hat{H}} = e^{-\frac{\mathrm{i}}{\hbar}\Delta t\hat{H}} e^{-\frac{\mathrm{i}}{\hbar}\Delta t\hat{H}} e^{-\frac{\mathrm{i}}{\hbar}\Delta t\hat{H}} \ldots$$

Note that no CBH commutators occur because $\Delta t\hat{H}$ commutes with itself.

Having this, we can rewrite the propagator as (we omit the hat from the operators)

$$U(x_\mathrm{f}, t_\mathrm{f}; x_\mathrm{i}, t_\mathrm{i}) = \int dx_1 \ldots dx_{N-1} \left\langle x_\mathrm{f} | e^{-\mathrm{i}\Delta t H/\hbar} | x_{N-1} \right\rangle \left\langle x_{N-1} | e^{-\mathrm{i}\Delta t H/\hbar} | x_{N-2} \right\rangle \cdots \left\langle x_1 | e^{-\mathrm{i}\Delta t H/\hbar} | x_\mathrm{i} \right\rangle.$$

Now that the operators occurring in the exponents can be made arbitrarily small by taking Δt very small, we can evaluate the matrix elements explicitly:

$$\left\langle x_j | e^{-\mathrm{i}\Delta t H} | x_{j+1} \right\rangle = \left\langle x_j | e^{-\frac{\mathrm{i}\Delta t}{\hbar}[p^2/(2m)+V(x)]} | x_{j+1} \right\rangle = e^{-\mathrm{i}\Delta t V(x_j)/\hbar} \left\langle x_j | e^{-\frac{\mathrm{i}\Delta t}{\hbar}p^2/(2m)} | x_{j+1} \right\rangle.$$

The last matrix element can be evaluated by inserting two unit operators formulated in terms of integrals over the complete sets $|p\rangle$:

$$\left\langle x | e^{-\frac{\mathrm{i}\Delta t}{\hbar}p^2/(2m)} | x' \right\rangle = \int\int \langle x|p\rangle \langle p| \, e^{-\frac{\mathrm{i}\Delta t}{\hbar}\hat{p}^2/(2m)} \, |p'\rangle \langle p'|x'\rangle \; dp\, dp'.$$

We have seen that $\langle x|p\rangle = \exp(\mathrm{i}px/\hbar)/\sqrt{2\pi\hbar}$. Realising that the operator $\exp\left[-\frac{\mathrm{i}\Delta t}{\hbar}\hat{p}^2/(2m)\right]$ is *diagonal* in p space, we find, after integrating over p:

$$\left\langle x | e^{-\frac{\mathrm{i}\Delta t}{\hbar}p^2/(2m)} | x' \right\rangle = \sqrt{\frac{m}{2\pi\hbar\Delta t}} \exp\left[\mathrm{i}m(x-x')^2/(2\Delta t\hbar)\right].$$

All in all we have

$$\left\langle x_j|e^{-\mathrm{i}\Delta tH/\hbar}|x_{j+1}\right\rangle = \sqrt{\frac{m}{2\pi\hbar\Delta t}}e^{-\mathrm{i}\Delta tV(x_j)/\hbar}\exp\left[m\mathrm{i}(x-x')^2/(2\Delta t\hbar)\right].$$

Note that we have evaluated matrix elements of operators. The result is expressed completely in terms of numbers, and we no longer have to bother about commutation relations. Collecting all terms together we obtain

$$U(x_\mathrm{f},t_\mathrm{f};x_\mathrm{i},t_\mathrm{i}) = \left(\frac{m}{2\pi\hbar\Delta t}\right)^{(N-1)/2}\int dx_1\dots dx_{N-1}\exp\left\{\frac{\mathrm{i}}{\hbar}\sum_{j=0}^{N}\left[\frac{m(x_{j+1}-x_j)^2}{2}-V(x_j)\right]\Delta t\right\}.$$

The expression in the exponent is the discrete form of the Lagrangian; the integral over all intermediate values x_j is the sum over all paths. We therefore have shown that the time evolution operator from x_i to x_f is equivalent to the sum of the phase factors of all possible paths from x_i to x_f.

In conclusion, we have seen that the idea that the probability to a find a particle starting off from x_0 at t_0 at x_1 at time t_1 is given by a sum of all the phase factors corresponding to all paths from the starting to the end point, gives us the quantum theory. On the other hand, once we have the Hamiltonian of a quantum theory, we can reformulate this as a path integral involving the corresponding Lagrangian.

3.5 SUMMARY

In this chapter, you have hopefully learned to appreciate the close relation between classical and quantum mechanics. In fact, the mathematical structure of both theories is more similar than you may have thought after your introductory courses on quantum mechanics. Let us list some of the relations we have found.

- Both classical and quantum mechanics allow us – in principle – to follow the evolution of a mechanical system by solving a first-order differential equation in the time. In both, the Hamiltonian is the agent governing this evolution.

  In classical mechanics, the time evolution equations read

  $$\frac{\partial p_j}{\partial t} = -\frac{\partial H}{\partial q_j};$$
  $$\frac{\partial q_j}{\partial t} = \frac{\partial H}{\partial p_j}.$$

  Writing $z_j = (p_j, q_j)$, we may write this in the form

  $$\frac{\partial z_j}{\partial t} = J\nabla_j H,$$

  where $\nabla_j = (\partial/\partial p_j, \partial/\partial q_j)$ and

  $$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

  In quantum mechanics, the Hamilton equations are replaced by the time-dependent Schrödinger equation:

  $$\mathrm{i}\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = \hat{H}\left|\psi(t)\right\rangle.$$

- In quantum mechanics, measurements disrupt the time evolution and bring stochastic elements into the theory.

- The Poisson bracket of classical mechanics is replaced by the commutator in quantum mechanics:
$$\{A,B\} \to -\mathrm{i}\hbar\left[\hat{A},\hat{B}\right].$$
- Th path integral allows us to work out the quantum mechanical time evolution as an infinite sum over paths, where each path is weighted by the factor
$$\exp\left(\mathrm{i}S_{\text{path}}/\hbar\right),$$
where S_{path} is the *classical action* of the path:
$$S_{\text{path}} = \int_{\text{path}} L(q_j,\dot{q}_j)\,dt,$$
with L the classical Lagrangian of the path.
- The smallness of $\hbar$ ensures that the paths with a value of the action deviating not more than $\mathcal{O}(\hbar)$ from the stationary value(s) of the action, yield a major contribution to the path integral (time evolution). For $\hbar \to 0$, only the stationary (i.e. the classical) path(s) survive(s).
- The time-dependent Schrödinger equation can be derived from the path integral formalism.

3.6 PROBLEMS

1. The product of the exponentials of two non-commuting operators X and Y can be written as
$$e^X e^Y = \exp(X+Y+[X,Y]/2+\ldots)$$
This is the so-called *Campbell Baker Hausdorff formula* (or CBH formula). The aim of this problem is to derive this formula.
   - (a) First, show that $\log(1+x) \approx x - x^2/2 + x^3/3 - \ldots$.
   - (b) Now expand the formula $\log(e^X e^Y)$ to second order in X and Y to derive the CBH formula.
   - (c) If you're brave, you may try to find the third order expansion to find the next term.
2. We consider a particle moving in one dimension. The state of the particle is $|\psi\rangle$.
   - (a) Show that
$$\langle x+a|\psi\rangle = \left\langle x\left|e^{\mathrm{i}a\hat{p}/\hbar}\right|\psi\right\rangle$$
by inserting a unit operator using completeness of the momentum basis $|p\rangle$.
   - (b) Demonstrate the same result by writing $\hat{p}$ as the operator $-\mathrm{i}\hbar d/dx$ and Taylor expanding ψ around x.
   - (c) Show that for a wave function in three dimensions:
$$\langle\mathbf{r}'|\psi\rangle = \left\langle\mathbf{r}|e^{\mathrm{i}\alpha\hat{L}_z/\hbar}|\psi\right\rangle,$$
where $\mathbf{r}'$ is the vector $\mathbf{r}$ rotated about an angle α around the z-axis.
   - (d) (The following two parts were already addressed in problem 9).
Show that the operator $\exp(\mathrm{i}\alpha\sigma_z/2)$ rotates a spin-1/2 state about an angle α around the z-axis.
   - (e) Give the rotation operator for the wave function describing a spin-1/2 particle.

3. The Lagrangian of a harmonic oscillator is $L = (m/2)(\dot{x}^2 - \omega^2 x^2)$. Show that the classical action is:

$$S_{\mathrm{cl}} = \frac{m\omega}{2\sin\omega T}\left[(x_{\mathrm{i}}^2 + x_{\mathrm{f}}^2)\cos\omega T - 2x_{\mathrm{i}}x_{\mathrm{f}}\right], \tag{3.1}$$

where $T = t_{\mathrm{f}} - t_{\mathrm{i}}$

Hint: the definition of the action is: $S = \int_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} L(\dot{x}, x, t)\, dt$. Make use of a classical trajectory of the harmonic oscillator: $x(t) = A\cos\omega t + B\sin\omega t$, which satisfies boundary conditions: $x(t_{\mathrm{i}}) = x_{\mathrm{i}}$ and $x(t_{\mathrm{f}}) = x_{\mathrm{f}}$.

4. The path integral formalism expresses the probability to move from a point x_{i} at time t_{i} to a point x_{f} at time t_{f} in terms of a sum over all paths:

$$K(x_{\mathrm{f}}, t_{\mathrm{f}}; x_{\mathrm{i}}, t_{\mathrm{i}}) = \sum_{\mathrm{Paths}} \exp(\mathrm{i}S_{\mathrm{path}}/\hbar).$$

We write the path as a sum of the *classical* path $x_{\mathrm{class}}(t)$ and a fluctuation:

$$x(t) = x_{\mathrm{class}}(t) + \delta x(t),$$

where $\delta x(t_{\mathrm{i}}) = \delta x(t_{\mathrm{f}}) = 0$, since the positions at the beginning and at the end of the path are fixed.

We consider a free particle.

(a) Give the action for the classical path for a free particle (that is, $V(x) = 0$ for all x).

(b) Show that the action for a general path can be written as

$$S = S[x_{\mathrm{class}}] + \frac{m}{2}\int \delta\dot{x}^2(t)\, dt.$$

Also show that

$$K(x_{\mathrm{f}}, t_{\mathrm{f}}; x_{\mathrm{i}}, t_{\mathrm{i}}) = K(0, t_{\mathrm{f}}; 0, t_{\mathrm{i}}) \exp\left[\frac{\mathrm{i}}{\hbar} S(x_{\mathrm{class}})\right].$$

(c) The properly normalised form of the discretised path integral is given by

$$K(0, t_{\mathrm{f}}; 0, t_{\mathrm{i}}) = \lim_{N\to\infty} \left(\frac{m}{2\pi\mathrm{i}\epsilon\hbar}\right)^{(N+1)/2} \int \prod_{n=1}^{N+1} dx_n \exp\left[\frac{\mathrm{i}}{\hbar}\epsilon \sum_{n=1}^{N} L(x, \dot{x})\right].$$

In the exponent, $\dot{x}$ means $(x_n - x_{n-1})/\epsilon$; $\epsilon = (t_{\mathrm{f}} - t_{\mathrm{i}})/(N+1)$. Furthermore, $x_0 \equiv x_{\mathrm{i}}$ and $x_{N+1} \equiv x_{\mathrm{f}}$.
The integration over all possible paths δx can be performed in the discretised path integral:

$$I_N = \int \prod_{i=1}^{N+1} dx_n \exp\left(\mathrm{i}\frac{\epsilon}{\hbar} \sum_{n=1}^{N} \frac{m\dot{x}^2}{2}\right).$$

It can be shown (using e.g. induction) that

$$I_N = \frac{1}{\sqrt{N+1}} \left(\frac{2\pi\mathrm{i}\epsilon\hbar}{m}\right)^{N/2}.$$

If you're in for a challenge, you may try to prove this, but that is not required here. Show, using this result, that

$$K(x_{\mathrm{f}}, t_{\mathrm{f}}; x_{\mathrm{i}}, t_{\mathrm{i}}) = \sqrt{\frac{m\hbar}{2\pi\mathrm{i}(t_{\mathrm{f}} - t_{\mathrm{i}})}} \exp\left(\frac{\mathrm{i}m}{2\hbar}\frac{(x_{\mathrm{f}} - x_{\mathrm{i}})^2}{(t_{\mathrm{f}} - t_{\mathrm{i}})}\right).$$

(d) Find this propagator directly, using

$$K(x_{\mathrm{f}}, t_{\mathrm{f}}; x_{\mathrm{i}}, t_{\mathrm{i}}) = \left\langle x_{\mathrm{f}} | e^{-\mathrm{i}(t_{\mathrm{f}} - t_{\mathrm{i}})p^2/2m\hbar} | x_{\mathrm{i}} \right\rangle.$$

Hint: insert two unit operators in this expression, formulated as integrals over p.

4

THE VARIATIONAL METHOD FOR THE SCHRÖDINGER EQUATION

4.1 VARIATIONAL CALCULUS

Quantum systems are governed by the Schrödinger equation. In particular, the solutions to the stationary form of this equation determine many physical properties of the system at hand. The stationary Schrödinger equation can be solved analytically in a very restricted number of cases – examples include the free particle, the harmonic oscillator and the hydrogen atom. In most cases we must resort to computers to determine the solutions to this equation. It is of course possible to integrate the Schrödinger equation using discretisation methods but in realistic electronic structure calculations for example we would need a huge number of grid points, leading to important computer time and memory requirements. On the other hand, the variational method enables us to solve the Schrödinger equation much more efficiently in many cases. In this chapter we introduce the variational method for solving the Schrödinger equation.

In the variational method, the possible solutions are restricted to a subspace of the Hilbert space, and in this subspace we seek the best possible solution (below we shall define what is to be understood by the 'best' solution). To see how this works, we first show that the stationary Schrödinger equation can be derived by a stationarity condition of the functional

$$E[\psi] = \frac{\int dX \psi^*(X) H \psi(X)}{\int dX \psi^*(X)\psi(X)} = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}, \tag{4.1}$$

which is recognised as the expectation value of the energy for a stationary state ψ (in order to keep the analysis general we are not specific about the form of the generalised coordinate X – it may include the space and spin coordinates of a collection of particles). The stationary states of this energy functional are defined by postulating that if such a state is changed by an arbitrary $\delta\psi$, the corresponding change in E vanishes to first order. Formally, this means that

$$\frac{d}{d\lambda}\left(E\left[\psi + \lambda\delta\psi\right] - E[\psi]\right)\bigg|_{\lambda=0} \equiv 0 \tag{4.2}$$

for *all* normalised vectors $\delta\psi$. Defining

$$\begin{aligned} P &= \langle\psi|H|\psi\rangle \qquad \text{and} \\ Q &= \langle\psi|\psi\rangle, \end{aligned} \tag{4.3}$$

we can write the change δE in the energy to first order in $\delta\psi$ as

$$\begin{aligned}\delta E &= \frac{\langle \psi+\delta\psi|H|\psi+\delta\psi\rangle}{\langle \psi+\delta\psi|\psi+\delta\psi\rangle} - \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \\ &\approx \frac{\langle\delta\psi|H|\psi\rangle - \frac{P}{Q}\langle\delta\psi|\psi\rangle}{Q} + \frac{\langle\psi|H|\delta\psi\rangle - \frac{P}{Q}\langle\psi|\delta\psi\rangle}{Q}.\end{aligned} \tag{4.4}$$

As this should vanish for an *arbitrary* but small change in ψ, we find, using $E = P/Q$:

$$H\psi = E\psi, \tag{4.5}$$

together with the Hermitian conjugate of this equation, which is equivalent.

In variational calculus, stationary states of the energy functional are found *within a subspace* of the Hilbert space. An important example is linear variational calculus, in which the subspace is spanned by a finite set of basis vectors $|\chi_p\rangle$, $p = 1,\dots,N$, that we take to be orthonormal at first, that is,

$$\langle\chi_p|\chi_q\rangle = \delta_{pq}, \tag{4.6}$$

where δ_{pq} is the Kronecker delta-function which is 0 unless $p = q$, in which case it is 1.

For a state

$$|\psi\rangle = \sum_{p=1}^{N} C_p\,|\chi_p\rangle, \tag{4.7}$$

the energy functional is given by

$$E = \frac{\sum_{p,q=1}^{N} C_p^* C_q H_{pq}}{\sum_{p,q=1}^{N} C_p^* C_q \delta_{pq}}, \tag{4.8}$$

with

$$H_{pq} = \langle\chi_p|H|\chi_q\rangle. \tag{4.9}$$

The stationary states follow from the condition that the derivative of this functional with respect to the C_p vanishes, which leads to

$$\sum_{q=1}^{N}\left(H_{pq} - E\delta_{pq}\right)C_q = 0 \quad \text{for} \quad p = 1,\dots,N. \tag{4.10}$$

Equation (4.10) is an eigenvalue problem which can be written in matrix notation:

$$\mathbf{HC} = E\mathbf{C}. \tag{4.11}$$

This is the Schrödinger equation formulated for a finite, orthonormal basis.

Linear parametrisations are often used because the resulting method is simple, allowing for numerical matrix diagonalisation techniques to be used. The lowest eigenvalue of (4.11) is always higher than or equal to the ground state energy of Eq. (4.5), as the ground state is the minimal value assumed by the energy functional over the full Hilbert space. If we restrict ourselves to a part of this space, then the minimum value of the energy functional must always be higher than or equal to the ground state of the full Hilbert space. Adding more basis functions to our set, the subspace becomes larger, and consequently the minimum of the energy functional will decrease (or stay the same). For the specific case of linear variational calculus, this result can be generalised to stationary states at higher energies: the higher eigenvalues are always higher than the equivalent solution to the full problem, but approximate the latter better with increasing basis set size. The formal statement of this is the Hylleraas-Undin-MacDonald theorem (see for example Springer Handbook of Atomic, Molecular, and Optical Physics, Volume 1, Gordon Drake (ed.), 2006). The behaviour of the spectrum found by solving (4.11) with increasing basis size is depicted in Figure 4.1.

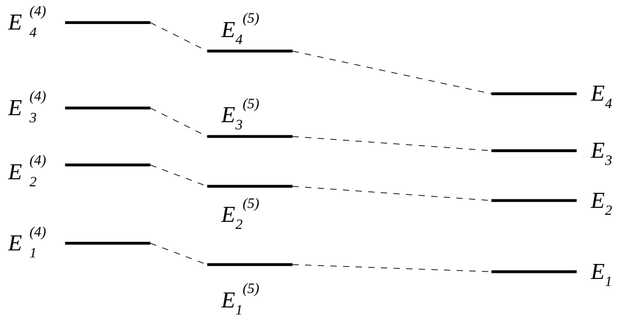


FIGURE 4.1: The behaviour of the spectrum of Eq. (4.11) with increasing basis set size in linear variational calculus. The upper index is the number of states in the basis set, and the lower index labels the spectral levels.

We now describe how to proceed when the basis consists of non-orthonormal basis functions, as is often the case in practical calculations. In that case, we must reformulate (4.11), taking care of the fact that the *overlap matrix* **S**, whose elements S_{pq} are given by

$$S_{pq} = \langle \chi_p | \chi_q \rangle \tag{4.12}$$

is not the unit matrix. This means that in Eq. (4.8) the matrix elements δ_{pq} of the unit matrix, occurring in the denominator, have to be replaced by S_{pq}, and we obtain

$$\mathbf{HC} = E\mathbf{SC}. \tag{4.13}$$

This looks like an ordinary eigenvalue equation, the only difference being the matrix **S** in the right hand side. It is called a *generalised eigenvalue equation* and there exist computer programs for solving such a problem.

4.2 LINEAR VARIATIONAL CALCULATIONS

In this section, we describe two quantum mechanical problems that can be analyzed numerically with a linear variational calculation. In both cases, a generalised matrix eigenvalue problem (4.13) must be solved, which can easily be done using a program like MATLAB.

4.2.1 THE INFINITELY DEEP POTENTIAL WELL

The potential well with infinite barriers is given by

$$V(x) = \begin{cases} \infty & \text{for} \quad |x| > |a|, \\ 0 & \text{for} \quad |x| \le |a|. \end{cases} \tag{4.14}$$

It forces the wave function to vanish at the boundaries of the well ($x = \pm a$). The exact solution for this problem is known and treated in every textbook on quantum mechanics (see for example Griffiths). Here we discuss a linear variational approach to be compared with the exact solution. We take $a = 1$ and use natural units such that $\hbar^2/2m = 1$.

As basis functions we take simple polynomials that vanish on the boundaries of the well:

$$\psi_n(x) = x^n(x-1)(x+1),\ n = 0, 1, 2, \ldots \tag{4.15}$$

The reason for choosing this particular form of basis functions is that the relevant matrix elements can easily be calculated analytically. We start with the matrix elements of the overlap matrix, defined by

$$S_{mn} = \langle \psi_n | \psi_m \rangle = \int_{-1}^{1} \psi_n(x)\psi_m(x)dx. \tag{4.16}$$

There is no complex conjugate with the ψ_n in the integral because we use real basis functions. Working out the integral gives

$$S_{mn} = \frac{2}{n+m+5} - \frac{4}{n+m+3} + \frac{2}{n+m+1} \tag{4.17}$$

TABLE 4.1: Energy levels of the infinitely deep potential well. The first four columns show the variational energy levels for various numbers of basis states N. The last column shows the exact values. The exact levels are approached from above as in Figure 4.1.

$N=5$	$N=8$	$N=12$	$N=16$	Exact
2.4674	2.4674	2.4674	2.4674	2.4674
9.8754	9.8696	9.8696	9.8696	9.8696
22.2934	22.2074	22.2066	22.2066	22.2066
50.1246	39.4892	39.4784	39.4784	39.4784
87.7392	63.6045	61.6862	61.6850	61.6850

for $n+m$ even; otherwise $S_{mn}=0$.

We can also calculate the Hamilton matrix elements – you can check that they are given by:

$$\begin{aligned} H_{mn} = \langle \psi_n | p^2 | \psi_m \rangle &= \int_{-1}^{1} \psi_n(x) \left(-\frac{d^2}{dx^2} \right) \psi_m(x) dx \\ &= -8 \left[\frac{1-m-n-2mn}{(m+n+3)(m+n+1)(m+n-1)} \right] \end{aligned} \tag{4.18}$$

for $m+n$ even, else $H_{mn}=0$.

The exact solutions are given by

$$\psi_n(x) = \begin{cases} \cos(k_n x) & n \text{ odd} \\ \sin(k_n x) & n \text{ even and positive} \end{cases} \tag{4.19}$$

with $k_n = n\pi/2$, $n = 1, 2, \ldots$, with corresponding energies

$$E_n = k_n^2 = \frac{n^2\pi^2}{4}. \tag{4.20}$$

For each eigenvector $\mathbf{C}$, the function $\sum_{p=1}^{N} C_p \chi_p(x)$ should approximate an eigenfunction (4.19). The variational levels are shown in table 4.1, together with the analytical results.

4.2.2 VARIATIONAL CALCULATION FOR THE HYDROGEN ATOM

As we shall see in further on in this course, one of the main problems of electronic structure calculations is the treatment of electron–electron interactions. Here we develop a scheme for solving the Schrödinger equation for an electron in a hydrogen atom for which the many-electron problem does not arise, so that a direct variational treatment of the problem is possible which can be compared to the analytic solution.

The electronic Schrödinger equation for the hydrogen atom reads:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r} \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{4.21}$$

where the second term in the square brackets is the Coulomb attraction potential of the nucleus. The mass m is the reduced mass of the proton–electron system which is approximately equal to the electron mass. The ground state is found at energy

$$E = -\frac{m}{\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \approx -13.6058 \, e\text{V} \tag{4.22}$$

and the wave function is given by

$$\psi(\mathbf{r}) = \frac{1}{a_0^{3/2}\sqrt{\pi}} e^{-r/a_0}, \tag{4.23}$$

in which a_0 is the Bohr radius,

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} \approx 0.529\,18\,\text{Å}. \tag{4.24}$$

When performing a calculation for such an equation, it is convenient to use units such that equations take on a simple form, involving only coefficients of order 1. Standard units in electronic structure physics are so-called *atomic units*: the unit of distance is the Bohr radius a_0, masses are expressed in the electron mass m_e and the charge is measured in unit charges (e). The energy is finally given in 'Hartrees' (E_H), given by $m_e c^2 \alpha^2$ (α is the fine-structure constant and m_e is the electron mass) which is roughly equal to 27.212 eV. In these units, the Schrödinger equation for the hydrogen atom assumes the following simple form:

$$\left[-\frac{1}{2}\nabla^2 - \frac{1}{r}\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{4.25}$$

We try to approximate the ground state energy and wave function of the hydrogen atom in a linear variational procedure. We use *Gaussian basis functions*. For the ground state, we only need angular momentum $l = 0$ functions (s-functions) – they have the form:

$$\chi_p(r) = e^{-\alpha_p r^2} \tag{4.26}$$

centred on the nucleus (which is thus placed at the origin). We have to specify the values of the exponents α_p. A large α_p defines a basis function which is concentrated near the nucleus, whereas small α_p characterises a function with a long tail. Optimal values for the exponents α_p can be found by solving the *non-linear* variational problem including the linear coefficients C_p *and* the exponents α_p. Several numerical methods for solving such non-linear optimisation problem exist and the solutions can be found in textbooks or documentation of quantum chemical software packages.

We shall use known, fixed values of the exponents:

$$\begin{aligned} \alpha_1 &= 13.007\,73 \\ \alpha_2 &= 1.962\,079 \\ \alpha_3 &= 0.444\,529 \\ \alpha_4 &= 0.121\,949\,2, \end{aligned} \tag{4.27}$$

but relax the values of the coefficients C_p. The wave function therefore has the form

$$\langle \mathbf{r}|\psi\rangle = \sum_{p=1}^{4} C_p e^{-\alpha_p r^2}$$

with the α_p listed above. We now discuss how to find the best values of the linear coefficients C_p. To this end, we need the overlap and Hamiltonian matrix. The advantage of using Gaussian basis functions is that analytic expressions for these matrices can be found. In particular, it is not so difficult to show that the elements of the overlap matrix **S**, the kinetic energy matrix **T** and the Coulomb matrix **A** are given by:

$$\begin{aligned} S_{pq} &= \int d^3r e^{-\alpha_p r^2} e^{-\alpha_q r^2} = \left(\frac{\pi}{\alpha_p + \alpha_q}\right)^{3/2}, \\ T_{pq} &= -\frac{1}{2}\int d^3r e^{-\alpha_p r^2} \nabla^2 e^{-\alpha_q r^2} = 3\frac{\alpha_p \alpha_q \pi^{3/2}}{(\alpha_p + \alpha_q)^{5/2}}, \\ A_{pq} &= -\int d^3r e^{-\alpha_p r^2} \frac{1}{r} e^{-\alpha_q r^2} = -\frac{2\pi}{\alpha_p + \alpha_q}. \end{aligned} \tag{4.28}$$

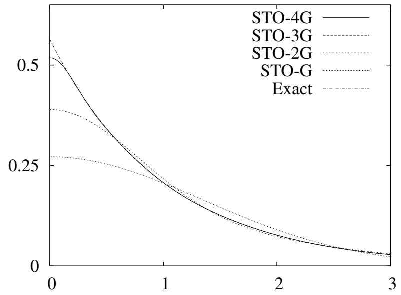


FIGURE 4.2: The best fit of a Gaussian basis set consisting of 1, 2, 3 and 4 basis functions, to the exact Slater solution $e^{-r/a}$ of the hydrogen atom. The mnemonic STO-nG denotes a fit using n Gaussian functions to a Slater Type Orbital (STO).

These expressions can be put into a computer program which solves the generalised eigenvalue problem. The resulting ground state energy is -0.499278 Hartree, which is amazingly close to the exact value of $-1/2$ Hartree, which is -13.6058 eV. We conclude that four Gaussian functions can be linearly combined into a form which is surprisingly close to the exact ground state wave function which is known to have the so-called *Slater-type* form $\exp(-r/a)$ rather than a Gaussian! This is shown in figure 4.2.

4.2.3 EXPLOITING SYMMETRY

We have seen that the solution of a stationary quantum problem using linear variational calculus, in the end, boils down to solving a (generalised) matrix eigenvalue problem. Finding the eigenvalues (and eigenvectors) of an $N \times N$ matrix, or solving a generalised eigenvalue problem, requires a number of floating point operations in the computer proportional to N^3. This means that if we double the size of the basis set used in the variational analysis, the computer time goes up by a factor of 8. As it turns out, we are often interested in problems having some symmetry. We shall now briefly sketch how this can be used to significantly reduce the computer time for variational calculations.

In subsection 4.2.1, we considered a problem having a very simple symmetry: replacing x by $-x$ does not change the potential, and therefore the Hamiltonian is insensitive to this transformation. Let us denote the operation $x \to -x$ by $\mathscr{R}$. Because flipping the sign of x twice leaves the space invariant, we have

$$\mathscr{R}^2 = \mathbb{1},$$

where $\mathbb{1}$ is as usual the identity operator. Let us consider the eigenvalues λ of this operator. From $\mathscr{R}^2 = \mathbb{1}$ we have that $\lambda^2 = 1$. Therefore, $\lambda = \pm 1$. Furthermore, $\mathscr{R}$ commutes with the Hamiltonian:

$$\mathscr{R}H - H\mathscr{R} = 0,$$

since the Hamiltonian is not affected by $\mathscr{R}$. We know (or should know!) that if an operator commutes with H we can always find eigenvalues which are eigenvalues of H *and* of that operator. This means that we can divide the eigenfunctions of H into two classes: one of symmetric eigenfunctions (symmetric meaning having eigenvalue $\lambda = +1$ when acting on it with $\mathscr{R}$) and one of antisymmetric eigenfunctions ($\lambda = -1$).

Now suppose we construct our variational basis set such that it can be divided into two classes, that of symmetric and that of anti-symmetric basis functions. Let us calculate the inner product of a symmetric and an anti-symmetric eigenfunction. Using antisymmetry of the product of the two may immediately convince you that this vanishes. To illustrate the more general procedure, we consider two eigenfunctions, $|\phi_1\rangle$ and $|\phi_2\rangle$ with *different* eigenvalues λ_1 and λ_2 for the symmetry operation $\mathscr{R}$. Then we can write

$$\langle \phi_1 | \mathscr{R} | \phi_2 \rangle = \lambda_1 \langle \phi_1 | \phi_2 \rangle = \lambda_2 \langle \phi_1 | \phi_2 \rangle,$$

where we first let $\mathscr{R}$ act on the left, and then on the right function. The fact that $\lambda_1 \neq \lambda_2$ leads to the well-known theorem saying that two eigenvectors of a Hermitian operator with *different* eigenvalues, are orthogonal (if the eigenvalues are the same, the wave functions are either identical or they can be chosen orthogonal).

The key result is that a similar conclusion can be drawn for the expectation value of the Hamiltonian.

$$\langle \phi_1 | \mathscr{R} H | \phi_2 \rangle = \lambda_1 \langle \phi_1 | H | \phi_2 \rangle = \langle \phi_1 | H \mathscr{R} | \phi_2 \rangle = \lambda_2 \langle \phi_1 | H | \phi_2 \rangle,$$

which, as $\lambda_1 \neq \lambda_2$ directly gives

$$\langle \phi_1 | H | \phi_2 \rangle = 0.$$

For an orthonormal basis set, we see that, if we would order the basis functions in our set with respect to their eigenvalue of $\mathscr{R}$, the Hamiltonian becomes *block-diagonal.* For our simple reflection-symmetric example, denoting the symmetric basis functions by $|\phi_{ps}\rangle$ and the antisymmetric ones by $|\phi_{pa}\rangle$ where p runs from 1 to $M = N/2$, we have

$$H = \left(\begin{array}{ccc|ccc} H_{1s,1s} & \ldots & H_{1s,Ms} & 0\ldots & 0 & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ H_{Ms,1s} & \ldots & H_{Ms,Ms} & 0\ldots & 0 & \\ \hline 0 & \ldots & 0 & H_{1a,1a} & \ldots & H_{1a,Ma} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & H_{Ma,1a} & \ldots & H_{Ma,Ma} \end{array} \right)$$

We can diagonalise the two blocks on the diagonal independently. This takes $2(N/2)^3$ steps (up to a multiplicative constant), which is 4 times less than N^3 (up to the same constant) required to diagonalise the full Hamiltonian! If there are additional symmetries, they can be used to reduce the work required even further.

If the basis is non-orthogonal, it still holds that basis functions having different eigenvalues under the symmetry-operator $\mathscr{R}$ are orthogonal and that the matrix elements of the Hamiltonian between them vanishes. This means that the Hamiltonian matrix *and* the overlap matrix have the same block-diagonal structure. Therefore, the respective generalised eigenvalues for the blocks can be dealt with independently of each other and we achieve the same speed-up.

What we have touched upon is an example of the application of group theory in physics, which is an important topic on its own.

4.3 EXAMPLES OF NON-LINEAR VARIATIONAL CALCULUS

Linear variational calculus leads to a (generalised) matrix eigenvalue problem to be solved. Variational methods form a much wider class, including trial functions which depend non-linearly on the variational parameters. Numerically this is quite complicated to solve, but several analytic non-linear variational calculations exist which give quite good results even if one only one or two variational parameters are used. A nice example is the hydrogen atom, which we try to solve using a variational wave function ('trial function') $|\psi_{\mathrm{T}}\rangle$ of the form

$$\langle \mathbf{r} | \psi_{\mathrm{T}} \rangle \propto e^{-r/a}.$$

You may note that this is the form of the *exact* ground state of the hydrogen atom. However, to illustrate the variational method, we first *relax* the value of the parameter a and then vary it to minimise the expectation value of the energy. We should then find the exact ground state wave function and energy.

The Schrödinger equation was already given in atomic units in the previous section:

$$\left[-\frac{1}{2}\nabla^2 - \frac{1}{r} \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{4.29}$$

It is useful to first normalise the trial wave function:

$$4\pi \int r^2 e^{-2r/a} dr = \pi a^3,$$

so that we have

$$\langle \mathbf{r}|\psi_{\mathrm{T}}\rangle = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}.$$

Its is now easy to calculate the expectation value of the kinetic energy. Using the fact that for a function $\psi(\mathbf{r})$ in 3D which only depends on r,

$$\nabla^2 \psi(\mathbf{r}) = \frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{d}{dr}\psi(r)\right),$$

we have

$$\nabla^2 \psi_{\mathrm{T}}(\mathbf{r}) = \frac{1}{\sqrt{\pi a^3}}\left(\frac{1}{a^2} - \frac{2}{ar}\right) e^{-r/a},$$

we find, after some calculation that

$$-\left\langle \psi_{\mathrm{T}} \left| \frac{1}{2}\nabla^2 \right| \psi_{\mathrm{T}} \right\rangle = \frac{1}{2\pi a^3}\int \left(-\frac{1}{a^2} + \frac{2}{ar}\right) e^{-2r/a} 4\pi r^2 dr = \frac{1}{2a^2}.$$

For the potential energy, we find

$$-\frac{1}{\pi a^3}\int e^{-2r/a} 4\pi r dr = -\frac{1}{a}.$$

Therefore, the expectation value of the energy for the trial wave function is given by

$$E_{\mathrm{T}} = \frac{1}{2a^2} - \frac{1}{a}.$$

The minimum of this expression is found at $a = 1$ and yields an energy of $E_{\mathrm{T}} = -1/2$ in units of 27.212 eV, which is the correct ground state energy of -13.6058 eV.

Now we turn to a more complicated problem: the helium atom, which (when the nucleus is considered not to move because of its large mass) is described by the Hamiltonian

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|}.$$

In atomic units, this becomes, with $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$:

$$H = -\frac{1}{2}\frac{d^2}{dr_1^2} - \frac{1}{2}\frac{d^2}{dr_2^2} - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}.$$

For the trial wave function we use the form

$$\langle \mathbf{r}_1, \mathbf{r}_2|\psi\rangle = \exp\left[-2(r_1 + r_2)/a\right].$$

This function is chosen such that it yields the ground state of two *noninteracting* electrons moving in the field of the helium nucleus (the nuclear charge $Z = 2$ leads to a scaling of 2 in the exponent). So the trial wave function is simply the successful trial wave function (in the sense that it contains the exact solution) for the independent-electron case. In particular, when an electron approaches the nucleus, its behaviour is properly described by this wave function as the electron-nucleus interaction largely dominates the electron-electron interaction in that case.

We have taken the wave function to be symmetric in the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$. The two electrons should however form an antisymmetric wave function as they are fermions. The antisymmetry is taken care of by the spin wave function

$$\frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right).$$

Later we shall go much deeper into the structure of many-body wave functions – here we just mention that, as the Hamiltonian does not contain any spin dependence, we can forget about the spin part of the wave function and can safely assume that the ground state wave function is symmetric in $\mathbf{r}_1$ and $\mathbf{r}_2$.

We first must normalise this solution. This can be done for $\mathbf{r}_1$ and $\mathbf{r}_2$ independently and we obtain

$$4\pi \int r^2 dr e^{-4r/a} = \frac{\pi a^3}{8},$$

so that the normalised solution reads

$$\langle \mathbf{r}_1 \mathbf{r}_2 | \psi \rangle = \frac{8}{\pi a^3} \exp\left[-2(r_1 + r_2)/a\right]. \tag{4.30}$$

In order to find the expectation value of the Hamiltonian for this wave function, it is convenient to write it in the form

$$H = H_1 + H_2 + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|},$$

where

$$H_i = \frac{p_i^2}{2m} - \frac{2}{r_i}, \qquad i = 1,2.$$

This is the Hamiltonian for an electron moving in the helium potential.

We calculate the kinetic energy following the method used for the hydrogen atom. The result is (for the two electrons together)

$$E_\mathrm{K} = \frac{4}{a^2}.$$

A quick way to arrive at this result is by taking the kinetic energy $1/(2a^2)$ for the hydrogen atom, replacing $a \to a/2$ and multiplying by 2 because we now have two electrons. The potential energy due to the attraction between the nucleus and the electrons is found to be

$$E_\text{n-e} = \frac{-8}{a},$$

as can be verified by taking the hydrogen result $1/a$, replacing $a \to a/2$ and multiplying by 2 because the nuclear charge is twice as large as for the hydrogen atom and again by 2 because we now have two electrons.

Now we must add to this the contribution from the electron repulsion:

$$\int \frac{64 e^{-4(r_1+r_2)/a}}{\pi^2 a^6} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3 r_1 \, d^3 r_2.$$

If we fix $\mathbf{r}_1$, we can evaluate the integral over $\mathbf{r}_2$. Choosing the z-axis to be the direction of $\mathbf{r}_1$, we have (without the prefactor):

$$\int e^{-4(r_1+r_2)/a} \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta}} 2\pi \sin\theta d\theta r_2^2 dr_2.$$

We first evaluate the integral over θ. Choosing $\cos\theta = u$, this is of the form $\int du/\sqrt{p - qu}$ and we are left with

$$2\pi \int e^{-4(r_1+r_2)/a} \frac{1}{r_1 r_2} \left(\sqrt{r_1^2 + r_2^2 + 2r_1 r_2} - \sqrt{r_1^2 + r_2^2 - 2r_1 r_2} \right) r_2^2 dr_2 =$$

$$2\pi \int e^{-4(r_1+r_2)/a} \left[\frac{1}{r_1 r_2} (r_1 + r_2 - |r_1 - r_2|) \right] r_2^2 dr_2 \tag{4.31}$$

The term in square brackets equals $2/\max(r_1, r_2)$ where the function $\max(x, y)$ returns the largest of the two numbers x and y. Therefore, we need to split the integral over r_2 into a part running from 0 to r_1 and a part running from r_1 to ∞:

$$4\pi e^{-4r_1/a}\left[\frac{1}{r_1}\int_0^{r_1} e^{-4r_2/a} r_2^2\, dr_2 + \int_{r_1}^{\infty} e^{-4r_2/a} r_2\, dr_2\right] =$$
$$\frac{\pi a^3}{8r_1} e^{-4r_1/a}\left[1 - e^{-4r_1/a} - 2\frac{r_1}{a} e^{-4r_1/a}\right].$$

Now it remains to multiply this by the normalisation factor $64/(\pi^2 a^6)$ and by $\exp(-4r_1/a)$, and then integrate over $4\pi r_1^2 dr_1$ (there is no dependence on the two angular variables for $\mathbf{r}_1$). All the integrals are straightforward and the final result is

$$\left\langle \psi_\mathrm{T} \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \psi_\mathrm{T} \right\rangle = \frac{5}{4a}$$

in units of Hartree = 27.212 eV. So the total energy is given by

$$E_T = \frac{4}{a^2} - \frac{-8}{a} + \frac{5}{4a} = \frac{4}{a^2} - \frac{27}{4a}.$$

Taking $a = 1$, i.e. assuming that both electrons are in the ground state of the atom with the electron-electron interaction switched off, yields an energy of

$$4 - 27/4 = -11/4 \text{ Hartree} = 74.833 \text{ eV}.$$

The experimental value is −79 eV, so although our result is not extremely bad, it is not impressively accurate either.

Now let us relax a and find the minimum of the trial energy. This is found at $a = 32/27 =$ 1.1815 Bohr radii. The energy is then found at −2.8477 Hartree= 77.49 eV.

4.4 SUMMARY

In this chapter, we have studied a successful method for approximating the ground state solution of a complicated quantum mechanical problem: variational calculus. This method simply consists of finding the minimum of the expectation value of the Hamiltonian in a *restricted* space, where the term 'restricted' indicates that the space is a subset of the full Hilbert space of the quantum problem. It is trivial to see that this variational solution yields an energy equal to or larger than the exact value.

The space in which we search for the solution usually is a set which we can represent by parametrised trial solutions, $\psi_\mathrm{T} = \psi_\mathrm{T}(\alpha_1, \alpha_2, \dots, \alpha_N)$. The parameters α_j enter in a possibly non-linear way into the wave function. Solving non-linear minimisation problems is non-trivial, but numerical routines are available for this. Analytical non-linear variational problems are usually restricted to one or two variational parameters with respect to which the expectation value of the Hamiltonian is to be minimised.

A special case is the one in which the trial wave functions depend *linearly* on the variational parameters, which we now call C_p. That is, we can write the trial wave function as

$$|\psi_\mathrm{T}\rangle = \sum_{p=1}^{N} C_p |\chi_p\rangle.$$

In that case, the minimisation problem of the energy reduces to solving the *generalised eigenvalue problem*:

$$\mathbf{HC} = E\mathbf{SC},$$

where **C** stands for the *vector* with elements C_p and the matrices **H** and **S** have elements

$$H_{pq} = \langle \chi_p | H | \chi_q \rangle;$$
$$S_{pq} = \langle \chi_p | \chi_q \rangle.$$

For the special case of an orthonormal basis χ_p, $S_{pq} = \delta_{pq}$, the Dirac delta-function, and the generalised eigenvalue problem reduces to an ordinary eigenvalue problem:

$$\mathbf{HC} = E\mathbf{C}.$$

4.5 PROBLEMS

1. (a) We consider the ground state of an electron in the hydrogen atom. Approximate the ground-state wave function by

   $$\langle \mathbf{r}|\psi\rangle = e^{-\alpha r^2}$$

   and find an upper bound to the ground-state energy using variational calculus.

   (b) Approximate the ground state of the one-dimensional harmonic oscillator using a trial wave function

   $$\langle x|\psi\rangle = \begin{cases} \alpha^2 - x^2 & \text{for } |x| \le \alpha, \\ 0 & \text{elsewhere.} \end{cases}$$

2. The attractive potential felt by an electron in an atom is sometimes taken to be the screened potential

   $$V(r) = \frac{Ae^2}{r} e^{-r/\xi}.$$

   Consider trial wave functions of the form

   $$\psi(\mathbf{r}) = \exp(-r/\rho).$$

   Minimise the variational energy for this wave function.

3. Consider a one-dimensional potential

   $$V(x) = \begin{cases} \lambda x & \text{for } 0 < x < a; \\ \infty & \text{for } x \le 0 \text{ and } x \ge a. \end{cases}$$

   Find the ground state for this Hamiltonian using variational calculus. Take a second order polynomial as a trial function. The polynomial should obviously satisfy the correct boundary conditions at $x = 0$ and $x = a$.

4. We consider a particle moving in one dimension in a 'quartic potential'. The Hamiltonian is given as

   $$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{b}{4}x^4,$$

   with b some positive constant. Now take a trial wave function of the form

   $$\langle x|\psi\rangle = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-x^2/(2\sigma^2)}.$$

   Calculate the variational energy and minimise it to obtain the variational ground state energy.

5. Consider a system of two particles with equal masses m and momenta $\mathbf{p}_1$ and $\mathbf{p}_2$, interacting via a potential $V(r_{12})$ where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$.

   (a) Write the Hamiltonian $\hat{H}$ of the system in terms of the momenta

   $$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2; \tag{4.32}$$

   $$\mathbf{p} = \frac{\mathbf{p}_1 - \mathbf{p}_2}{2} \tag{4.33}$$

(all momentum vectors are operators) and of r_{12}. Also use the total mass $M + m_1 + m_2$ and the reduced mass $\mu = m_1 m_2/(m_1 + m_2)$.
Show that $\hat{H}$ can be written in the form

$$\hat{H} = \frac{P^2}{2M} + \hat{H}_{12}.$$

(b) We denote by $E^{(2)}$ the ground state energy of $\hat{H}_{12}$. Give the expression for $E^{(2)}$ when $V(r) = -b^2/r$ and for $V(r) = \kappa r^2/2$.

(c) Consider a system of three particles of equal mass m with pairwise interactions:

$$V = V(r_{12}) + V(r_{23}) + V(r_{31}).$$

Show that

$$3(p_1^2 + p_2^2 + p_3^2) = (\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3)^2 + (\mathbf{p}_1 - \mathbf{p}_2)^2 + (\mathbf{p}_2 - \mathbf{p}_3)^2 + (\mathbf{p}_3 - \mathbf{p}_1)^2,$$

and that the Hamiltonian of the three-body Hamiltonian can be written as

$$H^{(3)} = \frac{\mathbf{P}^2}{6m} + H^{(3)}_{\text{rel}},$$

where

$$H^{(3)}_{\text{rel}} = H_{12} + H_{23} + H_{31},$$
$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3$$

and where H_{ij} contains a kinetic part with a reduced mass μ'. Express μ' in terms of m.

(d) Check whether the H_{ij} commute amongst each other. What can you say about the energy spectrum if this were the case?

(e) Show that the three-body ground state energy $E^{(3)}$ is related to the ground state energy $E^{(2)}$ of the two-body problem described by H_{ij} by the inequality

$$E^{(3)} \geq 3E^{(2)}.$$

Note that the latter depends on μ' calculated in (c).
Hint: write $E^{(3)}$ as the expectation value of $H^{(3)}$ for the ground state $|\Omega\rangle$. Then use the fact that $\left\langle \Omega \middle| H^{(2)} \middle| \Omega \right\rangle \geq E^{(2)}$.

(f) Give the lower bounds for the case where $V(r) = -b/r^2$ and where $V(r) = \kappa r^2/2$. How does this lower bound for the first case compare with the numerical result

$$E^{(3)} \approx -1.067 m b^4/\hbar^2?$$

6. *Linear variational calculus for the Cooper-pair box.*

The Cooper-pair box consists of a small piece of superconductor (called the island) coupled to a larger piece (called the reservoir) via a Josephson junction: a thin, insulating layer. Superconductivity will be addressed later in this course, and Josephson junctions not at all. Detailed knowledge however is not needed to do this exercise. In a superconductor, *Cooper pairs*, pairs of electrons with opposite spin and momentum, form a condensate which requires a finite energy to create excitations. The Josephson junction is thin enough for allowing Cooper pairs to tunnel through it, and this tunneling generates a particular coupling between the two superconducting volumes connected by the junction.

We use the charge basis $|n\rangle$, where n denotes the number of Cooper pairs that have tunneled through the junction, and therefore the number of charges that have moved to the small superconductor (see figure below). This island has an electrostatic capacity C, which means that every electron contributes an amount $E_C = e^2/2C$ to the energy. The Hamiltonian reads

$$H = 4E_C \sum_n (n - n_g)^2 |n\rangle\langle n| - \frac{E_J}{2} \sum_n (|n+1\rangle\langle n| + |n\rangle\langle n+1|),$$

with $n = \ldots -1, 0, 1, \ldots$ and n_g is some charge offset. The first term reflects the electrostatic charging, and the second reflects tunneling (also called *hopping*). The parameter E_J is the so-called *Josephson energy*. Crucially, note that n_g is a constant (not an operator) under the control of the experimentalist. Note that the charge basis is orthonormal, $\langle n|m\rangle = \delta_{n,m}$.

We will now use linear variational calcultus to estimate the ground and first-excited state energies and wave functions. Caution: please do not attempt to solve this problem analytically. Rather, use your favorite mathematical software: Mathematica, Matlab, anything you like! Please print out your code.

(a) Write the matrix H_{qp} in the charge basis. The matrix is of course infinite dimensional, show just a subset of it, revealing its basic structure.

(b) Write the matrix S_{qp} also in the charge basis.

(c) Consider $E_C = 10E_J$. Restrict your trial functions to the subspace $n = -N, \ldots, 0, \ldots + N$. Plot the ground and first-excited state energies as a function of n_g in the range $n_g \in [-1, 1]$. Do this for $N = 1$, $N = 2$ and $N = 3$.

(d) Repeat for $E_J = 10E_C$.

(e) For $E_J = 10E_C$ and $n_g = 0.5$, plot the ground and first-excited state energies as a function of N in the range $N \in [1, 5]$. Confirm McDonald's theorem: show that the energies decrease monotonically as you increase N.

(f) For $E_J = 10E_C$, plot the ground and first-excited state wave functions in the charge basis. That is, plot the coefficient c_n in the expansion

$$|\Psi\rangle = \sum_{n=-N}^{N} c_n |n\rangle$$

Do this for $n_g = 0$ and $n_g = 0.5$. What choice of N would you say is accurate enough?

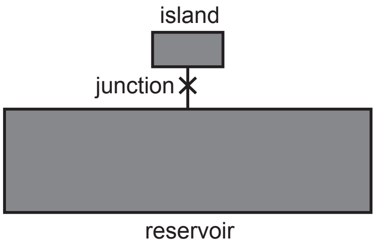


FIGURE 4.3: The Cooper-pair box

5

THE WKB APPROXIMATION

5.1 INTRODUCTION

The WKB approximation provides a simple way of obtaining reasonable values for different aspects of the solution to a quantum problem in one dimension. In this chapter we briefly describe how this approximation works. You are advised to read the material up in other books containing more extensive treatments. A good reference is chapter 8 of Griffiths' book.

5.2 THE WKB ANSATZ

When electrons move in one dimension through some constant potential, their state can be written as

$$\langle x|\psi\rangle = A\exp(\mathrm{i}kx)$$

with a wave vector

$$k=\sqrt{\frac{2m(E-V)}{\hbar^2}}.$$

The wave runs towards the right (the positive x-direction), as can be seen by looking at the full time-dependent wave function:

$$\langle x|\psi\rangle = A\exp\left[\mathrm{i}\left(kx-\omega t\right)\right],$$

with $\omega = E/\hbar$. A wave running towards the left would then be described by the stationary wave function

$$\langle x|\psi\rangle = A\exp(-\mathrm{i}kx).$$

What happens if the wave is incident on a non-constant potential, such as a step, a well or barrier? Your first guess would probably be: part of the wave is reflected, and another part is transmitted. That is correct as demonstrated by figure 5.1(a), which shows the time evolution of a Gaussian wave packet incident on a rectangular potential barrier. However, the rectangular barrier is special in the sense that at it varies *abruptly* from 0 to a finite value, and then as abruptly back again.

Now let us consider a wave packet incident on a *smooth* barrier, described by a Gaussian shape – see figure 5.1(b). The shape of the well is chosen such that its rise from 0 to its maximum value extends over a few wavelengths of the packet (of course, the packet contains a continuum of wavelengths, but they are close to an average value). We see that the packet is just transmitted and not reflected! So we conclude that

- A wave incident on a potential which varies substantially over length scales smaller than a wave length, is reflected and transmitted.
- If the potential is *smooth*, i.e. it varies substantially on length scales (much) larger than a wave length, an incident wave is not reflected, just transmitted.

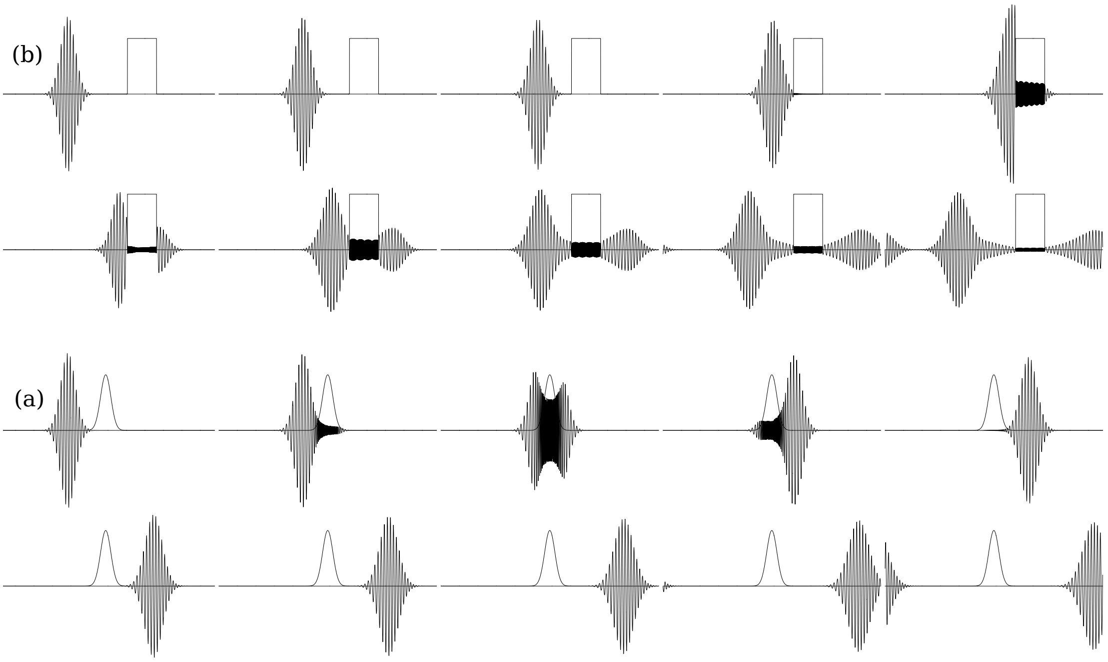


FIGURE 5.1: (a) Reflection and transmission of a wave packet by a rectangular potential well. Periodic boundary conditions are used, so the transmitted part disappearing off the right edge reappears at the left edge. (b) Transmission of a wave packet by a Gaussian potential well. Reflection is virtually absent. Periodic boundary conditions are applied.

This notion is the basis of the Wentzel-Kramers-Brillouin, or WKB approximation. In this section, we introduce this approximation in a hand-waving way.

We assume that $E > V(x)$, so that running waves are possible. Any wave function is then characterised by an amplitude and a phase:

$$\langle x|\psi\rangle = A(x)e^{\mathrm{i}\varphi(x)},$$

where both A and φ are real. We focus on the phase φ. Moving over a distance Δx, the phase of the wave function changes by $k\Delta x$ where

$$k = \sqrt{\frac{2m}{\hbar^2}(E-V)}$$

where we assume that the potential V is constant on the interval Δx. This assumption is valid provided Δx is taken sufficiently small. Now consider the following interval where the potential has a different value V'. Then the total phase picked up by the wave function on the two intervals is

$$k\Delta x + k'\Delta x = (k+k')\Delta x,$$

where obviously $k' = \sqrt{(2m/\hbar^2)(E-V')}$. Note that this only holds if we *only allow right-moving waves, i.e. we neglect left movers, arising from scattering.* This assumption is only valid when the potential is smooth. Dividing up a larger interval into N small ones Δx, a wave function which is real at x_0 will have the form

$$\langle x|\psi\rangle \approx A(x)\exp\left(\mathrm{i}\sum_{i=0}^{N-1} k(x_i)\Delta x\right); \quad x_i = x_0 + i\Delta x; \quad k(x) = \sqrt{\frac{2m}{\hbar^2}(E-V(x))}.$$

Taking the limit $N \to \infty$, we may write this as

$$\langle x|\psi\rangle \approx A(x)\exp\left(\mathrm{i}\int_{x_0}^{x} k(x')dx'\right).$$

This fixes the phase factor. In order to find the amplitude $A(x)$, we use a conservation principle. Remember that the absolute square of the wave function represents the probability to find the particle in a particular quantum state, which in our context means a particle located at a specific position. For our wave function, the density is given by $|A(x)|^2$. The speed at which the particles travel at position x is (approximately) given by

$$v = \frac{\hbar k(x)}{m},$$

therefore the flux is given as

$$j(x) = \frac{\hbar k(x)}{m} |A(x)|^2.$$

The crucial step is now that, since this wave function represents a stationary flow, the flux *must be constant through space*. This implies that

$$A(x) \propto \frac{1}{\sqrt{k(x)}}.$$

This then leads to the WKB wave function Ansatz:

$$\langle x|\psi\rangle \approx \frac{1}{\sqrt{k(x)}} \exp\left(\mathrm{i}\int_{x_0}^{x} k(x')dx'\right); \qquad k(x) = \sqrt{\frac{2m}{\hbar^2}(E - V(x))},$$

which fixes the form of the wave function up to a suitable normalization constant.

5.3 THE WKB ANSATZ II

In the previous section we have considered the WKB wave function which was based on the notion that a wave inciding on a smooth potential is not scattered back, so that the phase is simply the sum of the infinitesimal phases picked up on many short intervals. In this section we start from the form

$$\langle x|\psi\rangle = A(x)\exp[\mathrm{i}\varphi(x)], \tag{5.1}$$

for the solution of the one-dimensional Schrödinger equation, where *both* A and φ are *real* functions. Inserting this into the Schrödinger equation, we derive the form for φ and V under the assumption that V varies slowly. Note that any one-dimensional wave function can be written in this form, so we have not imposed any restriction here.

Let us assume that in the interval where we describe the wave function by this form, $E > V(x)$. Defining

$$k(x) = \sqrt{\frac{2m(E - V(x))}{\hbar^2}},$$

we can write the Schrödinger equation in the form (using $\psi(x) = \langle x|\psi\rangle$):

$$\psi''(x) = -k^2(x)\psi(x),$$

where ψ'' is the second derivative of ψ with respect to x.

Now we put our wave function (5.1) into this equation. This leads straightforwardly to

$$A'' + 2\mathrm{i}A'\varphi' + \mathrm{i}A\varphi'' - A\varphi'^2 = -k^2 A.$$

Note that this yields two real equations, one following from the imaginary, and one from the real part. They are, respectively:

$$2A'\varphi' + A\varphi'' = 0$$

and

$$A'' - A\varphi'^2 = -k^2 A.$$

The first equation can be recast into the form

$$\frac{1}{A}\frac{d}{dx}\left(A^2\varphi'\right) = 0,$$

from which we immediately have

$$A = \frac{\text{Const}}{\sqrt{\varphi'}}. \tag{5.2}$$

For a slowly varying potential, we anticipate that $\varphi(x) \approx k(x)x$, so that $k(x) \approx \varphi'(x)$. From the relation (5.2) between A and φ, we can then infer that $A''/A \propto [V'/(E-V)]^2, V''/(E-V)$. We now neglect these terms, i.e. we set $A'' \ll k^2 A$. We then have

$$\varphi'(x) = \pm k(x),$$

and we write the solution in the form

$$\varphi(x) = \pm \int_{x_0}^{x} k(x')dx'.$$

This is the WKB approximation. Note that the approximation essentially consists of neglecting the term A'' in comparison with the term $\varphi'^2 \approx k^2$, i.e. we neglect variations in the potential with a wavelength much larger the wavelength $k(x)$ of the particles due to their kinetic energy.

Remember that in this analysis we have required that $E > V$ and that the variation of V over a few wavelengths is small in comparison with $E - V(x)$. If $V(x)$ gets close to E, the wavelength grows larger and larger, and this condition no longer holds. We shall see in section 5.6 how this regime can be dealt with. Another regime is that where the energy E is substantially smaller than $V(x)$. This can be dealt with in the same way as above. Now we can choose the wave function to be real (this can also be done in the $E > V$ regime, but there it is inconvenient). Defining

$$\kappa(x) = \sqrt{\frac{2m(V(x)-E)}{\hbar^2}},$$

the Schrödinger equation turns into

$$\psi'' = \kappa^2\psi.$$

Plugging a solution

$$Ae^{\pm\varphi(x)}$$

where $\varphi'(x) = \kappa(x)$ into this equation yields

$$\left(A'' \pm 2A'\kappa \pm A\kappa' + \kappa^2\right)e^{\pm\varphi(x)} = \kappa^2 A e^{\pm\varphi(x)}.$$

Neglecting again the term proportional to A'', we are left with the condition that the second and third terms in the brackets on the left hand side vanish. Casting this into the form

$$\frac{1}{A}\frac{d}{dx}\left(A^2\kappa\right) = 0,$$

we find

$$A(x) = \frac{1}{\sqrt{\kappa(x)}},$$

leading to the WKB wave function:

$$\psi_{\text{WKB}}(x) = \frac{1}{\sqrt{\kappa(x)}}e^{\pm\int_{x_0}^{x}\kappa(x')dx'}.$$

Summary so far
We have considered an approximation to the solution of the stationary, one-dimensional Schrödinger equation in those regions where the wave function oscillates much more rapidly than the potential; in formula:

$$\frac{\psi''}{\psi} \gg \frac{V''}{V}, \left(\frac{V'}{V}\right)^2.$$

The solution to the stationary Schrödinger equation in those regions is approximated by

$$\psi_{\text{WKB}}(x) = \sqrt{\left|\frac{1}{k(x)}\right|} e^{\pm\int_{x_0}^{x} k(x')dx'},$$

where

$$k(x) = \pm\sqrt{\frac{2m\,[V(x)-E]}{\hbar^2}}.$$

For $E < V(x)$, the WKB solution represents a decaying or exponentially growing wave function; for $E > V(x)$, $k(x)$ is purely imaginary, and we obtain an oscillatory wave function.

5.4 THE WKB ANSATZ III

In this section we consider the WKB approximation once again, but using an alternative derivation. This derivation is performed in Problem 8.2 of Griffiths' book.

We consider again a Schrödinger equation in one dimension with a slowly varying potential. Slow means that the potential does not vary significantly on the scale of a wavelength of the solution.

$$\frac{-\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x).$$

Using $\kappa(x) = \sqrt{\frac{2m}{\hbar^2}\,[V(x)-E]}$, we can write this in the form

$$\psi''(x) = \kappa^2(x)\psi(x).$$

Note that κ^2 can be *positive or negative*. When it is positive, we are in the tunnelling region; if it is negative, we are in the classical region (the region where classical mechanics allows the particle to be).

We write the wave function in the form

$$\langle x|\psi\rangle = \exp[u(x)],$$

where $u(x)$ can be complex. The Schrödinger equation can now be worked out for this form in terms of u:

$$u'' + u'^2 = \kappa^2$$

(note that κ and u are functions of x). Note that, so far, *no approximation has been made.*

In order to make progress, we define a reference solution u_0 by

$$u_0'(x) = \pm\kappa(x).$$

It can easily be seen that for κ independent of x, the reference solution satisfies the Schrödinger equation:

$$u_0' = \pm\kappa; \qquad u_0'' = 0 \text{ for } \kappa \text{ constant}.$$

Now we write the *exact* solution u as

$$u(x) = u_0(x) + \delta u(x).$$

Putting this into the Schrödinger equation (for u as formulated above), we obtain

$$u'' + u'^2 = u_0'' + \delta u'' + \left(u_0'\right)^2 + 2u_0'\delta u' + \left(\delta u'\right)^2 = \kappa^2(x).$$

Now we use $u_0' = \pm\kappa(x)$ and neglect the second and the last term in the second expression as they are second order in the inverse wavelength of the variation of the potential and hence expected to be much smaller than the remaining terms. We then obtain:

$$u_0'' + 2u_0'\delta u' = 0;$$

hence

$$\delta u' = -\frac{1}{2}\frac{d}{dx}\ln u_0' = -\frac{1}{2}\frac{d}{dx}\ln\kappa$$

as can easily be verified.

Now we have

$$u(x) = u_0(x) + \delta u(x) = \pm\int_{x_0}^{x}\kappa(x')\,dx' + \ln\left(\frac{1}{\sqrt{\kappa(x)}}\right).$$

Translating this back into the original wave function ψ, we obtain

$$\langle x|\psi\rangle = \frac{1}{\sqrt{\kappa(x)}}\exp\left(\pm\int_{x_0}^{x}\kappa(x')dx'\right)$$

up to a normalisation constant. This is the WKB wave function.

In the classical region, $\kappa^2 < 0$, and we take $\kappa(x) = \mathrm{i}k(x)$, the same as the solution found at the beginning of the previous section. Similarly, for the tunnelling region, taking $\kappa(x)$ real we obtain the same result as in the previous section for $E < V(x)$.

5.5 TUNNELLING IN THE WKB APPROXIMATION

The WKB is often used to solve tunnelling problems. The standard way in which such problems are formulated is by considering a one-dimensional potential as in figure 5.2.

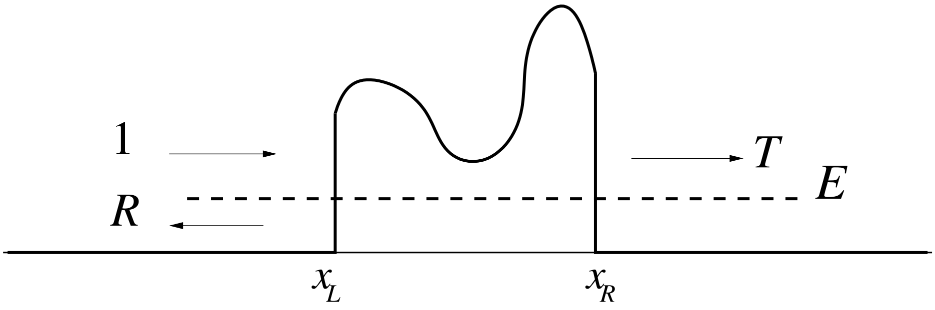


FIGURE 5.2: A typical tunnelling problem: A wave incident from the left splits into a part which bounces back and one which tunnels through the classically forbidden region.

Consider a wave $\exp(\mathrm{i}kx)$ incident from the left, where the potential is zero, at energy E. The wave vector $k = \sqrt{2m[E - V(x)]/\hbar^2}$. The wave is incident on a potential barrier between x_L and x_R where $V(x) > E$, that is, classically the particle cannot not enter this region. Part of the wave bounces back from this barrier and part tunnels through the classically forbidden potential. On the right hand side, the transmitted wave is given by $\exp(\mathrm{i}kx)$. All in all the wave function is

$$\langle x|\psi\rangle = \begin{cases} e^{\mathrm{i}kx} + Re^{-\mathrm{i}kx} & \text{for } x < x_L \\ u(x) & \text{for } x_L \le x \le x_R \\ Te^{\mathrm{i}kx} & \text{for } x > x_R. \end{cases}$$

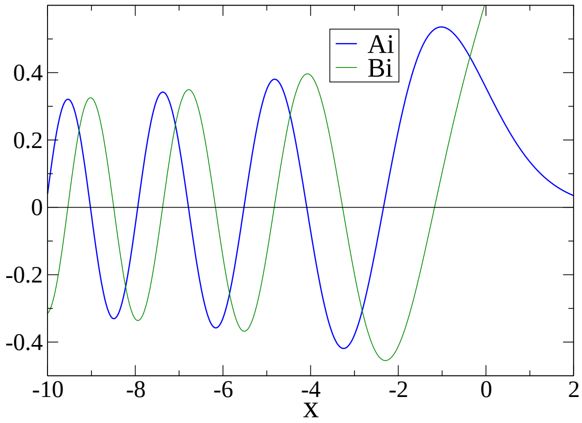


FIGURE 5.3: The Airy functions Ai and Bi. These are the solutions to the Schrödinger equation with a crossing point at $x = 0$. Left of this crossing point we see the oscillatory behaviour characteristic for $E > V(x)$ – the classically allowed region. Right of that point, the curve either decays (Ai) or it increases indefinitely (Bi).

For the solution in the classically forbidden region we use the WKB approximation:

$$u(x) = \frac{A}{\sqrt{\kappa(x)}} e^{\int_{x_L}^{x} \kappa(x')dx'} + \frac{B}{\sqrt{\kappa(x)}} e^{-\int_{x_L}^{x} \kappa(x')dx'}.$$

The matching conditions lead straightforwardly to a set of equations connecting A, B, T and R. The analysis simplifies considerably if the tunneling amplitude is small. This happens when the potential is wide and/or much higher than the energy of the incident wave. The wave function is then much smaller in amplitude at the right edge of the barrier than it is on the left. The exponent $\gamma = \int_{x_L}^{x_R} \kappa(x')dx'$ is mainly responsible for this difference – the prefactor $1/\sqrt{\kappa(x)}$ varies much less than the exponent. We therefore neglect the prefactor. It is then easy to see that matching $|T|$ to the solution yields $|T| \propto \exp(-\gamma)$.

The ratio between the tunnelling and the incident currents is proportional to $|T|^2$ as the velocity of the particles at the left and the right of the barrier is the same, so the current is dominated by the density, which is proportional to the square of the amplitude of the wave function. This ratio is usually denoted as the *transmission* $\mathcal{T}$. This transmission therefore satisfies

$$\mathcal{T} \propto e^{-2\gamma}; \qquad \gamma = \int_{x_L}^{x_R} \kappa(x')dx'.$$

5.6 THE CONNECTION FORMULAE

The WKB approximation is valid only when the wavelength of the solution is (much) smaller than the variation of the potential. Now consider a problem in which the potential varies continuously. Then at the ‘classical turning point’ x_t, where $E = V(x_\mathrm{t})$, the wavelength $1/\kappa(x)$ diverges! Hence the WKB approximation fails miserably in this case. So what should we do? Well, we know that the exact solution behaves smoothly at and close to the turning point. Let us therefore look at the exact solution close to the turning point and integrate from the turning point to the left and to the right. Then we match this left and right solution onto the WKB approximation when we are at some distance from the turning point. Now we describe how this is done. Close to the turning point, which we take at $x_\mathrm{t} = 0$, we approximate the potential by

$$V(x) = E + \frac{\hbar^2}{2m}\alpha^3 x.$$

The factor α^3 is the slope of the potential at the turning point. We have taken it to be α^3 just for convenience. Then the Schrödinger equation has the following form near the turning point:

$$\psi'' = \alpha^3 x \psi.$$

There are two solutions to this equation (as it is a second order differential equation): they are called the *Airy functions* $Ai(\alpha x)$ and $Bi(\alpha x)$ – they are shown in figure 5.3. They are not known in terms of standard functions, but it is known that for large positive and negative arguments they assume the forms:

$$Ai(z) \approx \begin{cases} \frac{1}{\sqrt{\pi}(-z)^{1/4}} \sin\left[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}\right] & \text{for } z \ll 0 \\ \frac{1}{2\sqrt{\pi} z^{1/4}} \exp\left(-\frac{2}{3} z^{3/2}\right) & \text{for } z \gg 0; \end{cases}$$

$$Bi(z) \approx \begin{cases} \frac{1}{\sqrt{\pi}(-z)^{1/4}} \cos\left[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}\right] & \text{for } z \ll 0 \\ \frac{1}{\sqrt{\pi} z^{1/4}} \exp\left(\frac{2}{3} z^{3/2}\right) & \text{for } z \gg 0. \end{cases}$$

The WKB form, which should also be valid once we are far enough from the turning point, is given by ($\alpha > 0$):

$$\psi_{\text{WKB}}(x < 0) = \frac{1}{\sqrt{|k(x)|}} \left[A e^{-\mathrm{i}\int_x^0 k(x')\,dx'} + B e^{\mathrm{i}\int_x^0 k(x')\,dx'} \right]$$

and

$$\psi_{\text{WKB}}(x > 0) = \frac{1}{\sqrt{|k(x)|}} \left[C e^{-\int_0^x k(x')\,dx'} + D e^{\int_0^x k(x')\,dx'} \right].$$

Both forms can be worked out analytically for $k(x) = \sqrt{\alpha^3 |x|}$:

$$\psi_{\text{WKB}}(x < 0) = \frac{1}{\alpha^{3/4}(-x)^{1/4}} \left[A e^{-\mathrm{i}\frac{2}{3}(-\alpha x)^{3/2}} + B e^{\mathrm{i}\frac{2}{3}(-\alpha x)^{3/2}} \right]$$

and

$$\psi_{\text{WKB}}(x > 0) = \frac{1}{\alpha^{3/4} x^{1/4}} \left[C e^{-\frac{2}{3}(\alpha x)^{3/2}} + D e^{\frac{2}{3}(\alpha x)^{3/2}} \right].$$

We see that the forms of the WKB and the Airy functions for large argument are the same (as it should) – we just need to match the coefficients.

A salient feature of the Airy functions is the factor 2 occurring in the denominator of Ai for positive arguments, and the lack of this factor of 2 in Bi. This shows up in the matching for positive x: writing

$$\langle x | \psi \rangle = a\, Ai(x) + b\, Bi(x),$$

we must have from the matching at positive x:

$$a = 2\sqrt{\pi/\alpha}\, C \qquad \text{and } b = \sqrt{\pi/\alpha}\, D.$$

For negative argument we therefore have a WKB solution

$$\frac{2C}{\sqrt{|k(x)|}} \sin\left(\int_x^0 k(x')\,dx' + \frac{\pi}{4} \right) + \frac{D}{\sqrt{|k(x)|}} \cos\left(\int_x^0 k(x')\,dx' + \frac{\pi}{4} \right).$$

We now express the matching condition for a turning point x_{t} as follows:

$$\psi(x \ll x_{\text{t}}) = \frac{2C}{\sqrt{|k(x)|}} \sin\left(\int_x^{x_{\text{t}}} k(x')\,dx' + \frac{\pi}{4} \right) + \frac{D}{\sqrt{|k(x)|}} \cos\left(\int_x^{x_{\text{t}}} k(x')\,dx' + \frac{\pi}{4} \right) \leftrightarrow$$

$$\frac{C}{\sqrt{|k(x)|}} \exp\left(-\int_{x_{\text{t}}}^x k(x')\,dx' \right) + \frac{D}{\sqrt{|k(x)|}} \exp\left(\int_{x_{\text{t}}}^x k(x')\,dx' \right);\ [x \gg x_{\text{t}}].$$

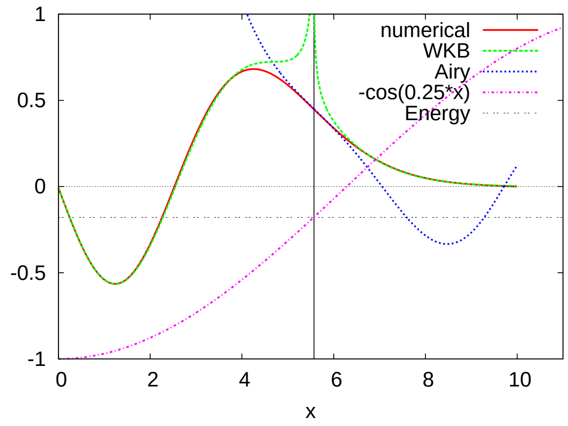


FIGURE 5.4: The matching procedure for a particle of unit mass, moving in a potential $-\cos(x/4)$. The red curve is an accurate numerical solution. The bound state energy, found as -0.17891517, is shown as the black horizontal line – it crosses the potential at the turning point (black vertical line). The green line is a WKB solution. The part left of the turning point was calculated starting off with the correct boundary condition $\psi(x=0)=0$ at $x=0$. The part right of the turning point started of near $x=10$ as a decaying solution. Both WKB parts fail close to the turning point. The blue curve is the Airy function solution matched to the exact solution at the turning point. It describes the solution well near that point but deviates from it far away from the turning point.

with the appropriate expression for $k(x)$ which is always real.

For a turning point with a classical region on the right rather than on the left we have:

$$\psi(x \ll x_t) = \frac{C}{\sqrt{|k(x)|}} \exp\left(-\int_x^{x_t} k(x')\,dx'\right) + \frac{D}{\sqrt{|k(x)|}} \exp\left(\int_x^{x_t} k(x')\,dx'\right) \leftrightarrow$$

$$\frac{2C}{\sqrt{|k(x)|}} \sin\left(\int_{x_t}^{x} k(x')\,dx' + \frac{\pi}{4}\right) + \frac{D}{\sqrt{|k(x)|}} \cos\left(\int_{x_t}^{x} k(x')\,dx' + \frac{\pi}{4}\right);\ [x \gg x_t].$$

In summary, we have a WKB solution which is accurate far away from the turning point. The Airy function solution is accurate close to the turning point. The exact solution satisfies the two limiting cases (close to and far away from) the turning point. The situation is represented in figure 5.4.

Now we can apply these matching expressions to several potentials. Suppose we have a potential well which is bounded by an infinite potential barrier on the left hand side, and a continuous turning point on the right hand side. For large x, we do not want the solution to explode, hence we have $D=0$ there. So we are left with the solution

$$\frac{2C}{\sqrt{|k(x)|}} \sin\left(\int_x^{x_t} k(x')\,dx' + \frac{\pi}{4}\right)$$

within the well and

$$\frac{C}{\sqrt{|k(x)|}} \exp\left(-\int_{x_t}^{x} k(x')\,dx'\right).$$

outside. Note that these solutions need not match at $x = x_t$ as these are the approximate WKB forms which should only hold far enough from the turning point. Now we must require the solution to vanish near the infinite wall. There, the WKB solution must be valid (within the WKB approximation). This boundary condition directly leads to

$$\int_x^{x_t} k(x')\,dx' + \pi/4 = n\pi.$$

If we have a potential well bounded by a turning point on the left and one on the right hand side, the solution inside the well is given by

$$\frac{2C}{\sqrt{|k(x)|}}\sin\left(\int_x^{x_{\mathrm{R}}} k(x')\,dx' + \frac{\pi}{4}\right)$$

and by

$$\frac{D}{\sqrt{|k(x)|}}\cos\left(\int_{x_{\mathrm{L}}}^{x} k(x')\,dx' + \frac{\pi}{4}\right)$$

where the first expression is based on the right turning point x_{R} and the second one on the left turning point x_{L}. Requiring both solutions to be identical leads to

$$\int_x^{x_{\mathrm{R}}} k(x')\,dx' + \frac{\pi}{4} = \pm\int_{x_{\mathrm{L}}}^{x} k(x')\,dx' + \frac{\pi}{4} + (n+1/2)\pi,$$

which follows directly from the notion that $\sin x = \cos(\pi/2 - x)$. The last equation should hold for all x and this can be the case only for the + sign before the first term on the right hand side. We then obtain the condition

$$\int_{x_{\mathrm{L}}}^{x_{\mathrm{R}}} k(x')\,dx' = (n+1/2)\pi.$$

Final Summary

The WKB approximation leads to practical schemes for approximating tunneling amplitudes and bound state energies. For tunnelling, we have found that the transmission probability T for tunnelling through a barrier $V(x)$, is given as

$$T = e^{-2\gamma},$$

where

$$\gamma = \int_{x_L}^{x_R} \kappa(x')dx', \qquad \kappa(x) = \sqrt{\frac{2m}{\hbar^2}(V(x)-E)},$$

and $x_{\mathrm{L,R}}$ are the left and right classical turning points where $E = V(x)$.

We have concluded that the WKB form, which has a limited domain of validity, can be extended with an Airy function solution close to the points where $E = V(x)$. This leads to rather simple conditions for finding the bound state energies of quantum particles in one dimension (see the figure below).

1. For a particle in a well with two vertical walls (i.e. jumps in the potential at x_{L} and x_{R}, both crossing the value E), the condition for having a bound state is

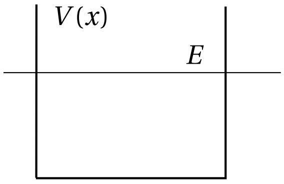
$$\int_{x_{\mathrm{L}}}^{x_{\mathrm{R}}} k(x)dx = n\pi, \qquad n = 1,2,3,\ldots, \qquad \kappa(x) = \sqrt{\frac{2m}{\hbar^2}(E - V(x))}.$$

2. For a particle in a well with a single vertical wall (i.e. a jump in the potential energy at x_{L} crossing the value E, and a slope which reaches the value E at $x = x_{\mathrm{R}}$), the WKB condition for a bound state is:

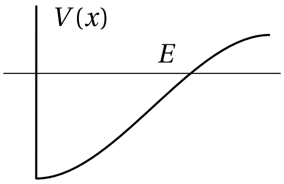
$$\int_{x_{\mathrm{L}}}^{x_{\mathrm{R}}} k(x)dx = \left(n - \frac{1}{4}\right)\pi, \qquad n = 1,2,3,\ldots$$

3. For a particle in a well with two sloping walls crossing the energy at x_{L} and x_{R}, both crossing the value E), the condition for having a bound state is

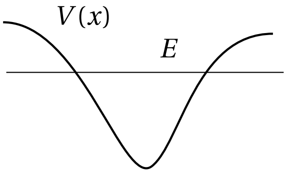
$$\int_{x_{\mathrm{L}}}^{x_{\mathrm{R}}} k(x)dx = \left(n - \frac{1}{2}\right)\pi, \qquad n = 1,2,3,\ldots$$



1



2



3

5.7 PROBLEMS

1. Consider a one-dimensional potential

$$V(x) = \begin{cases} \lambda x & \text{for } 0 < x < a; \\ \infty & \text{for } x \leq 0 \text{ and } x \geq a. \end{cases}$$

Show that the ground state energy of a particle with mass m in this potential is, in the WKB approximation, given as the solution to the implicit equation

$$E^{3/2} - (E - E_0)^{3/2} = \frac{3}{2}\frac{\pi\hbar\lambda}{\sqrt{2m}}$$

where $E_0 = \lambda a$. Note that for $E_0 = 1 = \frac{3}{2}\frac{\pi\hbar\lambda}{\sqrt{2m}}$, the solution to this equation is given by $E = 1$. Please comment on the suitability of the WKB method in this limit. Compare this result with that obtained in problem 3 of chapter 4.

2. Consider a particle moving in one dimension in a 'quartic potential'. The Hamiltonian is given as

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{b}{4}x^4,$$

with b some positive constant.

Solve the spectrum using the WKB approximation. Compare your result with that of problem 4 of chapter 4.

3. Using the WKB approximation, derive a formula for the energies of the bound s-states of a particle of a particle of mass m in a potential $V(r) = -V_0 \exp(-r/R)$ with V_0 and R both positive.

4. Use the WKB approximation to find the allowed energies (E_n) of an infinite square well with a 'shelf' of height V_0, extending half-way across:

$$V(x) = \begin{cases} V_0 & \text{if } 0 < x < a/2, \\ 0 & \text{if } a/2 < x < a, \\ \infty & \text{otherwise.} \end{cases}$$

Express your answer in terms of V_0 and $E_n^0 \equiv (n\pi\hbar)^2/(2ma^2)$ (the n^{th} allowed energy for the 'unperturbed' infinite square well, with *no* shelf). Assume that $E_1^0 > V_0$, but do *not* assume that $E_n \gg V_0$. Compare your result for the same problem in first-order perturbation theory:

$$E_n = E_n^0 + \frac{V_0}{2}.$$

Note that they are in agreement if either V_0 is very small (perturbative regime) or when n is very large (semi-classical WKB regime).

5. In WKB, when we have a system with two infinite walls at x_L and x_R with a classical regime $[E > V(x)]$ in between, we have the quantisation condition

$$\int_{x_\mathrm{L}}^{x_\mathrm{R}} p(x)\, dx = n\pi\hbar.$$

For a system with an infinite wall at x_L and a single classical turning point at x_R, this quantisation condition is replaced by

$$\int_{x_\mathrm{L}}^{x_\mathrm{R}} p(x)\, dx = (n - 1/4)\pi\hbar$$

for positive, integer n.

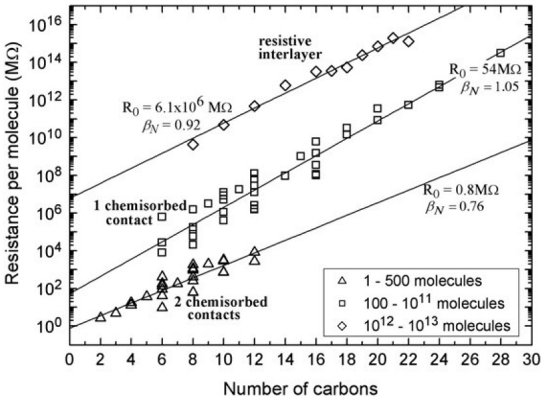
Now consider a potential given by

$$V(x) = \begin{cases} mgx & \text{if } x > 0, \\ \infty & \text{if } x < 0. \end{cases}$$

This describes a ball of mass m moving along a straight sloped track, against a wall on the left end of the track. Find the turning point for an energy E. Find the energy levels from the WKB quantisation condition. Give the first three levels in units of $(mg^2\hbar^2)^{1/3}$. Compare your results with the exact values (units of $(mg^2\hbar^2)^{1/3}$):

$$E_0 = 1.8558; \qquad E_1 = 3.2446 \qquad \text{and} \qquad E_2 = 4.3817.$$

6. The graph below shows the relation between the resistance and the length, measured in numbers of carbon atoms, of alkane chains. In the experimental setups used, molecules of different lengths were connected to two electrodes.



A small bias voltage was then applied across the electrodes, and the current flow was recorded.

(a) Explain the fact that for each experiment, a straight line connects the data in this graph.

(b) The tunneling takes place via a so-called ‘molecular orbital’: this is an electronic state on the (uncoupled) molecule at a definite energy.

Find approximate values of the location of the molecular orbitals in the experiment with respect to the Fermi energy of the gold.

You will need the length per CH_2 unit in an alkane chain. The carbon atoms are arranged in a zigzag pattern with an angle of about 109^o between successive atoms. The distance between those atoms is 1.5 Å. A good guess for the distance is therefore about 2 Bohr radii.

The literature value for the difference between the potential and the fermi energy is 0.2 eV. If your value differs, can you give arguments why this would be the case?

6

GREEN'S FUNCTIONS IN QUANTUM MECHANICS

6.1 INTRODUCTION

Green's functions are the workhorses of theoretical quantum mechanics. They are used in many subfields of quantum mechanics because they are very powerful. Nevertheless, to many researchers, Green's functions often seem abstract and difficult. Sometimes this is right, but the use of Green's functions for quantum systems in which the interactions between particles are not explicitly considered is not so complicated. And they can be useful even for such noninteracting systems. In this chapter, we shall explain what the Green's function is and how it can be used for analysing different types of problems. In chapter 7 we shall use Green's functions when discussing scattering theory.

6.2 DEFINITION OF THE GREEN'S FUNCTION

The Green's function of a system described by a Hamiltonian H is defined as

$$(z - H)\, G = \mathbb{1}, \tag{6.1}$$

where the right hand side is the unit operator. This operator can have different forms, depending on the structure of the Hilbert space. If that space is a finite-dimensional vector space, the unit operator can be written as the *unit matrix*

$$\mathbb{1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

In the case where the Hilbert space consists of the normalizable complex functions defined on the real axis, the matrix elements of the unit operator are given in terms of a delta function

$$\left\langle x|\mathbb{1}|x'\right\rangle = \delta(x - x').$$

For the Hilbert space of particles moving in 3D, i.e. the space consisting of normalizable in $\mathbb{R}3$ (i.e. $L2$-functions) in $\mathbb{R}^3$, the matrix elements are given by the three-dimensional delta-function $\delta^{(3)}(\mathbf{r} - \mathbf{r}')$.

The Green's function may seem a rather arbitrary object – it is not immediately clear why this function could be useful in any way. Moreover, it is defined as the inverse of an operator, and that is usually difficult to find. It is in particular not obvious what information we could obtain from this inverse, while we could instead diagonalize H (which is, numerically, equally difficult as inversion). In order to give some insight into these questions, we must recall an important result from complex function theory (see the end of chapter 1):

$$\lim_{\epsilon\downarrow 0}\frac{1}{x+\mathrm{i}\epsilon}=\mathscr{P}\left(\frac{1}{x}\right)-\mathrm{i}\pi\delta(x).$$

This turns out very useful as can be seen by expanding the Green's function for a system with a *discrete* spectrum in the basis consisting of the eigenstates $|\phi_n\rangle$ of H:

$$G(z)=\frac{1}{z-H}=\sum_n|\phi_n\rangle\frac{1}{z-E_n}\langle\phi_n|$$

Here, z can in principle be any complex number, but we decide to choose it close to, but above the real axis. We then obtain the *retarded Green's function*

$$G^{\mathrm{r}}(E)=\frac{1}{E-H+\mathrm{i}\eta}$$

where E is real and η is considered to be small and positive. We only give E as an argument of the Green's function; the superscript 'r' indicates that we have moved the energy slightly upward (i.e. to the positive imaginary part) in the complex plane.

We have

$$G^{\mathrm{r}}(E)=\sum_n|\phi_n\rangle\mathscr{P}\left(\frac{1}{E-E_n}\right)\langle\phi_n|-\mathrm{i}\pi\sum_n|\phi_n\rangle\delta(E-E_n)\langle\phi_n|.$$

This Green's function is an operator depending on E (or more generally, on z), and we would like to work with a simpler object. We therefore study the trace of the Green's function. The trace of an operator is defined as the sum over the diagonal elements of that operator:

$$\mathrm{Tr}(\hat{A})=\sum_n\langle\phi_n|\hat{A}|\phi_n\rangle.$$

where the vectors $|\phi_n\rangle$ form an orthonormal basis. It can be shown that the trace is independent of the particular basis chosen – for the Green's function, we take the basis consisting of the eigenstates of the Hamiltonian and find:

$$\mathrm{Tr}G(z)=\sum_n\frac{1}{z-E_n}.$$

We see that the trace of the Green's function has a simple pole on the real axis at every energy eigenvalue E_n.

We have learned two important things: (i) the trace of the Green's function is a complex function which has poles on the real axis which correspond to the eigenvalues of H and (ii) at these poles the imaginary part of the Green's function (not its trace) is proportional to $|\phi_n\rangle\langle\phi_n|$, which is a projection operator onto the corresponding eigenstate ϕ_n. We see that having the Green's function is equally useful as having the eigenstates and eigenfunctions of the Hamiltonian. The reason why we often use Green's functions is that it is often possible to obtain them for systems for which the Hamiltonian cannot be diagonalised. An example is formed by a closed rather than an open system, as we shall see below.

We can also conclude that the trace of the imaginary part of the (retarded) Green's function gives a series of δ-functions, one for each energy:

$$\lim_{\eta\downarrow 0}\mathrm{Tr}\left[\mathrm{Im}G(E+\mathrm{i}\eta)\right]=-\pi\sum_n\delta(E-\epsilon_n).$$

This is an example of a general result which says that the imaginary part of the trace of the Green's function is proportional to the *density of states* of a system.

The Green's function is often powerful in studying quantum systems, as we already mentioned in the introduction. To be specific, (i) the Green's function is useful for working out perturbation series, (ii) it plays a major role in scattering theory (again when scattering is formulated as a pertubative problem) and (iii) the Green's function has a *local* character: we can evaluate it for a particular region, and it encodes the influence which this region has on adjacent regions. In this chapter we shall briefly go into these applications of Green's functions.

6.3 GREEN'S FUNCTIONS AND PERTURBATIONS

There exists a very important equation that is quite simple but turns out very powerful for perturbative problems. To obtain this equation, let us first formulate a perturbative quantum problem by splitting its Hamiltonian as

$$H = H_0 + V,$$

where V is 'small' in some sense. Usually we mean small with respect to the typical distance between the energy eigenvalues of H_0 or, in the case of a continuous spectrum, small with respect to the typical eigenenergy of H_0 measured with respect to the ground state energy. We define G_0 as the Green's function of the unperturbed Hamiltonian H_0:

$$(z - H_0) G_0 = \mathbb{1}, \tag{6.2}$$

where the unit operator on the right hand side is the same as above, i.e. its form depends on the Hilbert space of the system.

Now it is very easy to obtain the following result:

$$z - H = G^{-1} = z - H_0 - V = G_0^{-1} - V.$$

Multiplying the second and the fourth form of this equation from the *left* with G_0 and from the *right* with G, we obtain, after some reorganisation:

$$G = G_0 + G_0 V G. \tag{6.3}$$

It seems that we have not made much progress, as this is an implicit equation for the Green's function G. However, V is small, and this inspires us to take the expression for G and plug it into the right hand side of this equation. We then obtain:

$$G = G_0 + G_0 V G_0 + G_0 V G_0 V G.$$

The second term on the right hand side contains one V, and the third terms contains *two* V's. This means that for small V, the third term is a lot smaller than the second one, and if we neglect this very small term, we obtain an *explicit* equation for G:

$$G = G_0 + G_0 V G_0.$$

We can iterate further and further, each time replacing the G on the right hand side by the full right hand side, and thus obtain an infinite perturbation series:

$$G = G = G_0 + G_0 V G_0 + G_0 V G_0 V G_0 + G_0 V G_0 V G_0 V G_0 + \dots \tag{6.4}$$

The terms on the right hand side contain increasingly higher-order contributions in V to the unperturbed Green's function. Eq. (6.3) is the famous *Dyson equation.* It is a very important equation which is used in many fields of physics. Eq. (6.4) is called the *Born series.* Cutting this of after the first-order term (in V) is called the *first Born approximation*, after the second order term in V it is the *second Born approximation*, etcetera.

6.3.1 SYSTEMS WITH DISCRETE SPECTRA

In order to clarify the relation between the Born series and standard perturbation theory which you have learned in your previous quantum mechanics course, let us have a look at the first-order equation and check whether this reproduces the results of standard perturbation theory which is set up without Green's functions (see e.g. Griffiths Chapter 6). For simplicity, we take a nondegenerate Hamiltonian which has the unperturbed Green's function:

$$G_0(z) = \sum_n |\phi_n\rangle \frac{1}{z - E_n} \langle \phi_n| .$$

Now we switch on the perturbation V. The first order approximation to the Green's function is

$$G(z) = G_0 + G_0 V G_0.$$

We can easily evaluate the trace of this first order Green's function, using the basis of the eigenstates $|\phi_n\rangle$ of the *unperturbed* Hamiltonian:

$$\mathrm{Tr}\,[G(z)] = \sum_n \frac{1}{z-E_n} + \sum_n \frac{1}{z-E_n} \langle\phi_n| V |\phi_n\rangle \frac{1}{z-E_n}.$$

Here we have used that G_0 is diagonal in the basis $|\phi_n\rangle$. For small V, we note that

$$\frac{1}{z-E_n-\langle\phi_n| V |\phi_n\rangle} \approx \frac{1}{z-E_n} + \frac{1}{(z-E_n)^2}\langle\phi_n| V |\phi_n\rangle + \frac{1}{(z-E_n)^3}\langle\phi_n| V |\phi_n\rangle^2 + \ldots, \tag{6.5}$$

showing that, to *first* order in V, the trace of the Green's function may be written as the left hand side of the last equation, which has poles at $E_n + \langle\phi_n| V |\phi_n\rangle$. This is equivalent to the result of standard first order perturbation theory which says that the energy correction to E_n due to a perturbative potential V is given by $\langle\phi_n| V |\phi_n\rangle$ (see also Desai Eq. (16.26) on page 280, or Griffiths, Ch. 6).

Now we consider the second order expansion:

$$G = G_0 + G_0 V G_0 + G_0 V G_0 V G_0.$$

Taking the trace gives (we use the obvious notation $\langle\phi_n| V |\phi_m\rangle \equiv V_{nm}$):

$$\mathrm{Tr}G(z) = \sum_n \frac{1}{z-E_n} + \sum_n \frac{1}{z-E_n} V_{nn} \frac{1}{z-E_n} + \sum_{nm} \frac{1}{z-E_n} V_{nm} \frac{1}{z-E_m} V_{mn} \frac{1}{z-E_n}.$$

We first note that for $n = m$ in the sum in the rightmost term we obtain the *second order* term of the expansion of $1/(z-E_n-V_{nn})$; see Eq. (6.5). Therefore, we only need to add the contributions $n \neq m$ in the last term to first order in a Taylor expansion of the Green's function, and obtain:

$$\mathrm{Tr}G(z) = \sum_n \frac{1}{z-E_n-V_{nn}-V_{nn}^{(2)}}$$

with

$$V_{nn}^{(2)} = \sum_{m\neq n} \frac{V_{nm}V_{mn}}{z-E_m}.$$

The poles of the Green's function will occur at E_n plus a first order correction in V. So, we can replace z in the last expression by E_n and obtain the result that the pole will be shifted in second order by

$$\sum_{m\neq n} \frac{V_{nm}V_{mn}}{E_n-E_m}.$$

This is then recognised as the second order correction known from standard perturbation theory (see e.g. Desai, Eq. (16.28) or Griffiths, Ch. 6).

We see that the Born series allows us to easily find higher-order corrections to the energies. To find the corrections to the corresponding wave functions is a bit more complicated. We return to that problem in section 6.3.3.

6.3.2 SYSTEMS WITH CONTINUOUS SPECTRA

Now we turn to a system with a continuous spectrum. The operator H_0 is chosen such as to be easily diagonalizable:

$$H_0 |\phi_k\rangle = E_k |\phi_k\rangle,$$

where k is now a continuous index.

We take the same energy E_k and define the eigenstate of the full Hamiltonian at that energy $|\psi_k\rangle$:

$$H|\psi_k\rangle = E_k|\psi_k\rangle.$$

Note that we assume that the energy E_k is in the continuous spectrum of both the unperturbed and the perturbed Hamiltonian. This is the case for systems which are of interest to us.

The unperturbed Hamiltonian H_0 was assumed to be diagonalized, and we anticipate that in general the $|\psi_k\rangle$ are difficult to find. We use perturbation theory for this problem. We can write the last equation in the form

$$(E_k - H_0)|\psi_k\rangle = V|\psi_k\rangle.$$

and combine this with

$$(E_k - H_0)|\phi_k\rangle = 0.$$

Now we can write

$$|\psi_k\rangle = |\phi_k\rangle + (E_k - H_0)^{-1}V|\psi_k\rangle.$$

(From now on, we leave out the subscript k to the wave functions ψ and ϕ.) We have obtained an implicit equation for the wave function $|\psi\rangle$, similar to the Dyson equation found in the previous section. We recognize the Green's function of the unperturbed system as the operator in front of V on the right hand side of this equation:

$$|\psi\rangle = |\phi\rangle + G_0 V|\psi\rangle.$$

Note that the second term on the right hand side is the perturbation of the wave function $|\phi\rangle$ at a *fixed* eigenenergy due to the presence of the perturbation V. Similar to the approach taken for the Dyson equation, we can try to solve this equation iteratively. The lowest order approximation is by replacing ψ on the right hand side by ϕ:

$$|\psi\rangle = |\phi\rangle + G_0 V|\phi\rangle.$$

This is the so-called *first Born approximation* often abbreviated as 'Born approximation'. We may also replace the full expression for ψ into the equation and obtain:

$$|\psi\rangle = |\phi\rangle + G_0 V|\phi\rangle + G_0 V G_0 V|\psi\rangle.$$

etcetera. This expression just given is the second Born approximation. This Born approximation is frequently used in scattering theory.

6.3.3 DISCRETE SPECTRA REVISITED

Now that we have analysed the energy corrections for discrete spectra in section 6.3.1 and the wave function corrections for continuous spectra in the previous section, we now combine the techniques used in both in order to analyse the change in the wave functions for discrete problems.

We start from the unperturbed Schrödinger equations

$$H_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle.$$

and its 'full' version:

$$H|\psi_n\rangle = E_n|\psi_n\rangle, \tag{6.6}$$

where $H = H_0 + V$. The energy has shifted: $E_n = E_n^{(0)} + \delta E_n$ and the wave function has too:

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + |\delta\psi_n\rangle.$$

Therefore, Eq. (6.6) can be written in the form

$$(H_0 + V)\left(\left|\psi_n^{(0)}\right\rangle + \left|\delta\psi_n\right\rangle\right) = (E_0 + \delta E_n)\left(\left|\psi_n^{(0)}\right\rangle + \left|\delta\psi_n\right\rangle\right).$$

Using the unperturbed Schrödinger equation, this can directly be reworked to give

$$(E_0 - H_0)\left(\left|\psi_n\right\rangle - \left|\psi_n^{(0)}\right\rangle\right) = (V - \delta E)\left|\psi_n\right\rangle.$$

Multiplying this equation by $G_0 = (E_n^{(0)} - H_0 + \mathrm{i}\eta)^{-1}$ directly gives

$$\left|\psi_n\right\rangle = \left|\psi_n^{(0)}\right\rangle + G_0\,(V - \delta E)\left|\psi_n\right\rangle.$$

The right hand side can then be expanded as a Born series:

$$\left|\psi_n\right\rangle = \left|\psi_n^{(0)}\right\rangle + G_0\,(V - \delta E)\left|\psi_n^{(0)}\right\rangle + G_0\,(V - \delta E)\,G_0\,(V - \delta E)\left|\psi_n^{(0)}\right\rangle + \ldots$$

Now we can use the first order correction to the eigenvalue: $\delta E = \left\langle \psi_n^{(0)} |V| \psi_n^{(0)} \right\rangle$ found in section 6.3.1, in order to find the first order correction to $\left|\psi_n^{(0)}\right\rangle$:

$$\left|\psi_n\right\rangle = \left|\psi_n^{(0)}\right\rangle + G_0\,(V - \delta E)\left|\psi_n^{(0)}\right\rangle = \left|\psi_n^{(0)}\right\rangle + \sum_m \left|\psi_m^{(0)}\right\rangle \frac{1}{E_n - E_m + \mathrm{i}\eta} \left\langle\psi_m^{(0)}\right|(V - \delta E)\left|\psi_n^{(0)}\right\rangle.$$

First note that the term for $m = n$ vanishes (why?). For all other terms, the part proportional to δE vanishes too (why?) and we are left with

$$\left|\psi_n\right\rangle = \left|\psi_n^{(0)}\right\rangle + \sum_{m \neq n} \left|\psi_m^{(0)}\right\rangle \frac{\left\langle \psi_m^{(0)} \middle| V \middle| \psi_n^{(0)} \right\rangle}{E_n - E_m}.$$

This result is well known from stationary perturbation theory.

6.4 GREEN'S FUNCTIONS AND BOUNDARIES

Green's functions are very useful for calculations on very complex systems, containg bulk-like and microscopic parts. As an illustration, we consider an atom near a crystal surface, where the atom has a single energy level ϵ_a. The energy spectrum of the bulk crystal is continuous (it is a band structure). We can write the full Hamiltonian as

$$H = H_\mathrm{B} + H_\mathrm{a} + H_\mathrm{T},$$

where H_B is the Hamiltonian of the bulk crystal, H_a is that of the atom and H_T is the coupling Hamiltonian between the atom and the crystal. The Hilbert space of the problem can be written as

$$\mathscr{H} = \mathscr{H}_\mathrm{a} \oplus \mathscr{H}_\mathrm{B},$$

where $\mathscr{H}_a$ is the one-dimensional space spanned by the atomic eigenstate. In matrix form, this Hamiltonian reads:

$$H = \left(\begin{array}{c|c} \epsilon_\mathrm{a} & -\tau \\ \hline -\tau^\dagger & H_\mathrm{B} \end{array}\right).$$

Here, τ is the coupling between atom and bulk, and we take for the energy of the isolated atom the value ϵ_a. More generally, we can write

$$H = \left(\begin{array}{c|c} H_\mathrm{S} & -\tau \\ \hline -\tau^\dagger & H_\mathrm{B} \end{array}\right).$$

where the subscript S stands for 'system'. This Hamiltonian is supposed to be small enough that we can diagonalize it. Denoting the dimension of the Hilbert spaces of the system and the bulk system by D_{S} and D_{B}, we see that H_{B} is a $D_{\mathrm{B}} \times D_{\mathrm{B}}$ matrix, H_{S} is a $D_{\mathrm{S}} \times D_{\mathrm{S}}$-sized matrix and τ is of size $D_{\mathrm{S}} \times D_{\mathrm{B}}$.

The solution to this problem seems very difficult. Perhaps we can solve for the bulk spectrum and states using Bloch's theorem, and the atom by itself is easy (only a single level with a known energy). But when we couple the two, the problem seems to get hopelessly complicated. However, we shall now show that it is possible to find the atomic part of the Green's function of the coupled system!

We write the Green's function of the coupled form as the matrix

$$G(z) = \left(\begin{array}{c|c} G_{\mathrm{S}} & A^\dagger \\ \hline A & G_{\mathrm{B}} \end{array}\right)$$

All the submatrices depend on z but we refrain from indicating that explicitly.

The Green's function is the inverse of $z - H$ so we can write:

$$\left(\begin{array}{c|c} z - H_{\mathrm{S}} & \tau \\ \hline \tau^\dagger & z - H_{\mathrm{B}} \end{array}\right) \left(\begin{array}{c|c} G_{\mathrm{S}} & A^\dagger \\ \hline A & G_{\mathrm{B}} \end{array}\right) = \left(\begin{array}{c|c} \mathbb{1}_{\mathrm{S}} & 0 \\ \hline 0 & \mathbb{1}_{\mathrm{B}} \end{array}\right)$$

We extract the following two equations from this:

$$(z - H_{\mathrm{S}})\, G_{\mathrm{S}} + \tau A = \mathbb{1}_{\mathrm{S}},$$

$$\tau^\dagger G_{\mathrm{S}} + (z - H_{\mathrm{B}})\, A = 0.$$

(Note that the 0 on the right hand side is a $D_{\mathrm{B}} \times D_{\mathrm{S}}$ matrix). The second of these equations allows us to eliminate A:

$$A = (z - H_{\mathrm{B}})^{-1}\, \tau^\dagger G_{\mathrm{S}},$$

which we can use to remove A from the first equation to obtain

$$G_{\mathrm{S}} = (z - H_{\mathrm{S}} - \Sigma)^{-1},$$

where

$$\Sigma = \tau g_{\mathrm{B}} \tau^\dagger$$

and

$$g_{\mathrm{B}} = (z - H_{\mathrm{B}})^{-1}.$$

By considering the sizes of the matrices g_{B} ($D_{\mathrm{B}} \times D_{\mathrm{B}}$) and τ ($D_{\mathrm{S}} \times D_{\mathrm{B}}$) it is found that σ is a matrix of size $D_{\mathrm{S}} \times D_{\mathrm{S}}$. This matrix encodes all the influence of the bulk onto the system.

The operator Σ is called the *self-energy*. This operator is not Hermitian. It can be written as

$$\Sigma = \Lambda + \mathrm{i}\Gamma,$$

where Λ and Γ are Hermitian. This is very important to interpret how the presence of a bulk system influences the discrete spectrum of a system. The effect of Λ is to shift the discrete energy levels. The effect of Γ is to *broaden* the discrete level into a Lorentzian peak. To see this, let us consider a system with a one-dimensional Hilbert space, e.g. an atom with a single energy level ϵ_{a}. The Green's function of the atom close to the bulk system is then

$$G_{\mathrm{S}} = \left(z - \epsilon_{\mathrm{a}} - \lambda - \mathrm{i}\gamma\right)^{-1},$$

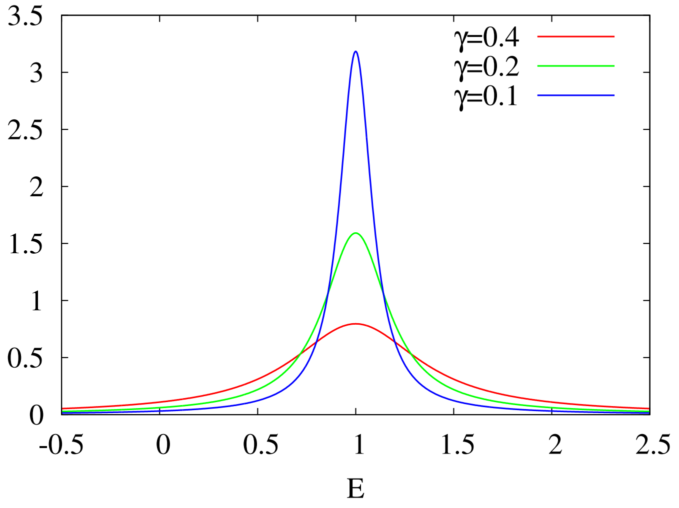


FIGURE 6.1: The Lorentz curve for $\epsilon_a + \lambda = 1$ and different values of γ.

where the use of the lower case letters λ and γ indicates that they correspond to the small system. Calculating the density of states yields

$$\mathrm{DOS}(E) = \frac{1}{\pi}\mathrm{Im}\left[\frac{1}{E-\epsilon_a-\lambda-\mathrm{i}\gamma}\right] = \frac{1}{\pi}\frac{\gamma}{(E-\epsilon_a-\lambda)^2+\gamma^2}.$$

Figure 6.1 shows this function for various values of γ.

6.5 SUMMARY

In this chapter, we have encountered different uses of Green's functions. First, they are convenient for analysing different forms of perturbation theory. We have used Green's functions for calculating energy and wave function corrections for discrete spectra, and analysed the wave functions for continuous spectra. All these methods were based on the Dyson equation, which relates the Green's function G for a system described by a Hamiltonian

$$H = H_0 + V$$

to the Green's function G_0 of the unperturbed system:

$$G = G_0 + G_0 V G.$$

Iterating this equation gives the *Born series*

$$G = G_0 + G_0 V G_0 + G_0 V G_0 V G_0 + \dots$$

Finally we have used the Green's function concept to study the behaviour of a system which is *embedded* into another, usually larger, system. We have seen that the effect of the environment on a system which is coupled to that environment, can be captured by a *self-energy* Σ, which is in general a non-hermitian operator. The Green's function of the embedded system S is then given as

$$G(z) = (z - H_S + \Sigma)^{-1},$$

where, for an environmen with Green's function g_B, Σ is given by

$$\Sigma = \tau g_B \tau^\dagger.$$

Here τ is the part of the Hamiltonian which contains the coupling between system and bath.

For a single level, it is easy to see that the self-energy has two effects: its real part shifts the energy, and its imaginary parts broadens the level to a Lorentzian density of states.

Green's functions are very important in the study of interacting many-body systems. In that case, they are defined in a more general way. The interaction is treated in a perturbative way and the equations satisfied by the Green's functions can be illustrated in a more or less straightforward way by diagrams – these are called *Feynman diagrams*.

6.6 PROBLEMS

1. Consider a system of N noninteracting particles in a box. The particle states in the box are denoted $|\phi_n\rangle$, $n = 1, 2, \ldots$. The eigenenergies of these states are (in ascending order) ϵ_n, $n = 1, 2, \ldots$. We assume that these energies are nondegenerate.

   (a) Write the Green's function of the one-particle system in terms of the single particle states ϕ_n and their eigenenergies ϵ_n.

   (b) The ground state of the N-particle system is obtained by filling the *lowest N* eigenstates $\phi_n(\mathbf{r})$. The density is then given by

   $$n(\mathbf{r}) = \sum_{n=1}^{N} |\phi_n(\mathbf{r})|^2.$$

   Now consider the Green's function $G(z)$. A famous result from complex analysis (the *residue theorem*, see ch. 1 of the lecture notes) tells us that the line integral over a closed contour Γ in the complex plane of a complex function $f(z)$ is given by

   $$\oint_\Gamma f(z)dz = \sum_k 2\pi\mathrm{i}\ \mathrm{res}_{a_k} f(z),$$

   where a_i are all the singularities ('poles') of $f(z)$ *within* the contour Γ and $\mathrm{res}_a f(z) = \lim_{z\to a}(z-a)f(z)$. Using the residue theorem, show that the density can be written as

   $$\frac{1}{2\pi\mathrm{i}} \oint_\Gamma \langle \mathbf{r}|G(z)|\mathbf{r}\rangle\, dz = n(\mathbf{r}),$$

   where the contour Γ encloses the energy eigenvalues ϵ_n, $n = 1, \ldots, N$ on the real axis. This equation is frequently used in quantum transport.

2. We define a Green's function

   $$\left\langle k \left| e^{-\mathrm{i}H(t-t_0)/\hbar} \right| j \right\rangle = \mathrm{i}G(k, t; j, t_0),$$

   where $|k\rangle$ and $|j\rangle$ are basis functions.

   (a) How would you interpret this expression?

   (b) Show that the Green's function generates the time evolution, which, for a particle moving in three dimensions, using $|\mathbf{r}\rangle$ as basis states, reads:

   $$\psi(\mathbf{r}', t') = \mathrm{i}\int G(\mathbf{r}', t'; \mathbf{r}, t)\psi(\mathbf{r}, t)d^3r.$$

   (c) Using the fact that $\psi(\mathbf{r}, t)$ satisfies Schrödinger's equation, show that the Green's function satisfies

   $$(\mathrm{i}\hbar\partial_t - H)\, G(\mathbf{r}', t'; \mathbf{r}, t) = 0 \text{ for } t' > t$$

   and that

   $$\mathrm{i}G(\mathbf{r}', t; \mathbf{r}, t) = \delta^3(\mathbf{r} - \mathbf{r}').$$

(d) Show that for

$$H = -\frac{\hbar^2}{2m\nabla^2},$$

the Green's function is

$$G(\mathbf{r}', t'; \mathbf{r}, t) = -\mathrm{i}\left[\frac{m}{2\pi\mathrm{i}\hbar\,(t-t')}\right]^{3/2} \exp\left(\frac{\mathrm{i}m\left|\mathbf{r}'-\mathbf{r}\right|^2}{2\hbar(t-t')}\right).$$

7

SCATTERING IN CLASSICAL AND IN QUANTUM MECHANICS

Scattering experiments are perhaps the most important tool for obtaining detailed information on the structure of matter, in particular the interaction between particles. Examples of scattering techniques include neutron and X-ray scattering for liquids, atoms scattering from crystal surfaces, elementary particle collisions in accelerators. In most of these scattering experiments, a beam of incident particles hits a target which also consists of many particles. The distribution of scattering particles over the different directions is then measured, for different energies of the incident particles. This distribution is the result of many individual scattering events. Quantum mechanics enables us, in principle, to evaluate for an individual event the probabilities for the incident particles to be scattered off in different directions; and this probability is identified with the measured distribution.

Suppose we have an idea of what the potential between the particles involved in the scattering process might look like, for example from quantum mechanical energy calculations (programs for this purpose will be discussed in the next few chapters). We can then *parametrise* the interaction potential, i.e. we write it as an analytic expression involving a set of constants: the parameters. If we evaluate the scattering probability as a function of the scattering angles for different values of these parameters, and compare the results with experimental scattering data, we can find those parameter values for which the agreement between theory and experiment is optimal. Of course, it would be nice if we could evaluate the scattering potential directly from the scattering data (this is called the *inverse problem*), but this is unfortunately very difficult (if not impossible) as many different interaction potentials can have similar scattering properties as we shall see below.

Many different motivations for obtaining accurate interaction potentials can be given. One is that we might use the interaction potential to make predictions about the behaviour of a system consisting of many interacting particles, such as a dense gas or a liquid.

Scattering might be *elastic* or *inelastic*. In the former case the energy is conserved, in the latter energy disappears. This means that energy transfer takes place from the scattered particles to degrees of freedom which are not included explicitly in the system (inclusion of these degrees of freedom would cause the energy to be conserved). In this chapter we shall consider elastic scattering.

7.1 CLASSICAL ANALYSIS OF SCATTERING

A well known problem in classical mechanics is that of the motion of two bodies attracting each other by a gravitational force whose value decays with increasing separation r as $1/r^2$. This analysis is also correct for opposite charges which feel an attractive force of the same form (Coulomb's law). When the force is repulsive, the solution remains the same – we only have to change the sign of the parameter A which defines the interaction potential according to $V(r) = A/r$. One of the key experiments in physics which led to the notion that atoms consist of small but heavy kernels, surrounded by a cloud of light electrons, is *Rutherford scattering*. In this experiment, a thin gold sheet was bombarded with α-particles (i.e. helium-4 nuclei) and the scattering of the latter was analysed using detectors behind the gold film. In this section, we shall first formulate some new quantities for describing scattering processes and then calculate those quantities for the case of Rutherford scattering.

Rutherford scattering is chosen as an example here – scattering problems can be studied more generally; see Griffiths, chapter 11, section 11.1.1 for a nice description of classical scattering.

We consider scattering of particles incident on a so-called 'scattering centre', which may be another particle. The scattering centre is supposed to be at rest. This might not always justified in a real experiment, but in a standard approach in classical mechanics, the full two-body problem is reduced to a one-body problem with with a reduced mass, which is the present case (in problem 1 we will perform the same procedure for a quantum two-body system). The incident particles interact with the scattering centre located at $\mathbf{r} = \mathbf{0}$ by the usual scalar two-point potential $V(r)$ which satisfies the requirements of Newton's third law. Suppose we have a beam of incident particles parallel to the z-axis. The beam has a homogeneous density close to that axis, and we can define a *flux*, which is the number of particles passing a unit area perpendicular to the beam, per unit time. Usually, particles close to the z-axis will be scattered more strongly than particles far from the z-axis, as the interaction potential between the incident particles and scattering centre falls off with their separation r. An experimentalist cannot analyse the detailed orbits of the individual particles – instead a detector is placed at a large distance from the scattering centre and this detector counts the number of particles arriving at each position. You may think of this detector as a photographic plate which changes colour to an extent related to the number of particles hitting it. The theorist wants to predict what the experimentalist measures, starting from the interaction potential $V(r)$ which governs the interaction process.

7

In figure 7.1, the geometry of the process is shown. In addition a small cone, spanned by the spherical polar angles $d\vartheta$ and $d\varphi$, is displayed. It is assumed here that the scattering takes place in a small neighbourhood of the scattering centre, and for the detector the orbits of the scattered particles all seem to be directed radially outward from the scattering centre. The surface dA of the intersection of the cone with a sphere of radius R around the scattering centre is given by $dA = R^2 \sin\vartheta d\vartheta d\varphi$. The quantity $\sin\vartheta d\vartheta d\varphi$ is called *spatial angle* and is usually denoted by $d\Omega$. This $d\Omega$ defines a cone like the one shown in figure 7.1. Now consider the number of particles which will hit the detector within this small area per unit time. This number, divided by the total incident flux (see above) is called the *differential scattering cross section*, $d\sigma/d\Omega$:

$$\frac{d\sigma(\Omega)}{d\Omega} = \frac{\text{Number of particles leaving the scattering centre through the cone } d\Omega \text{ per unit time}}{\text{Flux of incident beam}}. \tag{7.1}$$

The differential cross section has the dimension of area (length^2).

First we realise ourselves that the problem is symmetric with respect to rotations around the z-axis, so the differential scattering cross section only depends on ϑ. The only two relevant parameters of the incoming particle then are its velocity and its distance b from the z-axis. This distance is called the *impact parameter* – it is also shown in figure 7.1.

We first calculate the scattering angle ϑ as a function of the impact parameter b. We perform this calculation for the example of Rutherford scattering, for which we have the standard Kepler solution which is now a hyperbola (see your classical mechanics lecture course). The potential for the Kepler problem is $V(r) = -A/r$. The orbits are given by specifying $r(t)$, $\vartheta(t)$. However, for the Kepler problem it is more convenient to focus on the *shape* of the orbitals, which is given as

$$r = \lambda \frac{1+\epsilon}{\epsilon\cos(\vartheta - C) - 1} \tag{7.2}$$

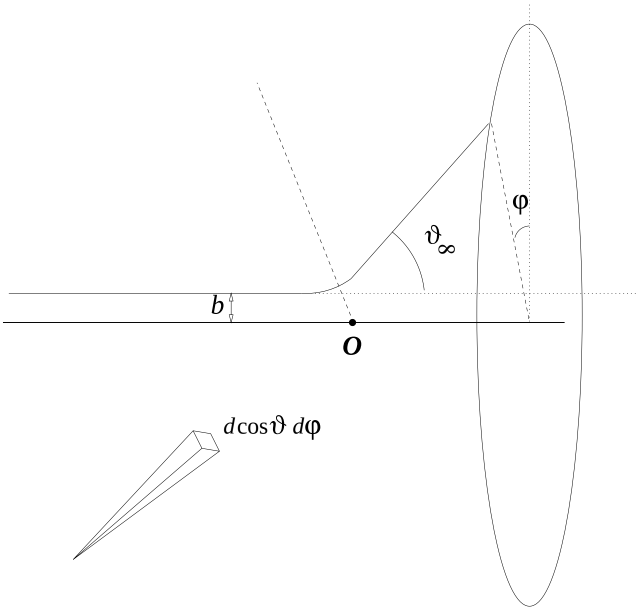


FIGURE 7.1: Geometry of the scattering process. b is the impact parameter and φ and ϑ are the angles of the orbit of the outcoming particle.

with

$$\epsilon = \sqrt{1 + \frac{2E\ell^2}{\mu A^2}}; \tag{7.3}$$

this parameter is called *eccentricity* – for a hyperbola, we have $|\epsilon| > 1$. Here, ℓ is the angular momentum and μ the reduced mass. The integration constant C reappears in the cosine because we have not chosen $\vartheta = 0$ at the perihelion – the closest approach occurs when the particle crosses the dashed line in figure 7.1 which bisects the in- and outgoing particle direction.

We know that for the incoming particles, for which $\vartheta = \pi$, $r \to \infty$, we have

$$\cos(\pi - C) = 1/\epsilon. \tag{7.4}$$

Because of the fact that cosine is even [$\cos x = \cos(-x)$] we can infer that the other value of ϑ for which r goes to infinity, and which corresponds to the outgoing direction occurs when the argument of the cosine is $C - \pi$, so that we find

$$\vartheta_\infty - C = C - \pi, \tag{7.5}$$

or $\vartheta_\infty = 2C - \pi$. The subscript ∞ indicates that this value corresponds to $t \to \infty$. From the last two equations we find the following relation between the scattering angle ϑ_∞ and ϵ:

$$\sin(\vartheta_\infty/2) = \cos(\pi/2 - \vartheta/2) = \cos(\pi - C) = 1/\epsilon. \tag{7.6}$$

We want to know ϑ_∞ as a function of b rather than ϵ however. To this end we note that the angular momentum is given as

$$\ell = \mu v_{\text{inc}} b, \tag{7.7}$$

where 'inc' stands for 'incident', and the total energy as

$$E = \frac{\mu}{2} v_{\text{inc}}^2, \tag{7.8}$$

so that the impact parameter can be found as

$$b = \frac{\ell}{\sqrt{2\mu E}}. \tag{7.9}$$

Using Eq. (7.3) and the fact that $\cot(x) = \sqrt{1-\sin^2(x)}/\sin(x)$, we can finally write (7.6) in the form:

$$\cot(\vartheta_\infty/2) = \sqrt{\epsilon^2 - 1} = \frac{2Eb}{|A|}. \tag{7.10}$$

From the relation between b and ϑ_∞ we can find the differential scattering cross section. The particles scattered with angle between ϑ and $\vartheta + d\vartheta$, must have approached the scattering centre with impact parameters between particular boundaries b and $b + db$. The number of particles flowing per unit area through the ring segment with radius b and width db is given as $j 2\pi b db$, where j is the incident flux. We consider a segment $d\varphi$ of this ring. Hence:

$$d\sigma(\Omega) = b(\vartheta) db d\varphi. \tag{7.11}$$

Relation (7.10) can be used to express the right hand side in terms of ϑ_∞:

$$d\sigma(\Omega) = \left(\frac{A}{2E}\right)^2 \cot(\vartheta/2)\, d\cot(\vartheta/2)\, d\varphi = \left(\frac{A}{2E}\right)^2 \cot(\vartheta/2) \frac{d\cot(\vartheta/2)}{d\vartheta} \frac{d\vartheta}{d\cos\vartheta} d\cos\vartheta d\varphi. \tag{7.12}$$

This can be worked out straightforwardly to yield:

$$\frac{d\sigma(\Omega)}{d\Omega} = \left(\frac{A}{4E}\right)^2 \frac{1}{\sin^4 \vartheta/2}. \tag{7.13}$$

This is known as the famous *Rutherford formula*.

7.2 QUANTUM SCATTERING

In quantum scattering, we know parts of the wave function describing the particles that are scattered off some other particles. Just as for classical scattering, we consider a two-particle collision, which we can reduce to a single particle problem in which an incident particle scatters off a static potential which is nonzero only near the origin – see problem 1. Assuming that the potential vanishes indeed outside of a sphere with radius $r_{\max}$ centered at the origin, we know the wave functions for the incident beam along the z-axis and the scattered waves. So, outside of the sphere, we have

$$\psi \propto e^{ikz} + f(\vartheta, \varphi) \frac{e^{ikr}}{r}.$$

We shall justify the form of the second term below. The main message for us at this stage is that we have an incident wave and a scattered wave. The amplitude of the scattered wave (the second term) depends on the polar angles ϑ, φ that are defined with respect to the z-axis along which the incident particles approach the scattering centre.

The detection of particles far from the scattering centre provides information about the amplitude $f(\vartheta, \varphi)$. Knowing that the outgoing wave e^{ikr}/r represents a unit flux per solid angle $d\Omega = \sin\vartheta d\vartheta d\varphi$, we conclude that

$$\frac{d\sigma}{d\Omega} = \left|f(\theta, \varphi)\right|^2.$$

Therefore, we need to find $f(\vartheta, \varphi)$ in order to determine the experimentally relevant quantity $d\sigma/d\Omega$.

In order to get a handle on this quantity, we start from the Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}).$$

For $V(\mathbf{r}) \equiv 0$, the solution to this equation would be an incoming plane wave. It turns out possible to write the solution to the Schrödinger equation with a potential formally as an integral expression. This is done using the Green's function formalism, discussed in chapter 6.4. The

Green's function is an *operator* – in this case that means that it depends on two positions $\mathbf{r}$ and $\mathbf{r}'$ – it is defined by

$$\left[E + \frac{\hbar^2}{2m}\nabla^2 - V(\mathbf{r})\right] G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}').$$

In fact, these are the matrix elements of the Green's function operator:

$$G(\mathbf{r},\mathbf{r}') = \langle \mathbf{r}|\hat{G}|\mathbf{r}'\rangle.$$

You may view the delta function on the right hand side as a unit operator, so that G may be called the *inverse* of the operator $E\mathbb{1}-\hat{H}$, where $\mathbb{1}$ is the unit operator, all in line with the theory of chapter 6.4. Now we want to calculate the unperturbed Green's function, i.e. the one for $V(\mathbf{r}) \equiv 0$, which we denote as usual as G_0:

$$\left[E + \frac{\hbar^2}{2m}\nabla^2\right] G_0(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}').$$

Before calculating G_0 let us assume we have it at our disposal. We then may write the solution to the *full* Schrödinger equation, i.e. including the potential V, in terms of a solution $\phi(\mathbf{r})$ to the 'bare' Schrödinger equation, that is, the Schrödinger equation with potential $V \equiv 0$:

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) + \int G_0(\mathbf{r},\mathbf{r}')V(\mathbf{r}')\psi(\mathbf{r}')\, d^3r', \tag{7.14}$$

which is an explicit form of the equation we met in chapter 6.4:

$$|\psi\rangle = |\phi\rangle + G_0 V|\psi\rangle.$$

For this case, Eq. (7.14) can easily be checked by substituting the solution into the full Schrödinger equation and using the fact that $E\mathbb{1} - \hat{H}_0$, acting on the Green's function $\hat{G}_0$, gives a delta-function:

$$(E\mathbb{1} - H_0)\,\psi(\mathbf{r}) = 0 + \int \delta(\mathbf{r}-\mathbf{r}')V(\mathbf{r}')\psi(\mathbf{r}')d^3r' = V(\mathbf{r})\psi(\mathbf{r}),$$

showing that $\psi(\mathbf{r})$ satisfies the Schrödinger equation for the full Hamiltonina $H_0 + V$.

Now we consider the scattering problem with an incoming beam of the form $\phi(\mathbf{r}) = \exp(\mathrm{i}\mathbf{k}_\mathrm{i}\cdot\mathbf{r})$ (the subscript 'i' denotes the incoming wave vector; do not confuse it with $\mathrm{i} = \sqrt{-1}$!). We see from Eq. (7.14) that this wave persists but that it is accompanied by a scattering term which is the integral on the right hand side. Now the wavefunction $\psi(\mathbf{r})$ is still very difficult to find, as it occurs in Eq. (7.14) in an implicit form. We can make the equation explicit if we assume that the potential $V(\mathbf{r})$ is small, so that the scattered part of the wave is much smaller than the wavefunction of the incoming beam. In a first approximation we might then replace $\psi(\mathbf{r}')$ on the right hand side of Eq. (7.14) by $\phi(\mathbf{r}')$ which is a plane wave:

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) + \int G_0(\mathbf{r},\mathbf{r}')V(\mathbf{r}')\phi(\mathbf{r}')\, d^3r' = e^{\mathrm{i}\mathbf{k}_\mathrm{i}\cdot\mathbf{r}} + \int G_0(\mathbf{r},\mathbf{r}')V(\mathbf{r}')e^{\mathrm{i}\mathbf{k}_\mathrm{i}\cdot\mathbf{r}'}\, d^3r'.$$

The key to the scattering amplitude is given by the notion that it must always be possible to write the solution (7.14) in the form:

$$\psi(\mathbf{r}) = e^{\mathrm{i}\mathbf{k}_\mathrm{i}\cdot\mathbf{r}} + f(\vartheta,\varphi)\frac{e^{\mathrm{i}kr}}{r}.$$

At this moment we hardly recognise this form in the expression obtained for the wavefunction. We first must find the explicit expression for the Green's function G_0. Without going through the derivation (see for example Griffiths, section 11.4[1]) we give it here:

$$G_0(\mathbf{r},\mathbf{r}') = -\frac{2m}{\hbar^2}\frac{e^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}$$

---

[1] The free Schrödinger equation is also known as the *Helmholtz equation.* You will probably derive the Green's function for this in your math course under that name.

with $k = \sqrt{2mE/\hbar^2}$.

We are interested in the wave function at the detector – this position is **r**. Therefore we can take **r** far from the origin. As the range of the potential is finite, we know that only contributions with $r' \ll r$ have to be taken into account. Taylor expanding the exponent occuring in the Green's function:

$$\left|\mathbf{r}-\mathbf{r}'\right| = \sqrt{r^2 - 2\mathbf{r}\cdot\mathbf{r}' + r'^2} \approx r - \frac{\mathbf{r}\cdot\mathbf{r}'}{r}$$

leads to

$$G(\mathbf{r},\mathbf{r}') = -\frac{2m}{\hbar^2}\frac{e^{\mathrm{i}kr}}{4\pi r}e^{-\mathrm{i}k\mathbf{r}\cdot\mathbf{r}'/r}.$$

The denominator does not have to be taken into account as it gives a much smaller contribution to the result for $r \gg 1/k$. Now we define $\mathbf{k}_\mathrm{f} = k\mathbf{r}/r$, i.e. $\mathbf{k}_\mathrm{f}$ is a wave vector corresponding to an outgoing wave from the scattering centre to the point **r** at the detector. We have

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) - \frac{2m}{\hbar^2}\frac{e^{\mathrm{i}kr}}{4\pi r}\int V(\mathbf{r}')e^{-\mathrm{i}\mathbf{k}_\mathrm{f}\cdot\mathbf{r}'}e^{\mathrm{i}\mathbf{k}_\mathrm{i}\cdot\mathbf{r}'}d^3r'.$$

This is precisely of the required form provided we set

$$f(\vartheta,\varphi) = -\frac{m}{2\pi\hbar^2}\int V(\mathbf{r}')e^{\mathrm{i}(\mathbf{k}_\mathrm{i}-\mathbf{k}_\mathrm{f})\cdot\mathbf{r}'}\,d^3r'.$$

This is the so-called *first Born approximation.* It is valid for weak scattering – higher order approximations can be made by iterative substitution for $\psi(\mathbf{r}')$ in the integral occurring in Eq. (7.14). In the first order Born approximation, the scattering amplitude $f(\vartheta,\varphi)$ is in fact a Fourier transform of the scattering potential.

For future reference, we note that the exact expression for the scattering amplitude is given by

$$f(\vartheta,\varphi) = -\frac{m}{2\pi\hbar^2}\int V(\mathbf{r}')e^{\mathbf{i}\mathbf{k}\cdot\mathbf{r}}\psi(\mathbf{r})\,d^3r'. \tag{7.15}$$

where **k** is the outgoing wave with polar angles ϑ,φ. This is the equation one would obtain from (7.14) without the Born approximation.

As an example, we consider the Coulomb potential which is strictly speaking not weak, but we pretend that we can use the Born approximation for this. The Coulomb potential has the form

$$V(r) = \frac{q_1q_2}{4\pi\epsilon_0}\frac{1}{r}.$$

The Fourier transform of this potential reads

$$V(k) = \frac{q_1q_2}{\epsilon_0}\frac{1}{k^2}.$$

Therefore, we immediately find for $f(\vartheta)$:

$$f(\vartheta) = -\frac{mq_1q_2}{2\pi\epsilon_0\hbar^2(\mathbf{k}_\mathrm{i}-\mathbf{k}_\mathrm{f})^2}.$$

The angle ϑ is hidden in $\mathbf{k}_\mathrm{i}-\mathbf{k}_\mathrm{f}$, the norm of which is equal to $2k\sin(\vartheta/2)$. The result therefore is, using $E = \hbar^2k^2/(2m)$:

$$\frac{d\sigma}{d\Omega} = \left[\frac{q_1q_2}{16\pi\epsilon_0E\sin^2(\vartheta/2)}\right]^2.$$

This is precisely the classical Rutherford formula, which also turns out to be the correct quantum mechanical result. This could not possibly be anticipated beforehand, but it is a happy coincidence.

7.3 THE OPTICAL THEOREM

We conclude this chapter by deriving an important theorem which results from the conservation of matter. This is the *optical theorem*, which also exists in classical optics (there it derives from total energy conservation). To study the optical theorem, we should therefore study the conservation of matter in quantum mechanics. For the Schrödinger equation we can derive a material conservation law a follows. Suppose we have a volume V in three-dimensional space which is bounded by a surface Γ. We calculate the rate of change of the amount of material in that space. That amount of material is given by

$$Q = \int_V |\psi|^2 \, d^3 r.$$

We then find for the rate of change (the dot indicates derivative with respect to time):

$$\dot{Q} = \int_V \langle \dot{\psi} | \psi \rangle \, d^3 r + \int_V \langle \psi | \dot{\psi} \rangle \, d^3 r.$$

The time-derivatives of the wave function are given by the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \hat{H} \psi(\mathbf{r}, t)$$

and its hermitian conjugate

$$-i\hbar \frac{\partial}{\partial t} \psi^*(\mathbf{r}, t) = \hat{H} \psi^*(\mathbf{r}, t).$$

Taking for the Hamiltonian the form $-\hbar^2/(2m)\nabla^2 + V(\mathbf{r})$, we obtain

$$\dot{Q} = -\frac{\hbar}{2im} \int_V \left[\psi^*(\mathbf{r}) \nabla^2 \psi(\mathbf{r}) - \psi(\mathbf{r}) \nabla^2 \psi^*(\mathbf{r}) \right] d^3 r$$

(note that the terms containing V cancel). Using Green's second identity, the volume integral can be transformed into a surface integral:

$$\dot{Q} = -\frac{\hbar}{2im} \int_\Gamma \left[\psi^*(\mathbf{r}) \nabla \psi(\mathbf{r}) - \psi(\mathbf{r}) \nabla \psi^*(\mathbf{r}) \right] d\mathbf{a}.$$

Here, $d\mathbf{a}$ is a vector, the norm of which is that of a small surface segment, and directed along the outward surface normal. The change in material can in this case only be caused by flow of material through the boundary. In particular, we have

$$\dot{Q} = -\int \mathbf{j} \cdot d\mathbf{a}.$$

so that we see that the expression for the particle flux is

$$\mathbf{j} = \frac{\hbar}{2im} \left[\psi^*(\mathbf{r}) \nabla \psi(\mathbf{r}) - \psi(\mathbf{r}) \nabla \psi^*(\mathbf{r}) \right]$$

Note that using the diverence theorem, the equation relating the material change to the surface integral can be related to the continuity equation:

$$\dot{Q} = \int_V \dot{\rho} d^3 r = -\int \mathbf{j} \cdot d\mathbf{a} = \int_V \nabla \cdot \mathbf{j} d^3 r.$$

As the volume is arbitrary, this equation can only hold when

$$\dot{\rho} + \nabla \cdot \mathbf{j} = 0.$$

Let's now turn again to the scattering problem, where we are dealing with an incoming and an outgoing wave:

$$\psi = \phi_{\text{in}} + \phi_{\text{out}}.$$

The total current in this wave is then given as

$$\mathbf{j} = \frac{\hbar}{2\mathrm{i}m}\left[\left(\phi_{\mathrm{in}}^* + \phi_{\mathrm{out}}^*\right)\nabla\left(\phi_{\mathrm{in}} + \phi_{\mathrm{out}}\right) - \left(\phi_{\mathrm{in}} + \phi_{\mathrm{out}}\right)\nabla\left(\phi_{\mathrm{in}}^* + \phi_{\mathrm{out}}^*\right)\right].$$

As this expression consists of a complex number minus its complex conjugate, we can write this in the form

$$\mathbf{j} = \frac{\hbar}{m}\mathrm{Im}\left[\left(\phi_{\mathrm{in}}^* + \phi_{\mathrm{out}}^*\right)\nabla\left(\phi_{\mathrm{in}} + \phi_{\mathrm{out}}\right)\right].$$

This can be rewritten as

$$\mathbf{j} = \mathbf{j}_{\mathrm{in}} + \mathbf{j}_{\mathrm{out}} + \frac{\hbar}{m}\mathrm{Im}\left(\phi_{\mathrm{in}}^*\nabla\phi_{\mathrm{out}} + \phi_{\mathrm{out}}^*\nabla\phi_{\mathrm{in}}\right) = \mathbf{j}_{\mathrm{in}} + \mathbf{j}_{\mathrm{out}} + \frac{\hbar}{m}\mathrm{Im}\left(\phi_{\mathrm{in}}^*\nabla\phi_{\mathrm{out}} - \phi_{\mathrm{out}}\nabla\phi_{\mathrm{in}}^*\right).$$

In a scattering problem we look at the scattered flux generated by a stationary flux of incoming particles. In the stationary limit, there is no generation or absorption of new matter: the particles flowing into some sphere centred around the scattering potential should also come out at the same rate. Therefore we must have:

$$\int_S \mathbf{j}\cdot d\mathbf{a} = 0,$$

as this expression calculates the total matter flux through the sphere's surface.

As we have written the total flux as a sum over an outgoing, an ingoing and a mixing term, we can also divide the total flux through the surface up into these three contributions. The flux of the incoming wave turns out to be zero! This is because the beam described by $\exp(\mathrm{i}kz)$ gives an incoming flux on one side of the sphere, and an *equal* outgoing flux on the other side. The outgoing flux through the sphere is given by

$$\int_S \mathbf{j}_{\mathrm{out}}\cdot d\mathbf{a} = \int \left|f(\theta)\right|^2 \frac{\hbar k}{m}\frac{1}{r^2}2\pi r^2 \sin\theta\, d\theta = \frac{\hbar k}{m}\sigma_{\mathrm{tot}}.$$

Therefore, we obtain

$$-k\sigma_{\mathrm{tot}} = \mathrm{Im}\int\left(\phi_{\mathrm{in}}^*\nabla\phi_{\mathrm{out}} - \phi_{\mathrm{out}}\nabla\phi_{\mathrm{in}}^*\right)\cdot d\mathbf{a}.$$

Using Green's theorem again, the expression on the right hand side can be reworked:

$$-k\sigma_{\mathrm{tot}} = \mathrm{Im}\int\left(\phi_{\mathrm{in}}^*\nabla^2\phi_{\mathrm{out}} - \phi_{\mathrm{out}}\nabla^2\phi_{\mathrm{in}}^*\right) d^3r.$$

Writing $\psi = \phi_{\mathrm{in}} + \phi_{\mathrm{out}}$, this can be rewritten as

$$-k\sigma_{\mathrm{tot}} = \mathrm{Im}\int_V\left[\phi_{\mathrm{in}}^*\nabla^2\left(\psi - \phi_{\mathrm{in}}\right) - \left(\psi - \phi_{\mathrm{in}}\right)\nabla^2\phi_{\mathrm{in}}^*\right] d^3r = \mathrm{Im}\int_V\left[\phi_{\mathrm{in}}^*\nabla^2\left(\psi\right) - \psi\nabla^2\left(\phi_{\mathrm{in}}^*\right)\right] d^3r. \tag{7.16}$$

Now we use that the full wave function satisfies the full Schrödinger equation

$$\nabla^2\psi = -k^2\psi + \frac{2mV(\mathbf{r})}{\hbar^2}\psi,$$

and the incoming wave satisfies the Schrödinger equation with potential 0:

$$\nabla^2\phi_{\mathrm{in}} = -k^2\phi_{\mathrm{in}}.$$

Putting this into (7.16) gives:

$$-k\sigma_{\mathrm{tot}} = \mathrm{Im}\int_V\left(-\phi_{\mathrm{in}}^*k^2\psi + \phi_{\mathrm{in}}^*\frac{2mV(r)}{\hbar^2}\psi + \psi k^2\phi_{\mathrm{in}}^*\right) d^3r.$$

The first and the third term in the integral cancel and we have

$$-k\sigma_{\mathrm{tot}} = \mathrm{Im}\int_V \phi_{\mathrm{in}}^*\frac{2mV(r)}{\hbar^2}\psi\, d^3r.$$

The last term is recognised as the exact scattering amplitude for $\theta = 0$ (up to a negative prefactor) – see Eq. (7.15).

Therefore, we find the *optical theorem*:

$$\sigma_{\text{tot}} = \text{Im}\frac{4\pi f(\theta = 0)}{k}.$$

If the wave is scattered, we see an attenuation in the forward direction compared to the case where the incoming particles would not be scattered. The forward scattering amplitude $f(0)$ is therefore related to the scattering of the particles (i.e. σ_{tot}). It is important to realise that the optical theorem holds exactly; in the first Born approximation, the theorem does not hold.

7.4 SUMMARY

In this chapter, we have analysed scattering of particles by a potential localised in a finite region around some point, which we take as the origin. The starting point of the quantum mechanical analysis is the wave function far from the scattering centre, which reads:

$$\psi(\mathbf{r}) = e^{\mathrm{i}\mathbf{k}_\mathrm{i}\cdot\mathbf{r}} + f(\vartheta,\varphi)\frac{e^{ikr}}{r}.$$

The first term represents the incoming wave (the wave vector $\mathbf{k}_\mathrm{i}$ is usually taken along the z-axis) and the second term represents the scattered wave whose amplitude can be measured as a function of ϑ and φ at a detector – this amplitude is given as the *differential cross section*:

$$\frac{d\sigma}{d\Omega} = \left|f(\vartheta,\varphi)\right|^2.$$

The shape of the function $f(\vartheta,\varphi)$ is determined by the interaction potential $V(\mathbf{r})$, which is often taken to be spherically symmetric: $V(\mathbf{r}) = V(r)$. The total cross section is the integral of the differential cross section:

$$\sigma_{\text{tot}} = \int \frac{d\sigma}{d\Omega}\sin\vartheta d\vartheta d\varphi.$$

The calculation of $f(\vartheta,\varphi)$ from the scattering potential V can be performed quite easily provided the interaction potential is weak. In that case, we can use the exact Green's function solution:

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) + \int G_0(\mathbf{r},\mathbf{r}')V(\mathbf{r}')\psi(\mathbf{r}')\, d^3r'$$

and its *first Born approximation*:

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) + \int G_0(\mathbf{r},\mathbf{r}')V(\mathbf{r}')\phi(\mathbf{r}')\, d^3r'.$$

From this last equation, the form of $f(\vartheta,\varphi)$ can readily be derived:

$$f(\vartheta,\varphi) = -\frac{m}{2\pi\hbar^2}\int V(\mathbf{r}')e^{\mathrm{i}(\mathbf{k}_\mathrm{i}-\mathbf{k}_\mathrm{f})\cdot\mathbf{r}'}\, d^3r'.$$

We usually set $\mathbf{q} = \mathbf{k}_\mathrm{i} - \mathbf{k}_\mathrm{f}$. For a spherically symmetric potential, the integral depends only on the length q of this vector. This length is given as

$$q = 2k\sin(\vartheta/2),$$

where k is the length of the wave vector of the in- and outgoing waves, and ϑ is the scattering angle.

The Born approximation yields the *exact* result for scattering off a Coulomb potential (the Rutherford formula):

$$\frac{d\sigma}{d\Omega} = \left[\frac{q_1 q_2}{16\pi\epsilon_0 E\sin^2(\vartheta/2)}\right]^2.$$

Finally, we have derived the *optical theorem* as a consequence of the conservation of matter:

$$\sigma_{\text{tot}} = \text{Im}\frac{4\pi f(\vartheta = 0)}{k}.$$

7.5 PROBLEMS

1. Show that for two identical particles, 1 and 2, with coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$, the kinetic energy can be written as

$$T = -\frac{\hbar^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) = -\frac{\hbar^2}{2M}\nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2,$$

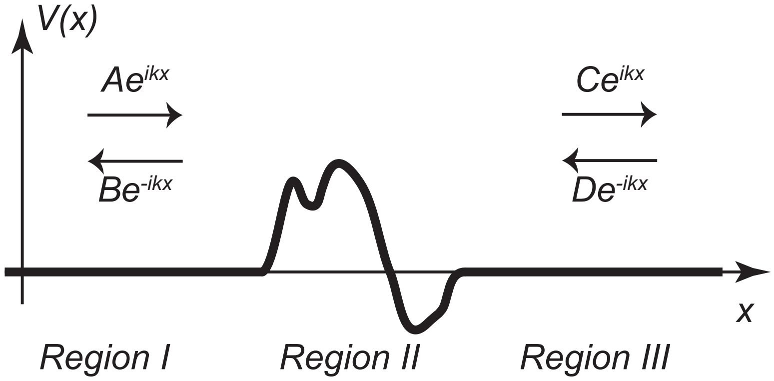
where $\nabla_{\mathbf{r}}$ denotes a gradient with respect to $\mathbf{R} = \frac{\mathbf{r}_1+\mathbf{r}_2}{2}$ and $\nabla_{\mathbf{r}}$ a gradient with respect to $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Finally, $M = 2m$ and $\mu = m/2$.

The Hamiltonian then becomes

$$H = \left[-\frac{\hbar^2}{2M}\nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2 + V(r)\right]\psi(\mathbf{r},\mathbf{R}) = E\psi(\mathbf{r},\mathbf{R}).$$

Also show that $\psi(\mathbf{r},\mathbf{R})$ can be written as $\phi(\mathbf{r})\chi(\mathbf{R})$. Find suitable eigen-equations for ϕ and χ (this is separation of variables).

2. Consider the 1-D scattering problem illustrated in the figure below, with an arbitrary localised potential (without any particular spatial symmetry) in Region II, and $V(x) = 0$ in Regions I and III.



In Regions I and III, the solutions to the time-independent Schrodinger equation take on the form

$$\langle x|\psi\rangle = \begin{cases} Ae^{\mathrm{i}kx} + Be^{-\mathrm{i}kx} & \text{in Region I} \\ Ce^{\mathrm{i}kx} + De^{-\mathrm{i}kx} & \text{in Region III} \end{cases},$$

where $k = \sqrt{2mE}/\hbar$. For scattering from left to right, $\{A,B,C,D\} = \{1, S_{\mathrm{LL}}, S_{\mathrm{RL}}, 0\}$. For scattering from right to left, $\{A,B,C,D\} = \{0, S_{\mathrm{LR}}, S_{\mathrm{RR}}, 1\}$.

(a) Show that, if $\langle x|\psi\rangle$ is a solution to the stationary Schrödinger equation, so is its complex conjugate $\langle x|\psi\rangle^*$. This is a consequence of time reversal invariance.

(b) Calculate the flux for a wave function $\psi(x) = \exp(\mathrm{i}kx)$. Show that particle number conservation implies that

$$|A|^2 - |B|^2 = |C|^2 - |D|^2.$$

(c) Show $S_{\mathrm{LL}}S_{\mathrm{RL}}^* = -S_{\mathrm{RR}}^*S_{\mathrm{LR}}$ (Hint: use the linearity of the Schrodinger equation).

(d) Using the linearity of the Schrödinger equation and the conservation law found in (b), show that S_{RL}=S_{LR}. Thus, the transmission amplitude through the potential is symmetric, even though the potential has no left-right symmetry.

(e) Show that the *scattering matrix*, defined by

$$S = \begin{pmatrix} S_{\mathrm{LL}} & S_{\mathrm{LR}} \\ S_{\mathrm{RL}} & S_{\mathrm{RR}} \end{pmatrix},$$

is unitary.

(f) Are (c) and (d) necessary or simply sufficient conditions for S to be unitary?

3. Obtain the low-energy cross-section for the potential given by

$$V(r) = \begin{cases} -V_0 & \text{for } r < a; \\ 0 & \text{for } r > a. \end{cases}$$

4. Consider scattering off a spherical cage (such as a bucky ball), which is described by a δ-function

$$V(r) = g\delta(r-a).$$

Find the scattering amplitude.

5. In time-dependent perturbation theory, we can calculate the probability to move from some initial state to some final state as a result of a perturbation which was turned on for a finite time. Specifically, we write the Hamiltonian in the form

$$H = H_0 + H'(t),$$

where H_0 is the unperturbed Hamiltonian with eigenstates ϕ_n and eigenenergies E_n.

The time-dependent solutions of H_0 have the form

$$\left|\psi_0(t)\right\rangle = \sum_n c_n e^{-\mathrm{i}E_n t/\hbar} \left|\phi_n\right\rangle,$$

where the c_n are time-independent expansion coefficients.

For the full Hamiltonian, we write the time-dependent solutions in the form

$$\left|\psi(t)\right\rangle = \sum_n c'_n(t) e^{-\mathrm{i}E_n t/\hbar} \left|\phi_n\right\rangle.$$

Note that the expansion coefficients $c'_n(t)$ are now time-dependent. We have solved the time-dependent problem if we know the explicit time-dependence of these coefficients.

(a) From now on, we will drop the prime from the coefficients c of the full solution. Show that the c_n satisfy the equation

$$\mathrm{i}\hbar\dot{c}_k = \sum_n \langle\phi_k|H'|\phi_n\rangle\, e^{\mathrm{i}\omega_{kn}t} c_n = \sum_n H'_{kn} e^{\mathrm{i}\omega_{kn}t} c_n, \tag{7.17}$$

where $\omega_{kn} = (E_k - E_n)/\hbar$.

(b) Now suppose that we start off at $t \to -\infty$ with a state $\left|\phi_l\right\rangle$ for a single n. We want to know the probability to end up in some other state $\left|\phi_m\right\rangle$ with $m \neq l$. Show that the coefficient can be solved by requiring that c_n on the right hand side of (7.17) can be replaced by δ_{nl}. Which condition on H' would justify this approach? Show that now

$$c_m = \frac{1}{\mathrm{i}\hbar}\int_{-\infty}^{t} H'_{ml}(t') e^{\mathrm{i}\omega_{ml}t'} dt'.$$

(c) Now suppose that the Hamiltonian H' is switched on at $-T$ and switched off again at T where $T \gg \hbar\omega_{ml}$, and that it can be assumed to be constant in between these two times. Show that then

$$c_m = \frac{2}{\mathrm{i}\hbar} H'_{ml} \frac{\sin(\omega_{ml}T)}{\omega_{ml}}.$$

From the fact that

$$\lim_{x\to\infty} \sin^2(ax)/(a^2x) = \frac{\pi}{2}\delta(a),$$

derive that the rate at which the probability to find the particle in state m after T increases, is given by

$$\frac{2\pi}{\hbar^2}|H'_{ml}|^2\delta(\omega_{ml}).$$

(d) Now consider a scattering problem, where the unperturbed states are given by plane waves with wave vector $\mathbf{k}$, which are as usual denoted by $|\mathbf{k}\rangle$. The transition rate from an initial state $|\mathbf{k}_\mathrm{I}\rangle$ to *any* final state $|\mathbf{k}_\mathrm{F}\rangle$ is then given by

$$R = \int \frac{2\pi}{\hbar}\left|\langle\mathbf{k}_\mathrm{F}|\, H'\, |\mathbf{k}_\mathrm{I}\rangle\right|^2 \delta(E_\mathrm{I} - E_\mathrm{F})d^3k_\mathrm{F}.$$

Show from this that the differential cross section for scattering is given by

$$\frac{d\sigma}{d\Omega} = \frac{2\pi m^2}{\hbar^4}\left|\langle\mathbf{k}_\mathrm{F}|\, H'\, |\mathbf{k}_\mathrm{I}\rangle\right|^2,$$

i.e., we have recovered the first Born approximation. In the expressions, we assume the normalisation

$$\langle\mathbf{k}|\mathbf{r}\rangle = \frac{e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}}.$$

6. Calculate the differential cross section $d\sigma/d\Omega$ in the first Born approximation for a potential of the form

$$V(r) = -V_0 e^{-r/R}.$$

Hint: $\int_0^\infty e^{-ar}r\,dr = 1/a^2$. Note that a may be complex in this expression!

7. *Scattering off a charge distribution*
We consider scattering of particles with charge e off a charge distribution $\rho(\mathbf{r})$ located near the origin. The electrostatic potential felt by the incoming particles is given as the solution to the Poisson equation (in SI units):

$$-\nabla^2\varphi(\mathbf{r}) = \rho(\mathbf{r})/\epsilon_0.$$

(a) Show that in Fourier space, the Poisson equation takes the form

$$k^2\tilde{\varphi}(\mathbf{k}) = \tilde{\rho}(\mathbf{k})/\epsilon_0,$$

where $\tilde{\varphi}$ and $\tilde{\rho}$ denote the Fourier transforms of $\varphi(\mathbf{r})$ and $\rho(\mathbf{r})$.
The potential energy is given by $e\varphi(\mathbf{r})$.

(b) Show that the scattering amplitude is given by

$$f(\theta,\phi) = -\frac{me}{2\pi\hbar^2}\frac{\tilde{\rho}(\mathbf{q})}{\epsilon_0 q^2}.$$

8. In the Born approximation, the forward-scattering amplitude (the scattering angle $\vartheta = 0$) is real, yielding $\sigma_\mathrm{tot} = 0$, and therefore it appears to be in contradiction with the optical theorem.

(a) Resolve this contradiction.

(b) Show that the second-order Born approximation yields a perturbative correction to the scattering amplitude which, for $\vartheta = 0$, is given by

$$f^{(2)}(\vartheta = 0) = \left(\frac{m}{2\pi\hbar^2}\right)^2 e^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\int V(\mathbf{r})\frac{e^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}V(\mathbf{r}')\,d^3r\,d^3r'.$$

(c) Show that the optical theorem is satisfied to second order in the potential V.

8

SYSTEMS OF HARMONIC OSCILLATORS – PHONONS AND PHOTONS

In this chapter, we start with one of the main topics of this course: dealing with many-particle systems. These systems are particularly hard to treat when the particles interact. In this chapter, we will stick to systems of non-interacting particles. Furthermore, we shall restrict ourselves to systems with boson character. Notwithstanding these restrictions, you will find the material sometimes complicated. It may therefore be helpful to always keep the general idea behind the analysis in mind: we try to decompose the systems at hand into normal modes, which leads us to view them as composed of non-interacting harmonic oscillators. A good understanding of the quantum description of the harmonic oscillator is therefore the starting point from which we proceed – this description is given in the next section. We shall apply the general approach to two types of systems: phonons (vibrational modes occurring in solids) and photons (vibrational excitation of the electromagnetic field). The last system leads us to consider an exciting observable quantum phenomenon: the Casimir effect.

8.1 CREATING AND ANNIHILATING QUANTA IN THE HARMONIC OSCILLATOR

In this section we review the analysis of a well-known quantum mechanical problem: the harmonic oscillator in one dimension, which is described by the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2.$$

To solve for the spectrum, we introduce the operators

$$a = \sqrt{\frac{m\omega}{2\hbar}} \left(x + \mathrm{i}\frac{p}{m\omega} \right); \qquad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(x - \mathrm{i}\frac{p}{m\omega} \right).$$

It is easy to see that $a^\dagger$ is indeed the hermitian conjugate of a as the notation suggests. The commutation relation between x and p:

$$[x, p] = \mathrm{i}\hbar$$

can be used to derive the commutation relation between the a's:

$$[a, a^\dagger] = 1 \qquad \text{check this!.}$$

Furthermore, the Hamiltonian can be formulated in terms of the a's:

$$H = \left(a^\dagger a + 1/2 \right) \hbar\omega$$

which can also be checked by substituting the expressions for the a-operators in terms of x and p.

The spectrum can now be found from the commutation relations between the a's. Suppose we have an eigenstate $|\psi_n\rangle$ with energy $\hbar\omega(n+1/2)$. The notation here suggests that n is integer, but we leave this open for the moment. We now show that $a|\psi_n\rangle$ is proportional to another eigenstate of the Hamiltonian with energy $\hbar\omega(n-1/2)$:

$$Ha|\psi_n\rangle = \hbar\omega\left(a^\dagger a + 1/2\right) a|\psi_n\rangle = a\hbar\omega\left(a^\dagger a - 1/2\right)|\psi_n\rangle$$

where we have used the commutation relation between the a's in order to move the a in front of the ket vector to the left. This decreases the term in parentheses by 1. In a similar fashion, one can show that $a^\dagger|\psi_n\rangle$ is an eigenstate of H with eigen-energy $\hbar\omega\left(a^\dagger a + 3/2\right)$. Thus, a has the effect of *decreasing* the energy by $\hbar\omega$ and $a^\dagger$ will *increase* the energy when acting on an eigenstate.

In order to find the spectrum, we use a physical argument. The spectrum must be bounded from below as the potential does not assume infinitely negative values. Therefore, if we start with some $|\psi_E\rangle$ and act successively on it with the lowering operator a, we must have at some point:

$$a^n|\psi_E\rangle = 0, \tag{8.1}$$

because otherwise the spectrum would not be bounded from below. Let us call $a^{n-1}|\psi_E\rangle = |\psi_G\rangle$. Then $a|\psi_G\rangle = 0$. Therefore,

$$a^\dagger a|\psi_G\rangle = 0, \tag{8.2}$$

that is, $|\psi_0\rangle$ is an eigenstate of $\hat{n} = a^\dagger a$ with eigenvalue $n = 0$, or

$$|\psi_G\rangle = |\psi_0\rangle.$$

8

This eigenstate has energy eigenvalue $\hbar\omega/2$, which is found when acting on it with the Hamiltonian. Acting with $a^\dagger$ on $|\psi_0\rangle$ we obtain an eigenstate $|\psi_1\rangle$ (up to a constant) with energy eigenvalue $3\hbar\omega/2$ etc. Acting n times with $|a^\dagger\rangle$ on $|\psi_0\rangle$, we obtain an eigenstate $|\psi_n\rangle$ (up to a constant) with energy $\hbar\omega(n+1/2)$.

The operator $a^\dagger a$ is called *number operator*, denoted by $\hat{n}$, and H can now be written as $\hbar\omega(\hat{n}+1/2)$. $|\psi_n\rangle$ is an eigenstate of $\hat{n}$ with eigenvalue n. The norm of $a^\dagger|\psi_n\rangle$ can be expressed in that of $|\psi_n\rangle$:

$$\left\langle a^\dagger\psi_n\middle| a^\dagger\psi_n\right\rangle = \left\langle\psi_n\middle| aa^\dagger\psi_n\right\rangle = \left\langle\psi_n\middle|(a^\dagger a+1)\psi_n\right\rangle = (n+1)\left\langle\psi_n\middle|\psi_n\right\rangle. \tag{8.3}$$

Therefore, if $|\psi_n\rangle$ is normalised, $a^\dagger|\psi_n\rangle/\sqrt{n+1}$ is normalised too, and normalised states $|\psi_n\rangle$ can be constructed from a normalised state $|\psi_0\rangle$ according to:

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n|\psi_0\rangle. \tag{8.4}$$

Using the commutation relations for a, $a^\dagger$, it is also possible to show that states belonging to *different* energy levels are mutually orthogonal:

$$\left\langle\psi_n\middle|\psi_m\right\rangle \propto \left\langle\psi_0\middle| a^n\left(a^\dagger\right)^m\middle|\psi_0\right\rangle. \tag{8.5}$$

Moving the a's to the right and the $a^\dagger$ to the left by application of the commutation relations leads to a form involving the lowering operator a acting on $|\psi_0\rangle$ and/or $\langle\psi_0|$ followed by $a^\dagger$'s, which both vanish.

Exercise: show that $\langle\psi_2|\psi_3\rangle$ vanishes indeed.

We have succeeded in finding the energy spectrum but it might seem that we have not made any progress in finding the form of the eigenfunctions $|\psi_n\rangle$. However, we have a simple differential equation defining the ground state $|\psi_0\rangle$:

$$\sqrt{\frac{m\omega}{2\hbar}}\left(x + \mathrm{i}\frac{p}{m\omega}\right)\langle x|\psi\rangle = 0. \tag{8.6}$$

We now transform to the coordinates

$$\tilde{x} = \sqrt{\frac{m\omega}{\hbar}}x; \qquad \tilde{p} = \frac{1}{\sqrt{m\hbar\omega}}p,$$

and find that the operator $\tilde{p}$ can be related to $\tilde{x}$ as follows:

$$\tilde{p} = \frac{1}{\sqrt{m\omega\hbar}}\frac{\hbar}{\mathrm{i}}\frac{d}{dx} = \frac{1}{\mathrm{i}}\frac{d}{d\tilde{x}}.$$

We have

$$a\langle\tilde{x}|\psi_0\rangle = \frac{1}{\sqrt{2}}(\tilde{x} + \mathrm{i}\tilde{p})\langle\tilde{x}|\psi_0\rangle = 0 \tag{8.7}$$

or:

$$\left(\tilde{x} + \frac{d}{d\tilde{x}}\right)\psi_0(\tilde{x}) = 0 \tag{8.8}$$

The solution can immediately be found as:

$$\psi(x) = \text{Const. } e^{-\tilde{x}^2/2} \tag{8.9}$$

in accordance with the result obtained in the direct method which can be found in any quantum mechanics textbook. The normalisation constant is found as

$$\text{Const. } = (m\omega/\hbar\pi)^{1/4} \tag{8.10}$$

(check this for yourself!).

Using (8.4), we can find the solution for general n from:

$$\langle\tilde{x}|\psi_n\rangle = \left(\frac{m\omega}{\hbar}\pi\right)^{1/4}\frac{1}{\sqrt{n!2^n}}\left(\tilde{x} + \mathrm{i}\tilde{p}\right)^n e^{-\tilde{x}^2/2}. \tag{8.11}$$

We have seen that the harmonic oscillator has an *equidistant* spectrum, and if the energy is given as $\hbar\omega(n+1/2)$, we say that there are n quanta in the system. However, there is nothing which prevents us from viewing those quanta as particles, each of which carry the same energy $\hbar\omega$. The operators a and $a^\dagger$ annihilate, respectively create a particle. In the next chapter we shall see that we can derive creation and annihilation operators for electrons and other massive particles, within an algebraic structure which is very similar to that of one of more harmonic oscillators.

The fact that we do not distinguish between particles and energy quanta is the reason why in physics we often mix the two: photons and phonons are sometimes viewed as energy quanta related to harmonic oscillator excitations, but often we speak about them as if they were particles.

In the next sections we consider the two major examples of harmonic oscillator systems, lattice vibrations and the electromagnetic field, in some detail. We shall see that lattice vibrations are conveniently analysed in terms of normal modes: vibrational excitations in which all the degrees of freedom oscillate at the same frequency. These are the well-known *normal* modes. From classical mechanics, we know that the Hamiltonian can be written as a sum of *independent* harmonic oscillators, one for each normal mode. Similarly, the Hamiltonian of the electromagnetic field in vacuum can be viewed as a sum of independent harmonic oscillators. The programme of this chapter is then to turn all these classical harmonic oscillators into quantum oscillators which we then can formulate in terms of their creation and annihilation operators. The energy quanta of the lattice vibrations are called phonons, and those of the electromagnetic field are the photons.

First, however we consider particular states of the quantum harmonic oscillator which can be related to classical oscillatory excitations.

8.1.1 COHERENT STATES

The quantum solutions of the harmonic oscillator seem completely unrelated to those we know from classical mechanics. In this section, we briefly describe eigenfunctions that are closer to the classical solutions. These *coherent states* are not eigenstates of the Hamiltonian, but more general: they can (for each time t) be expanded in terms of these eigenstates.

A coherent state $|\alpha\rangle$ is by definition an eigenstate of the annihilation operator a:

$$a|\alpha\rangle = \alpha|\alpha\rangle$$

where it is important to note that a is an operator and α a number.

The properties of coherent states will be addressed in problems 1 and 2. Here we summarise a few:

- Coherent states are minimum-uncertainty states. Minimum-uncertainty states are states satisfying
$$\Delta x \Delta p = \hbar/2,$$
i.e. they match the lower bound of the Heisenberg uncertainty relation.
- They represent minimal-uncertainty wave packets that oscillate back and forth in time, just like a classical harmonic oscillator.
- If external driving is applied, a system evolves towards a coherent state. In problem 3 it will be shown that a coherent state solves the time-dependent Schrödinger equation for the harmonic oscillator.

8.2 QUANTIZATION OF THE LINEAR CHAIN OF ATOMS

In this section, we study phonon systems. Rather than doing this in detail for crystals with different symmetries, we study a homogeneous one-dimensional chain of particles, placed at a distance d from each other and moving along the chain. The particles are coupled by springs. If we take the distance between these particles to be small while reducing the mass accordingly, we can study the continuum limit, which is the elastic string.

To be specific, we formulate the Hamiltonian of this system

$$H = \sum_{n=0}^{N-1} \frac{p_n^2}{2m} + \frac{K}{2} \sum_{n=0}^{N-1} \left(q_{n+1} - q_n\right)^2,$$

where q_n denotes the deviation from the equilibrium position of oscillator n. We assume periodic boundary conditions, so that $q_N \equiv q_0$. The distance between the oscillators is a. We first consider this problem classically.

A Fourier basis is formed by the waves

$$u_{k_j n} = \frac{1}{\sqrt{N}} e^{\mathrm{i}k_j na}; \qquad k_j = \frac{2\pi j}{Na}, \quad j = 1, \dots, N,$$

and this is used to define (sums over n and m run from 0 to $N-1$):

$$q_{k_j} = \frac{1}{\sqrt{N}} \sum_n e^{\mathrm{i}k_j na} q_n.$$

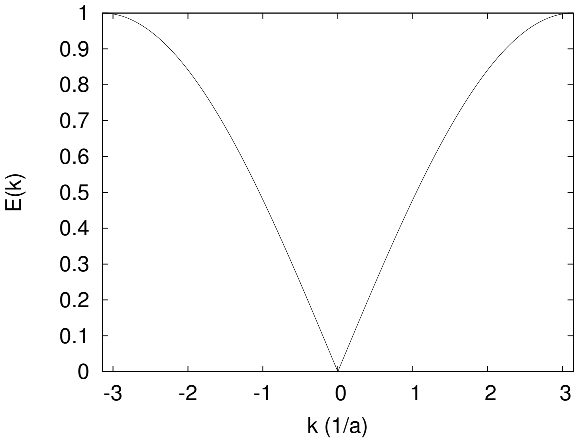


FIGURE 8.1: The dispersion relation for the harmonic chain.

From the definition of the waves u_{kn}, it follows that $u^*_{k_j n} = u_{-k_j n} = u_{-nk_j}$. We can define the inverse Fourier transform (assuming that N is even) as

$$q_n = \frac{1}{\sqrt{N}} \sum_{j=-N/2}^{N/2-1} e^{-\mathrm{i}k_j na} q_{k_j} = \sum_{j=-N/2}^{N/2-1} u^*_{k_j n} q_{k_j} = \sum_{j=-N/2}^{N/2-1} u_{-k_j n} q_{k_j}.$$

Here we have used that $u_{k_j n} = u_{k_{j+mN} n}$ for any integer m in order to let j run from $-N/2$ to $N/2-1$ rather than from 0 to $N-1$. We may verify that applying the Fourier transform and its inverse does not change q_n:

$$q_n = \frac{1}{\sqrt{N}} \sum_{j=-N/2}^{N/2-1} e^{-\mathrm{i}k_j na} q_{k_j} = \frac{1}{N} \sum_{j=-N/2}^{N/2-1} e^{-\mathrm{i}k_j na} \sum_m e^{\mathrm{i}k_j ma} q_m = \frac{1}{N} \sum_m \sum_{j=-N/2}^{N/2-1} e^{\mathrm{i}k_j (m-n)a} q_m = q_n.$$

Here we have used

$$\frac{1}{N} \sum_{j=-N/2}^{N/2-1} e^{\mathrm{i}k_j na} = \delta_{n,0}$$

in the last equality. From now on we shall use the short hand notation

$$\sum_{j=-N/2}^{N/2-1} f(k_j) \equiv \sum_k f(k).$$

The kinetic energy $T = \sum_n p_n^2/(2m)$ can then be written as $\sum_k p_k p_{-k}/(2m)$ by applying the same Fourier transformation to p_n as we have done for q_n (exercise!). The potential energy $V = \frac{K}{2} \sum_{n=0}^{N-1} \left(q_{n+1} - q_n\right)^2$ can furthermore be written as (another exercise!)

$$V = 2K \sum_k \sin^2\left(\frac{ka}{2}\right) q_k q_{-k},$$

so that the full Hamiltonian can be written in the form

$$H = \sum_k \left[\frac{p_k p_{-k}}{2m} + \frac{m\omega_k^2}{2} q_k q_{-k} \right]$$

where $\omega_k = 2\sqrt{K/m} \sin(|ka|/2)$. This relation between the energy ω_k and **k** is called the *dispersion relation.* It is represented in figure 8.1.

Now we consider the quantum version of this harmonic chain. This means that we consider the masses that move as quantum (point) particles. Their coordinates p_n, q_n then become operators with commutation relations $[p_n, p_{n'}] = [q_n, q_{n'}] = 0$; $[p_n, q_{n'}] = -\mathrm{i}\hbar\delta_{nn'}$.

From these commutation relations, their counterparts for the Fourier-transforms can be formulated.

$$\left[p_k, q_{k'}\right] = -\mathrm{i}\hbar\delta_{k,-k'}.$$

To obtain this result, we write out p_k and q_k as Fourier sums:

$$[p_k, q_{k'}] = \frac{1}{N}\sum_{nm} e^{\mathrm{i}kna} e^{\mathrm{i}k'ma}[p_n, q_m] = -\frac{\mathrm{i}\hbar}{N}\sum_n e^{\mathrm{i}(k+k')n} = -\mathrm{i}\hbar\delta_{k,-k'}. \tag{8.12}$$

Now we define new operators a_k and $a_k^\dagger$ as follows:

$$a_k = \sqrt{\frac{m\omega_k}{2\hbar}} q_k + \mathrm{i}\sqrt{\frac{1}{2\hbar m\omega_k}} p_k \tag{8.13}$$

and

$$a_k^\dagger = \sqrt{\frac{m\omega_k}{2\hbar}} q_{-k} - \mathrm{i}\sqrt{\frac{1}{2\hbar m\omega_k}} p_{-k}. \tag{8.14}$$

It is readily seen that the Hamiltonian can be written as

$$H = \sum_{k=-\pi/a}^{\pi/a}\left(a_k^\dagger a_k + \frac{1}{2}\right)\hbar\omega_k.$$

where the operators a_k and $a_k^\dagger$ satisfy commutation relations that can easily be derived from (8.12):

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}.$$

From these relations we find

$$\left[H, a_k^\dagger\right] = \hbar\omega_k a_k^\dagger; \qquad [H, a_k] = -\hbar\omega_k a_k.$$

Denoting by $|0_k\rangle$ the ground state of the harmonic oscillator, any state can then be constructed by acting on the ground state with the creation operators $a_k^\dagger$:

$$|\{n_k\}\rangle = \prod_k \frac{1}{\sqrt{n_k!}}\left(a_k^\dagger\right)^{n_k}|0\rangle.$$

We see that for the simplest possible elastic medium, i.e. a one-dimensional harmonic chain, quantization leads to a formulation in terms of creation and annihilation operators. It turns out that the standard quantum field theories can all be formulated along these lines. We shall however not go into details here.

Note that the creation and annihilation operators create and annihilate quantum excitations of the system, respectively. These excitations can be viewed as particles. Such particles are called *phonons*. We can also study phonons in three-dimensional grids. In that case, we have not only a quantum number **k** which now has become a three dimensional vector, indicating the direction in which the phonon wave travels, but also three components of the excitations: two transversal, and one longitudinal mode. This means that the creation and annihilation operators must have three components, corresponding to the polarization (i.e. the direction in which the atoms move) *in addition* to **k**: we therefore have operators $a_{\mathbf{k},\alpha}$ and $a_{\mathbf{k},\alpha}^\dagger$, where two of the α's correspond to the two transverse, and the third one to the longitudinal mode. We can then also denote these operators as vector operators $\mathbf{a}_\mathbf{k}$, $\mathbf{a}_\mathbf{k}^\dagger$, where the boldface indicates that, for example, $\mathbf{a}_\mathbf{k}$ contains three components α. We shall come back to phonons when dealing with electron-phonon coupling in the next chapter.

8.3 THE QUANTUM THEORY OF ELECTROMAGNETIC RADIATION

In previous courses, you have covered the classical theory of electromagnetism, which is summarized in the Maxwell equations. On the other hand you know that light has quantum properties, as is demonstrated by the existence of light quanta – the *photons*. In this section, we first review the classical theory of light in vacuum and then we formulate a quantum theory for this. As we shall see, the procedure is quite analogous to that which was applied to crystal vibrations, the phonons, in the previous section.

8.3.1 CLASSICAL ELECTROMAGNETISM

The Maxwell equations read (in SI units):

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \rho/\epsilon_0;\\
\nabla \cdot \mathbf{B} &= 0;\\
\nabla \times \mathbf{B} - \mu_0\epsilon_0 \frac{\partial E}{\partial t} &= \mu_0 \mathbf{j};\\
\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0.
\end{aligned}$$

For our purposes, it is convenient not to start from the fields **E** and **B** but from the vector and scalar potentials, **A** and ϕ. Note that fields and potentials all depend on position and time, e.g. $\mathbf{A} = \mathbf{A}(\mathbf{r}, t)$. The electric and magnetic fields are obtained from the potentials as follows:

$$\begin{aligned}
\mathbf{B} &= \nabla \times \mathbf{A};\\
\mathbf{E} &= -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}.
\end{aligned}$$

The potentials enjoy a gauge freedom: changing them according to

$$\begin{aligned}
\mathbf{A} &\to \mathbf{A} + \nabla\chi;\\
\phi &\to \phi - \frac{\partial \chi}{\partial t}
\end{aligned}$$

leaves the electric and magnetic fields unchanged. We choose χ such that

$$\nabla \cdot \mathbf{A} = 0,$$

which is known as the *Coulomb* or *transverse gauge*. Substituting the expression for the electric field in the first of the Maxwell equations given above, and using the Coulomb gauge condition, we obtain:

$$\nabla \cdot \left(-\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}\right) = -\nabla^2\phi = \rho/\epsilon_0.$$

We want to describe the electromagnetic field in vacuum, so $\rho \equiv 0$ and therefore $\phi =$ must be constant in space and time.

Now we use the expressions for both **E** and **B** in the third Maxwell equation, to obtain

$$\nabla \times (\nabla \times \mathbf{A}) - \mu_0\epsilon_0 \frac{\partial}{\partial t}\left(-\frac{\partial \mathbf{A}}{\partial t}\right) = 0,$$

where we have used the fact that ϕ is constant. The first term on the left hand side can be rewritten as $\nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$ so that, applying the gauge condition once again, we obtain a wave equation for **A**, which, using $\mu_0\epsilon_0 = 1/c^2$, takes the form:

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = 0.$$

We know the solutions to such a wave equation: they read

$$\mathbf{A} = \mathbf{d}_\mathbf{k} e^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)} \text{ and } \mathbf{d}'_\mathbf{k} e^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r} + \omega t)},$$

where, always, $\omega_{\mathbf{k}} = c|\mathbf{k}|$ (in vacuum). The first solution describes a wave running in the direction of $\mathbf{k}$ and the second a wave running in the opposite direction. The pre-factors $\mathbf{d}_{\mathbf{k}}$ and $\mathbf{d}'_{\mathbf{k}}$ specify the direction of the vector-potential field. Because of the Coulomb gauge $\nabla\cdot\mathbf{A} = 0$, we have

$$\mathbf{k}\cdot\mathbf{d}_{\mathbf{k}} = \mathbf{k}\cdot\mathbf{d}'_{\mathbf{k}} = 0,$$

which says that the vector potential is oriented perpendicular to the propagation direction. Therefore we can write both pre-factors as a linear combination of two basis vectors perpendicular to $\mathbf{k}$:

$$\mathbf{d}_{\mathbf{k}} = \sum_{\alpha=1,2} d_{\alpha,\mathbf{k}}\hat{\boldsymbol{\epsilon}}_\alpha$$

and similar for $\mathbf{d}'_{\mathbf{k}}$. Let's keep this in the back of our minds for later on.

The general solution can be written as a linear superposition of the monochromatic waves:

$$\mathbf{A} = \sum_{\mathbf{k}} \left(\mathbf{d}_{\mathbf{k}} e^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} + \mathbf{d}'_{\mathbf{k}} e^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}+\omega_{\mathbf{k}}t)} \right).$$

The two terms in the sum are related by the fact that $\mathbf{A}$ is a *real* field. Its complex conjugate is given by

$$\mathbf{A}^* = \sum_{\mathbf{k}} \left(\mathbf{d}^*_{\mathbf{k}} e^{-\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} + \mathbf{d}'^*_{\mathbf{k}} e^{-\mathrm{i}(\mathbf{k}\cdot\mathbf{r}+\omega_{\mathbf{k}}t)} \right).$$

We focus on the time dependence of the different terms in $\mathbf{A}$ and $\mathbf{A}^*$ and try to equate those. This leads to

$$\mathbf{d}_{\mathbf{k}} e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}} = \mathbf{d}'^*_{\mathbf{k}} e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$$

and the complex conjugate of this equality. At first sight, the left and right hand side seem incompatible as the equality of $\mathbf{A}$ and $\mathbf{A}^*$ should hold throughout space. However, realizing that $\omega_{\mathbf{k}} = \omega_{-\mathbf{k}}$ we can replace $\mathbf{k}$ on the right hand side by $-\mathbf{k}$, which directly leads to the conclusion

$$\mathbf{d}'_{\mathbf{k}} = \mathbf{d}^*_{-\mathbf{k}}$$

Therefore, $\mathbf{A}$ can be written as

$$\mathbf{A} = \sum_{\mathbf{k}} \left(\mathbf{d}_{\mathbf{k}} e^{-\mathrm{i}\omega_{\mathbf{k}}t} e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} + \mathbf{d}^*_{-\mathbf{k}} e^{\mathrm{i}\omega_{\mathbf{k}}t} e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \right) = \sum_{\mathbf{k}} \left(\mathbf{d}_{\mathbf{k}}(t) + \mathbf{d}^*_{-\mathbf{k}}(t) \right) e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}},$$

where, in the last expression, we have incorporated the time dependence of the waves into the $\mathbf{d}_{\mathbf{k}}(t)$'s. In the sequel, we shall need the time derivative of $\mathbf{A}$, and we therefore note that

$$\dot{\mathbf{d}}_{\mathbf{k}} = -\mathrm{i}\omega_{\mathbf{k}}\mathbf{d}_{\mathbf{k}} \text{ and } \dot{\mathbf{d}}^*_{-\mathbf{k}} = \mathrm{i}\omega_{\mathbf{k}}\mathbf{d}_{\mathbf{k}},$$

which we should also keep in our short-term memory. In an infinite volume, the $\mathbf{k}$ form a continuous set, and the sum over them becomes an integral:

$$\mathbf{A}(\mathbf{r},t) = \int \left[\mathbf{d}(\mathbf{k},t) + \mathbf{d}^*(-\mathbf{k},t) \right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \frac{d^3k}{(2\pi)^{3/2}}$$

The classical Hamiltonian is found as the integral over the energy density of an electromagnetic field:

$$H = \frac{1}{2}\int \left(\epsilon_0\mathbf{E}^2 + \frac{1}{\mu_0}\mathbf{B}^2 \right) d^3r.$$

As we have seen, the electric field can be found as $-\dot{\mathbf{A}}$ (remember $\phi \equiv 0$) and the magnetic field as $\mathbf{B} = \nabla\times\mathbf{A}$. We can therefore write the electric field as:

$$\mathbf{E}(\mathbf{r},t) = \frac{\mathrm{i}}{(2\pi)^{3/2}} \int \omega_{\mathbf{k}} \left[\mathbf{d}(\mathbf{k},t) - \mathbf{d}^*(-\mathbf{k},t) \right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k =$$

Using the fact that **E** is real, we can write it also in the form

$$\mathbf{E}(\mathbf{r},t) = -\frac{\mathrm{i}}{(2\pi)^{3/2}}\int \omega_{\mathbf{k}}\left[\mathbf{d}^*(\mathbf{k},t)-\mathbf{d}(-\mathbf{k},t)\right]e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}d^3k.$$

For the magnetic field, we obtain the expression:

$$\begin{aligned}\mathbf{B}(\mathbf{r},t) &= \frac{\mathrm{i}}{(2\pi)^{3/2}}\int \mathbf{k}\times\left[\mathbf{d}(\mathbf{k},t)+\mathbf{d}^*(-\mathbf{k},t)\right]e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}d^3k\\ &= -\frac{\mathrm{i}}{(2\pi)^{3/2}}\int \mathbf{k}\times\left[\mathbf{d}^*(\mathbf{k},t)+\mathbf{d}(-\mathbf{k},t)\right]e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}d^3k.\end{aligned}$$

Using these equations for the electric and magnetic field, we can write the Hamiltonian in terms of the coefficients **d**. For the first term, defined in terms of the electric field, we obtain, using the orthogonality of the $\exp(\mathrm{i}\mathbf{k}\cdot\mathbf{r})$:

$$\begin{aligned}&\frac{\epsilon_0}{2}\int \mathbf{E}\cdot\mathbf{E}^*\,d^3r = \\ &\frac{\epsilon_0}{2}\frac{1}{(2\pi)^3}\int\int \omega_{\mathbf{k}}\left[\mathbf{d}(\mathbf{k},t)-\mathbf{d}^*(-\mathbf{k},t)\right]e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}d^3k\int \omega_{\mathbf{k}'}\left[\mathbf{d}^*(\mathbf{k}',t)-\mathbf{d}(-\mathbf{k}',t)\right]e^{-\mathrm{i}\mathbf{k}'\cdot\mathbf{r}}d^3k'\,d^3r = \\ &\qquad\frac{\epsilon_0}{2}\int \omega_{\mathbf{k}}^2\left[2\left|\mathbf{d}(\mathbf{k},t)\right|^2-\mathbf{d}(\mathbf{k},t)\mathbf{d}(-\mathbf{k},t)-\mathbf{d}^*(\mathbf{k},t)\mathbf{d}^*(-\mathbf{k},t)\right]d^3k.\end{aligned}$$

For the term involving **B** we find a similar result, using that $[\mathbf{k}\times\mathbf{d}(\mathbf{k})]\cdot[\mathbf{k}\times\mathbf{d}^*(-\mathbf{k})] = k^2\left|\mathbf{d}\right|^2$, since **k** is perpendicular to both $\mathbf{d}(\mathbf{k})$ and $\mathbf{d}^*(-\mathbf{k})$ (see above):

$$\frac{1}{2\mu_0}\mathbf{B}^2 = \frac{1}{2\mu_0}\int k^2\left[2\left|\mathbf{d}(\mathbf{k},t)\right|^2+\mathbf{d}(\mathbf{k},t)\mathbf{d}(-\mathbf{k},t)+\mathbf{d}^*(\mathbf{k},t)\mathbf{d}^*(-\mathbf{k},t)\right]d^3r.$$

Adding the two contributions to the Hamiltonian gives, with $\omega_{\mathbf{k}} = c\left|\mathbf{k}\right|$ and $\epsilon_0\mu_0 = 1/c^2$:

$$H = 2\epsilon_0\int \omega_{\mathbf{k}}^2\left|\mathbf{d}(\mathbf{k},t)\right|^2d^3k = 2\epsilon_0\sum_{\alpha=1,2}\int \omega_{\mathbf{k}}^2\left|d_\alpha(\mathbf{k},t)\right|^2d^3k = \sum_{\alpha=1,2}\int H_\alpha(\mathbf{k})d^3k. \tag{8.15}$$

8.3.2 QUANTIZATION OF THE ELECTROMAGNETIC FIELD

The form of the Hamiltonian we have arrived at suggests that it can be written as a sum of harmonic oscillators as each oscillator **d** evolves in time as $\exp(\mathrm{i}\omega_{\mathbf{k}}t)$. However, there is an important difference: a classical harmonic oscillator is described in terms of *real* variables x and p. Our $d_\alpha(\mathbf{k},t)$ are complex, so they cannot play the role of an x and p. We know however how to arrive from the **d** at real variables proportional to $\cos(\omega t)$ or $\sin(\omega t)$: we take the real or imaginary part of the **d**, with additional proportionality factors. These factors are chosen such as to arrive at a suitable form of the Hamiltonian, as we shall see shortly:

$$\begin{aligned}x_\alpha(\mathbf{k},t) &= \sqrt{\epsilon_0}\left[d_\alpha(\mathbf{k},t)+d_\alpha^*(\mathbf{k},t)\right],\\ p_\alpha(\mathbf{k},t) &= -\mathrm{i}\omega_{\mathbf{k}}\sqrt{\epsilon_0}\left[d_\alpha(\mathbf{k},t)-d_\alpha^*(\mathbf{k},t)\right].\end{aligned}$$

These relations can simply be inverted, to yield:

$$d_\alpha(\mathbf{k},t) = \frac{1}{2\sqrt{\epsilon_0}}\left[x_\alpha(\mathbf{k},t)+\mathrm{i}\frac{p_\alpha(\mathbf{k},t)}{\omega_{\mathbf{k}}}\right]. \tag{8.16}$$

Furthermore, the Hamiltonian for each harmonic oscillator

$$H_\alpha(\mathbf{k}) = \frac{1}{2}p_\alpha^2(\mathbf{k},t)+\frac{1}{2}\omega_{\mathbf{k}}^2x_\alpha^2(\mathbf{k},t),$$

precisely equals $H_\alpha(\mathbf{k})$ occuring in the last expression of (8.15).

The foregoing analysis may seem a bit arbitrary and unnatural: why would x_α and p_α be defined this way? To convince yourself that the resulting expressions make sense, let's see how x and p are related, by using the Maxwell equations for the classical fields. Let's therefore consider the expressions for **E** and **B** in terms of plane waves again. The electric field contains a term (up to a constant)

$$\mathrm{i}\omega_{\mathbf{k}}\hat{\boldsymbol{\epsilon}}_\alpha d_\alpha e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$$

and the same mode in the **B** field is (up to the same constant):

$$\mathrm{i}\omega_{\mathbf{k}}\hat{\mathbf{n}}\times\hat{\boldsymbol{\epsilon}}_\alpha d_\alpha e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}.$$

Now let us apply the Maxwell equation

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

to these modes. We then find

$$\mathrm{i}^2\omega_{\mathbf{k}}\mathbf{k}\times\hat{\boldsymbol{\epsilon}}_\alpha\left(\omega_{\mathbf{k}}x_\alpha(\mathbf{k},t)+\mathrm{i}p_\alpha(\mathbf{k},t)\right)e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} = -\mathrm{i}\mathbf{k}\times\hat{\boldsymbol{\epsilon}}_\alpha\left(\dot{x}_\alpha(\mathbf{k},t)+\mathrm{i}\dot{p}_\alpha(\mathbf{k},t)\right)e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}.$$

Equating the imaginary and real parts on the left and right hand side, we immediately see that

$$\frac{\partial x_\alpha}{\partial t} = p_\alpha; \qquad \frac{\partial p_\alpha}{\partial t} = -\omega_{\mathbf{k}}^2 x_\alpha. \tag{8.17}$$

These equations demonstrate that the coordinates x and p which we have introduced obey the classical equations of motion for the harmonic oscillator! So it seems they really deserve their names x and p, and we conclude again that the electromagnetic field is correctly described by a superposition of harmonic oscillators.

We can now readily write down the quantum version of the Hamiltonian: we just impose a commutation relation between the $x_\alpha(\mathbf{k})$ and $p_\alpha(\mathbf{k})$:

$$[x_\alpha(\mathbf{k}), p_{\alpha'}(\mathbf{k}')] = \mathrm{i}\hbar\delta(\mathbf{k}-\mathbf{k}')\delta_{\alpha\alpha'}.$$

The usual creation and annihilation operators can now be defined in terms of these:

$$a_\alpha(\mathbf{k},t) = \sqrt{\frac{\omega_{\mathbf{k}}}{2\hbar}}\left[x_\alpha(\mathbf{k},t)+\mathrm{i}\frac{p_\alpha(\mathbf{k},t)}{\omega_{\mathbf{k}}}\right],$$
$$a_\alpha^\dagger(\mathbf{k},t) = \sqrt{\frac{\omega_{\mathbf{k}}}{2\hbar}}\left[x_\alpha(\mathbf{k},t)-\mathrm{i}\frac{p_\alpha(\mathbf{k},t)}{\omega_{\mathbf{k}}}\right]$$

and the Hamiltonian can then be formulated as

$$H = \sum_{\alpha=1,2}\int\frac{\hbar\omega_{\mathbf{k}}}{2}\left[a_\alpha(\mathbf{k},t)a_\alpha^\dagger(\mathbf{k},t)+a_\alpha^\dagger(\mathbf{k},t)a_\alpha(\mathbf{k},t)\right]d^3k =$$
$$\sum_{\alpha=1,2}\int\hbar\omega_{\mathbf{k}}\left[a_\alpha^\dagger(\mathbf{k},t)a_\alpha(\mathbf{k},t)+\frac{1}{2}\right]d^3k. \tag{8.18}$$

Interestingly, we note that the definition of the a-operators in terms of the x and p operators is very similar to the relation between the d's and the x- and p operators [(8.16)] – the only difference is the constant factor $\sqrt{\hbar/(2\epsilon_0\omega_{\mathbf{k}})}$:

$$d_\alpha(\mathbf{k},t)\leftrightarrow\sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}}}a_\alpha(\mathbf{k},t).$$

We can always use the classical expressions derived in the previous subsection and then perform this transformation and impose the standard commutation relation for the operators $a_\alpha(\mathbf{k})$ and $a_\alpha^\dagger(\mathbf{k})$:

$$[a_\alpha(\mathbf{k}), a_{\alpha'}^\dagger(\mathbf{k}')] = \delta_{\alpha\alpha'}\delta(\mathbf{k}-\mathbf{k}')$$

to arrive at the quantum version of electromagnetic field theory.

We see that the modes are characterized by the wave vector **k** and by their polarization direction $\hat{\boldsymbol{\epsilon}}$, which represents two quantum numbers labelled by α. The energy quanta are called *photons*. Inspection of the expression of the Hamiltonian (8.18) reveals a peculiar feature: even when there are *no* photons present, there is an energy due to the factor 1/2. This energy is infinite as we must integrate over *all* possible wave vectors **k**! We may of course make it finite by placing he system in a finite box and imposing a maximum on the possible energies (i.e. a short-wavelength cut-off). A finite offset energy is no problem as physical processes are driven only by energy differences, and not the energy values themselves. Therefore the term 1/2 is neglected when calculating the energy of the electromagnetic field. It is however important to realize that it is there, and why it is there: as the field is composed of infinitely many modes, and each mode has a lowest energy of $\hbar\omega_{\mathbf{k}}/2$ due to the fact that there are always quantum fluctuations present in each mode, even at the lowest possible energies, we have to accept that there is an infinite offset energy.

We conclude this section by providing the electric and magnetic fields and vector potential, expressed in terms of the the creation and annihilation operators. This procedure involves replacing $d_\alpha(\mathbf{k},t)$ by $a_\alpha(\mathbf{k},t)\sqrt{\hbar/(2\epsilon_0\omega_{\mathbf{k}})}$ and similar for the Hermitian (complex) conjugates. Collecting the vector operator $\sum_\alpha \hat{\boldsymbol{\epsilon}}_\alpha a_\alpha$ into a vector **a**, and similarly for $\mathbf{a}^\dagger$, we find the expansions:

$$\mathbf{A}(\mathbf{r},t) = \int \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}}}\left[\mathbf{a}(\mathbf{k},t) + \mathbf{a}^\dagger(-\mathbf{k},t)\right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k.$$

$$\mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{A}}{\partial t} = -\frac{\mathrm{i}}{(2\pi)^{3/2}}\int\sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0}}\left[\mathbf{a}(\mathbf{k},t) - \mathbf{a}^\dagger(-\mathbf{k},t)\right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k.$$

$$\mathbf{B}(\mathbf{r},t) = \nabla\times\mathbf{A} = \frac{\mathrm{i}}{(2\pi)^{3/2}}\int\sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}}}\mathbf{k}\times\left[\mathbf{a}(\mathbf{k},t) + \mathbf{a}^\dagger(-\mathbf{k},t)\right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k.$$

8.3.3 SOME PROPERTIES OF THE ELECTROMAGNETIC FIELD

Using the expressions for the fields, we can calculate the total momentum of the field. The expression for this quantity is in classical electrodynamics given as an integral over the *Poynting vector*:

$$\begin{aligned}\mathbf{P} &= \epsilon_0\int \mathbf{E}\times\mathbf{B}^*\, d^3r = \\ &\frac{\hbar}{2(2\pi)^3}\int\sqrt{\frac{\omega_{\mathbf{k}}}{\omega'_{\mathbf{k}}}}\left[\mathbf{a}(\mathbf{k},t) - \mathbf{a}^\dagger(\mathbf{k},t)\right]\times\mathbf{k}'\times\left[\mathbf{a}(-\mathbf{k}',t) + \mathbf{a}^\dagger(\mathbf{k}',t)\right] e^{\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} d^3k\, d^3k' d^3r \\ &= \frac{\hbar}{2}\int \mathbf{k}\left[\mathbf{a}(\mathbf{k},t)\cdot\mathbf{a}^\dagger(\mathbf{k},t) - \mathbf{a}^\dagger(-\mathbf{k},t)\cdot\mathbf{a}(-\mathbf{k},t) + \mathbf{a}(\mathbf{k},t)\cdot\mathbf{a}(-\mathbf{k},t) - \mathbf{a}^\dagger(\mathbf{k},t)\cdot\mathbf{a}^\dagger(-\mathbf{k},t)\right] d^3k.\end{aligned}$$

To go from the first to the second line, we have used the fact that the **a**'s are all perpendicular to **k**. Noting that the third and fourth term are odd under the substitution **k** → −**k**. On the other hand under the same substitution, the first term turns into the second, but also the **k** in front of the square bracket changes sign, so we are left with

$$\mathbf{P} = \frac{\hbar}{2}\int \mathbf{k}\sum_\alpha \hat{\boldsymbol{\epsilon}}_\alpha\cdot\hat{\boldsymbol{\epsilon}}_\alpha\left[a_\alpha^\dagger(\mathbf{k},t)a_\alpha(\mathbf{k},t) + a_\alpha(\mathbf{k},t)a_\alpha^\dagger(\mathbf{k},t)\right] d^3k.$$

Using the commutation relations for the operators to swap the last product in this integral yields a term 1 inside the square brackets, similar to that occurring in the Hamiltonian. However, in this case that term is not dangerous as the integral vanishes due to the antisymmetry of **k**. The final result is thus

$$\mathbf{P} = \int \hbar\mathbf{k}\sum_\alpha N_\alpha(\mathbf{k}) d^3k.$$

We could have guessed this form: we sum the photon momenta for each occupied quantum. Indeed, it is well known from classical electrodynamics that the Poynting vector equals the product of the electromagnetic energy density (which is $N_\alpha(\mathbf{k})\hbar\omega$) and the speed of light. For a mode $\mathbf{k}, \alpha$, indeed:

$$\mathbf{S} = N_\alpha(\mathbf{k})\hbar\omega c\hat{\mathbf{k}} = \hbar\mathbf{k}\sum_\alpha N_\alpha(\mathbf{k})\hbar\mathbf{k}.$$

Given the fact that the momentum density is given by the Poynting vector $\mathbf{E}\times\mathbf{B}/(4\pi c)$, we also have an expression for the angular momentum of the field:

$$\mathbf{J} = \epsilon_0 \int \mathbf{r}\times(\mathbf{E}\times\mathbf{B})\, d^3r.$$

We calculate the i-th component of this field:

$$J_i = \epsilon_0 \int d^3r \sum_{jklm} \epsilon_{ijk} r_j \epsilon_{klm} E_l B_m.$$

Using the fact that $\mathbf{B} = \nabla\times\mathbf{A}$, and the formula

$$\sum_k \epsilon_{klm}\epsilon_{kij} = \delta_{li}\delta_{mj} - \delta_{lj}\delta_{mi},$$

we obtain

$$\begin{aligned} J_i &= \epsilon_0 \int \sum_{jkl} \epsilon_{ijk}\left(r_j E_l \partial_k A_l - r_j E_l \partial_l A_k\right) \\ &= \epsilon_0 \int \left[\sum_l E_l(\mathbf{r}\times\mathbf{A}) - \sum_{jkl}\epsilon_{ijk}(\partial_l(r_j E_l A_k) - (\partial_l r_j)E_l A_k - r_j(\partial_l E_l)A_k\right] d^3r. \end{aligned} \tag{8.19}$$

The second term is a gradient, so it should be evaluated at the surface. Assuming vanishing fields at infinity, we can neglect this term. The last term vanishes as a result of $\nabla\cdot\mathbf{E} = 0$. In conclusion, we can write $\mathbf{J} = \mathbf{L} + \mathbf{S}$, where

$$\mathbf{L} = \epsilon_0 \int \sum_l E_l(\mathbf{r}\times\nabla)A_l d^3r,$$

and

$$\mathbf{S} = \epsilon_0 \int \mathbf{E}\times\mathbf{A}\, d^3r.$$

The term $\mathbf{L}$ has the form of

$$\frac{\mathrm{i}}{\hbar}\mathbf{r}\times\mathbf{p},$$

acting as a diagonal operator between the components of the $\mathbf{E}$ and $\mathbf{A}$ fields. The term $\mathbf{r}\times\mathbf{p}$ is recognized as the mechanical orbital momentum.

Writing the second term with the same prefactor between the electric field and vector potential components, we obtain the form

$$S^{(i)}_{jk} = -\mathrm{i}\hbar\epsilon_{ijk}.$$

It can be shown that $\mathbf{S}$ is the component of $\mathbf{J}$ along $\mathbf{k}$. It is easy to show that the eigenvalues of this operator are ±1. Note that the last result tells us that the photon is a spin one particle. The two possibilities ±1 correspond to the two circular polarizations.

8.3.4 THE CASIMIR EFFECT

In the previous sections we have seen that the electromagnetic field can be seen as a collection of simple harmonic oscillators, one for each polarization α and wave vector $\mathbf{k}$. We also noted that a harmonic oscillator has a *zero-point energy* $\hbar\omega_{\mathbf{k}}/2$, which is the ground state energy of that oscillator. For an infinitely large volume, there are infinitely many $\mathbf{k}$-vectors, so the ground state of the electromagnetic field is infinite. We do not care about this in general, as we know that physical processes are driven by energy differences and not by the actual values of the energies involved.

The Casimir effect is a rather dramatic manifestation of the existence of the vacuum energy. This statement already should be puzzling: we just argued that the vacuum energy is not noticeable as it is only the energy *differences* we should care about. The Casimir effect however is a manifestation of the *difference* between vacuum energies, that is, it shows us a difference between infinity and infinity! It was first postulated based on theoretical arguments in 1948 by H. B. G. Casimir, and then (somewhat tentatively) demonstrated experimentally for the first time by M. Sparnaay in 1958. Currently, there is great interest in measuring the Casimir effect using nanotechnology, see for example G. Bressi et al., Phys. Rev. Lett. 88 041804 (2002).

So let us see how we can find a force from the vacuum. We consider a segment of space enclosed between two parallel planes. We take the z-axis perpendicular to these planes. We introduce some cut-off to make the vacuum energy between these two planes finite. Let us first write up the vacuum energy in a box of size $L \times L \times L$ with periodic boundary conditions. The k-vectors in such a box are $2\pi(n_x, n_y, n_z)/L$. This means that the volume per k-point is $(2\pi/L)^3$. Using this, we can turn the sum over all k-points into an integral, in addition to a sum over the two possible polarizations α:

$$E_0 = \frac{1}{2}\sum_\alpha \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} = \sum_{\mathbf{k}} \hbar c\sqrt{k_x^2 + k_y^2 + k_z^2} = \left(\frac{L}{2\pi}\right)^3 \int \hbar c k d^3 k$$

with c the speed of light (the sum over the two polarizations α cancels the prefactor $1/2$). The cut-off is introduced with a parameter λ, the inverse of which sets the scale up to which the k-modes contribute to the energy:

$$E_0(\lambda) = \left(\frac{L}{2\pi}\right)^3 \int \hbar c k e^{-\lambda k} d^3 k.$$

After completing the calculation we shall put $\lambda \to 0$ to remove the cut-off – physically useful results should then be independent of λ. The integral can be worked out:

$$E_0(\lambda) = \frac{V}{2\pi^2} c\hbar \int_0^\infty k^3 dk e^{-\lambda k} = \frac{1}{\lambda^4}\frac{3Vc\hbar}{\pi^2}.$$

This implies a vacuum energy per volume of

$$e_0(\lambda) = \frac{E_0(\lambda)}{V} = \frac{3\hbar c}{\pi^2\lambda^4}.$$

So if we consider the space contained between two infinite parallel planes, separated by a distance a, then the vacuum energy contained in between those two planes per surface area A is

$$E_0/A = \frac{3a\hbar c}{\pi^2\lambda^4}.$$

Now we calculate the same energy in the case where these planes are perfect conductors. This has a significant influence on the vacuum energy contained between them, as a result of the fact that the field should vanish at the conducting planes. This means that the field must be proportional to $\sin(n\pi z/a)$, so that only the values $k_z = n\pi/a$ are allowed, and the modes along the z-direction now become discrete:

$$\left(\frac{L}{2\pi}\right)^3 \int d^3k \rightarrow \left(\frac{L}{2\pi}\right)^2 \int dk_x dk_y \sum_{k_z}.$$

The last sum can no longer be replaced by an integral when a is not large. The new vacuum energy per surface area now becomes

$$E_0'/A = \frac{1}{2}\sum_{\alpha}\frac{\hbar c}{2\pi}\int k_\parallel dk_\parallel \sum_{n_z=0}^{\infty}\sqrt{k_\parallel^2 + (\pi n_z/a)^2}\, e^{-\lambda\sqrt{k_\parallel^2+(\pi n_z/a)^2}}$$

Wait a minute! The sum on the right hand side does not take into account that there are two polarizations only for $n_z \neq 0$. This can be seen by actually writing the expressions for the electric field:

$$\begin{aligned}
E_x &= E_x^{(0)}\cos(k_x x)\sin(k_y y)\sin(n_z\pi z/a);\\
E_y &= E_y^{(0)}\sin(k_x x)\cos(k_y y)\sin(n_z\pi z/a);\\
E_z &= E_z^{(0)}\sin(k_x x)\sin(k_y y)\cos(n_z\pi z/a)
\end{aligned}$$

Two of the components E_x, E_y, E_z can be chosen freely, but the third is fixed by the Maxwell equations, as are the three magnetic field components. However, for $n_z = 0$ we see that only the amplitude of the z-component can be chosen freely, hence there is only one mode possible in that case.

Therefore it is useful to split off the $n_z = 0$ term:

$$E_0'/A = \frac{1}{2}\frac{\hbar c}{2\pi}\int k_\parallel dk_\parallel \left[k_\parallel e^{-\lambda k_\parallel} + 2\sum_{n_z=1}^{\infty}\sqrt{k_\parallel^2 + (\pi n_z/a)^2}\, e^{-\lambda\sqrt{k_\parallel^2+(\pi n_z/a)^2}}\right].$$

The first part of the integral can easily be evaluated:

$$\frac{\hbar c}{\pi}\int k_\parallel^2 dk_\parallel e^{-\lambda k_\parallel} = \frac{c\hbar}{2\pi\lambda^3}.$$

To evaluate the second part of the integral on the right hand side, we substitute

$$y = \sqrt{k_\parallel^2 + (\pi n_z/a)^2},$$

which turns the integral for n_z into

$$\int k_\parallel dk_\parallel \sqrt{k_\parallel^2 + (\pi n_z/a)^2}\, e^{-\lambda\sqrt{k_\parallel^2+(\pi n_z/a)^2}} = \int_{n_z\pi/a}^{\infty} y^2 e^{-\lambda y} dy = \frac{d^2}{d\lambda^2}\left(\frac{1}{\lambda}e^{-n_z\pi\lambda/a}\right).$$

All in all, the result for the energy is now

$$E_0'/A = \frac{c\hbar}{2\pi\lambda^3} + \frac{c\hbar}{2\pi}\frac{d^2}{d\lambda^2}\sum_{n_z=1}^{\infty}\frac{1}{\lambda}e^{-n_z\pi\lambda/a}.$$

Carrying out the sum on the right hand side, which contains a geometric series, we obtain

$$E_0'/A = \frac{c\hbar}{2\pi\lambda^3} + \frac{c\hbar}{2\pi}\frac{d^2}{d\lambda^2}\frac{1}{\lambda}\frac{1}{e^{\pi\lambda/a} - 1}.$$

The right hand side diverges for $\lambda \to 0$ as

$$E_{\text{Divergent}}'/A = \frac{3a\hbar c}{\pi^2\lambda^4},$$

which is precisely equal to the divergent vacuum energy in the absence of the two conductors.

Expanding the right hand side further in powers of λ gives

$$\frac{1}{e^z-1}=\frac{1}{z}-\frac{1}{2}+\frac{z}{12}-\frac{z^3}{720}.$$

We find that the next non-vanishing term is *independent of* λ and

$$\frac{E_0'-E_0}{A}=-\frac{c\hbar\pi^2}{720a^3}.$$

This implies that the two conductors attract each other with a force per unit area:

$$\frac{F}{A}=-\frac{c\hbar\pi^2}{240a^4}.$$

The principle of the Casimir effect seems general enough to apply in other cases where waves are confined in space. Indeed, it has been claimed that two ships would be attracted to each other by a Casimir effect of the water waves in between them. This conjecture however does not seem to have a firm ground.

8.4 SUMMARY

In this chapter we have analysed systems that can be considered as collections of harmonic oscillators. We started by formulating the harmonic oscillator Hamiltonian

$$H=\frac{p^2}{2m}+\frac{m}{2}\omega^2x^2$$

in terms of creation and annihilation operators, which are defined as:

$$a=\sqrt{\frac{m\omega}{2\hbar}}\left(x+\mathrm{i}\frac{p}{m\omega}\right);\qquad a^\dagger=\sqrt{\frac{m\omega}{2\hbar}}\left(x-\mathrm{i}\frac{p}{m\omega}\right).$$

These operators satisfy the commutation relation

$$[a,a^\dagger]=1.$$

In terms of the creation and annihilation operators, the Hamiltonian takes the form

$$H=\hbar\omega\left(a^\dagger a+1/2\right).$$

These ingredients lead to the following conclusions:

- The operator $\hat{n}=a^\dagger a$ takes on non-negative integer values – it is called the *number operator*.
- The eigenstates of the Hamiltonian are eigenstates $|\psi_n\rangle$ of the number operator with eigenvalue $n=0,1,2,\ldots$ with that operator. With the Hamiltonian, they have eigenvalues
$$H=\hbar\omega\,(n+1/2)\,.$$
- The creation and annihilation operators have the following effect when acting on a state $|\psi_n\rangle$:
$$a|\psi_n\rangle=\sqrt{n}\,|\psi_{n-1}\rangle;\qquad a^\dagger|\psi_n\rangle=\sqrt{n+1}\,|\psi_{n+1}\rangle\,.$$

Coherent states are eigenstates of the annihilation operator a. They represent minimum-uncertainty wave packets and lead to an oscillating wave packet when propagated in time.

Phonon and photon fields can be written as superpositions of independent harmonic oscillators. This implies that we can readily quantise them. The phonon modes are labelled by their wave vector **k** and a characterised by their frequency which is given in terms of **k** in the so-called *dispersion relation.* For a linear chain with nearest neighbour couplings with coupling strength K, the longitudinal modes have the dispersion relation

$$\omega_k = 2K\sin(|ka|/2).$$

The Hamiltonian is given by

$$H_{\text{Chain}} = \int \hbar\omega_k \left(a_k^\dagger a_k + 1/2\right) dk.$$

Also for the electromagnetic field, we have demonstrated that the Hamiltonian can be written as a sum over independent modes. These have dispersion relation $\omega_{\mathbf{k}} = c|\mathbf{k}|$. The Hamiltonian of the electromagnetic field is

$$H_{\text{EM}} = \sum_\alpha \int \frac{\hbar\omega_{\mathbf{k}}}{2}\left[a_\alpha^\dagger(\mathbf{k},t)a_\alpha(\mathbf{k},t) + a_\alpha(\mathbf{k},t)a_\alpha^\dagger(\mathbf{k},t)\right] d^3k.$$

Here, the $a_\alpha(\mathbf{k},t)$ and $a_\alpha^\dagger(\mathbf{k},t)$ are the time-dependent operators

$$a_\alpha(\mathbf{k},t) = e^{\mathrm{i}\omega_{\mathbf{k}}t} a_\alpha(\mathbf{k}),$$

with $a_\alpha(\mathbf{k})$ the time-independent annihilation operator and similar for the hermitian conjugate (creation operator). The subscript α denotes the two polarization directions perpendicular to the propagation direction **k**. The expressions used are always valid within the transverse gauge $\nabla\cdot\mathbf{A} = 0$. The energy quanta of this Hamiltonian are called *photons.*

We can also formulate the quantum expressions for the fields in terms of these creation and annihilation operators. These are:

$$\mathbf{A}(\mathbf{r},t) = \int \frac{1}{(2\pi)^{3/2}}\sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}}}\left[\mathbf{a}(\mathbf{k},t) + \mathbf{a}^\dagger(-\mathbf{k},t)\right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k.$$

$$\mathbf{E}(\mathbf{r},t) = -\frac{\partial\mathbf{A}}{\partial t} = \frac{\mathrm{i}}{(2\pi)^{3/2}}\int\sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0}}\left[\mathbf{a}(\mathbf{k},t) - \mathbf{a}^\dagger(-\mathbf{k},t)\right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k.$$

$$\mathbf{B}(\mathbf{r},t) = \nabla\times\mathbf{A} = \frac{\mathrm{i}}{(2\pi)^{3/2}}\int\sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}}}\mathbf{k}\times\left[\mathbf{a}(\mathbf{k},t) + \mathbf{a}^\dagger(-\mathbf{k},t)\right] e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} d^3k.$$

Finally, we can formulate the Poynting vector, which represents the energy flow of the EM field, as

$$\mathbf{P} = \int \hbar\mathbf{k}\sum_\alpha N_\alpha(\mathbf{k}) d^3k,$$

where $N_\alpha = a^\alpha(\mathbf{k})a(\mathbf{k})$ counts the number of photons with polarization alpha.

The quantum expression for the angular momentum carried by the electromagnetic field splits into two contributions: the orbital angular momentum and the spin.

$$\mathbf{J} = \mathbf{L} + \mathbf{S},$$

where

$$\mathbf{L} = \epsilon_0 \int \sum_l E_l(\mathbf{r}\times\nabla)A_l d^3r,$$

and

$$\mathbf{S} = \epsilon_0 \int \mathbf{E} \times \mathbf{A}\, d^3 r.$$

Finally, we have seen that the electric field energy density always tends to infinity, due to the fact that it is composed of an infinite amount of harmonic oscillators, leading to an infinite amount of zero-point energies $\hbar\omega_{\mathbf{k}}/2$. We never 'see' the infinite energy density, as processes are driven only by energy *differences*. The infinite energy density manifests itself however through the dramatic *Casimir effect*: two ideal, parallel conductors with separation a deform the field in between them due to the zero tangential field boundary condition at their surface, leading to a finite energy difference per surface area. This energy difference can be calculated to be

$$\frac{E_0' - E_0}{A} = -\frac{c\hbar\pi^2}{720 a^3}.$$

From this, we see that the two conductors attract each other with a force per unit area:

$$\frac{F}{A} = -\frac{c\hbar\pi^2}{240 a^4}.$$

This is the famous *Casimir force*.

8.5 PROBLEMS

1. (a) Derive the equations of motion for the operators $\hat{x}$ and $\hat{p}$ for the harmonic oscillator, described by the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2.$$

(b) Transforming the variables according to

$$X(t) = x(t)\sqrt{m\omega/\hbar}; \qquad P(t) = p(t)/\sqrt{\hbar m\omega},$$

find the equations of motion for X and P.

(c) Now we search for the eigenstates of the annihilation, or 'lowering' operator a. What is the form of this operator, expressed in terms of x and p? And in terms of X and P?

The eigenstates of a satisfy

$$a\,|\alpha\rangle = \alpha\,|\alpha\rangle.$$

We write $|\alpha\rangle$ in the form

$$|\alpha\rangle = \sum_{n=0}^{\infty} c_n\,|n\rangle,$$

where $|n\rangle$ is an eigenstate of the Hamiltonian with energy eigenvalue $\hbar\omega(n+1/2)$. Show that the fact that this is an eigenstate of the lowering operator leads to the condition

$$c_{n+1} = \frac{\alpha}{\sqrt{n+1}} c_n,$$

and show from this that

$$c_n = \frac{\alpha^n}{\sqrt{n!}} c_0.$$

(d) Find the normalization condition, and assuming that c_0 is real

$$c_0^2 = \frac{1}{\sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!}} = e^{-|\alpha|^2},$$

so that

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle.$$

These states are called *coherent states*.

(e) Show that this can be written as

$$|\alpha\rangle = e^{-|\alpha|^2/2} e^{\alpha a^\dagger} |0\rangle .$$

Hint: remember that

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle .$$

2. We continue with the coherent states considered in the previous problem. We now focus on the time dependence.

   (a) Show that the Heisenberg equation for a has as its solution

   $$a(t) = e^{-\mathrm{i}\omega t} a(0).$$

   (b) Show that the eigenstate of a with eigenvalue α at $t = 0$ will remain an eigenstate of a with eigenvalue $\alpha(t) = \exp(-\mathrm{i}\omega t)\alpha$.

   (c) From $X(t) = \left[a(t) + a^\dagger(t)\right]/\sqrt{2}$, show that the expectation value of X, for a system starting off in the state $|\alpha\rangle$, varies in time as

   $$\langle X\rangle(t) = X_m \cos(\omega t - \varphi).$$

   Express X_m and φ in terms of α. Find a similar equation for $P(t)$.

   (d) Calculate $\langle \alpha|\Delta^2 X|\alpha\rangle$ and $\langle \alpha|\Delta^2 P|\alpha\rangle$. How do these quantities evolve in time?

   (e) According to the minimum uncertainty principle, what is the lower bound on the product $\langle \Delta^2 X\rangle \langle \Delta^2 P\rangle$? (The expression $\langle \Delta^2 X\rangle$ denotes the variance $\langle X^2\rangle - \langle X\rangle^2$ and similar for $\langle \Delta^2 P\rangle$). Are coherent states thus minimum uncertainty states?

   (f) The coherent state $|\alpha\rangle$ is related to the ground state of the harmonic oscillator via

   $$|\alpha\rangle = D(\alpha)|0\rangle ,$$

   where

   $$D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}.$$

   Prove this.

   (g) Show that for a general operator A and a unitary operator U we have the relation

   $$U^\dagger e^A U = e^{U^\dagger A U}$$

   and use that to calculate the operator

   $$U_0^\dagger(t) D(\alpha) U_0(t),$$

   where $U_0(t)$ is the time evolution operator for the harmonic oscillator.

   Use this result to calculate the time evolution of an initial state $|\alpha\rangle$. Show that the initial state remains a coherent state.

3. In this problem, we show that a *driven* harmonic oscillator, starting off in a coherent state, remains in a coherent state. The Hamiltonian of the driven harmonic oscillator reads

   $$H = \frac{p^2}{2m} + \frac{m\omega^2}{2} x^2 + \sqrt{2m\omega^3\hbar} f(t) x.$$

   (a) Formulate this Hamiltonian in terms of creation and annihilation operators, and show that the equation of motion for $a(t)$ reads

   $$\dot{a} = -\mathrm{i}\omega a - \mathrm{i}\omega f.$$

   Write the solution to this equation as an integral involving f.

(b) Now consider a system starting off in a coherent state $|\alpha(0)\rangle$. Work out

$$a(t)\,|\alpha(0)\rangle\,.$$

Show that the time-dependent state is a coherent state

$$|\alpha(t)\rangle$$

where

$$\alpha(t)=\alpha(0)e^{-\mathrm{i}\omega t}-\mathrm{i}\omega\int_0^t e^{-\mathrm{i}\omega(t-t')}f(t')dt'.$$

4. Consider the harmonic oscillator in two dimensions (in convenient units):

$$H=\frac{1}{2}\left(p_x^2+p_y^2\right)+\frac{1}{2}(x^2+y^2). \tag{8.20}$$

The units furthermore imply $\hbar\equiv 1$, so that $[x_\alpha,p_\beta]=\mathrm{i}\delta_{\alpha\beta}$, for $\alpha,\beta=1,2$.

Define annihilation operators a_1 and a_2 as follows:

$$a_1=\frac{1}{\sqrt{2}}(x+\mathrm{i}p_x)\text{ and }a_2=\frac{1}{\sqrt{2}}(y+\mathrm{i}p_y). \tag{8.21}$$

Creation operators $a_1^\dagger$ and $a_2^\dagger$ are the Hermitian conjugates of these.

(a) Calculate all commutation relations for a_α and $a_\alpha^\dagger$, $\alpha=1,2$.

(b) Show that $N_1=a_1^\dagger a_1$ and $N_2=a_2^\dagger a_2$ form an observation maximum, i.e. a maximal set of commuting operators. (Capital N's are used for operators, and lower case n's for their eigenvalues). Show furthermore that

$$|n_1n_2\rangle=\frac{\left(a_1^\dagger\right)^{n_1}}{\sqrt{n_1!}}\frac{\left(a_2^\dagger\right)^{n_2}}{\sqrt{n_2!}}|0\rangle \tag{8.22}$$

are the corresponding eigenstates and give their degeneracies.
Now we define

$$J_x=\frac{1}{2}\left(a_2^\dagger a_1+a_1^\dagger a_2\right) \tag{8.23a}$$

$$J_y=\frac{i}{2}\left(a_2^\dagger a_1-a_1^\dagger a_2\right) \tag{8.23b}$$

$$J_z=\frac{1}{2}\left(a_1^\dagger a_1-a_2^\dagger a_2\right) \tag{8.23c}$$

(c) Show that $[J_k,J_l]=i\epsilon_{klm}J_m$ and that $J^2=N/2(N/2+1)$, $N=N_1+N_2$.

(d) Show that J^2 and J_z form an observation maximum and show that

$$\left|jm\right\rangle=\frac{\left(a_1^\dagger\right)^{j+m}}{\sqrt{(j+m)!}}\frac{\left(a_2^\dagger\right)^{j-m}}{\sqrt{(j-m)!}}|0\rangle \tag{8.24}$$

What are the eigenvalues of the Hamiltonian and their degeneracies?

(e) For a spin-less particle in a magnetic field $\mathbf{B}=(0,0,B_z)$ the Hamiltonian can be written as

$$H=H_1+H_2+H_3 \tag{8.25}$$

with H_1 the Hamiltonian of the two dimensional harmonic oscillator in the xy plane, H_2 the Hamiltonian for a free particle along the z axis and $H_3=-qBL_z/2m$. Show that $[H_i,H_j]=0, i,j=1,2,3$.

(f) Show that L_z can be expressed in J_y.

(g) How can the states of H be labelled? Give energy levels and degeneracies.

5. In this problem, we study the mean and variance of the electric field in the photonic ground state, which we denote as $|0\rangle$.

   (a) Show that $\langle 0|E(\mathbf{r},t)|0\rangle = 0$.

   (b) Show that, written as a sum over the modes in a $L \times L \times L$ cavity with periodic boundary conditions,

   $$\left\langle 0 \middle| E^2(\mathbf{r},t) \middle| 0 \right\rangle = \sum_{\mathbf{k},\alpha} \frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 L^3}.$$

   (c) Show, using the integral representation and introducing a cut-off wave vector k_c, that the uncertainty diverges as k_c^4.

   (d) In practice, we never measure the field at a single point but in a finite volume. If the linear size of that volume is ℓ, how does the uncertainty of the field scale with ℓ?

9

SECOND QUANTISATION

9.1 INTRODUCTION

The quantum mechanics you have been confronted with so far usually deals with single particles moving in a potential which is caused by the environment (e.g. gravity, an electric or a magnetic field). If we have more than one particle, each of them feels the other particles through some interaction (gravity, Coulomb interaction, spin-orbit interaction, ...). When the system is close to equilibrium, we can decompose its Lagrangian into its normal modes, and these can be described formally by independent harmonic oscillators – this is the approach we used in the previous chapter to analyse the phonons in a chain. Generally, however, this analysis is not possible, and the problem becomes tremendously difficult. As interacting many-body systems are so difficult to analyse thoroughly, we shall only scratch the surface of this interesting topic in this lecture course. However, treating the particles in a many body system as independently moving in one potential which reflects the effect of the interactions of all other particles, has proven to be extremely useful, and many phenomena in for example solids can be explained using this 'independent particle' picture. Even without treating the interactions explicitly, many-body systems are interesting, if only because of the statistics, which, for indistinguishable fermions, gives rise to a pressure in the degenerate limit (high density and low temperature), originating from the fact that two fermions having the same spin do not want to occupy the same position in space.

In the previous chapter we have dealt with phonons and photons. We know that these 'particles' ('excitations' would be a better term for them) do not preserve their number as phonons and photons are continuously created and annihilated. This happens for example in processes where electromagnetic radiation interacts with electrons, and it is a direct consequence of the fact that there is no 'mass energy' $E = mc^2$ to be paid for creating a photon – just its momentum determines the energy via $E = \hbar\omega = \hbar ck$, and this momentum may be (very) small.

The varying number of photons or phonons is an example of the general phenomenon that interactions cause the particle numbers of different species in the system to vary. Therefore we need a formalism in which the number of particles is not constant, but allowed to change. It will turn out that this description is also the most convenient one in a mathematical sense. We therefore shall be working in a space which is a direct sum of Hilbert spaces with fixed numbers of particles. This space is called the *Fock space* $\mathscr{F}$:

$$\mathscr{F} = \mathscr{H}(1) \oplus \mathscr{H}(2) \oplus \mathscr{H}(3) \oplus \mathscr{H}(4) \ldots$$

which is a direct sum of the Hilbert spaces $\mathscr{H}(n)$ for n particles, where n assumes any positive integer value. For completeness, we state that if the Hamiltonian of a one-particle system is

$\hat{h}$, then for an n-particle non-interacting system, it is given as

$$\hat{H}(n) = \hat{h}(1) + \hat{h}(2) + \cdots + \hat{h}(n) = \sum_{i=1}^{n} \hat{h}(i).$$

Here, i, as in $\hat{h}(i)$, denotes the appropriate quantum number(s) of particle i. Examples are the position $\mathbf{r}_i$ of a spin-less particle, or, if the particle has spin, $(\mathbf{r}_i, s_i)$ where s_i is the spin quantum number (corresponding to, say, the z-component of the spin) of the particle. An example of $\hat{h}$ is the kinetic energy:

$$\hat{h}(i) = \frac{p_i^2}{2m}.$$

From now on, we will omit the hat from operators when there is no ambiguity. If all particles feel the same external potential $U(\mathbf{r})$, $\hat{h}$ becomes

$$\hat{h}(i) = \frac{p_i^2}{2m} + U(\mathbf{r}_i).$$

The particles are all assumed to be of the same species for simplicity. For interacting particles, we have additional terms in the Hamiltonian. An important example is a two-particle interaction v:

$$H(n) = \sum_{i=1}^{n} h(i) + \sum_{i<i'\leq n} v(\mathbf{r}_i - \mathbf{r}_{i'}).$$

We need a basis of the Fock space. We start from some basis $|\phi_j\rangle$, $j = 1, \ldots$ of the one-particle Hilbert space. It is important to distinguish the *particles*, labelled by i, i' etc. from the one-particle *basis states*, which are labelled by j, k, m etc. A basis for the two-particle Hilbert space is then $|\phi_j \phi_k\rangle$, where the first particle is in the one-particle state j and the second in k. However, this is not a suitable basis when the particles are indistinguishable. In that case, the Hamiltonian is invariant under exchange of any pair of particles. This implies that H commutes with the exchange operator P_{jk} for particles j and k:

$$[P_{jk}, H] = 0.$$

From linear algebra we conclude that we can find eigenstates of H which are at the same time eigenfunctions of P_{jk}, since they are two commuting Hermitian operators. Furthermore we know that $P_{jk}^2 = \mathbb{1}$. Combined with the fact that P_{jk}, being Hermitian, has real eigenvalues, we see that only the eigenvalues $+1$ and -1 of P_{jk} are allowed, and therefore it is possible to have eigenstates of the Hamiltonian which satisfy $P_{jk}|\psi\rangle = \pm|\psi\rangle$ for each ψ which is a many body state including particles j and k. It turns out that a particle has *either* a $+1$ *or* a -1 as its eigenvalue of any exchange operator, i.e. it can not switch from being a particle with $+1$ under exchange, to -1. Particles belonging to the former class are called *bosons* and those belonging to the latter are *fermions*. The so-called *spin-statistics theorem* states that bosons have integer spin, whereas fermions have half-integer spin. This theorem was proven by M. Fierz (1939) and W. Pauli (1940) but the argument is rather tricky. Recently, alternative proofs have been formulated – we do not give the details in these notes.

We have:

> There are two types of particles. The first type is called bosons; these particles are *symmetric* under particle exchange; they have integer spin. The second type of particles, carrying half-integer spin, are called *fermions*. These particles are *anti-symmetric* under particle exchange.

Note that a general permutation of any number of particles, can always be constructed from a sequence of particle exchanges (pair-wise swaps). Although that sequence is not unique, the parity of the number of such exchanges in the sequence is always the same for a given permutation – therefore we distinguish *even* and *odd* permutations. The eigenvalue of the wave function for an operator which permutes the particles according to an even permutation is +1. For an odd permutation, the fermion particle wave function has then eigenvalue −1 (the boson wave function always has eigenvalue +1).

The requirement imposed by these symmetry properties implies that our basis functions should have the appropriate symmetry. For bosons, we can construct symmetric basis functions for n particles starting from single-particle states $|\phi_j\rangle$, as follows:

$$\langle 1,2,3,4,\dots,n|\psi_S\rangle = \mathcal{N}\sum_P \phi_1(P_1)\phi_2(P_2)\dots\phi_n(P_n).$$

where the sum is over all permutations P of the sequence $1,2,\dots,n$:

$$P(1,2,3,\dots) = (P_1,P_2,P_3,\dots)$$

and where $\mathcal{N}$ is an appropriate normalisation factor. Note that P is used as an operator, working on the set of numbers $(1,2,\dots,N)$, whereas the P_1, P_2 …, are numbers between 1 and N. For two particles we only have the two permutations

$$1,2 \to 1,2 \text{ and } 1,2 \to 2,1.$$

For two bosons, occupying two different one-particle states $|\phi_1\rangle$ and $|\phi_2\rangle$, we can construct the symmetric two-particle states as follows:

$$\langle 1,2|\psi_S\rangle = \frac{1}{\sqrt{2}}\left[\langle 1|\phi_1\rangle\langle 2|\phi_2\rangle + \langle 2|\phi_1\rangle\langle 1|\phi_2\rangle\right]. \tag{9.1}$$

It can easily be verified that this state is symmetric under exchange for particles 1 and 2. If the particles occupy the *same* one-particle state ($\phi_1 = \phi_2$) we have $\langle 1|\phi_1\rangle\langle 2|\phi_1\rangle$ as the two-particle state.

For fermions the situation is different. The antisymmetric wave function for two fermions is

$$\langle 1,2|\psi_{AS}\rangle = \frac{1}{\sqrt{2}}\left[\langle 1|\phi_1\rangle\langle 2|\phi_2\rangle - \langle 2|\phi_1\rangle\langle 1|\phi_2\rangle\right],$$

We see that this vanishes when ϕ_1 and ϕ_2 are the same. Indeed, for two fermions in the same state ϕ, we see that the antisymmetric property says:

$$\langle 1|\phi\rangle\langle 2|\phi\rangle = -\langle 2|\phi\rangle\langle 1|\phi\rangle = 0.$$

For more than two (say, n) fermions, occupying a set of one-particle states $|\phi_1\rangle$, $|\phi_2\rangle,\dots|\phi_n\rangle$, the antisymmetric wave function is given by a so-called *Slater determinant*

$$\langle 1,2,3,4,\dots,n|\psi_{AS}\rangle = \frac{1}{\sqrt{n!}}\begin{vmatrix} \langle 1|\phi_1\rangle & \langle 2|\phi_1\rangle & \dots & \langle n|\phi_1\rangle \\ \langle 1|\phi_2\rangle & \langle 2|\phi_2\rangle & \dots & \langle n|\phi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle 1|\phi_n\rangle & \langle 2|\phi_n\rangle & \dots & \langle n|\phi_n\rangle \end{vmatrix}$$

This wave function can be written in a form similar to the one used for bosons:

$$\langle 1,2,3,4,\dots,n|\psi_{AS}\rangle = \frac{1}{\sqrt{n!}}\sum_P \operatorname{sgn}(P)\langle P_1|\phi_1\rangle\langle P_2|\phi_2\rangle\dots\langle P_n|\phi_n\rangle,$$

where sgn(P) is the sign of the permutation: it is +1 if the permutation can be written as a product of an even number of transpositions (exchange operations) and −1 otherwise.

For a symmetric boson wave function in which the one-particle orbital ϕ_1 occurs n_1 times, ϕ_2 n_2 times etcetera, we use the notation

$$|\psi_S\rangle = |n_1 n_2 \dots\rangle. \tag{9.2}$$

We can do the same for a fermion system, for which n_j can take on the values 0 or 1 only. [1]

To conclude this section, we consider the normalisation of the boson states. In principle, we can sum over all permutations, which would yield $n!$ terms. However, if several states are multiply occupied, any permutation involving a reshuffling of particles within the same single particle state does not lead to a new state. This leads us to consider two different formulations for the symmetric boson states. The first involves a sum over *all* permutations, leading to an overcounting of identical states related by permutations of the particles *within each level.* This is compensated for in the normalisation factor:

$$\mathcal{N} = \frac{1}{\sqrt{n! n_1! n_2! \ldots}},$$

where there are n_1 particles in the single particle state 1, n_2 in the single particle state 2 and so on. Alternatively, we can sum over the *unique* permutations only, i.e. we take *only one* representative of each set of states obtained from each other by permutations of the particles within each level. Then the normalisation factor becomes:

$$\mathcal{N}_U = \sqrt{\frac{n_1! n_2! \ldots}{n!}}.$$

The formulation of a quantum theory which describes many-particle states in the Fock space is called *second quantisation.*

9.2 MOVING AROUND IN FOCK SPACE – CREATION AND ANNIHILATION OPERATORS

In the previous section we have introduced the Fock space – the direct sum of Hilbert spaces of different particle numbers. Using this space obviously only makes sense if we can change the particle numbers – otherwise, we could have used the N-particle Hilbert space. In the previous chapter we have encountered operators which can change energy quanta: these are the creation and annihilation operators. We also pointed out that the difference between quanta and particles only has an interpretational character to it; on the formal level, they are the same. We now extend the notion of creation and annihilation operators to any type of particle, massive or massless, fermions or bosons. We start with bosons.

9.2.1 MANY-BOSON SYSTEMS

For many-boson systems, annihilation operators remove a particle from one of the single-particle states in a many-body state of the form (9.1) (or (9.2)); the creation operators add a particle. The definition of these operators is tied to the set of one-particle states we work with. Often, these single-particle states are denoted as *orbitals*, or *spin-orbitals* when the particles have spin-degrees of freedom. The annihilation and creation operators are thus defined by their action on the many-body basis states which are conveniently used in the occupation number representation:

$$a_j \left| n_1 n_2 \ldots n_j \ldots \right\rangle = \sqrt{n_j} \left| n_1 n_2 \ldots n_j - 1 \ldots \right\rangle$$

and

$$a_j^\dagger \left| n_1 n_2 \ldots n_j \ldots \right\rangle = \sqrt{n_j + 1} \left| n_1 n_2 \ldots n_j + 1 \ldots \right\rangle.$$

It is then easy to see that $a_j^\dagger a_j$ is a Hermitian operator with eigenvalue n_j:

$$a_j^\dagger a_j \left| n_1 n_2 \ldots n_j \ldots \right\rangle = a^\dagger \sqrt{n_j} \left| n_1 n_2 \ldots n_j - 1 \ldots \right\rangle = n_j \left| n_1 n_2 \ldots n_j \ldots \right\rangle.$$

---

[1] Some textbooks use a special notation instead of the ‘ket’ vector, such as $|\ldots)$ or $|\ldots\}$, on the right hand side in order to indicate that the state is specified by the ‘occupation numbers’ n_1, n_2,... rather than the labels of the occupied basis states.

Note that we have only *defined* the creation and annihilation operators: they are not a natural consequence of an analysis of a harmonic oscillator as in the previous chapter. In fact, the physics of our bosons may have no relation to the harmonic oscillator at all.

We can also construct a normalised state containing n_j particles in orbital j by acting on the vacuum sufficiently often with the creation operator:

$$\left|000n_j0\dots\right\rangle = \frac{(a_j^\dagger)^{n_j}}{\sqrt{n_j!}}\left|00\dots00\right\rangle.$$

Interestingly, with this choice of the definition of the creation and annihilation operators, we have commutation relations for the creation and annihilation operators reminding us of those for the harmonic oscillator:

$$\left[a_j, a_j^\dagger\right] = 1.$$

This can be seen by acting with this operator on an arbitrary state with n_j particles in orbital j:

$$\left(a_j a_j^\dagger - a_j^\dagger a_j\right)\left|n_1\dots n_j\dots\right\rangle = a_j\sqrt{n_j+1}\left|n_1\dots n_j+1\dots\right\rangle - a_j^\dagger\sqrt{n_j}\left|n_1\dots n_j-1\dots\right\rangle = \\ (n_j+1)\left|n_1\dots n_j\dots\right\rangle - n_j\left|n_1\dots n_j\dots\right\rangle = \left|n_1\dots n_j\dots\right\rangle.$$

Moreover, following similar steps, it can be checked that

$$\left[a_j, a_k\right] = \left[a_j^\dagger, a_k^\dagger\right] = 0,$$

and that

$$\left[a_j, a_k^\dagger\right] = 0 \text{ for } k \neq j.$$

Proving these relations is a useful exercise.

9.2.2 MANY-FERMION SYSTEMS

We can define similar operators for the fermion case. Note that no two fermions can be in the same orbital: in that case, the antisymmetric wave function is always zero. To show that the antisymmetry of the many-body wave function requires a lot of care, we compare the states

$$|n_1 = 1, n_2 = 1, n_3 = 1\rangle \text{ and } |n_2 = 1, n_1 = 1, n_3 = 1\rangle.$$

Note the swap of the labels 1 and 2. By the convention imposed by the Slater determinant, one of the terms in the wave function on the left hand side is a product of the orbitals of particle 1, 2 and 3 in the same order as they appear in the ket vector, i.e. particle 1 is in orbital $\left|\phi_1\right\rangle$; particle 2 is in orbital $\left|\phi_2\right\rangle$ and particle 3 is in orbital $\left|\phi_3\right\rangle$. This term occurs with a plus (+) sign. In the state on the right hand side, the first two particles are swapped, and the same term (particles 1, 2 and 3 in orbitals 1, 2 and 3 respectively) occurs with a minus (−) sign. All other terms have their sign swapped too; therefore, due to the anti-symmetry of the wave function, the second wave function has a minus sign with respect to the first:

$$|n_1 = 1, n_2 = 1, n_3 = 1\rangle = -|n_2 = 1, n_1 = 1, n_3 = 1\rangle.$$

We want to introduce again an annihilation operator. We may try to define this as:

$$a_j\left|n_1\dots n_j = 1\dots\right\rangle = \left|n_1\dots n_j = 0\dots\right\rangle$$

and

$$a_j\left|n_1\dots n_j = 0\dots\right\rangle = 0.$$

This however leads to a problem. The point is that

$$a_2|n_1 = 1, n_2 = 1, n_3 = 1\rangle = |n_1 = 1, n_3 = 1\rangle$$

and

$$a_2 \,|n_2 = 1, n_1 = 1, n_3 = 1\rangle = |n_1 = 1, n_3 = 1\rangle .$$

where the states on the left hand side in these two equalities differ by a factor -1! So appears that acting with a_2 on two states differing by a minus sign, we get the *same* result. This can obviously not be right, and we must introduce a more subtle definition for the annihilation operator:

$$a_j \left|n_1 \dots n_j = 1 \dots\right\rangle = (-)^{\Sigma_j} \left|n_1 \dots n_j = 0 \dots\right\rangle$$

where

$$\Sigma_j = \sum_{k=1}^{j-1} n_k,$$

i.e. Σ_j counts the number of particles in the orbitals occurring *before* j in the occupation number state. It is now easy to check that the minus-sign problem is resolved with this definition. We can formulate the definition of a_j for the two cases where $n_j = 0$ or 1 in a concise way as follows:

$$a_j \left|n_1 \dots n_j \dots\right\rangle = n_j (-)^{\Sigma_j} \left|n_1 \dots 1 - n_j \dots\right\rangle .$$

A similar sign problem has to be dealt with when defining a creation operator. We may naively define

$$a_j^\dagger \left|n_1 \dots n_j = 0 \dots\right\rangle = \left|n_1 \dots n_j = 1 \dots\right\rangle$$

and require $a_j^\dagger$ to give zero when acting on a state with $n_j = 1$. In order to let this operator be the Hermitian conjugate of the correct annihilation operator, we must however have

$$a_j^\dagger \left|n_1 \dots n_j \dots\right\rangle = (1 - n_j)(-)^{\Sigma_j} \left|n_1 \dots 1 - n_j \dots\right\rangle .$$

The definitions introduced here have a nice consequence, as can be seen by calculating the *anti-commutation relation* relation between, say two annihilation operators for orbitals j and l ($j \neq l$). First, we note that the result of the anti-commutator acting on the state is zero if at least one of n_j and n_l is zero. If they are both 1, we have:

$$(a_j a_l + a_l a_j) \left|\dots n_j = 1 \dots n_l = 1 \dots\right\rangle = a_j (-)^{\Sigma_l} \left|\dots n_j = 1 \dots 0 \dots\right\rangle + a_l (-)^{\Sigma_j} |\dots 0 \dots n_l = 1 \dots\rangle .$$

In working out the last expression, we should realise ourselves that there is a difference of 1 between the Σ_l in the first term and the one arising from a_l in the second term, as in the latter, the particle in orbital j has already been *removed*. Therefore, we obtain two terms with opposite signs and the commutator vanishes. For $j = l$, the anti-commutator obviously vanishes, as we try to remove two particles from the same state. So we have found for each j, l:

$$\left\{a_j, a_l\right\} = a_j a_l + a_l a_j = 0,$$

where $\{\cdot,\cdot\}$ denotes the anti-commutator.

A similar anti-commutation relation can be found for the creation operators:

$$\left\{a_j^\dagger, a_l^\dagger\right\} = 0.$$

Now we analyse the anti-commutator between a creation and an annihilation operator:

$$(a_j a_j^\dagger + a_j^\dagger a_j) \left|\dots n_j \dots\right\rangle = a_j (1 - n_j)(-)^{\Sigma_j} \left|\dots 1 - n_j \dots\right\rangle + a_j^\dagger (-)^{\Sigma_j} n_j \left|\dots 0 \dots 1 - n_j \dots\right\rangle = \\ \left[(1 - n_j)^2 + n_j^2\right] \left|\dots n_j \dots\right\rangle ,$$

where, for $n_j = 0$ or 1, the result is always $\left|\dots n_j \dots\right\rangle$ preceded by 1, leading to

$$\left\{a_j, a_j^\dagger\right\} = 1.$$

It can be shown that, for $j \neq l$, $\left\{a_j, a_l^\dagger\right\} = 0$, so that we have

$$\left\{a_j, a_l^\dagger\right\} = \delta_{jl}.$$

All in all we have:

For bosons, we can define annihilation and creation operators for removing, respectively adding particles from or to orbitals. These annihilation and creation relation operators satisfy the following commutator algebra:

$$[a_j, a_l] = [a_j^\dagger, a_l^\dagger] = 0, \qquad [a_j, a_l^\dagger] = \delta_{jl}.$$

For fermions, we can define similar operators. However, in that case the commutation relations should be replaced by *anti-commutation* relations:

$$\{a_j, a_l\} = \left\{a_j^\dagger, a_l^\dagger\right\} = 0; \qquad \left\{a_j, a_l^\dagger\right\} = \delta_{jl}.$$

9.3 INTERACTING PARTICLE SYSTEMS

We consider a many-particle system with the Hamiltonian introduced in section 9.1:

$$H = \sum_{j=1}^{N} \frac{p_j^2}{2m} + \sum_{j=1}^{N} V(\mathbf{r}_j) + \frac{1}{2}\sum_{j,l=1}^{N} v(\mathbf{r}_j - \mathbf{r}_l) = \sum_{j=1}^{N} h(j) + \frac{1}{2}\sum_{j,l=1}^{N} v(\mathbf{r}_j - \mathbf{r}_l),$$

where $h(j) = \frac{p_j^2}{2m} + V(\mathbf{r}_j)$ is the single-particle Hamiltonian. This system describes point particles moving in an external potential V and interacting via a two-body potential v. In the latter, the factor $1/2$ is included with the omission of the condition that $j < l$ in the sum to compensate for double counting. We implicitly assume that $v = 0$ for $j = l$ (if $v \neq 0$ for $j = l$, we could include it into the external potential V). We now formulate this Hamiltonian in terms of creation and annihilation operators.

We therefore evaluate the matrix element of the Hamiltonian for two basis states which, for fermion systems, by convention, are two Slater determinants of the form:

$$\langle 1, 2, \ldots, N | \psi_{\text{AS}} \rangle = \frac{1}{\sqrt{N!}} \sum_P \text{sgn}(P) \langle 1 | \phi_{P_1} \rangle \langle 2 | \phi_{P_2} \rangle \langle 3 | \phi_{P_3} \rangle \cdots \langle N | \phi_{P_N} \rangle.$$

We first consider the matrix elements of the one-body potential h for two such basis states. In view of the indistinguishability of the particles, we may consider just $h(1)$ and multiply the result by N. We call the Slater determinants $\left|\psi_{\text{AS}}^A\right\rangle$ and $\left|\psi_{\text{AS}}^B\right\rangle$. If we expand the two Slater determinants in products of orbitals ϕ_j, we obtain terms of the form

$$\left\langle \phi_{A1}\phi_{A2}\cdots\phi_{AN} \right| h(1) \left| \phi_{B1}\phi_{B2}\cdots\phi_{BN} \right\rangle,$$

where anti-symmetrisation is *not* implicit in the bra- and ket states. Now note that for this expression not to vanish, we *must* have $A2 = B2$, $A3 = B3$ etcetera, as the set of orbitals ϕ_j is orthonormal. For a non-vanishing matrix element, only $A1$ and $B1$ may be different, as they occur on both sides of $h(1)$. We conclude that matrix elements of h between two Slater determinants are non-zero only if the orbitals from which the two Slater determinants have been composed, differ at most by one orbital. If indeed an orbital j occurs *only* in $\left|\psi_{\text{AS}}^A\right\rangle$ and, likewise, orbital l *only* occurs in $\left|\psi_{\text{AS}}^B\right\rangle$, while all the other orbitals in both Slater determinants are pairwise equivalent, the matrix element will be

$$\left\langle \phi_j \right| h \left| \phi_l \right\rangle.$$

The pre-factor is correct, since the orbitals j and l *must both* be occupied by particle 1; the other $N-1$ orbitals must be in the same order in both A and B, but that still leaves room for $(N-1)!$ permutations. Together with the factor N arising from the fact that we have looked at $h(1)$ only, this cancels the $1/N!$ from the normalisation of the two Slater determinants.

If the orbitals in $|\psi^A_{\rm AS}\rangle$ are the same as those in $|\psi^B_{\rm AS}\rangle$, we obtain

$$\sum_j \langle \phi_j | h | \phi_j \rangle$$

(note that the sum over j is the same for the set of orbitals in A as in B). Again, we can assign all the N orbitals j in the sets A and B to the first place (hence the sum over j) and for each choice we can permute the other orbitals in $(N-1)!$ ways, which, together with the prefactor N, cancels the normalisation of $1/N!$.

Now we claim that the operator

$$\sum_{jl} h_{jl} a^\dagger_j a_l \text{ with } h_{jl} = \langle \phi_j | h | \phi_l \rangle, \tag{9.3}$$

where the sum is over *all* possible indices j and l, is the correct matrix representation for this one-particle Hamiltonian. This can easily be seen by studying its matrix elements between two Slater determinants, which we shall now write in the occupation number representation. When the two states on the left- and right hand side contain exactly the same set of occupied orbitals, we have

$$\sum_{jl} h_{jl} \left\langle n_1 n_2 \ldots \middle| a^\dagger_j a_l \middle| n_1 n_2 \ldots \right\rangle = \sum_{j\in A} h_{jj} = \sum_j \langle \phi_j | \, h \, | \phi_j \rangle .$$

Now suppose that in the ket, $N-1$ occupied orbitals are identical to orbitals in the bra-vector so that these two vectors differ in only one orbital which is the orbital k in the bra, and m in the ket. Our operator gives again the right result:

$$\sum_{jl} h_{jl} \left\langle n_1 n_2 \ldots n_k = 1 \ldots \middle| a^\dagger_j a_l \middle| n_1 n_2 \ldots n_m = 1 \ldots \right\rangle = h_{km} = \langle \phi_k | \, h \, | \phi_m \rangle .$$

It is also easily seen that, when more than one orbital differ in the bra and the ket, the matrix element vanishes. As the operator $\sum_{jl} h_{jl} a^\dagger_j a_l$ gives the correct matrix elements between all possible basis vectors, it is the correct representation of the single particle Hamiltonian.

Now we turn to the two-particle interaction, which we evaluate again between $|\psi^A_{\rm AS}\rangle$ and $|\psi^B_{\rm AS}\rangle$. Similar to the analysis of the single-particle operator h, we can evaluate the action of this interaction between particles 1 and 2, and multiply the result by $N(N-1)/2$, which is the total number of distinct pairs we can make. Similar to the case of the single particle operator, we can argue that we should have two identical sets of orbitals in the bra- and ket vector, which should be occupied by particles 3 to N; only the orbitals of particle 1 and 2 are free in both. Interestingly, a similar analysis as above, tells us that for *identical* orbital sets in the two vectors the matrix element reduces to

$$\frac{1}{2} \sum_{jk\in A} \left(\langle jk | v | jk \rangle - \langle jk | v | kj \rangle \right),$$

where

$$\langle jk | v | lm \rangle = \int d^3 r_1 \int d^3 r_2 \, \phi^*_j(\mathbf{r}_1) \phi^*_k(\mathbf{r}_2) v(\mathbf{r}_1 - \mathbf{r}_2) \phi_l(\mathbf{r}_1) \phi_m(\mathbf{r}_2).$$

The pre-factor requires some care. We can select $N(N-1)$ orbital pairs jk. This should then be multiplied by the pre-factor $1/N!$ deriving from the normalisation factors of the many body wave functions. Furthermore, the remaining $N-2$ states (i.e. not j and k) can be ordered in $(N-2)!$ ways. Together with the prefactor of $1/2$ in the Hamiltonian this multiplies to exactly 1/2. Note that the sum over j and k is unrestricted: each pair occurs twice in this sum.

Now we consider the case where the two many-particle states differ in *one* orbital. In that case, the two different orbitals (which we denote $k \in A$ and $m \in B$ should involve particles 1 and 2, and we can choose one additional orbital (which we denote $|j\rangle$) of the set of $N-1$ shared orbitals. The matrix element then becomes

$$\frac{1}{2}\sum_{j\in A, j\neq k} \left(\langle jk|\,v\,|jm\rangle - \langle jk|\,v\,|mj\rangle\right).$$

The combinatorics: k and m are fixed, and the sum over j takes care of all the other orbitals. Once j is chosen, $N-2$ orbitals are left, which can be ordered in $(N-2)!$ ways. Swapping the other of two orbitals in both the left and right many body wave function gives a factor of 2, and combining all this with the pre-factor and the normalisation of the Slater determinants, we see that the correct pre-factor is 1.

Finally, if two orbitals on the bra have no counterpart in the ket, we have

$$\frac{1}{2}\left(\langle jk|v|lm\rangle - \langle jk|v|ml\rangle\right)$$

where j, k are the two labels of the bra differing from lm in the ket. All other matrix elements give zero.

We now show that the operator

$$\frac{1}{2}\sum_{jklm} \langle jk|\,v\,|lm\rangle\, a_j^\dagger a_k^\dagger a_m a_l$$

gives the same matrix elements in the occupation number representation. Let us consider the case where the orbitals in bra and ket are all identical. Then it is easy to see that when we put our operator between the two states, either $j = l$ and $k = m$ or $j = m$ and $k = l$, so we are left with

$$\frac{1}{2}\sum_{jk\in A}\Big(\langle jk|\,v\,|jk\rangle\langle \dots n_j = 1\dots n_k = 1\dots|\,a_j^\dagger a_k^\dagger a_k a_j\,|\dots n_j = 1\dots n_k = 1\dots\rangle +$$
$$\langle jk|\,v\,|kj\rangle\langle \dots n_j = 1\dots n_k = 1\dots|\,a_j^\dagger a_k^\dagger a_j a_k\,|\dots n_j = 1\dots n_k = 1\dots\rangle\Big)$$

which directly yields

$$\frac{1}{2}\sum_{jk\in A}\left(\langle jk|\,v\,|jk\rangle - \langle jk|\,v\,|kj\rangle\right),$$

where the minus sign on the right hand side follows from the anti-commutation relations of the fermion creation and annihilation operators. The correctness for the two other cases can be checked in a similar way, and is left as an exercise to the reader. We summarise the very important results obtained in this section:

For a many-body system described by a N-particle Hamiltonian

$$H = \sum_{j=1}^{N}\frac{p_j^2}{2m} + \sum_{j=1}^{N} V(\mathbf{r}_j) + \frac{1}{2}\sum_{j,l=1}^{N} v(\mathbf{r}_j - \mathbf{r}_l) = \sum_{j=1}^{N} h(j) + \frac{1}{2}\sum_{j,l=1}^{N} v(\mathbf{r}_j - \mathbf{r}_l),$$

in Fock-space, this Hamiltonian is:

$$\hat{H} = \sum_{jk} h_{jk} a_j^\dagger a_k + \frac{1}{2}\sum_{jklm}\langle jk|v|lm\rangle\, a_j^\dagger a_k^\dagger a_m a_l. \tag{9.4}$$

This is the formulation of the Hamiltonian in second quantisation.

9.4 CHANGE OF BASIS – FIELD OPERATORS

We have seen that the creation and annihilation operators $a_j^\dagger$ and a_j create or annihilate particles in an orbital $|\phi_j\rangle$. Now suppose we want creation and annihilation operators which create and annihilate particles in orbitals taken from a *different* set, called $|u_\alpha\rangle$. The orbitals sets $|\phi_j\rangle$ and $|u_\alpha\rangle$ both are basis sets of the *same* single particle Hilbert space $\mathscr{H}$. In order to work out the form of a creation operator $a_\alpha^\dagger$ creating a particle in the orbital u_α, we write the action of this operator in a particular way. When a creation operator $a_\alpha^\dagger$ acts on a N-particle Slater determinant $|\psi(N)\rangle$ which itself does not contain the orbital $|u_\alpha\rangle$, it yields a $N+1$-particle state:

$$a_\alpha^\dagger |\psi(N)\rangle = \hat{A}|u_\alpha \psi(N)\rangle,$$

where we have introduced the *anti-symmetrisation operator* $\hat{A}$, which constructs a Slater determinant consisting of the orbitals in $|\psi(N)\rangle$ and u_α. Now we can write

$$\hat{A}\sum_j \left(|\phi_j\rangle\langle\phi_j\,|\,u_\alpha\rangle\right)|\psi(N)\rangle.$$

As $\langle\phi_j\,|\,u_\alpha\rangle$ is a number, we can move that in front of the anti-symmetrisation operator:

$$a_\alpha^\dagger |\psi(N)\rangle = \sum_j \langle\phi_j|u_\alpha\rangle\,\hat{A}\left(|\phi_j\psi(N)\rangle\right) = \sum_j \langle\phi_j|u_\alpha\rangle\, a_j^\dagger |\psi(N)\rangle$$

showing that

$$a_\alpha^\dagger = \sum_j \langle\phi_j\,|\,u_\alpha\rangle\, a_j^\dagger. \tag{9.5}$$

For the annihilation operators we have the transformation rule

$$a_\alpha = \sum_j \langle u_\alpha\,|\,\phi_j\rangle\, a_j, \tag{9.6}$$

which is just the Hermitian conjugate of (9.5).

Finally, we introduce the concept of *field operators.* These are creation and annihilation operators of a particle located at **r**. Using the rule we have just derived, we see that the field operators should be defined as

$$\boldsymbol{\psi}^\dagger(\mathbf{r}) = \sum_j a_j^\dagger \langle\phi_j|\mathbf{r}\rangle; \qquad \boldsymbol{\psi}(\mathbf{r}) = \sum_j a_j \langle\mathbf{r}|\phi_j\rangle$$

These operators create or annihilate a particle in a state $|\mathbf{r}\rangle$.

We now write the Hamiltonian we have discussed in section 9.3 in terms of these field operators. Using the basis $|\mathbf{r}\rangle$, the one-particle Hamiltonian can directly be written as

$$H^{(1)} = \sum_i h(i) = \int d^3r \int d^3r' \boldsymbol{\psi}^\dagger(\mathbf{r})\,\langle\mathbf{r}|h|\mathbf{r}'\rangle\,\boldsymbol{\psi}(\mathbf{r}').$$

For a local potential, $V(\mathbf{r})$ in h, the central matrix element is diagonal and we obtain

$$V = \int d^3r\, \boldsymbol{\psi}^\dagger(\mathbf{r}) V(\mathbf{r}) \boldsymbol{\psi}(\mathbf{r}).$$

For the kinetic energy T, we have

$$T = \sum_i \frac{p_i^2}{2m} = \frac{1}{2m}\int d^3r \int d^3r' \boldsymbol{\psi}^\dagger(\mathbf{r})\,\langle\mathbf{r}|\hat{p}^2|\mathbf{r}'\rangle\,\boldsymbol{\psi}(\mathbf{r}').$$

Fourier expanding to transform the integrals over **r** we obtain

$$T = \frac{1}{2m}\int d^3p \int d^3r \int d^3r' \boldsymbol{\psi}^\dagger(\mathbf{r})\,\langle\mathbf{r}|\mathbf{p}\rangle\, p^2\,\langle\mathbf{p}|\mathbf{r}'\rangle\,\boldsymbol{\psi}(\mathbf{r}') = \frac{1}{2m}\int d^3p \boldsymbol{\psi}^\dagger(\mathbf{p}) p^2 \boldsymbol{\psi}(\mathbf{p}),$$

where we have used the inverse Fourier transform

$$\boldsymbol{\psi}(\mathbf{r}) = \int \langle \mathbf{r}|\mathbf{p}\rangle \langle \mathbf{p}|\boldsymbol{\psi}\rangle \, d^3p.$$

The expression over $p^2\boldsymbol{\psi}(\mathbf{p})$ is recognised as the Fourier transform of $-\nabla^2\psi(\mathbf{r})$, and transforming back, we have

$$T = \frac{-1}{2m}\int d^3r \; \boldsymbol{\psi}^\dagger(\mathbf{r})\nabla^2\boldsymbol{\psi}(\mathbf{r}).$$

All in all we therefore have

$$H^{(1)} = \int d^3r \boldsymbol{\psi}^\dagger(\mathbf{r})\hat{h}\boldsymbol{\psi}(\mathbf{r}).$$

Now the interaction term can be analysed in a similar way, leading to the result

$$\hat{v} = \frac{1}{2}\int d^3r\, d^3r' \; \boldsymbol{\psi}^\dagger(\mathbf{r})\boldsymbol{\psi}^\dagger(\mathbf{r}')\, v(\mathbf{r}-\mathbf{r}')\boldsymbol{\psi}(\mathbf{r})\boldsymbol{\psi}(\mathbf{r}').$$

9.5 EXAMPLES OF MANY-BODY SYSTEMS

9.5.1 MANY NON-RELATIVISTIC PARTICLES IN A BOX

The particle in a box is a very common problem which is covered in textbooks on elementary quantum mechanics. For convenience, we use a big box here with periodic boundary conditions, so that the solutions of the one-particle Hamiltonian are of the form

$$\frac{1}{\sqrt{V}}\exp(i\mathbf{k}\cdot\mathbf{r})$$

where

$$\mathbf{k} = 2\pi\left(\frac{n_x}{L_x}, \frac{n_y}{L_y}, \frac{n_z}{L_z}\right);$$

L_x, L_y and L_z define the size of the rectangular box. The prefactor $1/\sqrt{V}$, where $V = L_xL_yL_z$, ensures proper normalisation of the plane wave inside that box. Inside the box, the potential is constant – we take it to be 0. It is natural to introduce creation and annihilation operators $a^\dagger(\mathbf{p})$ and $a(\mathbf{p})$, which create and annihilate particle with momentum $\mathbf{p} = \hbar\mathbf{k}$.

Now suppose we want to count the number of particles in the box. The state of the system may be a superposition of Slater determinants of different sizes and involving different orbitals, but the number of particles is always found using the operator

$$N = \sum_{\mathbf{p}} a^\dagger_{\mathbf{p}} a_{\mathbf{p}}.$$

For a large box, the sum over $\mathbf{p}$ can be turned into an integral as we precisely know which vectors $\mathbf{k}$ are summed over:

$$\sum_{\mathbf{k}} \to \frac{V}{(2\pi)^3}\int d^3k,$$

where $V = L_xL_yL_z$, so, using $\mathbf{p} = \hbar\mathbf{k}$, we have

$$N = \frac{V}{(2\pi\hbar)^3}\int d^3p \; a^\dagger(\mathbf{p})a(\mathbf{p}). \tag{9.7}$$

Note that $a_{\mathbf{p}}$ is used for the *discrete* set of $\mathbf{p}$'s which fit into the periodic box, whereas the notation $a(\mathbf{p})$ is used for continuous $\mathbf{p}$'s.

We can do the same for the Hamiltonian

$$H = \frac{V}{(2\pi\hbar)^3}\int d^3p \, \frac{p^2}{2m} a^\dagger(\mathbf{p})a(\mathbf{p}).$$

Instead of plane wave orbitals, we can also use orbitals describing particles localised at **r**: the corresponding operators are the field operators introduced in section 9.4. We can construct those operators from our orthonormal plane wave basis set. This gives

$$\boldsymbol{\psi}^\dagger(\mathbf{r}) = \sum_{\mathbf{p}} \frac{e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}} a^\dagger_{\mathbf{p}} = \frac{\sqrt{V}}{(2\pi\hbar)^3}\int d^3p\, e^{-\mathrm{i}\mathbf{p}\cdot\mathbf{r}} a^\dagger(\mathbf{p})$$

and

$$\boldsymbol{\psi}(\mathbf{r}) = \frac{\sqrt{V}}{(2\pi\hbar)^3}\int d^3p\, e^{\mathrm{i}\mathbf{p}\cdot\mathbf{r}} a(\mathbf{p}).$$

We can now write the total number of particles as

$$N = \int d^3r\, \boldsymbol{\psi}^\dagger(\mathbf{r})\boldsymbol{\psi}(\mathbf{r}),$$

and we show that this expression is identical to (9.7). Writing out the expressions for $\boldsymbol{\psi}^\dagger$ and $\boldsymbol{\psi}$, we obtain:

$$N = \int d^3r\, \frac{V}{(2\pi\hbar)^6}\int d^3p \int d^3p'\, e^{-\mathrm{i}\mathbf{p}\cdot\mathbf{r}} a^\dagger(\mathbf{p}) e^{\mathrm{i}\mathbf{p}'\cdot\mathbf{r}} a(\mathbf{p}').$$

Integrating over **r** yields $(2\pi\hbar)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}')$, and the expression reduces to (9.7) indeed.

The Hamiltonian becomes

$$H = \int d^3r\, \boldsymbol{\psi}^\dagger(\mathbf{r})\frac{-\hbar^2}{2m}\nabla^2\boldsymbol{\psi}(\mathbf{r}).$$

You may ask at this stage what we can learn from the formalism just presented. In this form, it is not directly clear what is the use of this many-body formulation of free particles. In fact, this formalism is be very powerful when dealing with interacting particles. We will go a bit into this in chapter 10.

9.5.2 THE HEISENBERG MODEL AND THE JORDAN-WIGNER TRANSFORMATION

In this section we shall illustrate how a second-quantised form arises when analysing a one-dimensional chain of spin-1/2 particles. Consider such a chain of length N, where there is a spin-1/2 particle sitting on each site of the chain, and where there is a nearest neighbour interaction:

$$H = -\sum_i \left[J_\perp\left(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y\right) + J_z S_i^z S_{i+1}^z\right].$$

In this section, sums run from site 1 to N; periodic boundary conditions are assumed: $1 \equiv N+1$. The operators $S_i^{x,y,z}$ are the spin-operators, acting on site i. They satisfy the usual commutation relations

$$[S_i^\alpha, S_i^\beta] = \mathrm{i}\epsilon_{\alpha\beta\gamma} S_i^\gamma$$

where α, β and γ run over the Cartesian coordinates x, y, z and $\epsilon_{\alpha\beta\gamma}$ is the anti-symmetric tensor.

Let us turn this Hamiltonian into a fermion chain by identifying the spin-down state $|\downarrow\rangle$ with 'empty' (no particle) and $|\uparrow\rangle$ with 'occupied' (one particle):

$$|\uparrow\rangle \equiv |1\rangle$$
$$|\downarrow\rangle \equiv |0\rangle\,.$$

Note that we have left out the site index i for the time being. In the spin picture, we can switch between up and down via the raising and lowering operators

$$S^+ = S^x + \mathrm{i}S^y$$
$$S^- = S^x - \mathrm{i}S^y,$$

which have the form (in the basis $|\uparrow\rangle$, $|\downarrow\rangle$):

$$S^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \qquad S^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

In the language of the occupation numbers, these operators take the form of creation and annihilation operators

$$S^+ \leftrightarrow c^\dagger \qquad S^- \leftrightarrow c$$

where c and $c^\dagger$ are the usual creation and annihilation operators:

$$\begin{aligned} c^\dagger |0\rangle &= |1\rangle, & c|1\rangle &= |0\rangle; \\ c^\dagger |1\rangle &= 0, & c|0\rangle &= 0. \end{aligned}$$

It seems that we can now just replace the S_i^+ by the $c_i^\dagger$ and similarly $S_i^- \to c_i$. However, this poses a problem, as the spin operators at *different* sites commute, whereas, if the c_i, $c_i^\dagger$ would be real fermion operators, they should *anti-commute* at different sites. A small modification in the definition of the c_i and $c_i^\dagger$ however fixes this. Using $\Sigma_i \equiv \sum_{j<i} n_j$, where $n_j = 1$ for the occupied, and $n_j = 0$ for the unoccupied sites, we obtain correct fermion operators c_i and $c_i^\dagger$ according to

$$\begin{aligned} c_i^\dagger &= (-)^{\Sigma_i} S_i^+; \\ c_i &= (-)^{\Sigma_i} S_i^-. \end{aligned}$$

You should verify for yourself that the c_i and $c_i^\dagger$ all anti-commute at different sites!

Using the fact that $S_z = S_i^+ S_i^- - 1/2$ (we take $\hbar \equiv 1$), we can now formulate the Hamiltonian in terms of the c_i:

$$H = -\sum_i \left[\frac{J_\perp}{2} \left(c_i^\dagger c_{i+1} + c_{i+1}^\dagger c_i \right) + J_z \left(\frac{1}{4} - c_i^\dagger c_i + c_i^\dagger c_i c_{i+1}^\dagger c_{i+1} \right) \right].$$

We have arrived at a formulation of the Heisenberg Hamiltonian in terms of a fermion chain. Note that the fermions have no spin. The first part of the Hamiltonian describes *hopping*: the ability of fermions to move from site i to $i+1$ and back. The second term contains a contribution proportional to $n_i n_{i+1}$. This is interpreted as a 'Coulomb interaction' between neighbouring sites: it is non-zero only when the two sites are occupied by a fermion.

In problem 9 we shall solve the spectrum for the case where $J_z = 0$.

9.6 SUMMARY

In this chapter, we have introduced the Fock space, defined as

$$\mathscr{F} = \mathscr{H}(1) \oplus \mathscr{H}(2) \oplus \mathscr{H}(3) \oplus \mathscr{H}(4) \ldots$$

Here, $\mathscr{H}(N)$ is the Hilbert space for N-particles. Within each Hilbert space, the states are either symmetric under particle exchange – the particles are then called bosons – or anti-symmetric – then they are fermions. The spin statistics theorem tells us that particles with integer spin are bosons and those with half-integer spin are fermions. A convenient way to denote states with n_j particles in level j is the *occupation number representation*:

$$|\psi\rangle = |n_1, n_2, \ldots\rangle.$$

These states are symmetric or anti-symmetric many-body states for bosons / fermions respectively, with single-particle orbitals occupied with occupation numbers n_j.

Creation and annihilation operators move us from the Hilbert space $\mathscr{H}(N)$ to $\mathscr{H}(N+1)$ and $\mathscr{H}(N-1)$ respectively. They are defined as follows:

- For bosons:

$$a_j \left|n_1 n_2 \ldots n_j \ldots\right\rangle = \sqrt{n_j} \left|n_1 n_2 \ldots n_j - 1 \ldots\right\rangle$$

  and

$$a_j^\dagger \left|n_1 n_2 \ldots n_j \ldots\right\rangle = \sqrt{n_j + 1} \left|n_1 n_2 \ldots n_j + 1 \ldots\right\rangle.$$

  These operators satisfy the commutation relations

$$[a_j, a_l] = [a_j^\dagger, a_l^\dagger] = 0, \qquad [a_j, a_l^\dagger] = \delta_{jl}.$$

- For Fermions, we use $\Sigma_j = \sum_{k=1}^{j-1} n_j$ in the definition of the annihilation operator:

$$a_j \left|n_1 \ldots n_j \ldots\right\rangle = n_j (-)^{\Sigma_j} \left|n_1 \ldots 1 - n_j \ldots\right\rangle.$$

  and in that of the creation operator:

$$a_j^\dagger \left|n_1 \ldots n_j = 0 \ldots\right\rangle = \left|n_1 \ldots n_j = 1 \ldots\right\rangle.$$

  They satisfy the anti-commutation relations:

$$\{a_j, a_l\} = \left\{a_j^\dagger, a_l^\dagger\right\} = 0; \qquad \left\{a_j, a_l^\dagger\right\} = \delta_{jl}.$$

Moving from a orbital basis $\left|\phi_j\right\rangle$ to a new orbital basis $|u_\alpha\rangle$, changes the creation and annihilation operators as follows:

$$a_\alpha^\dagger = \sum_j \langle \phi_j | u_\alpha \rangle \, a_j^\dagger$$

and

$$a_\alpha = \sum_j \langle u_\alpha | \phi_j \rangle \, a_j.$$

The Hamiltonian of a many-particle system can be formulated in terms of creation and annihilation operators as follows:

$$H = \sum_{jk} h_{jk} a_j^\dagger a_k + \frac{1}{2} \sum_{jklm} \langle jk|v|lm\rangle \, a_j^\dagger a_k^\dagger a_m a_l.$$

9.7 PROBLEMS

1. In this problem, c and $c^\dagger$ are fermion operators; a and $a^\dagger$ are boson operators.

   (a) Calculate, or try to write in the most compact way (i.e., using the smallest number of operators):

$$[a, aa^\dagger]$$

$$[c, a]$$

$$[c, c^\dagger c]$$

$$\{c, c^\dagger c\}$$

$$[ca, a^\dagger]$$

   (b) Show that

$$e^c e^{c^\dagger} = 1 + c + c^\dagger + cc^\dagger.$$

2. Consider a system with a one-particle Hilbert space of dimension N.

(a) Denote by N_S the dimension of the Hilbert space for a system of two identical bosonic particles. Find N_S.

(b) Denote by N_{AS} the dimension of the Hilbert space for a system of two identical fermionic particles. Find N_{AS}.

(c) Show $N_S + N_{AS} = N^2$.

(d) Make a plot or table of N_S and N_{AS} as a function of N.

(e) Now consider a system where single electrons can occupy two orbital levels. What is the dimension N of the one-particle Hilbert space (don't forget that the electron is a spin-1/2 particle!)? What is the dimension N_{AS} of the two-electron Hilbert space?

(f) Denote the two orbital states $|\phi_1\rangle$ and $|\phi_2\rangle$. One can construct a basis (we'll call it basis 1) for the two-electron system by considering all possible combinations of symmetric (antisymmetric) orbital and antisymmetric (symmetric) spin states, so that each basis state is overall antisymmetric. Construct this basis.

(g) Another basis (basis 2) for the two-electron system is obtained by constructing all possible Slater determinants, starting from the four wave functions

$$|\phi_1,+\rangle, \quad |\phi_2,+\rangle, \quad |\phi_1,-\rangle, \quad |\phi_2,-\rangle.$$

Find basis 2.

(h) Express the two-electron wave functions in basis 1 in terms of those in basis 2.

(i) Express the two-electron wave functions in basis 1 and in basis 2 in second quantisation. i.e., using creation operators acting on the vacuum state.

3. Consider now a system of n identical particles with single-particle Hilbert space of dimension N.

(a) Find $N_S(n,N)$.

(b) Find $N_{AS}(n,N)$. Under what conditions is $N_{AS}(n,N) > 0$?

(c) In the previous problem, you showed that two-particle systems satisfy $N_S(2,N) + N_{AS}(2,N) = N^2$. Show that $N_S(n,N) + N_{AS}(n,N) = N^2$ only for $n = 2$. What can you say for $n > 2$?

4. Consider a chain of fermions described by the Hamiltonian

$$H = -t\sum_{j=1}^{N}\left(c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j\right)$$

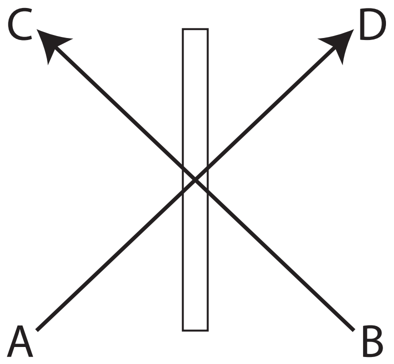
where periodic boundary conditions impose $N+1 \equiv 1$.

Use the Heisenberg equation of motion

$$\dot{A} = \mathrm{i}\,[H,A]$$

to find the time derivative of the operators $\dot{c}_j(t)$ and $\dot{c}_{j+1}(t)$.

5. We consider the scattering of particles at a beam splitter with input ports A and B and output ports C and D (see Figure). We assume throughout that no interactions take place at the beam splitter. Scattering is described by matrix elements $\langle\psi_c|U|\psi_a\rangle = r$, $\langle\psi_d|U|\psi_a\rangle = t$, $\langle\psi_c|U|\psi_b\rangle = t$, and $\langle\psi_d|U|\psi_b\rangle = -r$, where real coefficients r and t satisfy $t^2 + r^2 = 1$. The operator U describes the time evolution between the beginning of the experiment (when the particles are still in ports A and/or B) and the end (when the particles are in C and/or D).



SCHEMATIC OF A BEAM SPLITTER.

(a) Show that the matrix

$$S = \begin{pmatrix} r & t \\ t & -r \end{pmatrix}$$

is unitary. Why is this required?

(b) Consider two photons of identical polarisation and frequency, one in A and the other in B, incident on the beam splitter.Write the two-particle input wave function. Don't forget to symmetrise!

(c) Calculate the output wave function. What can you say for $r = t = 1/\sqrt{2}$? This quantum phenomenon is called *bunching*. What would you expect classically? If you wonder what we mean by classical, imagine two billiard balls scattering independently.

(d) Similarly, consider now the case of two electrons, one at each input, and both spin up. Write the two-electron input wave function. Don't forget to anti-symmetrise!

(e) Calculate the output wave function. Explain why this result is different from what you would expect classically. This phenomenon, you might have already guessed, is known as *anti-bunching*.

(f) Finally, Consider two incident electrons, again one on each arm, but now in a spin singlet configuration. Write the two-electron input wave function, and calculate the output wave function. Do the electrons bunch or anti-bunch? Therefore, does the overall particle symmetry (fermion, boson) dictate whether the two particles will bunch or anti-bunch upon scattering at a beam splitter?

6. We will now repeat the previous exercise, but using second quantisation. Define a and b as the annihilation operators of incident states $|\psi_a\rangle$ and $|\psi_b\rangle$, respectively. Similarly, define c and d as the annihilation operators of outgoing states $|\psi_c\rangle$ and $|\psi_d\rangle$.

(a) Using second quantisation, Write the input state for the case of two identical photons. Use a, b, c and d for the annihilation operators for particles in port A, B, C and D respectively.

(b) In the Heisenberg picture, The final output operators are related to the initial input operators by

$$\begin{pmatrix} c \\ d \end{pmatrix} = S \begin{pmatrix} a \\ b \end{pmatrix}.$$

Invert to find expressions for a and b in terms of c and d.

(c) Use this result and the bosonic commutation relations to calculate the output state. You should get the same result as in 5(c).

(d) Also using second quantisation, write the input state for the case of two electrons with identical spin. Calculate the output state by exploiting only the fermionic commutation relations. You should get the same result as in 5(e).

(e) Now consider the two-electron case in which the electrons are initially in a spin singlet. Note that for creating such a state, you need creation operators such as $a_\sigma^\dagger$, where
$\sigma = \uparrow, \downarrow$ denotes the spin. Write the input state in second quantisation. Calculate the output state. You should get the same result as in 5(f).

7. In this problem, we consider the scattering of coherent states at a beamsplitter. It is advised to use the operator formalism of the previous problem!

(a) Consider a coherent state $|\alpha\rangle$ incident in A and vacuum $|0\rangle$ incident in B. Calculate the output state. Show that it consists of a product of coherent states in C and D.

(b) Consider now coherent states incident at both inputs, $|\alpha\rangle$ in A, and $|\beta\rangle$ incident in B. Calculate the output state. Is there entanglement produced between the two output beams in this case?

8. Consider a particle moving on a periodic line of length L. The particle is subject to a (periodic) potential $V(x)$.

(a) Write up an equation for the expectation value of the energy of a state $|\psi\rangle$ which is normalised on the periodic line.

(b) Now we discretise the particle positions to a very narrowly spaced grid with grid constant $h = L/N$. Give the representation of the Laplace operator on that grid. Write the expectation value of the energy now in terms of a sum over the sites of the dense grid.

(c) Show that a Hamiltonian of the form

$$H = \sum_{i=0}^{N-1} \left[A\left(a_i^\dagger a_{i+1} + a_{i+1}^\dagger a_i\right) + B_i a_i^\dagger a_i \right]$$

gives the same expectation for the energy of a single particle for particular values of A and B_i. Calculate these values.

9. Consider the hopping Hamiltonian

$$H = -\sum_j \frac{J_\perp}{2} \left(c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j \right).$$

Now introduce the Fourier transform of the c-operators

$$c_k = \frac{1}{\sqrt{N}} \sum_j e^{\mathrm{i}kj} c_j,$$

where the sum runs over the sites j and k takes on the values $2\pi n/N$. Show that the Hamiltonian can be rewritten as

$$H = \sum_k \frac{J_\perp}{2} \omega_k c_k^\dagger c_k.$$

Calculate the dispersion relation ω_k.

10. For theoretically studying properties of electrons in a solid, often the so-called *Hubbard model* is considered. The partition function of this model was solved exactly by Lieb and Wu in the early sixties. It predicts for some parameter values a transition between a conductor and an insulator.

    The Hubbard model describes electrons that hop from atom to atom. Only nearest neighbour hopping is allowed. In the one-dimensional version, the hopping is described by the following term in the Hamiltonian:

    $$T = \sum_{i,s} \left(\tau c^\dagger_{i,s} c_{i+1,s} + \tau^* c^\dagger_{i+1,s} c_{i,s} \right)$$

    where i labels the sites on a one-dimensional chain of atoms (sites) where the electrons can reside. It runs from 1 to N and it is periodic, i.e. $1 \equiv N+1$. The label s is for the spin and can be + or −. The operators c and $c^\dagger$ are fermion creation and annihilation operators.

    (a) Explain why this term describes hopping along the chain.

    The second term in the Hamiltonian provides an energy penalty U for two electrons to be at the same atom:

    $$V = U \sum_i n_{i,-} n_{i,+}.$$

    (b) Show that the particle number is conserved.

    (c) Let $S_{i,\mu} = \frac{\hbar}{2} \sum_{ss'} c^\dagger_{i,s} (\sigma_\mu)_{ss'} c_{is'}$ be the spin operator at site i with σ_μ the Pauli matrices.

    Compute

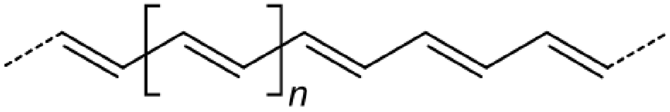
    $$|\mathbf{S}_i|^2 = \sum_{\mu=x,y,z} \left(S_{j,\mu} \right)^2 = \hbar^2 \left[\frac{3}{4} \left(n_{i,+} + n_{i,-} \right) - \frac{3}{2} n_{i,+} n_{i,-} \right]$$

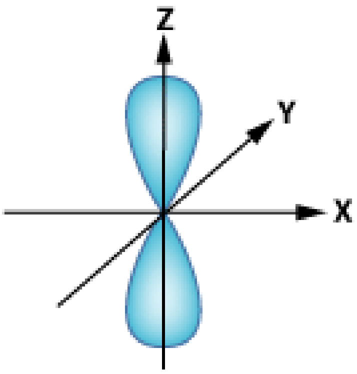
    ($n_{i,s}$ are number operators) and show that the Hubbard Hamiltonian can be expressed in terms of spin-operators (for real τ) as

    $$H^{\text{Hubbard}} = -\tau \sum_i \sum_s \left(c^\dagger_{i,s} c_{i+1,s} + c^\dagger_{i+1,s} c_{i,s} \right) - \frac{2}{3\hbar} U \sum_i |\mathbf{S}_i|^2 + \frac{U}{2} \hat{N}$$

    where $\hat{N} = \sum_i n_i$.

11. *Su-Shrieffer-Heeger model for acetylene*
    In this problem we study acetylene chains. Acetylene is a chain of C–H groups, arranges in a zig-zag form with angles of 120^o between successive bonds. The electronic properties are, just as in graphene, determined by the electrons in the p_z orbitals, where z is the direction perpendicular to the plane of the chain, see figure.





A particular feature of these chains is that some bonds may shrink a bit, whereas others stretch. We therefore include the displacement u_n of the atoms into the Hamiltonian. Here *positive* u_n is the displacement along the bond *right* of atom n, and negative u_n denotes a displacement along the *left* bond. We neglect the electron spin. The full Hamiltonian can be written as

$$H = -\sum_{n=0}^{N-2} [t - \alpha(u_{n+1} - u_n)] \left[c_n^\dagger c_{n+1} + h.c.\right] + \frac{\kappa}{2} \sum_{n=0}^{N-2} (u_{n+1} - u_n)^2,$$

where $h.c.$ stands for 'hermitian conjugate' as usual. N is odd.

The average length of the bonds is a, but the chains may *dimerise*: the double bonds will contract whereas the single bonds stretch a little bit. Some contemplation should convince you that this can be represented as

$$u_n = u(-1)^n,$$

where it is assumed that the left atoms of a double bond have even n, the atoms right of the double bond have odd n. The leftmost bond is taken to be a double bond.

(a) Show that the Hamiltonian can now be rewritten as

$$H = -\sum_{n=0}^{N-2} \left\{ \left[t + 2\alpha u(-1)^n\right] \left(c_n^\dagger c_{n+1} + h.c.\right) - 2\kappa u^2 \right\}.$$

(b) We now perform an important step by considering two electrons adjacent to one double bound as being part of one unit cell. We label these cells by m (and we assume that N is even). The *left* atoms of the double bond are labelled mA and the right atoms mB. Show that with this notation, the Hamiltonian reads

$$H = -\sum_{m=0}^{N/2-1} \left[t\left(c_{mA}^\dagger c_{mB} + c_{m+1,A}^\dagger c_{mB} + h.c.\right) + 2\alpha u\left(c_{mA}^\dagger c_{mB} - c_{m+1,A}^\dagger c_{mB} + h.c.\right) - 4\kappa u^2 \right] - 2\kappa u^2.$$

Setting $2u\alpha = t\Delta$ and neglecting the term $4\kappa u^2$, this can be rewritten in the form

$$H = -\sum_{m=0}^{N/2-1} \left[t(1+\Delta)\left(c_{mA}^\dagger c_{mB} + h.c.\right) + t(1-\Delta)\left(c_{m+1,A}^\dagger c_{mB} + h.c.\right)\right].$$

(c) Solve this Hamiltonian by trying the solution

$$\psi_m = \begin{pmatrix} u_A \\ u_B \end{pmatrix} e^{2ikma}.$$

This leads to a 2 × 2 matrix for each k. Diagonalise this matrix in order to show that

$$E(k) = \pm 2t\sqrt{1 + \left(\Delta^2 - 1\right)\sin^2(ka)}.$$

Does this solution match the boundary conditions? If not, how can you construct a solution which *does* match the correct boundary condition?

(d) Analyse the spectrum for $\Delta = 0$ and show that the Hamiltonian describes massless fermions in that case.

(e) Calculate the total energy by integrating over all eigenvalues. You may use the small-Δ approximation:

$$\int_{-\pi/2}^{\pi/2} \sqrt{1 - (1-\Delta^2)\sin^2(x)}\, dx \approx 2 + \left(a_1 - b_1 \ln \Delta^2\right)\Delta^2,$$

where a_1 and b_1 are (unspecified) numerical constants. Find the values for Δ for which a long chain is maximally unstable towards dimerisation. Hint: now you *should* include the term $4\kappa u^2$!

12. In this problem we consider a spin-1/2 particle in terms of fermions operators a, $a^\dagger$, along the lines of section 9.5.2. To this end, we identify a spin-'up' state as a particle, and a spin-'down' state as the vacuum state:

$$|\uparrow\rangle = |1\rangle = a^\dagger |0\rangle$$
$$|\downarrow\rangle = |0\rangle = a|1\rangle .$$

In this representation, the spin-raising and lowering operator can be written as

$$\sigma^+ = a^\dagger; \qquad \sigma^- = a.$$

In addition:

$$\sigma^z = a^\dagger a - 1/2.$$

(a) From their definitions in terms of the creation and annihilation operators, show that these operators satisfy the commutation relations

$$[\sigma^+, \sigma^-] = 2\sigma^z.$$

Now we consider a chain of fermions that are coupled to each other. The spin-operators therefore get a site label j, as do the fermion operators. So we have operators like a_j and σ_j^+ etcetera. Suppose we again represent the spins with our fermion operators. These operators anti-commute when they act on different sites, whereas the spin-operators (the σ's) on different sites commute. We therefore adjust the relation between the σ's and the a's to take this into account.

(b) We define

$$\sigma_j^+ = a_j^\dagger \exp\left[\mathrm{i}\pi \sum_{j'<j} a_{j'}^\dagger a_{j'}\right],$$
$$\sigma_j^- = \exp\left[-\mathrm{i}\pi \sum_{j'<j} a_{j'}^\dagger a_{j'}\right] a_j,$$
$$\sigma_j^z = a_j^\dagger a_j - 1/2.$$

Verify, using these definitions, that

$$\sigma_j^+ \sigma_{j+1}^- = a_j^\dagger a_{j+1}.$$

(c) The Hamiltonian of the anisotropic spin-1/2 chain reads:

$$H = -\sum_j \left[J_z \sigma_j^z \sigma_{j+1}^z + J_x \left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y\right)\right].$$

Rewrite this in terms of the creation and annihilation operators $a^\dagger$ and a.

13. Consider a single species of bosons with annihilation and creation operators a and $a^\dagger$ respectively. The Hamiltonian operator for this quantum many-body system is

$$\hat{H} = \omega\left(a^\dagger a + 1/2\right) + \frac{\Delta}{2}\left(a^\dagger a^\dagger + aa\right).$$

We take $\hbar = 1$ throughout this problem. The following transformation is useful to gain insight into the properties of this quantum system:

$$b = \lambda a + \mu a^\dagger$$
$$b^\dagger = \lambda^* a^\dagger + \mu^* a.$$

where λ and μ are complex numbers.

(a) Show that this transformation preserves the usual commutation relations (but now for b and $b^\dagger$) provided that $|\lambda|^2 - |\mu|^2 = 1$. In the remainder of this problem, you may write $\lambda = \cosh u$, $\mu = \sinh u$ if you're at ease with hyperbolic functions.

(b) Assuming λ and μ to be real and using the result of (a), show that, for a particular value of u, the transformation brings the Hamiltonian into the form

$$H = \tilde{\omega}\left(b^\dagger b + \frac{1}{2}\right).$$

Find an equation for λ and μ for which this form is obtained and. You do not have to solve explicitly for $\tilde{\omega}$, but it is necessary that the correct form of the Hamiltonian is obtained.

(c) If the bosons characterised by a and $a^\dagger$ are considered as excitations of a harmonic oscillator with Hamiltonian

$$H = \frac{\hat{p}^2}{2m} + m\frac{\omega}{2}\hat{x}^2,$$

then a is given by

$$a = \frac{1}{\sqrt{2}}\left(\sqrt{m\omega}\,\hat{x} + \frac{\mathrm{i}\hat{p}}{\sqrt{m\omega}}\right).$$

Express the Hamiltonian in terms of $\hat{x}$ and $\hat{p}$ for the special case $\Delta = \omega$. How would you interpret this result physically?

14. *Fermions and Majorana fermions*

Consider an electrically neutral solid at $T = 0$. The states up to the Fermi energy are filled, those above the Fermi energy are empty. An electron is described by fermion creation and annihilation operators $a^\dagger$ and a, respectively, satisfying the usual anti-commutation relations. An electron creation operator associated with an unoccupied orbital will put an electron in that orbital. An annihilation operator associated with an occupied orbital will remove an electron from that orbital. In solid state physics, we often say that removing an electron from an occupied orbital is equivalent to creating a hole. In a proper relativistic description of electrons in vacuum, the same structure is recovered: an operator which *creates* an electron can equivalently be viewed as an operator *annihilating* a positron and vice-versa. The positron is called the 'anti-particle' of the electron.

Starting from fermion creation and annihilation operators $a^\dagger$ and a, we define two new operators c_1 and c_2 as follows:

$$c_1 = a + a^\dagger, \qquad c_2 = \frac{a - a^\dagger}{\mathrm{i}}.$$

Show that these operators satisfy the relations

$$c_\alpha = c_\alpha^\dagger \text{ and } c_\alpha c_\beta + c_\beta c_\alpha = 2\delta_{\alpha\beta}.$$

The first condition is often formulated as 'the particle described by c_α is its own anti-particle'. Such a particle is generally called a *Majorana fermion*.

10

ELECTRONS AND PHONONS

10.1 THEORY OF THE ELECTRON GAS

If we want to understand the behaviour of the electrons in a solid, we face a formidable problem. We have to deal with the electrons, the nuclei and the interactions between all of these. A non-relativistic Hamiltonian which does not include magnetic interactions is already quite complicated. For a finite number (N) of electrons moving in the Coulomb potential field of K nuclei with charges $Z_n e$, the Hamiltonian reads

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{n=1}^{K} \frac{P_n^2}{2M_n} + \frac{1}{2} \sum_{i,j=1, i \neq j}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N} \sum_{n=1}^{K} \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|} + \sum_{n,m=1, n\neq m}^{K} \frac{Z_n Z_m e^2}{|\mathbf{R}_n - \mathbf{R}_m|}.$$

Here, the $\mathbf{p}_i$ are momenta of the electrons, $\mathbf{P}_n$ those of the nuclei, $\mathbf{r}_i$ are the positions of the electrons and $\mathbf{R}_n$ those of the nuclei, which have masses M_n and charges Z_n – the electrons have mass m. Needless to say, this Hamiltonian is impossible to solve if N and K are not very small, even on a powerful computer.

In order to make progress and at least partly understand the physics of this system, we must make approximations. A sensible approximation is the *Born-Oppenheimer* (BO) approximation which is based on the observation that the electron mass is at least about 2000 times smaller than the nuclear mass. If the kinetic energy is more or less evenly distributed over the electrons and the nuclei, this implies that the nuclei move much more slowly than the electrons, and these can therefore adapt their wave function at any time to the nuclear configuration as if that were stationary. The BO approximation can be formalised, but we restrict ourselves to this descriptive definition. This then leaves the (still formidable) task of calculating the wave function for the electrons with the nuclei standing still. Varying then the positions of the nuclei, we see that the ground state energy of the electrons varies, and the expression of the total energy as a function of the positions of the nuclei is called the *potential energy surface* (PES).

Even for stationary nuclei, the problem remains enormously difficult. This is due to the interactions between the electrons. In fact, the Hamiltonian within the BO approximation can be written as

$$H_{\mathrm{BO}} = \sum_{i=1}^{N} h(i) + \frac{1}{2} \sum_{i \neq j} v(\mathbf{r}_i - \mathbf{r}_j)$$

where $h(i) = p_i^2/(2m) + v_{\mathrm{ext}}(\mathbf{r}_i)$, with v_{ext} the electric potential energy felt by each electron individually and caused by the nuclei, and v is the electrostatic repulsion between the electrons. If that interaction were not there, the one-electron problem for h could be solved, and the total energy would simply be a sum of the energies of the occupied one-electron states, and the wave function would be a Slater determinant composed of those one-electron states (spin-orbitals).

In order to make further progress, we neglect the discrete structure of the nuclei and make the rather drastic assumption that their charge is smeared out evenly over space. Considering a solid in the thermodynamic limit, this means that we have a constant positive nuclear charge density, and hence a constant contribution to the potential felt by the electrons. It is as if the nuclei are transformed into a uniformly charged jelly, hence the name *jellium model* for this approximation.

We assume that the total system is electrically neutral, so that in any large volume V, the charge $-Ne$ of electrons in that volume is compensated by the positive jellium charge in that volume, which leads to a jellium charge density of $n_b = Ne/V$. The Hamiltonian for the jellium model can now be written as

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \frac{1}{2} \sum_{i \neq j=1}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V_{e-b}(\mathbf{r}_i),$$

where the letter b stands for 'background'. The potential representing the interaction between the electrons and the background is given by

$$V_{e-b}(\mathbf{r}) = -e^2 \int \frac{n_b}{|\mathbf{r} - \mathbf{r}'|} d^3 r'.$$

Furthermore, the background carries its own energy which is given by

$$E_{b-b} = \frac{e^2}{2} \int \frac{n_b^2}{|\mathbf{r} - \mathbf{r}'|} d^3 r \, d^3 r'.$$

Both of these terms are constants tending to infinity, which makes them delicate to evaluate. Let us therefore screen the potential with a screening length $1/\lambda$:

$$\frac{e^2}{r} \to \frac{e^2 \exp(-\lambda r)}{r},$$

which we can send to $\lambda = 0$ in the end. The Fourier transform of the potential can be directly found:

$$v_{\mathbf{q}} = \frac{4\pi e^2}{V(q^2 + \lambda^2)}.$$

In the calculation we *first* take the limit of the volume to infinity and *then* that of the screening length. This means that we can consider the screening length always small with respect to the volume. We can then split the integral into one over $\mathbf{r}$ and a second one over $\Delta\mathbf{r} = \mathbf{r}' - \mathbf{r}$. Note that for this to be possible we need $\lambda^3 \ll V$ indeed. We now have, with $n_b = N/V$:

$$E_{b-b} = \frac{e^2 N^2}{2V} 4\pi \int_0^\infty \frac{e^{-\lambda \Delta r}}{\Delta r} \Delta r^2 d\Delta r = \frac{e^2 N^2}{2V} \frac{4\pi}{\lambda^2}. \tag{10.1}$$

Now let us calculate the background potential V_{e-b} felt by the electrons. This term is independent of $\mathbf{r}$ and will therefore lead to a constant contribution to the energy of $V_{e-b}N$. It is easy to evaluate this term:

$$V_{e-b} = -n_b e \int \frac{e^{-\lambda r}}{r} d^3 r = -\frac{4\pi e N}{V\lambda^2}.$$

This term gives a constant contribution to the total energy of

$$E_{e-b} = eNV_{e-b} = -\frac{4\pi e^2 N^2}{V\lambda^2} \tag{10.2}$$

which scales with N and V the same as E_{b-b}, but it is negative and *twice as large*.

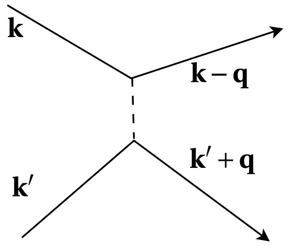


FIGURE 10.1: Electron-electron diagram.

Now we consider the electron-electron interaction. The Hamiltonian can again be written in its usual form $H = \sum_i h(i) + \frac{1}{2}\sum_{i\neq j} v(i,j)$. We have seen that in a many-body field theory, this Hamiltonian can be represented as an operator in Fock space. It is convenient to use a plane wave basis and work with the creation and annihilation operation operators for these basis functions. We neglect the spin for now to keep the analysis simple:

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2}\sum_{\mathbf{kqk'q'}} a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{q}} v_{\mathbf{kqk'q'}} a_{\mathbf{q'}} a_{\mathbf{k'}},$$

where $\epsilon_{\mathbf{k}} = \hbar^2 k^2/(2m)$. This form is precisely the one we formulated in the previous chapter [(9.4)] – we have used the plane wave labels **k**, **q** etcetera instead of j, k, l in that chapter.

Let us formulate the matrix element of v explicitly:

$$v_{\mathbf{kqk'q'}} = \frac{1}{V^2}\int e^{-\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_1} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} e^{-\mathrm{i}(\mathbf{q}-\mathbf{q}')\cdot\mathbf{r}_2} d^3r_1 d^3r_2.$$

Now we change to coordinates $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and $\mathbf{R} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$ and the integral transforms into

$$v_{\mathbf{kqk'q'}} = \frac{1}{V^2}\int e^{-\mathrm{i}(\mathbf{k}-\mathbf{k}'+\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}} d^3R \int e^{-\frac{1}{2}\mathrm{i}(\mathbf{k}-\mathbf{k}'-\mathbf{q}+\mathbf{q}')\mathbf{r}} \frac{e^2}{r} d^3r.$$

The integral over **R** gives a delta function

$$V\delta^{(3)}(\mathbf{k}-\mathbf{k}'+\mathbf{q}-\mathbf{q}')$$

and using this in the second integral leads to the form

$$v_{\mathbf{kqk'q'}} = \frac{1}{V}\delta(\mathbf{k}-\mathbf{k}'+\mathbf{q}-\mathbf{q}')\int e^{-\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} \frac{e^2}{r} d^3r.$$

The integral is the Fourier transform $V v_{\mathbf{k}-\mathbf{k}'}$. The delta-function expresses momentum conservation: the total 'incoming' momentum $\mathbf{k}' + \mathbf{q}'$ equals the 'outgoing' momentum $\mathbf{k} + \mathbf{q}$. Using **k** and **k**′ for the incoming momenta, we can write the outgoing ones in the form $\mathbf{k} - \mathbf{q}$ and $\mathbf{k}' + \mathbf{q}$ without loss of generality. Using these definitions, we can write interaction term of the Hamiltonian in the form

$$\frac{1}{2}\sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} a^\dagger_{\mathbf{k}-\mathbf{q}} a^\dagger_{\mathbf{k}'+\mathbf{q}} v_{\mathbf{q}} a_{\mathbf{k}'} a_{\mathbf{k}}.$$

This term can be represented pictorially, see figure 10.1. The final form of the Hamiltonian, now with the spin quantum numbers included, becomes:

$$H = \sum_\sigma \sum_{\mathbf{k}} \frac{\hbar^2 k^2}{2m} a^\dagger_{\mathbf{k},\sigma} a_{\mathbf{k},\sigma} + \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} a^\dagger_{\mathbf{k}-\mathbf{q},\sigma} a^\dagger_{\mathbf{k}'+\mathbf{q},\sigma'} v_{\mathbf{q}} a_{\mathbf{k}',\sigma'} a_{\mathbf{k},\sigma}.$$

At low temperatures, the electron gas will be in its ground state, so let's calculate the lowest possible energy. That is a difficult task – we can make progress by assuming the electron-electron interaction to be relatively weak so that we can use perturbation theory. We start by neglecting the interaction term altogether and minimize the kinetic energy. This can be done by putting the electrons by pairs in the lowest available momentum states. We fill up those states until we have exhausted all the electrons. The requirement that we fill the lowest momentum states means that the momenta are the set of points closest to the origin in reciprocal space – that is, they fill a *sphere* in reciprocal space; this is the *Fermi sphere.* The number of k-points inside the sphere is the volume of the sphere divided by the volume per k-points which, for a large $L \times L \times L$ volume, is $(2\pi/L)^3$. We therefore see that in the sphere with a radius which we call $k_{\rm F}$ we can store

$$N = 2\frac{L^3}{(2\pi)^3}\frac{4\pi}{3}k_{\rm F}^3$$

electrons. Note the factor of 2 which accounts for the two spin states. We see that the $k_{\rm F}$ can be calculated from the density $n = N/L^3$:

$$k_{\rm F} = (3\pi^2 n)^{1/3}.$$

The subscript 'F' stands for 'Fermi' and $k_{\rm F}$ is called the 'Fermi momentum'. It is straightforward to calculate the total ground state energy if we neglect the interactions:

$$E_G^{(0)} = 2\frac{V}{(2\pi)^2}\int_{k<k_{\rm F}} \frac{\hbar^2 k^2}{2m} d^3 k.$$

Evaluating the integral and dividing by the number of particles N, we obtain

$$\frac{E_G^{(0)}}{N} = \frac{3}{5}\frac{\hbar^2 k_{\rm F}^2}{2m} = \frac{3}{5}\epsilon_{\rm F},$$

where $\epsilon_{\rm F} = \hbar^2 k_{\rm F}^2/2m$ is the Fermi energy.

The next step is to take the electron interaction into account. We consider this term as a perturbation. Standard perturbation theory tells us that the first order correction to the ground state energy, due to a perturbation W to some Hamiltonian H_0 for which we know the ground state Φ, is given by

$$\Delta E = \langle \Phi | W | \Phi \rangle\,.$$

We use this result, taking for H_0 the kinetic energy:

$$H_0 = \sum_\sigma \sum_{\mathbf{k}} \frac{\hbar^2 k^2}{2m} a^\dagger_{\mathbf{k},\sigma} a_{\mathbf{k},\sigma},$$

for which the ground state $|\Phi\rangle$ is a Slater determinant built from all $\mathbf{k}$-vectors inside the Fermi sphere, and W is the electron-electron interaction

$$W = \frac{1}{2}\sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} a^\dagger_{\mathbf{k}-\mathbf{q}} a^\dagger_{\mathbf{k}'+\mathbf{q}} v_{\mathbf{q}} a_{\mathbf{k}'} a_{\mathbf{k}}.$$

So:

$$\Delta E = \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} \left\langle \Phi | a^\dagger_{\mathbf{k}-\mathbf{q},\sigma} a^\dagger_{\mathbf{k}'+\mathbf{q},\sigma'} v_{\mathbf{q}} a_{\mathbf{k}',\sigma'} a_{\mathbf{k},\sigma} | \Phi \right\rangle.$$

Now we consider which combinations of $\mathbf{k}$, $\mathbf{k}'$ and $\mathbf{q}$ yield nonzero matrix elements. The two annihilation operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}'}$ remove orbitals $\mathbf{k}$ and $\mathbf{k}'$ from the Fermi sphere. Any creation of orbitals *outside* the Fermi sphere gives zero, since $\langle\Phi|$ on the left does not contain these orbitals. Since all orbitals *within* the Fermi sphere, except for $\mathbf{k}$ and $\mathbf{k}'$, are occupied, we connot create orbitals other than these two there. So two possibilities remain:

- We have $\mathbf{q} = 0$, i.e. we annihilate $\mathbf{k}$ and then create it again, and similarly for $\mathbf{k}'$.

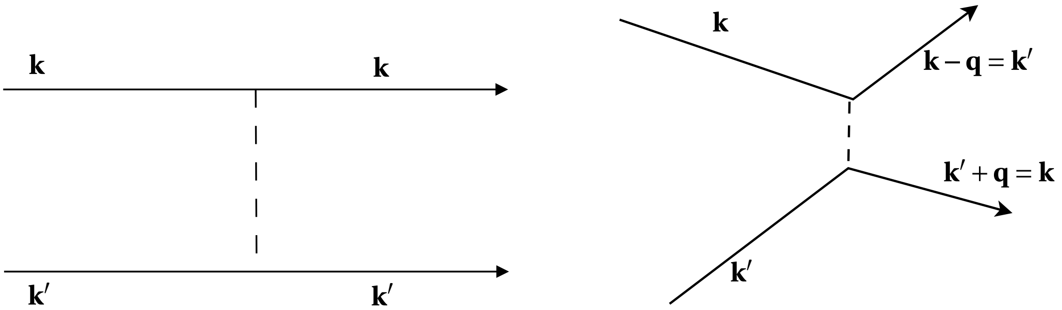


FIGURE 10.2: Electron-electron processes that conserve the momenta.

- The two incoming momenta are swapped by the interaction term, i.e.:

$$\mathbf{k}-\mathbf{q}=\mathbf{k}'.$$

These two processes are represented in figure 10.2. In the first process, $\mathbf{q}=0$, and we first analyze these $\mathbf{q}=0$ terms. These lead to

$$\frac{1}{2}\sum_{\sigma,\sigma'}\sum_{\mathbf{k},\mathbf{k}'} a^\dagger_{\mathbf{k},\sigma} a^\dagger_{\mathbf{k}',\sigma'} v_0 a_{\mathbf{k}',\sigma'} a_{\mathbf{k},\sigma}.$$

We would like to move the rightmost operator two places to the left, as we can then recognise two number operators in the expression. However, the anti-commutation relations lead to an extra 1 for the case where $\mathbf{k}=\mathbf{k}'$ and $\sigma=\sigma'$. We split these extra contributions off from the sum to obtain

$$\frac{1}{2}\left[\sum_{\sigma,\sigma'}\sum_{\mathbf{k},\mathbf{k}'} a^\dagger_{\mathbf{k},\sigma} a_{\mathbf{k},\sigma} a^\dagger_{\mathbf{k}',\sigma'} a_{\mathbf{k}',\sigma'} v_0 - \sum_{\sigma}\sum_{\mathbf{k}} a^\dagger_{\mathbf{k},\sigma} a_{\mathbf{k},\sigma} v_0\right] = \frac{1}{2}\left[\sum_{\sigma,\sigma'}\sum_{\mathbf{k},\mathbf{k}'} n_{\mathbf{k},\sigma} n_{\mathbf{k}',\sigma'} v_0 - \sum_{\sigma}\sum_{\mathbf{k}} n_{\mathbf{k},\sigma} v_0\right]. \tag{10.3}$$

Using $\sum_\sigma \sum_{\mathbf{k}} n_{\mathbf{k},\sigma} = N$ and substituting

$$v_0 = \frac{4\pi e^2}{V\lambda^2},$$

we obtain for the first term:

$$E_{e-e} = \frac{2\pi e^2 N^2}{V\lambda}$$

which is the ‘classical’ electron-electron interaction. Previously we have obtained self-interaction E_{b-b} of the positive background charge in Eq. (10.1) and the interaction energy of the electrons and the background in Eq. (10.2). We see that the all these terms cancel: $E_{b-b}+E_{e-b}+E_{e-e}=0$. The second term of (10.3) results in an energy

$$-\frac{N}{V}\frac{2\pi e^2}{\lambda^2}$$

which yields a zero contribution *per particle* for $V\to\infty$ (note the importance of taking *first* the limit for V to infinity and *then* $\lambda\to 0$). We conclude that the $\mathbf{q}=0$ term in the interaction potential cancels the background energies.

The terms in the interaction energy with $\mathbf{q}\neq 0$ should give us a more interesting contribution from the electron-electron interaction. It is

$$\Delta E = \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{\mathbf{k}\mathbf{k}'}\left\langle \Phi \middle| a^\dagger_{\mathbf{k}',\sigma} a^\dagger_{\mathbf{k},\sigma'} v_{\mathbf{k}-\mathbf{k}'} a_{\mathbf{k}',\sigma'} a_{\mathbf{k},\sigma} \middle| \Phi \right\rangle.$$

Now suppose $\sigma\neq\sigma'$. Then $a_{\mathbf{k},\sigma}$ removes an electron with wave vector $\mathbf{k}$ and spin σ from the ground state. The spin-orbital $\mathbf{k},\sigma'$ remains occupied. So the creation operator $a^\dagger_{\mathbf{k},\sigma'}$ will give zero. So, only terms with $\sigma = \sigma'$ give a non-zero result. If we now move in the operator $a_{\mathbf{k}',\sigma}$ one place to the left, the anti-commutation relation gives us the expression

$$\Delta E = -\frac{1}{2}\sum_{\sigma}\sum_{\mathbf{k}\mathbf{k}'} \langle \Phi | n_{\mathbf{k}',\sigma} n_{\mathbf{k},\sigma} v_{\mathbf{k}-\mathbf{k}'} | \Phi \rangle .$$

Using the fact that for $|\mathbf{k}| < k_\mathrm{F}$ all the states are occupied, we obtain

$$\Delta E = -\sum\nolimits^{(\mathrm{F})}_{\mathbf{k},\mathbf{k}'} v_{\mathbf{k}-\mathbf{k}'}$$

where $\sum^{(\mathrm{F})}$ denotes a sum over $\mathbf{k}$ and $\mathbf{k}'$ inside the Fermi sphere. The correction can now be calculated in the continuum limit. The expression for ΔE then reads

$$\Delta E = -\frac{V}{(2\pi)^6}\int \frac{4\pi e^2}{|\mathbf{k}-\mathbf{k}'|^2}\, d^3k\, d^3k'.$$

We write this in the form:

$$\Delta E = -\frac{V}{(2\pi)^3}\int \epsilon(\mathbf{k})\, d^3k,$$

where

$$\epsilon(\mathbf{k}) = \frac{1}{(2\pi)^3}\int_{|\mathbf{k}'|<k_\mathrm{F}} \frac{4\pi e^2}{|\mathbf{k}-\mathbf{k}'|^2} d^3k' = -\frac{e^2 k_\mathrm{F}}{\pi}\int_0^1 dx\, x^2 \int_{-1}^{1} \frac{du}{x^2+y^2-2xyu}.$$

Here we have put $x = k'/k_\mathrm{F}$ and $y = k/k_\mathrm{F}$ and u is the cosine of the angle between $\mathbf{k}$ and $\mathbf{k}'$. The integral can be done and we obtain:

$$\epsilon(\mathbf{k}) = -\frac{2e^2 k_\mathrm{F}}{\pi} f\left(\frac{k}{k_\mathrm{F}}\right),$$

with

$$f(x) = \frac{1}{2} + \frac{1-x^2}{4x}\ln\left|\frac{1+x}{1-x}\right|.$$

If we then sum (integrate) over the $\mathbf{k}$, we obtain the total energy correction, which turns out to be

$$\Delta E = -\frac{V}{4\pi^3} e^2 k_\mathrm{F}^4.$$

Let's discuss the results obtained. The correction term is negative. This may seem surprising as it is derived from the repulsive electron-electron energy. However, it is easy to understand if we realize that the $\mathbf{q} = 0$ term considered above takes the major part of the electron-electron interaction into account and the effect of the correction above is to reduce that large positive result. Why is it reduced? The calculation above showed that the mechanism responsible for the correction is momentum exchange. That is why this term is called *exchange interaction*. The negative sign is due to the anti-commutation relation between the creation and annihilation operators, and this in turn is due to the Fermi statistics obeyed by the electrons that are exchanged. Two electrons with the same spin cannot occupy the same position – the anti-symmetry of the wave function keeps them apart. The very fact that the electrons are kept apart by the anti-symmetry of the wave function causes a *reduction* of the electrostatic energy as this increases when the electrons are close together. The first-order perturbation analysis of the interacting electron gas is called the *Hartree-Fock theory*.

Several predictions of the Hartree-Fock theory are in conflict with experiments on metals: (i) the band width is predicted to be larger than that of free electrons, whereas in experiments it's usually found to be lower; (ii) the density of single-electron states is predicted to vanish logarithmically at the Fermi wave vector $k = k_\mathrm{F}$. In reality, a value close to the free-electron result is found. The reason for these failures of the Hartree-Fock is the fact that polarization has not been included: we have used *unperturbed* ground state orbitals to evaluate the exchange energy. In reality, the orbitals occupied by the electrons will deform as a result of the

Coulomb interaction. This has a major effect on the energies. Such effects can be taken into account within the *random phase approximation* (RPA), which yields much better values.

The same approach as the one followed here is often used in quantum chemistry to calculate the energies and to predict the excitation energies of molecules. It gives quite good results for those cases, indicating that polarization effects are more important in semi-conductors than in molecules.

Another approach is based on writing the exchange energy as a function of the local density which we shall now briefly sketch without going into details. Translating the exchange energy into a function of the density, we can derive a potential $V_{\rm xc}$ from it, which has the form

$$V_{\rm xc} = -2.95(a_0^3 n)^{1/3}{\rm Ry}$$

(1 Ry = 13.6 eV). This form has been used very often in calculations for electrons in solids. In these calculations, the Hartree-Fock form for the exchange energy is made a local function by replacing the average density n by the local density $n(\mathbf{r})$. Moreover, some extra terms are incorporated into the one-electron Hamiltonian, in particular the electrostatic energy of the electrons among themselves and the interaction with the nuclei, so that we arrive at the one-electron Hamiltonian:

$$H = \frac{p^2}{2m} + \sum_n \frac{-Z_n e^2}{|\mathbf{r}-\mathbf{R}_n|} + \int \frac{e^2 n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3 r' + V_{\rm xc}(\mathbf{r}).$$

The third term on the right hand side represents the electrostatic energy resulting from the electron cloud with density $n(\mathbf{r})$. In 1964, Hohenberg and Kohn proved a theorem which says that a similar (but not identical) Hamiltonian, with the same interaction with the nuclei and all other terms depending on the electron density only, gives the *exact* ground state energy. Although the exact form of the other terms is as yet unknown, many approximations to it exist, and allow for accurate ground state energy calculations of atoms, molecules and solids. The theory based on the Hohenberg-Kohn theorem is called *density functional theory*.

To systematically improve on the Hartree-Fock analysis, we may consider the second-order perturbation term. We only sketch briefly how this works. If we start from the ground state, $|\Phi\rangle$, we obtain for the energy correction the form

$$\Delta E^{(2)} = \sum_I \frac{\langle \Phi | v | \Phi_I \rangle \langle \Phi_I | v | \Phi \rangle}{E_I - E_G}$$

where $|\Phi_I\rangle$ is any intermediate state in which one or two electrons may be excited with respect to the ground state – these intermediate states are summed over. Figure 10.3 shows a picture of typical terms occurring in this expansion. Similar to the first order case, the initial and final momenta can either remain the same or they will be exchanged, whereas the intermediate states can have electrons excited to outside the Fermi sphere.

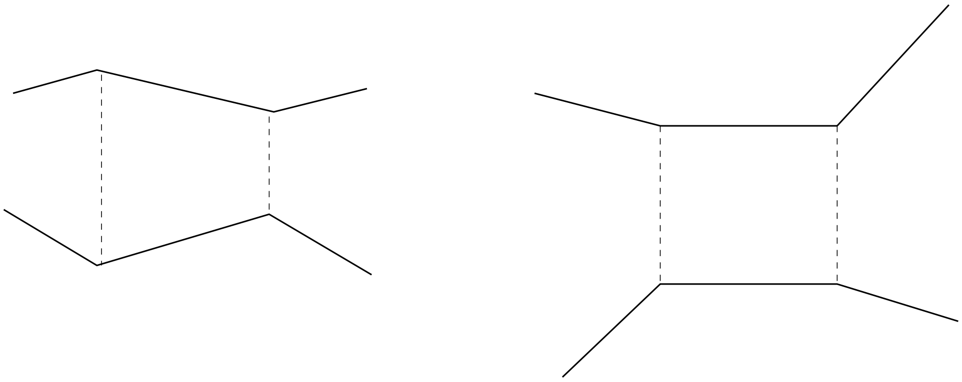


FIGURE 10.3: Second order electron-electron diagrams with momentum conservation.

Summary of the electron gas
In this section we have considered a gas of interacting electrons in a medium with a positive background charge. We have used perturbation theory to analyse this system, using the electron-electron interaction as a small term in the Hamiltonian. If the electron-electron interaction is neglected, we find a ground state consisting of a sphere of occupied states in reciprocal space, the *Fermi sphere.* In the ground states, all **k**-orbitals within the Fermi sphere are occupied by two electrons, one with spin-up and one with spin-down. The **k**-points outside the Fermi sphere are unoccupied in the ground state.

The first order correction to the ground state gives two contributions. The first is one which does not involve momentum exchange and represents the static repulsion between the electrons. This term cancels the interaction between the electrons and the positive background plus the self-energy of the positive background.

The second first-order contribution *lowers* represents the effect of *exchanging* the particles with two different momenta (but having the same spin) in the Fermi sphere. This term has the effect of *lowering* the energy of the ground state by an amount

$$\Delta E = -\frac{V}{4\pi^3} e^2 k_{\mathrm{F}}^4.$$

The lowering can be explained from the fact that this term is a direct consequence of the fermion character of the wave function, which effectively keeps the electrons apart so that their interaction energy is reduced. First-order perturbation theory is called *Hartree-Fock theory.*

A systematic way of improving on the Hartree-Fock theory is to include higher order terms in the analysis, which quickly becomes a lot more tricky due to the large amount of contributions needed. Other approximations which give better results than Hartree-Fock are density functional theory, as well as the random phase approximation.

10.2 ELECTRON-PHONON COUPLING

In the beginning of this chapter we mentioned the Born-Oppenheimer (BO) approximation in which the nuclei are standing still, and the electronic state can be calculated for any configuration of nuclei. This then gives rise to a potential felt by the nuclei, the so-called *potential energy surface* (PES). Given a PES, a classical calculation predicts the motion of the nuclei. This approach is called the *adiabatic* approximation as the nuclear positions are considered as external parameters which change very slowly in time. In the adiabatic approximation, the electrons will therefore always remain in the ground state for the actual configuration of nuclei. Now we want to improve on this approximation by introducing energy exchange between electrons and phonons more explicitly into the Hamiltonian. We have already studied the example of a simple phonon system in section 8.2: a linear chain of which we have considered the longitudinal modes. This analysis can be generalised straightforwardly to 3D lattices and elastic waves in the longitudinal and the two transverse directions. All in all, this leads, for a monatomic Bravais lattice, to the following Hamiltonian determining the motion of the nuclei:

$$H = \sum_\alpha \sum_{\mathbf{q}} \left(\frac{p_{\mathbf{q}\alpha} p_{-\mathbf{q}\alpha}}{2m} + \frac{m\omega^2_{\mathbf{q}\alpha} y_{\mathbf{q},\alpha} y_{-\mathbf{q}\alpha}}{2} \right).$$

Here $\mathbf{q}$ is a vector inside the Brillouin zone, and α denotes the longitudinal and transverse directions. The coordinates $y_{\mathbf{q}\alpha}$ are the displacements of the nuclei with respect to their equilibrium positions $\mathbf{R}_n$. The nuclear mass is m.

Now we want to emphasise that in this section, we use the notation:

- $\mathbf{q}, \mathbf{q}'$: wave vectors used in the Fourier transforms of nuclear displacements (phonons). For a monatomic Bravais lattice, these are inside the Brillouin zone.
- $\mathbf{k}, \mathbf{k}'$: wave vectors used for Fourier transforms of the orbital electron wave functions. They are inside the reciprocal lattice and are usually decomposed into a wave vector $\mathbf{k}$ inside the Brillouin zone plus a vector $\mathbf{K}$ of the reciprocal lattice.
- The operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$ create and annihilate electrons in a state $|\mathbf{k}\rangle$.
- Similarly, the operators $d_{\mathbf{q}\alpha}$ and $d^\dagger_{\mathbf{q}\alpha}$ create and annihilate phonons.

The operator $d_{\mathbf{q}\alpha}$ is defined as

$$d_{\mathbf{q},\alpha} = \sqrt{\frac{m\omega_{\mathbf{q},\alpha}}{2\hbar}}\, y_{\mathbf{q},\alpha} + \mathrm{i}\frac{1}{\sqrt{2m\hbar\omega_{\mathbf{q}\alpha}}} p_{\mathbf{q}\alpha}.$$

and

$$d^\dagger_{\mathbf{q},\alpha} = \sqrt{\frac{m\omega_{\mathbf{q},\alpha}}{2\hbar}}\, y_{-\mathbf{q},\alpha} - \mathrm{i}\frac{1}{\sqrt{2m\hbar\omega_{\mathbf{q}\alpha}}} p_{-\mathbf{q}\alpha}.$$

With these definitions, we obtain

$$H = \sum_{\mathbf{q}\alpha} \hbar\omega \left(d^\dagger_{\mathbf{q}\alpha} d_{\mathbf{q}\alpha} + 1/2 \right).$$

In the previous section we have already considered the Hamiltonian for the electrons. Here we shall consider the electrons in the independent particle approximation, in which the electrons move only in an external potential. This external potential may, in some approximation schemes, be generated (at least in part) by all the electrons in the system. The electron Hamiltonian is therefore given as

$$H_{\mathrm{el}} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}}.$$

With a homogeneous background, the sum runs over the all $\mathbf{k}$-vectors in reciprocal space. The fact that we use a homogeneous background may seem strange, since the electron-nucleus interaction is locally very strong. However, we may focus on the valence electrons in a metal, and they see the nuclei as screened by the core electrons, resulting in a weak and smooth potential. Which of the phonon states are occupied depends on the temperature (in equilibrium). In a periodic solid, the sum runs over the vectors $\mathbf{k}$ of the Brillouin zone and perhaps band labels, or other quantum numbers.

The electrons feel the attractive potential of the nuclei, and this potential is determined by the positions of the latter. We can write the interaction as

$$H_{\rm I} = \sum_{\mathbf{k},\mathbf{k}',n} \left\langle \mathbf{k}| v_{\text{e-n}}(\mathbf{r}-\mathbf{R}_n-\mathbf{y}_n)|\mathbf{k}'\right\rangle a^\dagger_{\mathbf{k}} a_{\mathbf{k}'},$$

where the index n runs over the nuclei, $\mathbf{R}_n$ is the equilibrium position of nucleus n, $v_{\text{e-n}}$ is the electron-nucleus interaction (which will be Coulomb-like), and $\mathbf{y}_n$ is the displacement of the nucleus with respect to equilibrium, which is supposed to be small. We expand the potential in terms of the small displacement $\mathbf{y}_n$:

$$v_{e-n}(\mathbf{r}-\mathbf{R}_n-\mathbf{y}_n) = v_{e-n}(\mathbf{r}-\mathbf{R}_n) - \mathbf{y}_n\cdot\nabla v_{e-n}(\mathbf{r}-\mathbf{R}_n).$$

We see that the effect of the nuclei on the electrons is two-fold: the first effect is a shift of the potential due to all the nuclei when they are at equilibrium, and the second is due to the displacements $\mathbf{y}_n$ of the nuclei with respect to that equilibrium. We consider both terms as external potentials for the electrons and use first order perturbation theory. This tells us that the first term results in a shift of the total energy given by

$$H_{\rm B} = \sum_{\mathbf{k},\mathbf{k}',n} \left\langle \mathbf{k}| v_{e-n}(n)|\mathbf{k}'\right\rangle a^\dagger_{\mathbf{k}} a_{\mathbf{k}'},$$

where $v_{e-n}(n)$ denotes the potential energy due to nucleus n. We can write out the matrix element:

$$H_{\rm B} = \sum_{\mathbf{k},\mathbf{k}',n} \frac{1}{V}\int e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} v_{e-n}(\mathbf{r}-\mathbf{R}_n)\, d^3r\; a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

We now make the substitution $\mathbf{r}\to\mathbf{r}+\mathbf{R}_n$ in the integral to obtain

$$H_{\rm B} = \sum_{\mathbf{k},\mathbf{k}',n} \frac{1}{V}\exp\left[\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_n\right]\int e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} v_{e-n}(\mathbf{r})\, d^3r\; a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

Carrying out the sum over $\mathbf{R}_n$ results in a 'modified' delta-function: since the sum is only over the discrete vectors $\mathbf{R}_n$, the argument of the 'delta'-function only forces the component inside the Brillouin zone to be equal – hence the argument may still be a reciprocal lattice vector rather than 0. The formal result is $\sum_n e^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}_n} = \sum_m \delta(\mathbf{k}+\mathbf{K}_m)$ where $\mathbf{K}_m$ are the reciprocal lattice vectors. We therefore have

$$H_{\rm B} = N\sum_{\mathbf{k},m} v_{e-n}(\mathbf{K}_m)\, a^\dagger_{\mathbf{k}+\mathbf{K}_m} a_{\mathbf{k}}.$$

This shift can be absorbed in the single-electron energies $\epsilon_{\mathbf{k}}$ – it is called the *Bloch term.*

The second term depends on the nuclear displacements. It reads

$$H_{\text{e-ph}} = -\sum_n \mathbf{y}_n\cdot\sum_{\mathbf{k},\mathbf{k}'}\left\langle\mathbf{k}|\nabla_n v_{e-n}(n)|\mathbf{k}'\right\rangle a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} = -\sum_n \mathbf{y}_n\cdot\sum_{\mathbf{k},\mathbf{k}'}\frac{1}{V}\int e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}}\nabla v_{e-n}(\mathbf{r}-\mathbf{R}_n)d^3r\, a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

We now focus on the integral in the right hand side. Using partial integration in the last integral, we obtain

$$\frac{1}{V}\int e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}}\nabla v_{e-n}(\mathbf{r}-\mathbf{R}_n)d^3r = -\frac{\mathrm{i}}{V}(\mathbf{k}'-\mathbf{k})\int e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} v_{e-n}(\mathbf{r}-\mathbf{R}_n)d^3r =$$
$$-\frac{\mathrm{i}}{V}(\mathbf{k}'-\mathbf{k})\exp\left[\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_n\right]\int e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} v_{e-n}(\mathbf{r})d^3r,$$

where we have made the substitution $\mathbf{r}\to\mathbf{r}+\mathbf{R}_n$ again in order to obtain the last form. The integral is recognized as the Fourier transform of v_{e-n}, and we can write the electron-phonon Hamiltonian as

$$H_{\text{e-ph}} = \mathrm{i}\sum_{\mathbf{k},\mathbf{k}',n}\mathbf{y}_n\cdot(\mathbf{k}'-\mathbf{k})\exp\left[\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}_n\right] v_{e-n}(\mathbf{k}-\mathbf{k}')a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

Now we substitute for $\mathbf{y}_n$ its Fourier transform

$$\mathbf{y}_n = \frac{1}{\sqrt{N}} \sum_{\mathbf{q}} \mathbf{y}_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{R}_n}$$

to obtain

$$H_{\text{e-ph}} = \frac{\mathrm{i}}{\sqrt{N}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q},n} e^{\mathrm{i}(\mathbf{k}'-\mathbf{k}+\mathbf{q})\cdot\mathbf{R}_n} (\mathbf{k}'-\mathbf{k}) \cdot \mathbf{y}_{\mathbf{q}} \, v_{\text{e-n}}(\mathbf{k}-\mathbf{k}') a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

Carrying out the sum over n forces $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ and we have

$$H_{\text{e-ph}} = \mathrm{i}\sqrt{N} \sum_{\mathbf{k},\mathbf{k}'} (\mathbf{k}'-\mathbf{k}) \cdot \mathbf{y}_{\mathbf{k}-\mathbf{k}'} \, v_{\text{e-n}}(\mathbf{k}-\mathbf{k}') a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

The result of the dot-product depends on the polarization of the phonon (longitudinal or transverse) – we shall not go into details here. It is important to again realize that the delta-function we have obtained in the sum over n needs to be interpreted with some care: if $\mathbf{k} - \mathbf{k}'$ lies inside the Brillouin zone, it is correct as given here; if $\mathbf{k} - \mathbf{k}'$ lies outside the first Brillouin zone, we have $\mathbf{q} = \mathbf{k}' + \mathbf{k} - \mathbf{K}$ where $\mathbf{K}$ is a reciprocal lattice vector. The processes where this is the case are called 'Umklapp' processes – they have a noticeable influence on the temperature dependence of the resistance. Note that the deviation $\mathbf{y}_{\mathbf{q}}$ has a polarization. In a homogeneous elastic medium, we can take two transverse and one longitudinal polarization. For a crystalline solid, this is in general not always be possible, but we shall not treat that case here. The term $\mathbf{q} \cdot \mathbf{y}_{\mathbf{q}}$ in the electron-phonon Hamiltonian shows that only longitudinal modes can interact with the electrons. The polarization can be along the direction $\hat{\mathbf{q}}$ or $-\hat{\mathbf{q}}$, and this obviously has its effect on the electron-phonon interactions, as a phonon can either hit an electron 'in the back' or 'head-on'. In the sum over $\mathbf{k}$ and $\mathbf{k}'$, both cases will occur, which is necessary for the Hamiltonian to Hermitian. From now on, we only consider the longitudinal component of the displacement and therefore we consider $y_{\mathbf{q}}$ as the amplitude of the displacement along $\mathbf{q}$: $y_{\mathbf{q}}$ is therefore no longer a vector.

The major step is now to express $\mathbf{y}$ in terms of $d^\dagger_{\mathbf{q}}$ and $d_{\mathbf{q}}$. Now we use Eqs. (8.13) and (8.14) to find

$$y_{\mathbf{q}} = \sqrt{\frac{\hbar}{2m\omega_{\mathbf{q}}}} (d^\dagger_{-\mathbf{q}} + d_{\mathbf{q}}),$$

which, using the definition

$$M_{\alpha,\mathbf{q}} = \mathrm{i}\sqrt{\frac{N\hbar}{2m\omega_{\mathbf{q}}}} \left|\mathbf{q}\right| v_{\text{e-n}}(\mathbf{q})$$

leads to

$$H_{\text{e-ph}} = \sum_{\mathbf{k},\mathbf{k}',\alpha} M_\alpha(\mathbf{k}-\mathbf{k}')(d^\dagger_{-\mathbf{q},\alpha} + d_{\mathbf{q},\alpha}) a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

where $\mathbf{q} = \mathbf{k} - \mathbf{k}'$, reduced to the first Brillouin zone. Now the total Hamiltonian can be given:

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}} d^\dagger_{\mathbf{k}} d_{\mathbf{q}} + \sum_{\mathbf{k},\mathbf{k}'} M(\mathbf{k}-\mathbf{k}')(d^\dagger_{-\mathbf{q}} + d_{\mathbf{q}}) a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

The last term has a very nice interpretation which can be visualised using diagrams: the term containing $d^\dagger_{-\mathbf{q}}$ describes a process in which the electron momentum $\mathbf{k}'$ is changed into $\mathbf{k}$ under *emission* of a phonon. The term containing $d_{\mathbf{q}}$ describes the *absorption* of a phonon. The total momentum and energy is to be conserved in these processes. They are represented in figure 10.4.

We can now evaluate the effect on the total energy of electron-phonon interactions. We again use perturbation theory for this purpose, just as in the case of the electron gas. The perturbation to second order is

$$E = E_0 + \left\langle \Phi | H_{\text{e-ph}} | \Phi \right\rangle + \left\langle \Phi | H_{\text{e-ph}} (E_0 - H_0)^{-1} H_{\text{e-ph}} | \Phi \right\rangle.$$

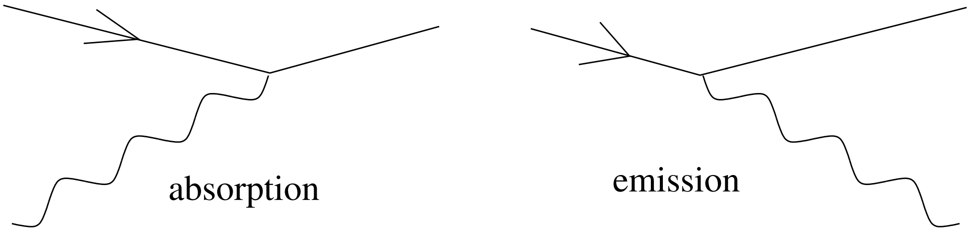


FIGURE 10.4: Electron-phonon diagrams.

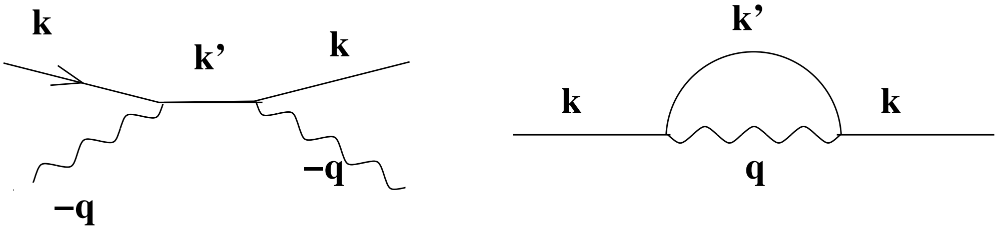


FIGURE 10.5: Second order electron-phonon diagrams.

Here $|\Phi\rangle$ is the ground state of a system of electrons and phonons. The first term in the perturbation expansion vanishes, as it changes the momentum of a single electron, and $|\Phi\rangle$ contains a Slater determinant composed of all **k**-waves within the Fermi sphere – changing one of the **k**-vectors gives a new state which is perpendicular to $|\Phi\rangle$. The second terms contains intermediate states that may be perpendicular to the ground state and we have

$$E - E_0 = \left\langle \Phi \left| \sum_{\mathbf{k},\mathbf{k}'} \left| M(\mathbf{k}-\mathbf{k}') \right|^2 \left[d^\dagger_{-\mathbf{q}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} (E_0 - H_0)^{-1} d_{-\mathbf{q}} a^\dagger_{\mathbf{k}'} a_{\mathbf{k}} + d_{\mathbf{q}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} (E_0 - H_0)^{-1} d^\dagger_{\mathbf{q}} a^\dagger_{\mathbf{k}'} a_{\mathbf{k}} \right] \right| \Phi \right\rangle .$$

The two processes in the right hand side can be visualised as in figure 10.5. Let us analyze the first term in this expression. Acting with $d_{-\mathbf{q}} a^\dagger_{\mathbf{k}'} a_{\mathbf{k}}$ on the ground state gives a state with energy $E_0 - \hbar\omega_{-\mathbf{q}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}}$. Therefore the term $(E_0 - H_0)^{-1}$ gives $1/(\hbar\omega_{\mathbf{q}} + \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'})$. If we furthermore move the creation and annihilation operators through this expression, we find, with a similar treatment of the second term:

$$E - E_0 = \sum_{\mathbf{k},\mathbf{k}'} \left| M(\mathbf{k}-\mathbf{k}') \right|^2 \langle n_{\mathbf{k}}(1 - n_{\mathbf{k}'}) \rangle \left(\frac{\langle n_{-\mathbf{q}} \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} + \hbar\omega_{-\mathbf{q}}} + \frac{\langle n_{\mathbf{q}} + 1 \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} - \hbar\omega_{\mathbf{q}}} \right).$$

Note that the $n_{\mathbf{k}}$ and $n_{\mathbf{k}'}$ are electron occupations, whereas $n_{\pm\mathbf{q}}$ denote phonon occupations. Also note that the +1 in the numerator of the second term represents an emitted phonon during the interaction.

Now we focus on the term in the interaction which has the form of a two-electron interaction:

$$E'_{\text{e-ph}} = -\sum_{\mathbf{k},\mathbf{k}'} \left| M(\mathbf{k}-\mathbf{k}') \right|^2 \langle n_{\mathbf{k}} n_{\mathbf{k}'} \rangle \left(\frac{\langle n_{\mathbf{q}} \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} + \hbar\omega_{\mathbf{q}}} + \frac{\langle n_{\mathbf{q}} + 1 \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} - \hbar\omega_{\mathbf{q}}} \right).$$

Note that we have replaced $-\mathbf{q} \to \mathbf{q}$: in equilibrium, there is no preference for the direction of the phonon propagation. The term proportional to $\langle n_{\mathbf{q}} \rangle$ can be cast into the form

$$E^{(1)}_{\text{e-ph}} = \sum_{\mathbf{k},\mathbf{k}'} \left| M(\mathbf{k}-\mathbf{k}') \right|^2 \langle n_{\mathbf{k}} n_{\mathbf{k}'} \rangle \langle n_{\mathbf{q}} \rangle \frac{2\hbar\omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) - \left(\hbar\omega_{\mathbf{q}}\right)^2}.$$

The term which is not proportional to $\langle n_{\mathbf{q}} \rangle$ has the form

$$E^{(2)}_{\text{e-ph}} = -\sum_{\mathbf{k},\mathbf{k}'} \left| M(\mathbf{k}-\mathbf{k}') \right|^2 \frac{\langle n_{\mathbf{k}} n'_{\mathbf{k}} \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} - \hbar\omega_{\mathbf{q}}},$$

which can be cast into a similar form as the first:

$$E^{(2)}_{\text{e-ph}} = -\frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'} \left|M(\mathbf{k}-\mathbf{k}')\right|^2 \frac{2\hbar\omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}'})^2-(\hbar\omega_{\mathbf{q}})^2} \langle n_{\mathbf{k}} n_{\mathbf{k}'}\rangle,$$

where we have used the fact that

$$\sum_{\mathbf{k},\mathbf{k}'} (\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}'})\, A(\mathbf{k}-\mathbf{k}') = 0$$

when A is symmetric in $\mathbf{k}-\mathbf{k}'$.

Now we study the possible signs of these terms. Although the electron energies are usually much higher than the phonon energies, *differences* between electron energies may become of the same order as the phonon energies. If that happens, we may get verry large values from the denominator in the expressions for the electron-interaction, with a positive or a negative sign. This contribution may become arbitrarily large as this denominator may even vanish! If the electrons are free to arrange themselves to minize their energy, it is like ly that they will try to profit from this negative (i.e. attractive) interaction. This is the mechanism behind the formation of *Cooper pairs*, which play a crucial role in superconductivity.

Summary for electron-phonon interactions
In this section, we have seen that the electron-phonon interaction can be represented by the following term in the Hamiltonian:

$$H_{\text{e-ph}} = \sum_{\mathbf{k},\mathbf{k}',\alpha} M_\alpha(\mathbf{k}-\mathbf{k}')(d^\dagger_{-\mathbf{q},\alpha} + d_{\mathbf{q},\alpha})\, a^\dagger_{\mathbf{k}} a_{\mathbf{k}'}.$$

where $\mathbf{q} = \mathbf{k}-\mathbf{k}'$, reduced to the first Brillouin zone. The amplitude $M_{\alpha,\mathbf{q}}$ appearing in this expression is given by:

$$M_{\alpha,\mathbf{q}} = \mathrm{i}\sqrt{\frac{N\hbar}{2m\omega_{\mathbf{q}}}}\,|\mathbf{q}|\, v_{\text{e-n}}(\mathbf{q}).$$

The fact that the interaction Hamiltonian contains *single* operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$ means that it vanishes when calculated in first-order perturbation theory: the interaction either fills a state outside the Fermi sphere with an electron, or it empties a state within the Fermi sphere. Both processes yield a new many-body state which is perpendicular to the ground state. The second order contribution however will in general be non-zero. This contribution can assume positive or negative values, indicating that the electron-phonon can increase or lower the energy of the electron system.

10.3 PROBLEMS

1. (a) Give the kinetic and the potential energy of the homogeneous electron gas within the Hartree-Fock approximation, both in terms of the Fermi wave vector k_{F}.

   (b) Calculate the *density* at which these two energies have the same (absolute) value in units of electrons/Å^3.

   (c) Which of these two energies dominates for higher densities?

2. Consider a system consisting of a large number N of spinless interacting fermions in a large one-dimensional box of length L. There are periodic boundary conditions. The particles interact via a delta-function potential, so the Hamiltonian is

$$H = \sum_k Ak^2 a^\dagger_k a_k + \frac{V}{L}\sum_{k,k',q} a^\dagger_{k-q} a^\dagger_{k'+q} a_{k'} a_k$$

with A and V constants. The sums run over all permitted values of k, k' and q, including $q = 0$.

(a) Show that the Hamiltonian given above is correct for particles with delta-function interaction.

(b) Calculate the ground state energy of the noninteracting system.

(c) Calculate the ground state energy of the interacting system within the Hartree-Fock approximation.

(d) Explain why the solution found in (c) must be the exact solution to the problem.

3. We have seen that the total energy of a system consisting of electrons occupying energy levels $\epsilon_{\mathbf{k}}$ and coupled to phonons with frequency spectrum $\omega_{\mathbf{k}}$, is given by

$$E - E_0 = \sum_{\mathbf{k},\mathbf{k}'} \left|M(\mathbf{k}-\mathbf{k}')\right|^2 \langle n_{\mathbf{k}}(1-n_{\mathbf{k}'})\rangle \left(\frac{\langle n_{-\mathbf{q}}\rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} + \hbar\omega_{-\mathbf{q}}} + \frac{\langle n_{\mathbf{q}}+1\rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} - \hbar\omega_{\mathbf{q}}} \right).$$

In the sum, $\mathbf{k}' = \mathbf{k} - \mathbf{q}$ and E_0 is the ground state energy for a system consisting of electrons and phonons:

$$H_0 = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} n_{\mathbf{k}} + \sum_{\mathbf{q}} \hbar\omega_{\mathbf{q}} n_{\mathbf{q}}$$

where the $n_{\mathbf{k}}$ is the number operator for electrons, and $n_{\mathbf{q}}$ that of the phonons.

(a) Explain that if the electrons are in the ground state of H_0, we only find contributions when $\mathbf{k}$ is inside, and $\mathbf{k}'$ outside the Fermi sphere.

(b) Use $\omega_{\mathbf{q}} = \omega_{-\mathbf{q}}$, $\langle n_{\mathbf{q}}\rangle = \langle n_{-\mathbf{q}}\rangle$ and to rearrange this equation to find

$$E - E_0 = \sum_{\mathbf{k},\mathbf{k}'} \left|M(\mathbf{k}-\mathbf{k}')\right|^2 \langle n_{\mathbf{k}}\rangle \left[2(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) \frac{\langle n_{\mathbf{q}}\rangle}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'})^2 - (\hbar\omega_{-\mathbf{q}})^2} + \frac{\langle 1-n_{\mathbf{k}'}\rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'} - \hbar\omega_{\mathbf{q}}} \right]$$

where the term proportional to $\langle n_{\mathbf{k}} n_{\mathbf{k}'}\rangle$ has vanished due to antisymmetry in the sum over $\mathbf{k}$ and $\mathbf{k}'$.

(c) The last expression can be used to find the chemical potential defined by the energy difference associated with adding or subtracting a phonon with wave vector $\mathbf{q}$ from the phonon bath. Calling this chemical potential $\hbar\omega_{\mathbf{q}}^{(p)}$, show that

$$\hbar\omega_{\mathbf{q}}^{(p)} = \hbar\omega_{\mathbf{q}} + \sum_{\mathbf{k}} \left|M(\mathbf{k}-\mathbf{k}')\right|^2 \frac{2\langle n_{\mathbf{k}}\rangle(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'})}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'})^2 - (\hbar\omega_{-\mathbf{q}})^2}$$

where, still, $\mathbf{k}' = \mathbf{k} - \mathbf{q}$. We see that the electron-phonon energy renormalizes the phonon frequencies.

(d) Note that for $\epsilon_{\mathbf{k}} = \epsilon_{\mathbf{k}'}$ we may have a singularity in the phonon spectrum. Usually this divergence disappears after integrating over $\mathbf{k}$ and $\mathbf{k}'$. This singularity is however at its 'worst' when $\mathbf{k}$ and $\mathbf{k}'$ are diametrically opposite at the surface of the Fermi sphere.

We assume the sound velocity to be so small that the term $\hbar\omega$ in the denominator can be neglected with respect to the $\epsilon_{\mathbf{k}}$. Take $\mathbf{q}$ along the z-axis and show that the integral over $\mathbf{k}$ on the Fermi sphere in the expression for the renormalized phonon energy is proportional to

$$\int \frac{dy}{q^2 - 2k_{\mathrm{F}} q y}.$$

Show that this integral gives a divergence when $q = 2k_{\mathrm{F}}$.

4. (a) Consider a Hamiltonian describing a one-dimensional superconductor:

$$H = -t\sum_{j=1}^{N-1}\left(a_j^\dagger a_{j+1} + a_{j+1}^\dagger a_j\right) - \mu\sum_{j=1}^{N} a_j^\dagger a_j + \sum_{j=1}^{N-1}\left(\Delta a_j a_{j+1} + \Delta^* a_{j+1}^\dagger a_j^\dagger\right).$$

Here, coefficients t and μ are real-valued and Δ is complex. Check that this Hamiltonian is Hermitian.

(b) We now set $\Delta = |\Delta|\, e^{\mathrm{i}\theta}$ (this is the superconducting gap). We define Majorana fermions c_j as follows:

$$c_{Aj} = e^{\mathrm{i}\theta/2} a_j + e^{-\mathrm{i}\theta/2} a_j^\dagger,$$
$$c_{Bj} = -\mathrm{i}e^{\mathrm{i}\theta/2} a_j + \mathrm{i}e^{-\mathrm{i}\theta/2} a_j^\dagger.$$

Show that the operators c_{Aj} and c_{Bj} describe Majorana fermions. From now on, we take $\theta = 0$.

(c) Show that the special case $\mu = 0$ and $t = \Delta$, the Hamiltonian, expressed in terms of Majorana operators, becomes

$$H = \mathrm{i}t\sum_{j=1}^{N-1} c_{Bj} c_{Aj+1}.$$

Note that this Hamiltonian does not depend on c_{A1} or c_{BN}.

(d) Now we define new fermionic operators:

$$b_j = \frac{1}{2}\left(c_{Bj} + \mathrm{i}c_{Aj+1}\right),\quad j = 1,\dots,N-1,\quad \text{and}\quad b_N = \frac{1}{2}\left(c_{BN} + \mathrm{i}c_{A1}\right).$$

Check that these operators satisfy the conventional fermion anti-commutation relations. Express the operators b_j in terms of the original fermion operators a_j.

(e) Show that H can be formulated in terms of the b's as follows:

$$H = 2t\sum_{i=1}^{N-1}\left(b_i^\dagger b_i - \frac{1}{2}\right).$$

(f) Discuss the eigenstates and the spectrum of H, paying particular attention to the degeneracy of the ground state. Argue whether this degeneracy is very sensitive to changes of the parameters away from $\mu = 0$, $\Delta = t$.

5. In chapter 9, we have seen that the Hamiltonian of a many-body system can be written as

$$H = \sum_{jk} a_j^\dagger a_k h_{jk} + \frac{1}{2}\sum_{jklm} a_j^\dagger a_k^\dagger v_{jklm} a_l a_m.$$

where

$$h_{jk} = \langle \psi_j | h | \psi_k \rangle$$

and

$$v_{jklm} = \int d^3r\, d^3r'\, \psi_j^\dagger(\mathbf{r})\psi_k^\dagger(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_l(\mathbf{r})\psi_m(\mathbf{r}').$$

Note that in the integral on the right hand side, the inner product of the spin-parts of the one-particle states is implicitly assumed, i.e. the spin of the states labelled j and l must be equal, and the same holds for the spins of states k and m. The Hamiltonian however does not affect the spin!

(a) Show that, for two particles in the *same* orbital $\langle \mathbf{r}|\phi\rangle$ and with opposite spin, the ground state for a wave function which has the form of a Slater determinant, is given by

$$E_0 = 2\langle\phi|h|\phi\rangle + \int d^3r\, d^3r' |\phi(\mathbf{r})|^2 |\phi(\mathbf{r}')|^2 \frac{1}{|\mathbf{r}-\mathbf{r}'|}.$$

Give a physical interpretation of the second term of this energy.

(b) For $\phi(\mathbf{r}) = \exp(-ar)$, find an analytic expression for the ground state energy for two electrons in the helium atom. In atomic units, these electrons are described by the Hamiltonian

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{-2}{r_1} + \frac{-2}{r_2} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

In atomic units, distances are calculated in units of the Bohr radius a_0, the electron mass and the charge are both 1, and the energy is expressed in Hartrees, 1 Hartree is 27.212 eV.

To find the expression for the energy, you must calculate quite a few integrals. There is one which is nontrivial:

$$\mathcal{N}^2 \int d^3r_1\, d^3r_2\, e^{-2ar_1} e^{-2ar_2} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

The prefactor $\mathcal{N}^2$ ensures normalisation of the orbitals $\exp(-ar)$. The result of this integral (including that prefactor) is $5a/8$.

Minimize the ground state with respect to a and compare the result with the known value for the ground state energy of a helium atom, which is $-78.975\ eV$.

6. Consider a many-body system consisting of N identical spin-1/2 particles with Hamilton operator:

$$H = \sum_{i=1}^{N} h^{(i)}; \qquad h^{(i)} = \frac{\mathbf{p}_i^2}{2m} + B\sigma_x^{(i)}.$$

Use the states $|\mathbf{k}\sigma\rangle$ as a basis (as usual, $\mathbf{p} = \hbar\mathbf{k}$, and σ is the eigenvalue of σ_z and therefore takes on the values +1 or −1).

(a) Show that the many-body energy operator H in the $\mathbf{k}\sigma$ representation is of the following form and determine f, g_+ and g_-:

$$H = \sum_{\mathbf{k},\sigma} f a^\dagger_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma} + \sum_{\mathbf{k}} \left(g_+ a^\dagger_{\mathbf{k},+} a_{\mathbf{k},-} + g_- a^\dagger_{\mathbf{k},-} a_{\mathbf{k},+} \right).$$

(b) Compute $\left[H, a_{\mathbf{k},\sigma}\right]$.

(c) Calculate the time-dependent operators $c_{\mathbf{k}}(t) = a_{\mathbf{k},+} + a_{\mathbf{k},-}$ and $d_{\mathbf{k}}(t) = a_{\mathbf{k},+} - a_{\mathbf{k},-}$. Give the physical meaning of these operators.

7. Consider the jellium model for electrons of density n in three dimensions. Assume that the interaction between two electrons has the form $U(r) = e^2/r$ as a function of the distance r between them. The ground state energy in the Hartree-Fock approximation is given by

$$E = 2\sum_{\mathbf{k}} \frac{\hbar k^2}{2m} + \frac{1}{2} \sum_{\mathbf{k},\mathbf{k}',\sigma} \left\langle a^\dagger_{\mathbf{k},\sigma} a^\dagger_{\mathbf{k}',\sigma} U_{\mathbf{k}-\mathbf{k}'} a_{\mathbf{k},\sigma} a_{\mathbf{k}',\sigma} \right\rangle.$$

The sums in this expression are all over the Fermi sphere. The second term contains an expectation value with respect to the ground state of the noninteracting system (i.e. a filled Fermi sphere). This second term has a form $\propto k_\mathrm{F}^4$ – it is the exchange energy of the electron gas.

We first consider the ground state with equal numbers of spin-up and -down electrons.

(a) Show that the kinetic energy contribution to E (per unit volume) is of the form

$$An^{\alpha}.$$

What is α? What is A?

(b) Show that the exchange contribution to E (per unit volume) is of the form

$$\frac{E_{\mathrm{ex}}}{V} = Bn^{\beta}.$$

What is β? How does B depend on V? What do you know about the sign of B?

(c) We now consider a spin-polarized system where all spins are pointing up. Find an equation for the density in terms of A and B where the spin-polarized state becomes stable with respect to the unpolarized state.

11

SUPERCONDUCTIVITY

11.1 INTRODUCTION

Superconductivity is the phenomenon that the conductance drops to zero below a threshold temperature, accompanied by a complete expulsion of the magnetic field from the interior of the superconductor (the 'Meissner effect'). In this chapter we shall consider theoretical approaches to this phenomenon. Note however, that we mainly discuss superconductivity as an illustration and culmination point of the formalism developed in the previous chapters, that is, you should not expect a course on superconductivity here.

The emphasis in this chapter is on the BCS theory of superconductivity. In the next section, we first introduce Cooper pairs which are held responsible for the phenomenon of superconductivity. In section 11.3 we shall discuss the form of the superconducting ground state wave function. This then allows us to formulate a Hamiltonian which is reduced to contain only the interaction terms relevant in the BCS wave function in section 11.4. This Hamiltonian will then be diagonalised by writing the wave function in the form of section 11.3. Finally, we give a brief account of the Landau-Ginzburg description of a superconductor in section 11.6.

11.2 COOPER PAIRS

In the previous chapter, we learned that the interaction between electrons and phonons can lead to an attractive effective interaction between electrons. This followed from the Hamiltonian

$$H = H_0 + \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} \left|M_{\mathbf{q}}\right|^2 \frac{\hbar\omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}})^2 - (\hbar\omega_{\mathbf{q}})^2} a^\dagger_{\mathbf{k}'+\mathbf{q}} a^\dagger_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}} a_{\mathbf{k}'}.$$

Lumping all the scalar (non-operator) terms in this interaction into a coupling constant $G_{\mathbf{q}}$, and realizing that we are dealing with an interaction which only depends on the distance between the electrons, we can write the Hamiltonian, now including spin, as [see (9.4)]:

$$H = H_0 + \sum_{\sigma,\sigma'} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}} G_{\mathbf{q}} a^\dagger_{\mathbf{k}'+\mathbf{q},\sigma'} a^\dagger_{\mathbf{k}-\mathbf{q},\sigma} a_{\mathbf{k},\sigma} a_{\mathbf{k}',\sigma'}.$$

The Hamiltonian H_0 describes non-interacting particles:

$$H_0 = \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k},\sigma} a_{\mathbf{k},\sigma}.$$

In principle, the full Hamiltonian H preserves the number of particles. Therefore, we can work in a N-electron Hilbert space. However, in the context of superconductivity it turns out useful to relax the constraint on the number of particles and control their number using a *chemical potential*, μ. This is defined as the total energy needed to add a particle to the system. From statistical mechanics we know that two systems in equilibrium at some temperature T with the possibility to exchange particles, have equal chemical potentials. If one of these systems (the 'bath') is much bigger than the other, the smaller system will loose or gain particles until its chemical potential is the same as that of the big system.

At zero temperature, if we neglect interactions, the ground state consists of a filled Fermi sphere, with the Fermi energy equalling the chemical potential μ. In the following, we take μ equal zero, so we should measure particle energies with respect to the chemical potential:

$$\epsilon_{\mathbf{k}} \to \epsilon_{\mathbf{k}} - \mu \equiv \tilde{\epsilon}_{\mathbf{k}}.$$

Neglecting the interactions, we see that the energy can be *lowered* by adding particles with negative energy, i.e. particles with $\epsilon_{\mathbf{k}} < \mu$, therefore μ is easily seen to control the particle number.

Superconductivity is the dramatic consequence of the fact that the electron-phonon coupling causes electrons to bind in pairs. The fact that the electron-phonon coupling plays an essential role in superconductivity was first suggested by the experimental discovery of the so-called *isotope effect*: the dependence of the critical temperature, above which superconductivity vanishes, on the mass of the nuclei. The electron-phonon coupling is the only mechanism by which nuclear mass can affect electronic behavior. Superconductivity was discovered in 1911 by Kamerlingh Onnes in Leiden, three years after he succeeded to liquefy helium for the first time. Following the 1933 discovery of the Meissner effect, which refers to the expulsion of a magnetic field inside a superconductor, and the description of this effect by the London equations in 1935, Landau and Ginzburg formulated in 1948 a successful phenomenological theory of superconductivity. Remarkably, it was not until 1957 that the first microscopic theory explaining the basic phenomena associated with superconductivity appeared. Named BCS theory after its creators Bardeen, Cooper and Schrieffer, we shall discuss it in some detail below.

We first want to gain insight into the structure of the ground state of the interacting system, which is the equilibrium state at zero absolute temperature. We start with a collection of particles occupying momentum states inside the Fermi sphere, and then add two particles to this system. The total energy of the system may be lowered when the particles form a bound state. We then expect these particles to have zero total momentum in order to minimize their centre of mass kinetic energy – hence the two particles are composed of plane waves with momenta $\mathbf{k}$ and $-\mathbf{k}$, both outside the Fermi sphere (the states inside the Fermi sphere are already occupied). In view of the overall anti-symmetry, either the orbital or the spin components of the two-particle wave function must be antisymmetric and the other symmetric. We anticipate that the symmetric orbital wave function will give us the lowest energy as the other option forbids the particles to approach each other and thereby take advantage of their attraction. These considerations lead to a wave function of the form

$$\Psi(x_1, x_2) = \sum_{k>k_{\mathrm{F}}} A_{\mathbf{k}} \cos(\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)) \frac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right),$$

where x_i denotes the combined orbital and spin coordinates $(\mathbf{r}_i, \sigma_i)$. Th presence of the cosine, and not the sine, in the sum is a consequence of the orbital symmetry. We can replace the cosine by an exponential function, and require $A_{\mathbf{k}} \equiv A_{-\mathbf{k}}$. We omit the spin part of the wave function in the following.

Inserting this Ψ into the Schrödinger equation, we arrive at

$$(E - 2\epsilon_{\mathbf{k}} + 2\mu) A_{\mathbf{k}} = \sum_{|\mathbf{k}'|>k_{\mathrm{F}}} G_{\mathbf{k}-\mathbf{k}'} A_{\mathbf{k}'}.$$

The coupling $G_{\mathbf{k}-\mathbf{k}'}$ is induced by phonons, and the typical maximum phonon frequency is the Debye frequency ω_{D}. The wave vector associated with these phonons is usually much smaller than the Fermi wave vector, and the electron-phonon coupling is relevant only when

$|\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}| < \hbar\omega_{\rm D}$. Within this range we approximate $G_{\mathbf{q}}$ by a constant, $-G$. We then have, replacing $\epsilon_{\mathbf{k}}$ by $\tilde{\epsilon}_{\mathbf{k}} + \mu$:

$$A_{\mathbf{k}} = G\frac{\sum_{\mathbf{k}'} A_{\mathbf{k}'}}{2\tilde{\epsilon}_{\mathbf{k}} - E},$$

where the sum over $\mathbf{k}'$ is understood to be in the narrow range corresponding to $\hbar\omega_{\rm D}$ and outside the Fermi sphere. Summing both sides over the same set of $\mathbf{k}$'s, we obtain

$$\frac{1}{G} = \sum_{\mathbf{k}} \frac{1}{2\tilde{\epsilon}_{\mathbf{k}} - E}.$$

Supposing that $N_{\rm F}$, the density of states near the Fermi level, is approximately constant over an energy range $\omega_{\rm D}$, we can replace the sum over $\mathbf{k}$ by an energy integral:

$$\frac{1}{G} = N_{\rm F}\int_0^{\hbar\omega_{\rm D}} \frac{d\tilde{\epsilon}}{2\tilde{\epsilon} - E} = \frac{1}{2}N_{\rm F}\ln\left(\frac{E - 2\hbar\omega_{\rm D}}{E}\right).$$

Assuming that the coupling is weak, $N_{\rm F}G \ll 1$, we can then write

$$E \approx -\hbar\omega_{\rm D} e^{-2/(N_{\rm F}G)}.$$

We see that, no matter how weak the electron-phonon interaction is, the two new electrons have a negative energy – hence they are in a bound state that will be absorbed into the ground state. This shows that the electrons form pairs, held together by the electron-phonon interaction – these are the *Cooper pairs*. Such pairs behave more like bosons than fermions, and for bosons it is known that above a certain density, they macroscopically occupy the ground state as in a Bose-Einstein condensate. We shall however not use that picture in these notes.

11.3 THE BCS WAVE FUNCTION

In the previous section, we have seen that the Fermi sphere is unstable against the formation of a bound (Cooper) pair of electrons in a spin-singlet state (antisymmetric spin state) and a symmetric orbital state. But what happens when we create a second Cooper pair, a third, etcetera? It may be that after creating a large number of Cooper pairs, the Fermi sea gets distorted and new Cooper pairs are no longer favourable. This turns out to be the scenario indeed, and we will end up with a Fermi sphere which is smaller than the sphere for non-interacting electrons, plus a 'layer' of Cooper pairs. This picture is conceivable and turns out to reflect the actual situation, but a quantitative description seems rather complicated. The description can be simplified by anticipating a particular form of the wave function, or BCS wave function [for the original paper, see J. Bardeen, L. N. Cooper and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957)]. Excellent descriptions can also be found in the books by De Gennes (*Superconductivity of Metals and Alloys*, Benjamin, New York, 1966) and Tinkham (*Introduction to Superconductivity*, McGraw Hill, 1975). We shall construct the BCS form from a mean-field approximation applied to the interaction term of the Hamiltonian. Let us finally anticipate the physical picture corresponding to the superconducting electron state. For small k, the electrons are still in the normal state which is appropriately described by a Slater determinant of independent plane waves. Near the Fermi wave vector, we have a layer of Cooper-pair states. Outside that layer, the electron states are unoccupied.

In the previous section we have seen that the electrons gain energy by forming Cooper pairs. This suggests choosing a particular form of wave function

$$\Phi_N = A\varphi(\mathbf{r}_1 - \mathbf{r}_2)\varphi(\mathbf{r}_3 - \mathbf{r}_4)\ldots\varphi(\mathbf{r}_{N-1} - \mathbf{r}_N)\,(\uparrow\downarrow\uparrow\downarrow\ldots\uparrow\downarrow).$$

Here, A in front of the expression is the anti-symmetrisation operator which ensures that the wave function has the appropriate fermion exchange anti-symmetry. This form deserves some explanation. First of all, we see that it *does not* have the form of a Slater determinant, as the electrons are *correlated in pairs* like $\varphi(\mathbf{r}_1 - \mathbf{r}_2)$ etc. There is no way to recast this into a single Slater-determinant form. Now remember that the Slater determinants represented the eigenstates of a Hamiltonian of *non-interacting* particles, then it is clear that the wave function given here carries the fingerprint of the electron-electron interactions – it is an *interacting wave function.*

The spins are collected together in the last part of this expression for the wave function. In this part, the spins are assumed to correspond to the particle ordering: 123.... From this we see that each orbital pair function $\varphi(\mathbf{r}_i - \mathbf{r}_j)$ is multiplied by a wave function of two opposite spins.

Let us now write

$$\varphi(\mathbf{r}) = \sum_{\mathbf{k}} g_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}.$$

The total wave function Φ can then be written as

$$\Phi_N = \sum_{\mathbf{k}_1,\dots,\mathbf{k}_{N/2}} g_{\mathbf{k}_1} g_{\mathbf{k}_2} \cdots g_{\mathbf{k}_{N-1}} g_{\mathbf{k}_N} e^{i\mathbf{k}_1(\mathbf{r}_1-\mathbf{r}_2)} A|1\uparrow;2\downarrow\rangle \dots e^{i\mathbf{k}_{N/2}(\mathbf{r}_{N-1}-\mathbf{r}_N)} |(N-1)\uparrow;N\downarrow\rangle.$$

We see that in this wave function, the spin-orbital $|\mathbf{k},\uparrow\rangle$ is paired with $|-\mathbf{k};\downarrow\rangle$. It is the job of the anti-symmetrisation operator A to turn this into an anti-symmetrized product. Note that each term in the sum over the wave vectors can be written as Slater determinant. This is not in contradiction with what was said above about $|\Phi_N\rangle$, as we have written this wave function as a *superposition* of Slater determinants. The anti-symmetrisation operation yields a Slater determinant of the form

$$|\mathbf{k}_1,\uparrow;-\mathbf{k}_1\downarrow;\dots\mathbf{k}_{N/2}\uparrow;-\mathbf{k}_{N/2}\downarrow\rangle_{\mathrm{F}},$$

where the subscript 'F" with the ket means that this is a state in Fock space, i.e. the anti-symmetrisation has been taken care of. We see that the **k**-vector and the spin have become *entangled* in this form: it is no longer possible to write these states as a product of a **k** part and a spin part.

We can write this state also in a different form, using fermion creation operators:

$$|\mathbf{k}_1,\uparrow;-\mathbf{k}_1,\downarrow;\dots\mathbf{k}_{N/2},\uparrow;-\mathbf{k}_{N/2},\downarrow\rangle_{\mathrm{F}} = a^\dagger_{\mathbf{k}_1,\uparrow} a^\dagger_{-\mathbf{k}_1,\downarrow} \dots a^\dagger_{\mathbf{k}_{N/2},\uparrow} a^\dagger_{-\mathbf{k}_{N/2},\downarrow} |0\rangle. \tag{11.1}$$

The BCS wave function is a superposition of these Slater determinants with expansion coefficients $g_{\mathbf{k}_1} \cdots g_{\mathbf{k}_{N/2}}$. Our job is now to vary N and the $g_{\mathbf{k}}$ in order to minimize the quantity $E - \mu N$, which is the free energy of the system.

It turns out rather difficult to handle the wave functions in the form given above due to the fact that it is a state for a *fixed* particle number, which requires minimization with respect to the g coefficients for any particle number N. A variable number is therefore preferred, however keeping the pair-wise coupling! It turns out that a more convenient form can be used where we replace the $g_{\mathbf{k}}$ by pairs of real numbers $u_{\mathbf{k}}$, $v_{\mathbf{k}}$:

$$|\tilde{\Phi}\rangle = \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow} \right) |0\rangle. \tag{11.2}$$

Expanding the product on the right hand side generates all possible wave functions with all possible particle numbers of the form (11.1), provided

$$\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}} = g_{\mathbf{k}} \text{ and } u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 = 1,$$

the latter condition ensuring proper normalisation of $\left|\tilde{\Phi}_N\right\rangle$. It is easy to see that for this wave function, the particle number is

$$\langle N\rangle = \sum_{\mathbf{k}} 2 v_{\mathbf{k}}^2,$$

as each combination $a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow}$ generates the two particles of a Cooper pair.

Let us make a hold for a second and summarize where we stand right now. We have first constructed the *interacting* wave function $|\Phi\rangle$, in which the electrons are pair-wise correlated. Then we have expanded this function in terms of Slater determinants and seen that in these determinants, the spin-orbitals $|\mathbf{k},\uparrow\rangle$ and $|-\mathbf{k},\downarrow\rangle$ always occur in pairs. Realising that this would urge us to solve the complicated problem of calculating first the energy for arbitrary N, followed by minimizing the free energy $E-\mu N$ with respect to both the expansion coefficients $g_{\mathbf{k}}$ and N, we have changed to a different form, (11.2), which encapsulates all expansions of our Slater determinants for arbitrary particle numbers.

Let us now carry out the remainder of our programme and evaluate the free energy for the state of (11.2). The Hamiltonian has the form [see (9.4)]

$$H=\sum_{\mathbf{k},\sigma}\epsilon_{\mathbf{k}}a^{\dagger}_{\mathbf{k},\sigma}a_{\mathbf{k},\sigma}+\sum_{\sigma,\sigma'}\sum_{\mathbf{k}\mathbf{k}'\mathbf{q}}G_{\mathbf{k},\mathbf{k}',\mathbf{q}}a^{\dagger}_{\mathbf{k}'+\mathbf{q},\sigma'}a^{\dagger}_{\mathbf{k}-\mathbf{q},\sigma}a_{\mathbf{k},\sigma}a_{\mathbf{k}',\sigma'}. \tag{11.3}$$

Note that the interaction term depends on *three* wave vectors, and not only on the momentum transfer $\mathbf{q}$. This is because the interaction depends on the lattice vibrations and is therefore not translationally invariant (we will return to this shortly). As already discussed above, the fact that we should minimize the free energy $E-\mu N$ can be taken care of by replacing

$$\epsilon_{\mathbf{k}}\to\epsilon_{\mathbf{k}}-\mu\equiv\tilde{\epsilon}_{\mathbf{k}}$$

and we define

$$H_0=\sum_{\mathbf{k}}\tilde{\epsilon}_{\mathbf{k}}\left(a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}+a^{\dagger}_{-\mathbf{k},\downarrow}a_{-\mathbf{k},\downarrow}\right).$$

The contribution of this term to the free energy is easily seen to be

$$\langle\tilde{\Phi}|\,H_0\,|\tilde{\Phi}\rangle=2\sum_{\mathbf{k}}v^2_{\mathbf{k}}\tilde{\epsilon}_{\mathbf{k}}.$$

The interaction term is more difficult to analyse. This term removes two particles from $|\tilde{\Phi}\rangle$ and then creates two particles, possibly in different states. We obviously have the two possibilities that we already used in the Hartree-Fock theory (see chapter 10): (i) $\mathbf{q}=0$ or (ii) $\mathbf{q}=\mathbf{k}-\mathbf{k}'$. The first term just calculates the interaction between the two particle distributions and the second one results from the anti-symmetry of the wave function. It turns out that these contributions can be incorporated into the single-particle energies $\epsilon_{\mathbf{k}}$, hence they do not cause Cooper pairing.

As the ground state does not have a fixed set of filled states, there is another possibility: the annihilation operators may remove a pair $\mathbf{k}\uparrow;-\mathbf{k},\downarrow$ and replace it by another pair $\mathbf{l}\uparrow;-\mathbf{l},\downarrow$, where $\mathbf{k}\neq\mathbf{l}$. Because in the state on the right hand side, the $\mathbf{l}$ pair should be empty and the $\mathbf{k}$ pair occupied, we have a term proportional to $u_{\mathbf{l}}$ and to $v_{\mathbf{k}}$ from the wave function on the right hand side of the expectation value (the 'ket'-part). Then we have a contribution with $v_{\mathbf{l}}$ and $u_{\mathbf{k}}$ from the left hand wave function (the 'bra' part), as in that state the $\mathbf{l}$ pair must be occupied and the $\mathbf{k}$ pair empty. The contribution obtained should therefore be proportional to $u_{\mathbf{k}}v_{\mathbf{k}}u_{\mathbf{l}}v_{\mathbf{l}}$. Furthermore, we need $\mathbf{k}=-\mathbf{k}'$ and $\mathbf{q}=\mathbf{k}-\mathbf{l}$. We then obtain a contribution in (11.3)

$$\langle\tilde{\Phi}|\,H_{\text{int}}\,|\tilde{\Phi}\rangle=\sum_{\mathbf{k},\mathbf{l}}G_{\mathbf{k},\mathbf{l}}u_{\mathbf{k}}v_{\mathbf{k}}u_{\mathbf{l}}v_{\mathbf{l}}.$$

So we obtain for the total free energy

$$\langle\tilde{\Phi}|\,H-\mu N\,|\tilde{\Phi}\rangle=2\sum_{\mathbf{k}}v^2_{\mathbf{k}}\tilde{\epsilon}_{\mathbf{k}}+\sum_{\mathbf{k},\mathbf{l}}G_{\mathbf{k},\mathbf{l}}u_{\mathbf{k}}v_{\mathbf{k}}u_{\mathbf{l}}v_{\mathbf{l}}. \tag{11.4}$$

We must find the minimum of this function with respect to the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, subject to the condition that $u^2_{\mathbf{k}}+v^2_{\mathbf{k}}=1$. We realise this constraint by parametrising $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ by a single variable $\theta_{\mathbf{k}}$:

$$u_{\mathbf{k}}=\sin\theta_{\mathbf{k}};\qquad v_{\mathbf{k}}=\cos\theta_{\mathbf{k}}.$$

In terms of $\theta_{\mathbf{k}}$, the expectation value of the free energy now reads (using $\sin\theta\cos\theta = \sin(2\theta)/2$):

$$\langle\tilde{\Phi}|\, H-\mu N\, |\tilde{\Phi}\rangle = 2\sum_{\mathbf{k}} \tilde{\epsilon}_{\mathbf{k}} \cos^2\theta_{\mathbf{k}} + \frac{1}{4}\sum_{\mathbf{k},\mathbf{l}} \sin(2\theta_{\mathbf{k}})\sin(2\theta_{\mathbf{l}})\, G_{\mathbf{k},\mathbf{l}}.$$

The minimum of this free energy is found by putting its derivatives with respect to the $\theta_{\mathbf{k}}$ to zero:

$$0 = \frac{\partial}{\partial\theta_{\mathbf{k}}} \langle\tilde{\Phi}_N|\, H_{\text{int}}\, |\tilde{\Phi}_N\rangle = -2\tilde{\epsilon}_{\mathbf{k}}\sin(2\theta_{\mathbf{k}}) + \sum_{\mathbf{l}} \cos(2\theta_{\mathbf{k}})\sin(2\theta_{\mathbf{l}})\, G_{\mathbf{k},\mathbf{l}},$$

which we can reduce to

$$\tilde{\epsilon}_{\mathbf{k}}\tan(2\theta_{\mathbf{k}}) = \frac{1}{2}\sum_{\mathbf{l}} G_{\mathbf{k},\mathbf{l}}\sin(2\theta_{\mathbf{l}}).$$

We now introduce the quantities:

$$\Delta_{\mathbf{k}} = -\sum_{\mathbf{l}} G_{\mathbf{k},\mathbf{l}} u_{\mathbf{l}} v_{\mathbf{l}};$$
$$E_{\mathbf{k}} = \sqrt{\tilde{\epsilon}_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}.$$

Using these, we can simplify the minimum condition to

$$\tan(2\theta_{\mathbf{k}}) = -\frac{\Delta_{\mathbf{k}}}{\tilde{\epsilon}_{\mathbf{k}}}, FACTOROF2!!!!!!!!!!!!!!!!$$

and this in turn enables us to write

$$2u_{\mathbf{k}}v_{\mathbf{k}} = \sin(2\theta_{\mathbf{k}}) = \frac{\Delta_{\mathbf{k}}}{E_{\mathbf{k}}} \tag{11.5}$$
$$-u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 = \cos(2\theta_{\mathbf{k}}) = -\frac{\tilde{\epsilon}_{\mathbf{k}}}{E_{\mathbf{k}}}. \tag{11.6}$$

Using the first of these equations in the definition for $\Delta_{\mathbf{k}}$ above, we arrive at an implicit equation for the $\Delta_{\mathbf{k}}$:

$$\Delta_{\mathbf{k}} = -\sum_{\mathbf{l}} G_{\mathbf{k},\mathbf{l}} \frac{\Delta_{\mathbf{l}}}{2E_{\mathbf{l}}}. \tag{11.7}$$

Let us now try to solve this self-consistency equation. The first solution we immediately recognise is $\Delta_{\mathbf{k}} = 0$ for all $\mathbf{k}$. This means that either $v_{\mathbf{k}} = 0$ or $u_{\mathbf{k}} = 0$. As in that case $E_{\mathbf{k}} = \tilde{\epsilon}_{\mathbf{k}}$ for $\tilde{\epsilon}_{\mathbf{k}} > 0$ and $E_{\mathbf{k}} = -\tilde{\epsilon}_{\mathbf{k}}$ for $\tilde{\epsilon}_{\mathbf{k}} < 0$, we see from (11.6) that

$$v_{\mathbf{k}} = \begin{cases} 1 \text{ for } \tilde{\epsilon}_{\mathbf{k}} < 0; \\ 0 \text{ for } \tilde{\epsilon}_{\mathbf{k}} > 0. \end{cases} \tag{11.8}$$

This is precisely a filled Fermi sphere, and this does not give us superconductivity.

In order to take advantage of the correlated structure of the BCS wave function to lower the energy, we must have a non-zero $\Delta_{\mathbf{k}}$, requiring a non-zero interaction. BCS chose:

$$G_{\mathbf{k},\mathbf{l}} = \begin{cases} -G & \text{for } \tilde{\epsilon}_{\mathbf{k}}, \tilde{\epsilon}_{\mathbf{l}} < \hbar\omega_{\text{D}}; \\ 0 & \text{otherwise}, \end{cases}$$

where ω_{D} denotes the Debye frequency as usual. Now the self-consistency equation (11.7) can be solved. First we realise that

$$\Delta_{\mathbf{k}} = \begin{cases} 0 & \text{for } \tilde{\epsilon}_{\mathbf{k}} > \hbar\omega_{\text{D}}; \\ \Delta & \text{independent of } \mathbf{k} \text{ for } \tilde{\epsilon}_{\mathbf{k}} < \hbar\omega_{\text{D}}. \end{cases}$$

Putting this into (11.7) and replacing sums by integrals by means of the density of states (see section 11.2) gives

$$\Delta = N_\mathrm{F} G \int_{-\hbar\omega_\mathrm{D}}^{\hbar\omega_\mathrm{D}} \Delta \frac{1}{2\sqrt{\Delta^2+\tilde{\epsilon}^2}} d\tilde{\epsilon},$$

from which we find

$$\frac{1}{N_\mathrm{F} G} = \int_0^{\hbar\omega_\mathrm{D}} \frac{1}{\sqrt{\Delta^2+\tilde{\epsilon}^2}} d\tilde{\epsilon} = \sinh^{-1}\left(\frac{\hbar\omega_\mathrm{D}}{\Delta}\right). GANDFACTOROF2MISSING!!!!!!!!!!!gPROBABLYINDENOMONLHS$$

Note that for this equation to hold, the interaction should be attractive, i.e. $G > 0$. We rework the result analogously to section 11.2 to obtain

$$\Delta = 2\hbar\omega_\mathrm{D} e^{-1/(N_\mathrm{F} G)}$$

when $N_\mathrm{F} G \ll 1$. In practice, $N_\mathrm{F} G$ only rarely exceeds 0.3.

It remains to check whether the energy is reduced with respect to that of the non-interacting wave function (i.e. a Slater determinant corresponding to a filled Fermi sphere). For this purpose, it is useful to calculate the $u_\mathbf{k}$ and $v_\mathbf{k}$. Combining (11.6) with the normalisation $u_\mathbf{k}^2 + v_\mathbf{k}^2 = 1$, it is easy to see that for the BCS state:

$$u_\mathbf{k}^2 = \frac{1}{2}\left(1 + \frac{\tilde{\epsilon}_\mathbf{k}}{\sqrt{\tilde{\epsilon}_\mathbf{k}^2 + \Delta^2}}\right)$$

and

$$v_\mathbf{k}^2 = \frac{1}{2}\left(1 - \frac{\tilde{\epsilon}_\mathbf{k}}{\sqrt{\tilde{\epsilon}_\mathbf{k}^2 + \Delta^2}}\right),$$

whereas the normal state is characterised by (11.8). These can be used directly in the expression for the total energy (11.4) to obtain

$$E_\mathrm{BCS} = \sum_\mathbf{k} \tilde{\epsilon}_\mathbf{k}\left(1 - \frac{\tilde{\epsilon}_\mathbf{k}}{\sqrt{\tilde{\epsilon}_\mathbf{k}^2 + \Delta_\mathbf{k}^2}}\right) + \sum_\mathbf{kl} G_\mathbf{k,l} \frac{\Delta_\mathbf{k}\Delta_\mathbf{l}}{4\sqrt{\tilde{\epsilon}_\mathbf{k}^2 + \Delta_\mathbf{k}^2}\sqrt{\tilde{\epsilon}_\mathbf{l}^2 + \Delta_\mathbf{l}^2}}. \tag{11.9}$$

In problem 1 it is shown that the energy difference results in

$$\langle \Phi_\mathrm{N} | \hat{H} | \Phi_\mathrm{N} \rangle - \langle \Phi_\mathrm{BCS} | \hat{H} | \Phi_\mathrm{BCS} \rangle = N_\mathrm{F} \frac{\Delta^2}{2}.$$

This means that the ground state energy of the BCS is *lower* than that of the normal state. The BCS state however only wins by a small amount: it is only of the order of $\Delta^2/E_\mathrm{F} \approx 10^{-2}$K per particle.

11.4 THE BCS HAMILTONIAN

From now on we shall denote single particle states by k which denotes the momentum $\mathbf{k}$ and the spin s. A Cooper pair can then be denoted as $(k, -k)$, which means that we have two particles with opposite spin and opposite momentum. The Hamiltonian containing only the interaction terms representing the replacement of one Cooper pair by another then has the form

$$H = \sum_k \epsilon_k \left(a_k^\dagger a_k + a_{-k}^\dagger a_{-k}\right) - \sum_{kl} G_{kl} a_l^\dagger a_{-l}^\dagger a_{-k} a_k.$$

Finding the ground state for our Hamiltonian is impossible without making approximations. A classic approximation scheme that predicts the behaviour of standard superconductors remarkably well is the *mean field approximation*. In this approximation, we neglect the deviations – to a certain order – of the operators $a_{-k}a_k$ and $a^\dagger_k a^\dagger_{-k}$ from their expectation values in the ground state (which still must be determined). To be specific, we write

$$a^\dagger_l a^\dagger_{-l} a_{-k} a_k = \left(a^\dagger_l a^\dagger_{-l} - \left\langle a^\dagger_l a^\dagger_{-l}\right\rangle + \left\langle a^\dagger_l a^\dagger_{-l}\right\rangle\right)\left(a_{-k}a_k - \langle a_{-k}a_k\rangle + \langle a_{-k}a_k\rangle\right)$$

and expand the products, neglecting the second-order fluctuation term $\left(a^\dagger_l a^\dagger_{-l} - \left\langle a^\dagger_l a^\dagger_{-l}\right\rangle\right)$ $(a_{-k}a_k - \langle a_{-k}a_k\rangle)$:

$$H_{\text{MFA}} = \sum_k \epsilon_k \left(a^\dagger_k a_k + a^\dagger_{-k}a_{-k}\right) - \sum_{kl} G_{kl}\left(a^\dagger_l a^\dagger_{-l}\langle a_{-k}a_k\rangle + a_{-k}a_k\left\langle a^\dagger_l a^\dagger_{-l}\right\rangle - \langle a_{-k}a_k\rangle\left\langle a^\dagger_l a^\dagger_{-l}\right\rangle\right).$$

Noticing that the expectation values are scalars, we replace them by numbers $C_k = \langle a_{-k}a_k\rangle$, $C^*_k = \left\langle a^\dagger_k a^\dagger_{-k}\right\rangle$. Furthermore, calling

$$\Delta_k = \sum_l G_{kl}C_l,$$

the Hamiltonian assumes the form

$$H = \sum_k \epsilon_k\left(a^\dagger_k a_k + a^\dagger_{-k}a_{-k}\right) + \sum_k \Delta_k C^*_k - \sum_k\left(\Delta^*_k a_{-k}a_k + \Delta_k a^\dagger_k a^\dagger_{-k}\right).$$

It turns out that the phase of Δ_k can be chosen arbitrarily (to show this, we need arguments related to gauge invariance which are beyond the scope of these notes) – we choose Δ_k to be real.

The resulting Hamiltonian is quadratic (in the sense that every term contains at most two operators). If the Hamiltonian would only contain products of a creation and an annihilation operator, the diagonalisation would have been straightforward, as H could then be completely formulated in terms of number operators. In our case, however, the products of two annihilation and two creation operators cause problems. Related to this is that the Hamiltonian in this form does not conserve particle number, although the original Hamiltonian did. This lack of particle number conservation is a consequence of the mean field approximation.

There is a trick for transforming the mean field Hamiltonian into a diagonalizable one. This trick is essentially a *linear transformation* of the operators a_k and $a^\dagger_k$ which *removes* the 'difficult' terms (i.e. the terms with two annihilation or two creation operators) and leaves only number operators. This transformation, which is closely related to the construction of the BCS wave function discussed in the previous section, is called the 'Bogoliubov-Valatin' transformation. The most general linear transformation of the a-operators is

$$\alpha_k = u_k a_k - v_k a^\dagger_{-k},$$
$$\beta_k = u_k a_{-k} + v_k a^\dagger_k.$$

Note that the u_k and v_k are numbers; the a_k, α_k, β_k and their Hermitian conjugates are operators. We would like the operators α_k and β_k to obey the usual anti-commutation relations for fermion operators. This imposes a condition on the coefficients u_k and v_k:

$$u^2_k + v^2_k = 1.$$

Inverting the transformation leads to

$$a_k = u_k\alpha_k + v_k\beta^\dagger_k,$$
$$a_{-k} = u_k\beta_k - v_k\alpha^\dagger_k.$$

Now we use these transformations to recast the original Hamiltonian into a form containing only the α_k and β_k (and their Hermitian conjugates). The non-interacting part then yields

$$a^\dagger_k a_k + a^\dagger_{-k}a_{-k} = 2v^2_k + \left(u^2_k - v^2_k\right)\left(\alpha^\dagger_k\alpha_k + \beta^\dagger_k\beta_k\right) + 2u_k v_k\left(\alpha^\dagger_k\beta^\dagger_k + \beta_k\alpha_k\right).$$

For the terms deriving from the interaction, we have

$$a_{-k}a_k + a_k^\dagger a_{-k}^\dagger = 2u_k v_k - 2u_k v_k\left(\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k\right) + \left(u_k^2 - v_k^2\right)\left(\alpha_k^\dagger \beta_k^\dagger + \beta_k \alpha_k\right).$$

Note that except for the normalization condition for the u_k and v_k, we still have the freedom to choose these parameters. We choose them such as to make the coefficient of the terms containing two annihilation or two creation operators zero. Such terms arising from the non-interacting and interacting parts of the Hamiltonian combine into

$$\sum_k \left[2\epsilon_k u_k v_k - \Delta_k\left(u_k^2 - v_k^2\right)\right]\left(a_k^\dagger \beta_k^\dagger + \beta_k \alpha_k\right).$$

Therefore we impose

$$2\epsilon_k u_k v_k - \Delta_k\left(u_k^2 - v_k^2\right) = 0.$$

Defining $g_k = v_k/u_k$, we find that the positive value of g_k is given by

$$g_k = \frac{E_k - \epsilon_k}{\Delta_k},$$

with

$$E_k = \sqrt{\epsilon_k^2 + \Delta_k^2}.$$

Combining this with $u_k^2 + v_k^2 = 1$, we have, similar to the result obtained in the previous section:

$$u_k^2 = \frac{1}{2}\left(1 + \frac{\epsilon_k}{E_k}\right) = \frac{1}{2}\left(1 + \frac{\epsilon_k}{\sqrt{\epsilon_k^2 + \Delta_k^2}}\right)$$

and

$$v_k^2 = \frac{1}{2}\left(1 - \frac{\epsilon_k}{E_k}\right) = \frac{1}{2}\left(1 - \frac{\epsilon_k}{\sqrt{\epsilon_k^2 + \Delta_k^2}}\right).$$

The part of the Hamiltonian that survives is

$$H_{\mathrm{MFA}} = 2\sum_k \epsilon_k v_k^2 - \sum_k \Delta_k u_k v_k + \sum_k \left[\epsilon_{\mathbf{k}}\left(u_k^2 - v_k^2\right) + 2\Delta_k u_k v_k\right]\left(\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k\right).$$

This has the form

$$H_{\mathrm{MFA}} = \sum_k \left[M_k + N_k(n_{\alpha,k} + n_{\beta,k})\right],$$

where $n_{\alpha,\beta}$ are the number operators for the α and β particles; M_k and N_k are complex numbers. The α and β 'particles' are in fact the excitations of the ground state with definite energy. This brings us to the notion of a 'quasi-particle'. If we were to add an electron to the system without changing the state of the 'resident' particles, we obtain a state in the Fock space which is not an eigenstate of the Hamiltonian. Only judicious combinations of real particles give excitations which are eigenstates of the Hamiltonian – these are the α and β particles. As they are not actual electrons or holes, they are called *quasi-particles*.

The vacuum state of the electron system satisfies

$$a_k|0\rangle = 0$$

for each k. Similarly, the quasi-particle ground state is defined as the state satisfying

$$\alpha_k|0\rangle_{\mathrm{quasi}} = \beta_k|0\rangle_{\mathrm{quasi}} = 0$$

for all k. From this, and using the definition of the quasi-particle operators α_k and β_k, we can express the quasi-particle vacuum state in terms of the electron vacuum state:

$$|0\rangle_{\text{quasi}} = \prod_k \left(u_k + v_k a_k^\dagger a_{-k}^\dagger \right) |0\rangle . \tag{11.10}$$

The fact that this is the quasi-particle vacuum can be verified by acting with α_k and β_k on it. As the vacuum state is unique, the given expression must be correct.

Let us have a closer look at (11.10). We run over all possible k values. With probability $|v_k|^2$ we create a pair of states $(k, -k)$ and with probability $|u_k|^2$ we do nothing, i.e. we leave an empty state empty. That is, the relative sizes of u_k and v_k are parameters controlling the particle number. For u_k close to 1 (and, consequently, v_k close to 0), there are no electrons with momentum $\mathbf{k}$. For u_k close to zero, hence v_k close to one, we have 'ordinary' electron pairs.

Let us return to the definition of Δ_k:

$$\Delta_k = \sum_l G_{kl} C_l,$$

with $C_l = \langle a_l a_{-l} \rangle$. Let us work out this expectation value for the ground state of the α and β particles:

$$C_l = {}_{\text{quasi}}\langle 0 | a_{-l} a_l | 0 \rangle_{\text{quasi}} = \prod_{kk'} \left\langle 0 \left| (u_{k'} + v_{k'} a_{-k'} a_{k'})\, a_{-l} a_l \left(u_k + v_k a_k^\dagger a_{-k}^\dagger \right) \right| 0 \right\rangle = u_l v_l .$$

In obtaining the right-hand side, we made use of the fact that for $k \neq l$, the terms of the quasi-particle ground state in the ket vector are orthonormal to the terms in the bra-vector. This forces the k'-s to be equal to the k-s. Thus we have

$$\Delta_k = \sum_l G_{kl} u_l v_l = \frac{1}{2} \sum_l \frac{G_{kl} \Delta_l}{\sqrt{\tilde{\epsilon}_l^2 + \Delta_l^2}} .$$

This is an implicit equation for Δ_k – it is known as the *gap equation.*

The quasi-particle Hamiltonian can now be written as

$$H = 2 \sum_{k>0} \tilde{\epsilon}_k v_k^2 - \sum_k \Delta_k u_k v_k + \sum_{k>0} E_k \left(\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k \right) .$$

We can simplify the analysis considerably by neglecting the kl dependence of G: $G_{kl} \equiv G$. Under this approximation, Δ becomes a constant independent of k and we can write:

$$H = 2 \sum_{k>0} \tilde{\epsilon}_k v_k^2 - \frac{\Delta^2}{G} + \sum_{k>0} E_k \left(\alpha_k^\dagger \alpha_k + \beta_k^\dagger \beta_k \right) .$$

The gap Δ is now given by the implicit equation

$$1 = \frac{G}{2} \sum_k \frac{1}{\sqrt{\tilde{\epsilon}_k^2 + \Delta^2}} .$$

Now that we have cast the Hamiltonian into a diagonisable form, we can study the structure of its ground-state wave function. Let us first consider the states at low energies. We assume that Δ is small with respect to ϵ_F. Low energy means that $\tilde{\epsilon}$ is close to $-\epsilon_F$. Then we find that

$$u_k^2 = \frac{1}{2} \left(1 + \frac{\tilde{\epsilon}_k}{\sqrt{\tilde{\epsilon}_k^2 + \Delta^2}} \right)$$

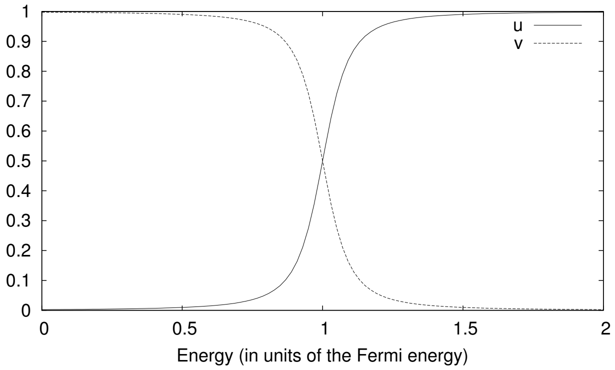


FIGURE 11.1: Variation of the coefficients $|u_k|^2$ and $|v_k|^2$ with energy. The width of the transition region is determined by the parameter Δ (here, $\Delta = 0.1\epsilon_F$).

is close to zero and

$$v_k^2 = \frac{1}{2}\left(1 - \frac{\tilde{\epsilon}_k}{\sqrt{\tilde{\epsilon}_k^2 + \Delta^2}}\right)$$

is close to one. As already mentioned, this creates independent $(k, -k)$ pairs, indistinguishable from ordinary electrons in the independent particle picture. These particles are said to be in the *normal state*.

At high energies, we have u_k close to 1 and v_k close to 0 – therefore, we have no particles. For an energy range Δ near $\epsilon_k = \epsilon_F$, we see that u_k and v_k vary between 0 and 1: this is the region where we find quasi-particles. Figure 11.1 shows the variation of u_k and v_k with energy, illustrating the behaviour described.

The ground state energy is given by

$$E_{BCS} = 2\sum_k \tilde{\epsilon}_{\mathbf{k}} v_k^2 - \sum_k \Delta_k u_k v_k.$$

Substituting the expressions for u_k and v_k, we can write the total energy as

$$E_{BCS} = -\sum_k \frac{(E_k - \tilde{\epsilon}_k)^2}{2E_k} < 0.$$

Creating a quasi-particle excitation $\alpha_k^\dagger |0\rangle_{\text{quasi}}$ requires an energy $E_k = \sqrt{\tilde{\epsilon}_k^2 + \Delta_k^2}$ which at the Fermi energy ($\tilde{\epsilon}_k = 0$) is equal to Δ. This gap is responsible for the exponential temperature dependence of the specific heat below the critical temperature.

The operator $a_k^\dagger a_k$ counts the number of electrons in the momentum-spin state k. The number of electrons is therefore given by

$$N = \sum_k \left\langle 0 \middle| a_k^\dagger a_k + a_{-k}^\dagger a_{-k} \middle| 0 \right\rangle_{\text{quasi}} = 2\sum_k v_k^2.$$

We finally note that the equation for the gap leads to the same relation between G and Δ as found for the single Cooper pair:

$$\Delta = 2\hbar\omega_D e^{-1/(N_F G)}.$$

11.5 SUMMARY OF BCS THEORY

In the previous sections, we have seen that an attractive interaction leads to bound states of electron pairs, which have opposite $\mathbf{k}$ and opposite spin – these are the Cooper pairs. In the section 11.2, we have seen that each Cooper pair lowers the energy by an amount

$$\Delta E = -\hbar\omega_D e^{-1/(N_F G)}.$$

Here, N_{F} is the density of states near the Fermi energy and G is the effective coupling strength by which electrons are attracted: its physical background is the electron-phonon interaction, which explains the presence of the pre-factor ω_{D} which is the Debye frequency, a cut-off of the phonon frequency.

In section 11.3, we have analysed the structure of the many-body wave function which contains two-particle correlations due to the Copper pair interaction. The BCS wave function is given as

$$|\tilde{\Phi}_N\rangle = \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow}\right) |0\rangle .$$

which contains wave functions for all possible particle numbers. These wave functions furthermore contain entanglement of the wave vector and spin [1]. It turns out that the corresponding Hamiltonian containing the terms which are exploited by the BCS wave function to lower the energy, reads

$$H = \sum_k \epsilon_k \left(a^\dagger_k a_k + a^\dagger_{-k} a_{-k}\right) - \sum_{kl} G_{kl} a^\dagger_l a^\dagger_{-l} a_{-k} a_k.$$

Here we have used the notation $k \equiv (\mathbf{k}, \uparrow)$ and $-k = (-\mathbf{k}, \downarrow)$.

This Hamiltonian was analysed in section 11.4 using mean-field theory. In this theory, products of two creation operators are replaced by their expectation values:

$$a_{-k} a_k \to \langle a_{-k} a_k \rangle \equiv C_k$$

and similar for creation operators. Fluctuations from these expectation values to second order are then removed from the Hamiltonian. This leads to a quadratic Hamiltonian (that is, a Hamiltonian containing products of at most two creation/annihilation operators) which still contains the unknown values C_k and their complex conjugates C_k^*. These have to be solved for self-consistently. Diagonalised proceeds using the Bogoliubov-Valatin transformation and it leads to *quasi-particles* destroyed by the operators α and β (and created by their Hermitian conjugates). The quasi-particles carry energies

$$E_{\mathbf{k}} = \sqrt{\tilde{\epsilon}^2_{\mathbf{k}} + \Delta^2}.$$

Here, Δ is the *superconducting gap*, its value is given by

$$\Delta = \sum_l G_{kl} v_k u_l.$$

11.6 LANDAU-GINZBURG THEORY AND THE LONDON EQUATIONS

In the previous section we have considered the BCS ground state and analyzed excitations from this. Seven years before BCS theory was developed, there already existed a successful *phenomenological* description of superconductivity: the *Landau Ginzburg theory*. A detailed discussion of this theory is beyond the scope of these notes. Instead, we shall give a short description of it. Landau and Ginzburg have written down a phenomenological expression for the free energy difference $\Delta\mathscr{F}$ between a superconductor and the normal state. This expression is

$$\Delta\mathscr{F} = \left[\int \psi^*(\mathbf{r}) \frac{1}{2m^*} \left(\frac{\hbar}{\mathrm{i}}\nabla - \frac{e^*}{c}\mathbf{A}\right)^2 \psi(\mathbf{r}) + \alpha|\psi(\mathbf{r})|^2 + \frac{\beta}{2}|\psi(\mathbf{r})|^4\right] d^3r.$$

Here, $\psi(\mathbf{r})$ is not an operator, but rather a complex field, the phase of which represents the phase of the wave function, and its norm gives the 'density of superconducting electrons', n_s:

$$n_s(\mathbf{r}) = \left|\psi(\mathbf{r})\right|^2.$$

---

[1] Entanglement will be formally defined in the next section; for now it means that the wave function of two electrons cannot be written as the product of a **k**-part and a spin-part

The form of this free energy can be understood as follows. The free energy will depend on the field ψ and its gradients. As the free energy is an extensive quantity, we can think of it as a sum of the free energies within small boxes into which we have divided our large system. This explains the integral over quantities depending only on $\mathbf{r}$. Assuming that ψ and its gradients are small, the integrand can be expanded in a Taylor expansion of the field ψ and its gradient. This expression must be real, and it should be symmetric under reversing the sign of ψ (for an isotropic system). This explains the powers of 2 and 4. In principle, higher powers will be present, but the ones used here are sufficient for describing the main features.

The free energy reaches its minimum value in equilibrium. We find this minimum energy by varying the field ψ:

$$\frac{1}{2m^*}\left(\frac{\hbar}{\mathrm{i}}\nabla - \frac{e^*}{c}\mathbf{A}\right)^2 \psi(\mathbf{r}) + \beta \left|\psi(\mathbf{r})\right|^2 \psi(\mathbf{r}) = -\alpha\psi(\mathbf{r}).$$

One might ask why we have taken a complex field, and not a real one. The reason is that we want to describe currents, and currents are related to the variation of the phase of the wavefunction. From the fact that $|\psi(\mathbf{r})|^2$ gives the superconducting density, the current must be given by

$$\mathbf{j} = \frac{e^*}{2m^*}\left[\psi^*(\mathbf{r})\left(\frac{\hbar}{\mathrm{i}}\nabla - \frac{e^*}{c}\mathbf{A}\right)\psi(\mathbf{r}) + \psi(\mathbf{r})\left(-\frac{\hbar}{\mathrm{i}}\nabla - \frac{e^*}{c}\mathbf{A}\right)\psi^*(\mathbf{r})\right].$$

Putting $e^* = 2e$ and $m^* = 2m$, as the ‘particles’ described by ψ are Cooper pairs, we obtain from the last equation:

$$\mathbf{j}_s = \frac{e}{4m}\left[\psi^*(\mathbf{r})\left(\frac{\hbar}{\mathrm{i}}\nabla - \frac{2e}{c}\mathbf{A}\right)\psi(\mathbf{r}) + \psi(\mathbf{r})\left(-\frac{\hbar}{\mathrm{i}}\nabla - \frac{2e}{c}\mathbf{A}\right)\psi^*(\mathbf{r})\right],$$

where the subscript ’s’ denotes that this is the current of the superconducting fraction of electrons. Neglecting the $\mathbf{r}$-dependence of ψ with respect to that of $\mathbf{A}$, we then obtain

$$\mathbf{j}_s = -\frac{e^2}{m} n_s \mathbf{A},$$

which, after taking the curl in the left and right hand side, gives

$$\nabla \times \mathbf{j}_s = -\frac{2e^2}{m} n_s \mathbf{B}.$$

This equation is well known for superconductors – it is called the *London equation.* Note that the approximation in which we have neglected the variation of ψ is justified in the case where the superconductor is homogeneous and the field is weak: in that case, ψ is approximately the equilibrium value with a small, space-dependent perturbation on top of it.

We now want to show that the London equation leads to the expulsion of magnetic fields from the superconductor. Using the Maxwell equation

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{j},$$

we have

$$\nabla \times \nabla \times \mathbf{B} = -k^2 \mathbf{B},$$

with

$$k^2 = \frac{8\pi e^2 n_s}{mc}.$$

But we also know from vector calculus that

$$\nabla \times \nabla \times \mathbf{B} = -\nabla^2 \mathbf{B},$$

so that we have

$$\nabla^2\mathbf{B} = k^2\mathbf{B}.$$

For a field which, at the $x-y$ surface of a superconductor, is homogeneous, the solution is

$$\mathbf{B} = e^{-kz}\mathbf{B}_0,$$

which tells us that the field decays inside the superconductor with a penetration depth of $-1/k = \sqrt{mc/(8\pi e^2 n_s)}$.

11.7 PROBLEMS

1. In this problem we consider the energy difference between the normal and the superconducting state.

   (a) Show that equation (11.9) can be written in the form

   $$\int_{-\infty}^{\infty} N(\tilde{\epsilon})\tilde{\epsilon}\left(1-\frac{\tilde{\epsilon}}{\sqrt{\tilde{\epsilon}^2+\Delta^2(\epsilon)}}\right)d\tilde{\epsilon} - \int N(\tilde{\epsilon})N(\tilde{\epsilon}')V\frac{\Delta(\tilde{\epsilon})\Delta(\tilde{\epsilon}')}{4\sqrt{\tilde{\epsilon}^2+\Delta(\tilde{\epsilon})}\sqrt{\tilde{\epsilon}'^2+\Delta(\tilde{\epsilon}')}}d\tilde{\epsilon}d\tilde{\epsilon}'.$$

   Here, $N(\tilde{\epsilon})$ is the density of states.

   (b) This expression can be strongly simplified by realising that $\Delta(\tilde{\epsilon}) = 0$ for $|\tilde{\epsilon}| > \hbar\omega$ and $\Delta_{\mathbf{k}} = \Delta$ for $|\epsilon_{\mathbf{k}}| < \hbar\omega_{\mathbf{k}}$. Use this to arrive at

   $$E_{\mathrm{BCS}} = 2\int_{-\infty}^{\hbar\omega_{\mathrm{D}}} N(\tilde{\epsilon})\tilde{\epsilon}d\tilde{\epsilon} + N_{\mathrm{F}}\int_{-\hbar\omega_{\mathrm{D}}}^{\hbar\omega_{\mathrm{D}}}\left(|\tilde{\epsilon}| - \frac{\tilde{\epsilon}}{\sqrt{\tilde{\epsilon}^2+\Delta^2(\epsilon)}}\right)d\tilde{\epsilon} +$$
   $$\int_{-\hbar\omega_{\mathrm{D}}}^{\hbar\omega_{\mathrm{D}}}\int_{-\hbar\omega_{\mathrm{D}}}^{\hbar\omega_{\mathrm{D}}} N(\tilde{\epsilon})N(\tilde{\epsilon}')V\frac{\Delta^2}{4\sqrt{\tilde{\epsilon}^2+\Delta}\sqrt{\tilde{\epsilon}'^2+\Delta}}d\tilde{\epsilon}d\tilde{\epsilon}'$$

   (c) Simplify this further by using $N(\tilde{\epsilon}) \approx N_{\mathrm{F}}$, to arrive at

   $$E_{\mathrm{BCS}} = \int_{-\infty}^{0} N(\tilde{\epsilon})\tilde{\epsilon}d\tilde{\epsilon} - N_{\mathrm{F}}\frac{\Delta^2}{2}.$$

   Hint: use the integrals

   $$\int\frac{1}{\sqrt{x^2+a^2}}dx = \sinh^{-1}(x/a); \qquad \int\frac{1}{\sqrt{x^2+a^2}}dx = \frac{x}{2}\sqrt{x^2+a^2} - \frac{a^2}{2}\sinh^{-1}(x/a).$$

2. In the lectures, we derived the BCS ground state

   $$\left|\Psi_{\mathrm{BCS}}^{(0)}\right\rangle = \prod_k\left(u_k + v_k a_k^\dagger a_{-k}^\dagger\right)|0\rangle.$$

   We also defined the Bogoliubov-Valatin operators

   $$\alpha_k = u_k a_k - v_k a_{-k}^\dagger$$
   $$\beta_k = u_k a_{-k} + v_k a_k^\dagger.$$

   (a) Show that $\alpha_k\left|\Psi_{\mathrm{BCS}}^{(0)}\right\rangle = 0$ for any k. Similarly, show $\beta_k\left|\Psi_{\mathrm{BCS}}^{(0)}\right\rangle = 0$ for any k.

   (b) We also showed that $\alpha_k^\dagger\left|\Psi_{\mathrm{BCS}}^{(0)}\right\rangle$ is an excited state of the BCS Hamiltonian. Evaluate that state explicitly and explain what it represents. Explain in particular what has happened to a Cooper pair by creating this excitation.

   (c) Repeat this calculation for $\beta_k^\dagger\left|\Psi_{\mathrm{BCS}}^{(0)}\right\rangle$.

3. *Hubbard model and superconductivity*
   The Hubbard model in two dimensions with negative U serves as a simple model for a superconductor. The Hamiltonian reads

$$H = -t \sum_{\langle nm\rangle,\sigma} c^\dagger_{n\sigma} c_{m\sigma} + U \sum_n c^\dagger_{n\uparrow} c_{n\uparrow} c^\dagger_{n\downarrow} c_{n\downarrow}$$

   with $U < 0$. The first sum in the Hamiltonian is over nearest neighbour lattice sites $\langle n, m\rangle$.

   (a) Derive the Hamiltonian in momentum space. Use the notation $k \equiv (\mathbf{k}, \uparrow)$ and $-k \equiv (-\mathbf{k}, \downarrow)$. Decouple the interaction term in a mean field approximation by assuming that only the expectation values $\left\langle c^\dagger_k c^\dagger_{-k} \right\rangle$ and $\langle c_{-k} c_k \rangle$ are nonzero, while other products of two fermion operators vanish.
   We introduce

$$\Delta = -\frac{U}{V} \sum_k \langle c_{-k} c_k \rangle .$$

   Derive the resulting mean-field Hamiltonian H_{BCS}.

   (b) Express H_{BCS} in terms of new operators α_k, β_k which are defined by the Bogoliubov transformation

$$c_k = u_{\mathbf{k}} \alpha_k + v_{\mathbf{k}} \beta^\dagger_k; \tag{11.11}$$

$$c^\dagger_{-k} = -v^*_{\mathbf{k}} \alpha_k + u^*_{\mathbf{k}} \beta^\dagger_k. \tag{11.12}$$

   (c) Show that, for α and β to be fermion annihilation operators, we must have $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$.

   (d) Derive the condition on $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ such that H_{BCS} becomes diagonal in the operators α_k and β_k.

   (e) Derive the energy spectrum of the quasi-particles described by the operators α_k and β_k and their hermitian conjugates.

4. *Quasiparticle excitations in a superconductor*
   In the lecture, we derived the BCS ground state

$$\left|\Psi^{(0)}_{\mathrm{BCS}}\right\rangle = \prod_k \left(u_k + v_k a^\dagger_k a^\dagger_{-k}\right) |0\rangle$$

   and showed that the Bogoliubov-Valatin operators

$$\alpha_k = u_k a_k - v_k a^\dagger_{-k}$$
$$\beta_k = u_k a_{-k} + v_k a^\dagger_k$$

   satisfy $\alpha_k \left|\Psi^{(0)}_{\mathrm{BCS}}\right\rangle = 0$ for any k.

   (a) Similarly, show $\beta_k \left|\Psi^{(0)}_{\mathrm{BCS}}\right\rangle = 0$ for any k.

   (b) We also showed that $\alpha^\dagger_k \left|\Psi^{(0)}_{\mathrm{BCS}}\right\rangle$ is an excited state of the BCS Hamiltonian. Evaluate that state explicitly and explain what it represents. Explain in particular what has happened to a Cooper pair by creating this excitation.

   (c) Repeat this calculation for $\beta^\dagger_k \left|\Psi^{(0)}_{\mathrm{BCS}}\right\rangle$.

5. Anderson's pseudospin formulation of the BCS theory starts by transforming from the pair creation operators $b^\dagger_{\mathbf{k}} = a^\dagger_k a^\dagger_{-k}$ to operators defined by $2s_x(\mathbf{k}) = b^\dagger_{\mathbf{k}} + b_{\mathbf{k}}; 2s_y(\mathbf{k}) = \mathrm{i}(b^\dagger_{\mathbf{k}} - b_{\mathbf{k}}); 2s_z(\mathbf{k}) = \mathbf{1} - \mathbf{n}_{\mathbf{k}} - \mathbf{n}_{-\mathbf{k}}$. Here we use the notation of Desai's book (and of the lecture notes) where k denotes $(\mathbf{k}, \uparrow)$ and $-k$ is $(-\mathbf{k}, \downarrow)$. Verify that in the units where $\hbar \equiv 1$ these operators obey the commutation relations of spins as defined by:

$$[s_x, s_y] = \mathrm{i}s_z; \qquad [s_y, s_z] = \mathrm{i}s_x; \qquad [s_z, s_x] = \mathrm{i}s_y$$

or, more concisely,

$$\boldsymbol{\sigma} \times \boldsymbol{\sigma} = \mathrm{i}\boldsymbol{\sigma}.$$

6. Verify that when the transformation of the previous problem is substituted in the BCS Hamiltonian:

$$H_{\mathrm{BCS}} = \sum_k \epsilon_{\mathbf{k}} \left(a^\dagger_k a_k + a^\dagger_{-k} a_{-k} \right) - \frac{1}{2} \sum_{k,l} G_{kl} a^\dagger_k a^\dagger_{-k} a_{-l} a_l$$

one finds a result of the form:

$$H_{\mathrm{BCS}} = -\sum_{\mathbf{k}} \mathbf{H}(\mathbf{k}) \cdot \mathbf{s}(\mathbf{k})$$

where **H** is an "effective pseudomagnetic field" given by

$$\mathbf{H}(\mathbf{k}) = \left(\sum_{\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'} s_x(\mathbf{k}'), \sum_{\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'} s_y(\mathbf{k}'), 2\hat{\epsilon}_{\mathbf{k}} \right)$$

Note that we have put $l = (\mathbf{k}', \sigma)$, G depends only on the relative wave vectors: $G_{kl} = G_{\mathbf{k},\mathbf{k}'}$. [The energy of this system can now be minimized by arguing in analogy with the theory of the domain walls in ferromagnets].

12

DENSITY OPERATORS — QUANTUM INFORMATION THEORY

12.1 INTRODUCTION

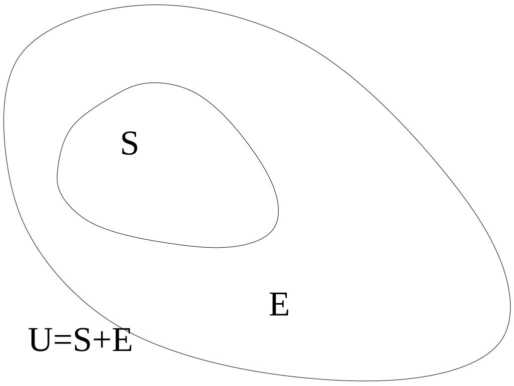
In this chapter, we extend the formalism of quantum physics to include non-isolated systems. Such systems can be described by an object called the *density operator*. Density operators can be used to capture the influence of the outside world on a particle without keeping track of the outside world's degrees of freedom. In the simplest case, the outside world might be just another particle. We shall see that this influence may lead to quantum states that do not have a classical analogue – these states are called *entangled*. Entanglement is used in novel technologies based on quantum mechanics. The most spectacular realization of this trend is quantum computing, which we briefly discuss at the end of this chapter.

12.2 THE DENSITY OPERATOR

Up to this point, we have dealt exclusively with quantum systems that are isolated, i.e., the Hilbert space in which we formulate our quantum system is considered to be all there is – no interactions with other systems are, or have ever been, present. Obviously this scenario is hardly ever met in experiments: interactions with the outside world are unavoidable, whether this is a different system, or degrees of freedom of our experimental apparatus that we would like to neglect. For example, we often treat the electrons in a solid as our quantum system. However, the electrons interact with the nuclei, and these form a quantum system by themselves, and the two are coupled by the electron-phonon interaction.

As we already know, the proper description of the state of an isolated system is the *wave function*. We now ask, *What is the most complete description of a non-isolated system?*

To begin answering this question, let us consider a system S that is coupled to an environment E. The system and environment together form the *universe* U (see figure). Do not take these names too literally – they just represent the roles of the different elements at stake quite well.



The universe U is itself an isolated system – hence it is described by a wave function $\left|\psi^{\rm U}\right\rangle$. We define the orthonormal basis sets $\left|\phi_j^{\rm S}\right\rangle$ in S and $\left|\chi_q^{\rm E}\right\rangle$ in E. A basis of U is then formed by combining these two bases:

$$\left|\phi_j^{\rm S}\right\rangle \otimes \left|\chi_q^{\rm E}\right\rangle . \tag{12.1}$$

The symbol ⊗ denotes a *tensor product*. This should not be considered mysterious or difficult, it simply denotes that we combine the two basis sets into one. In fact, we have done this all the time: when we write $|n, l, m_l, m_s\rangle$ to denote a basis state for an electron in the hydrogen atom, we mean by this the formal expression:

$$|n\rangle \otimes |l m_l\rangle \otimes |m_s\rangle . \tag{12.2}$$

In fact we could also denote the new basis as

$$\left|\phi_j^{\rm S}\chi_q^{\rm E}\right\rangle . \tag{12.3}$$

The reason why we dwell a little on this point is that tensor product notation is regularly used in this field, and it is worth to familiarise yourself with it.

Back to the problem. We seek a means to describe the state of S without including degrees of freedom of E. What is the 'best' description? That is a description which enables us to calculate the expectation value of *any* operator acting solely on S. For example, if we consider S to be the electrons in a solid, we would like to have a description of the electrons enabling us to calculate the expectation value of any physical quantity of the electrons, such as total spin, density, energy, ... The degrees of freedom of the nuclei should not be included in the description.

So, let us take an operator $\hat{A}$ acting on S. This means that the matrix elements $\langle\phi_j|\hat{A}|\phi_k\rangle$ *fully* represent the operator A. The expectation value of the operator $\hat{A}$ is known from the wave function of the universe:

$$\langle A\rangle = \left\langle\psi^{\rm U}\right|\hat{A}\left|\psi^{\rm U}\right\rangle . \tag{12.4}$$

This formal result however includes a state $\left|\psi^{\rm U}\right\rangle$, which lives partly on E and we want to restrict ourselves to S only. To make progress, we expand $\left|\psi^{\rm U}\right\rangle$ in terms of the basis which we have introduced above:

$$\left|\psi^{\rm U}\right\rangle = \sum_{jq} C_{jq}\left|\phi_j^{\rm S}\chi_q^{\rm E}\right\rangle . \tag{12.5}$$

We then obtain for the expectation value:

$$\langle A\rangle = \sum_{jq}\sum_{kr} C_{jq}^* C_{kr}\left\langle\phi_j^{\rm S}\chi_q^{\rm E}\right|\hat{A}\left|\phi_k^{\rm S}\chi_r^{\rm E}\right\rangle . \tag{12.6}$$

Now remember that A only acts on the degrees of freedom of the system S – hence we can write

$$\langle A\rangle = \sum_{jq}\sum_{kr} C_{jq}^* C_{kr}\left\langle\phi_j^{\rm S}\right|\hat{A}\left|\phi_k^{\rm S}\right\rangle\left\langle\chi_q^{\rm E}\middle|\chi_r^{\rm E}\right\rangle . \tag{12.7}$$

The orthonormality of the basis of E turns the last inner product into δ_{qr}, and we have

$$\langle A\rangle = \sum_{jk}\sum_{q} C_{jq}^* C_{kq}\left\langle\phi_j^{\rm S}\right|\hat{A}\left|\phi_k^{\rm S}\right\rangle . \tag{12.8}$$

Note that we have now arrived at a description involving only basis vectors of S. In order to cast this result into a more convenient form, we define

$$\sqrt{p_q}\left|\eta_q^{\rm S}\right\rangle = \sum_j C_{jq}\left|\phi_j^{\rm S}\right\rangle . \tag{12.9}$$

The real and positive factor p_q is chosen so as to normalize $\left|\eta_q^{\mathrm{S}}\right\rangle$: first the right hand side of this definition is worked out and then $\sqrt{p_q}$ is the norm of the result.

Now we can write the expectation value in a convenient form:

$$\langle A\rangle = \sum_q p_q \left\langle \eta_q^{\mathrm{S}} \middle| \hat{A} \middle| \eta_q^{\mathrm{S}} \right\rangle. \tag{12.10}$$

This is our final result. We have succeeded in finding a description of the state of system S in terms of vectors living only in S. Note that this description has a simple statistical interpretation. Suppose we are given a set of normalized states $\left|\eta_q^{\mathrm{S}}\right\rangle$, together with a probability p_q for every such state to occur. After some meditation, it will be clear that the final expression for the expectation value of A is precisely the one given in (12.10).

You may now think that the result we have obtained can be formulated as a superposition state:

$$\left|\phi^{\mathrm{S}}\right\rangle = \sum_q \sqrt{p_q}\left|\eta_q^{\mathrm{S}}\right\rangle. \qquad \text{WRONG!} \tag{12.11}$$

Note that writing out the expectation value of A for such a state would include matrix elements $\left\langle \eta_q^{\mathrm{S}}|\hat{A}|\eta_{q'}^{\mathrm{S}}\right\rangle$ for $q \neq q'$. Such elements do not occur in the correct expression (12.10) obtained above.

> It turns out that there is no description in terms of a *single* quantum state defined on S which gives the correct expectation value of an arbitrary operator $\hat{A}$ acting on S. The system S can only be described by a set (also called *ensemble* of quantum states $\left|\eta_q^{\mathrm{S}}\right\rangle$ with classical probabilities p_q.

This is the main result of this section. We can no longer describe the system S by a single quantum state. Instead, we can describe it as a set of quantum states that occur with a certain classical probability p_q.

To emphasize the difference between a so-called *mixed state*, consisting of an ensemble of states, and a *pure state* characterized by a single wave function, let us look at the simplest possible nontrivial system, described by a two-dimensional Hilbert state, e.g. a spin-1/2 particle. Suppose someone, Charlie, gives us an electron but he does not know its spin state. He does however know that there is no reason for the spin to be preferably up or down, so the probability to measure spin 'up' or 'down' is 1/2 for both. In the language used above, the p_q are 1/2 for both q ='up' and q ='down'. Can we describe this situation by a single quantum state? Well, you might guess that the state is

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right), \tag{12.12}$$

but why couldn't it be

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - |\downarrow\rangle\right)? \tag{12.13}$$

In fact the state of the system could be anything of the form

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + e^{\mathrm{i}\varphi}|\downarrow\rangle\right), \tag{12.14}$$

for any real φ. We see that we cannot assign a unique state to the spins we get from Charlie.

Although we do not know the wave function, we can evaluate the expectation value of the z-component of the spin: as we find $\hbar/2$ and $-\hbar/2$ with equal probabilities, the expectation value is 0. More generally, if we have a spin which is in the spin-up state with probability p and in the down state with probability $1-p$, the expectation value of the z-component of the spin is $\hbar(p-1/2)$. So expectation values can still be found, although we do not have complete information about the state of the system. This raises the question whether we may describe the ensemble of states we get from Charlie by some quantum wave function which would yield the same expectation values for measurements.

To answer this question we introduce the following states:

$$\left|\psi_1\right\rangle = |\uparrow\rangle; \tag{12.15a}$$

$$\left|\psi_2\right\rangle = |\downarrow\rangle; \tag{12.15b}$$

$$\left|\psi_3\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right); \tag{12.15c}$$

$$\left|\psi_4\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - |\downarrow\rangle\right); \tag{12.15d}$$

$$\left|\psi_5\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + \mathrm{i}|\downarrow\rangle\right); \tag{12.15e}$$

$$\left|\psi_6\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - \mathrm{i}|\downarrow\rangle\right). \tag{12.15f}$$

These states are recognized as the spin-up and -down states for the z, x and y directions. States $\left|\psi_3\right\rangle$ to $\left|\psi_6\right\rangle$ all give spin-up and -down with probabilities 1/2 along the z-axis.

Let us consider the most general quantum state which gives spin-up and -down along z with equal probabilities. This is the state $\left(|\uparrow\rangle + e^{\mathrm{i}\varphi}|\downarrow\rangle\right)/\sqrt{2}$. We now calculate the probability of finding in a measurement this particle in the state $\left|\psi_3\right\rangle$:

$$\left|\frac{1}{2}\left(\langle\uparrow| + \langle\downarrow|\right)\left(|\uparrow\rangle + e^{\mathrm{i}\varphi}|\downarrow\rangle\right)\right|^2 = \left|\frac{1+\exp(\mathrm{i}\varphi)}{2}\right|^2 = \frac{1}{2}\left(1+\cos\varphi\right). \tag{12.16}$$

If we evaluate the probability to find the particle in the state $\left|\psi_3\right\rangle$ in the case it was, before the measurement, in a so-called *mixed state* which is given with equal probabilities to be $|\uparrow\rangle$ and $|\downarrow\rangle$, we find 1/2, as can easily be verified. Calculating the probabilities for a particle to be found in the states $\left|\psi_1\right\rangle$ to $\left|\psi_6\right\rangle$ we find the following results.

State	$\left(\lvert\uparrow\rangle + e^{\mathrm{i}\varphi}\lvert\downarrow\rangle\right)/\sqrt{2}$	Equal mixture of $\lvert\uparrow\rangle$ and $\lvert\downarrow\rangle$
$\lvert\psi_1\rangle$	1/2	1/2
$\lvert\psi_2\rangle$	1/2	1/2
$\lvert\psi_3\rangle$	$1/2(1+\cos\varphi)$	1/2
$\lvert\psi_4\rangle$	$1/2(1-\cos\varphi)$	1/2
$\lvert\psi_5\rangle$	$1/2(1-\sin\varphi)$	1/2
$\lvert\psi_6\rangle$	$1/2(1+\sin\varphi)$	1/2

We see that there is no φ, i.e. no pure state, which leads to the same probabilities for *all* possible measurements.

Let us summarize what we have learned:

> A system can be either in a *pure* or *mixed* state. In the first case, the state of the system is completely described by a wave function. In the second case, we are not sure about the state, but we can ascribe a classical probability for the system to be in any state within a set. This situation can occur if our system is coupled to another system (although this does not exclude the possibility of having a pure state in that case). A mixed state cannot be represented by a single wave function.

Note that the uncertainty about the state of the particle is *classical.* Charlie can, for example, flip a coin and, depending on whether the result is head or tails, send us a spin-up or -down. We only know that with probability 1/2, the particle that we receive is in quantum state 'up', and similarly for the quantum state 'down'. This classical uncertainty in the quantum state of the spin should not be confused with the quantum uncertainty inherent to measurement outcomes.

We now turn to the general case of a non-isolated system that can be in either one of a set of normalized, but not necessarily orthogonal, states $|\psi_i\rangle$ (we leave out the superscript S in the sequel and replace the subscript q by i). The probability for the system to be in the state $|\psi_i\rangle$ is p_i, with obviously $\sum_i p_i = 1$. Suppose the expectation value of some operator $\hat{A}$ in state $|\psi_i\rangle$ is given by A_i. Then the expectation value of $\hat{A}$ for the system at hand is given by

$$\langle A\rangle = \sum_i p_i A_i = \sum_i p_i \langle\psi_i|\hat{A}|\psi_i\rangle. \tag{12.17}$$

We now introduce the *density operator*, which is in some sense the 'optimal' specification of the system. The density operator is defined as

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{12.18}$$

Suppose the set $|\phi_n\rangle$ forms a basis of the Hilbert space of the system under consideration. Then the expectation value of the operator $\hat{A}$ can be rewritten after inserting the identity operator $1 = \sum_n |\phi_n\rangle\langle\phi_n|$ as

$$\langle A\rangle = \sum_i p_i \langle\psi_i|\hat{A}|\psi_i\rangle = \sum_i p_i \langle\psi_i|\sum_n |\phi_n\rangle\langle\phi_n|\hat{A}|\psi_i\rangle =$$
$$\sum_n \langle\phi_n|\left[\sum_i p_i |\psi_i\rangle\langle\psi_i|\right]\hat{A}|\phi_n\rangle = \sum_n \langle\phi_n|\hat{\rho}|\phi_n\rangle = \mathrm{Tr}\,\left(\hat{\rho}\hat{A}\right). \tag{12.19}$$

Here we have used the *trace*, Tr , which sums all diagonal terms of an operator. For a general operator $\hat{Q}$:

$$\mathrm{Tr}\,\hat{Q} = \sum_n \langle\phi_n|\hat{Q}|\phi_n\rangle. \tag{12.20}$$

The trace is independent of the basis used — it is invariant under a basis transformation. We omit the hat from operators unless confusion may arise. Another property of the trace is

$$\mathrm{Tr}\,\left(|\psi\rangle\langle\chi|\right) = \langle\chi|\psi\rangle, \tag{12.21}$$

which is easily verified by writing out the trace with respect to a basis ϕ_n.

If a system is in a well-defined quantum state $|\psi\rangle$, we say that the system is in a *pure state*. In that case the density operator is

$$\rho = |\psi\rangle\langle\psi|. \tag{12.22}$$

If the system is not in a pure state, but if only the statistical weights p_i of the states $|\psi_i\rangle$ are known, we say that the system is in a *mixed state*. How can you assess if a system described by a given density operator is in a pure or mixed state? For a pure state we have $\rho^2 = \rho$, which means that ρ is a projection operator[1]:

$$\rho^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho, \tag{12.23}$$

where we have used the fact that $|\psi\rangle$ is normalized.

For a mixed state, such as

$$\rho = \alpha|\psi\rangle\langle\psi| + \beta|\phi\rangle\langle\phi|, \tag{12.24}$$

where $\langle\psi|\phi\rangle = 0$, we have

$$\rho^2 = \alpha^2|\psi\rangle\langle\psi| + \beta^2|\phi\rangle\langle\phi| \neq \rho. \tag{12.25}$$

---

[1] Recall that a projection operator P is an Hermitian operator satisfying $P^2 = P$.

Although we have considered a particular example here, it holds in general that, for a mixed state, ρ is not a projection operator.

Another way to see this is to look at the eigenvalues of ρ. For a pure state, $\rho = |\psi\rangle\langle\psi|$. Clearly, $|\psi\rangle$ is an eigenstate of ρ with eigenvalue 1, and all other eigenvalues are 0 (you can verify this by using that all other eigenstates are orthogonal to $|\psi\rangle$). These values for the eigenvalues are the only ones allowed by a projection operator. As

$$\mathrm{Tr}\rho = \sum_i p_i = 1, \tag{12.26}$$

we have

$$\sum_i \lambda_i = 1. \tag{12.27}$$

Now let us evaluate

$$\langle\phi|\,\rho\,|\phi\rangle = \sum_i p_i \left|\langle\psi_i|\phi\rangle\right|^2 \leq 1, \tag{12.28}$$

where the fact that $\left|\langle\psi_i|\phi\rangle\right| \leq 1$, combined with $\sum_i p_i = 1$ leads to the inequality. The condition $\sum_i \lambda_i = 1$ means that either one of the eigenvalues is 1 and the rest are 0, or they are all strictly less than 1. Thus, for an eigenstate ϕ of the density operator, we have

$$\langle\phi|\,\rho\,|\phi\rangle = \langle\phi|\lambda|\phi\rangle = \lambda < 1. \tag{12.29}$$

We see that a density operator has eigenvalues between 0 and 1.

For a pure state we know that $\hat{\rho}^2 = \hat{\rho}$. In view of $\mathrm{Tr}\,\hat{\rho} = 1$ we therefore have for a pure state that $\mathrm{Tr}\,\hat{\rho}^2 = 1$. If the state is not pure, $\mathrm{Tr}\,\hat{\rho}^2 < 1$. The quantity $\mathrm{Tr}\,\hat{\rho}^2$ is therefore called the *purity* of the density matrix $\hat{\rho}$.

In summary:

The sum of the eigenvalues of the density operator is 1. The special case where only one of these eigenvalues is 1 and the rest are 0 corresponds to a *pure state*.
If there are eigenvalues $0 < \lambda < 1$, then we are dealing with a mixed state.
The *purity* of a state can be quantified as $\mathrm{Tr}\,\rho^2$. This quantity lies is smaller than 1. The value 1 is reached for a pure state.

To summarize, if a system is in a mixed state, it can be characterized by a distribution of possible wave functions $|\psi_i\rangle$ with associated classical probabilities p_i. But a more compact way of representing our knowledge of the system is the *density operator*, which can be constructed when we know the possible states ψ_i and their probabilities p_i [see Eq. (12.18)]. The density operator can be used to calculate expectation values using the trace, see Eq. (12.19).

Let us consider an example. Take again the scenario where Charlie sends us a spin-up or -down particle with equal probabilities. For convenience, we denote these two states as $|0\rangle$ (spin up) and $|1\rangle$ (spin down). Then the density operator can be evaluated as

$$\rho = \frac{1}{2}\,|0\rangle\langle 0| + \frac{1}{2}\,|1\rangle\langle 1|. \tag{12.30}$$

This operator works in a two-dimensional Hilbert space – therefore it can be represented as a 2×2 matrix:

$$\rho = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}. \tag{12.31}$$

The matrix elements are evaluated as follows. The upper-left element is

$$\langle 0|\,\rho\,|0\rangle = \frac{1}{2}\langle 0|0\rangle\langle 0|0\rangle + \frac{1}{2}\langle 0|1\rangle\langle 1|0\rangle = 1/2 \tag{12.32}$$

as follows from (12.30) and from the orthogonality of the two basis states. The upper-right element is given by

$$\langle 0|\,\rho\,|1\rangle = \frac{1}{2}\,\langle 0|0\rangle\,\langle 0|1\rangle + \frac{1}{2}\,\langle 0|1\rangle\,\langle 1|1\rangle = 0 \tag{12.33}$$

as a result of orthogonality. The lower left element $\langle 1|\rho|0\rangle$ and the lower right $\langle 1|\rho|1\rangle$ are found similarly. Another interesting way to find the density matrix (i.e. the matrix representation of the density operator) is by directly using the vector representation of the states $|0\rangle$ and $|1\rangle$:

$$\rho = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0) + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}(0,1) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}. \tag{12.34}$$

Note the somewhat unusual order in which we encounter column and row vectors: the result is not a number, but an operator.

Another day, Charlie decides to send us particles which are either "up" or "down" along the x-axis. As you might remember, the eigenstates are

$$\frac{1}{\sqrt{2}}\,(|0\rangle + |1\rangle) \tag{12.35}$$

for spin-up (along x) and

$$\frac{1}{\sqrt{2}}\,(|0\rangle - |1\rangle) \tag{12.36}$$

for spin-down. You recognize these states as the states $|\psi_3\rangle$ and $|\psi_4\rangle$ given above. Now let us work out the density operator:

$$\rho = \frac{1}{4}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(1,1) + \frac{1}{4}\begin{pmatrix} 1 \\ -1 \end{pmatrix}(1,-1) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}. \tag{12.37}$$

We see that we obtain the *same* density matrix! Apparently, the particular axis used by Charlie does not affect what we measure at our end.

Another question we frequently ask ourselves when dealing with quantum systems is: What is the probability to find the system in a state $|\phi\rangle$ in a measurement?
The answer for a system in a pure state $|\psi\rangle$ is:

$$P_\phi = |\langle\phi\,|\,\psi\rangle|^2. \tag{12.38}$$

If the system can be in either one of a set of states $|\psi_i\rangle$ with respective probabilities p_i, the answer is

$$P_\phi = \sum_i p_i\,|\langle\phi\,|\,\psi_i\rangle|^2. \tag{12.39}$$

Another way to obtain the expression on the right hand side is by using the density operator:

$$\langle\phi\,|\,\rho\,|\,\phi\rangle = \sum_i p_i\,|\langle\phi\,|\,\psi_i\rangle|^2 = P_\phi. \tag{12.40}$$

12

This equation follows directly from the definition of the density operator.

Important examples of systems in a mixed state are statistical systems connected to a heat bath. Loosely speaking, the actual state of the system without the bath varies with time, and we do not know that state when we perform a measurement. We know however from statistical physics that the probability for the system to be in a state with energy E is given by the Boltzmann factor $\exp[-E/(k_B T)]$, so the density operator can be written as

$$\rho = \mathcal{N}\sum_i |\psi_i\rangle\, e^{-E_i/(k_B T)}\,\langle\psi_i| \tag{12.41}$$

where the ψ_i are eigenstates of the Hamiltonian. The prefactor $\mathcal{N}$ is adjusted such that $\mathcal{N}\sum e^{-E_i/(k_\mathrm{B}T)} = 1$ in order to guarantee that Tr $\rho = 1$. The density operator can also be written as

$$\rho = \mathcal{N} e^{-\hat{H}/(k_\mathrm{B}T)}, \tag{12.42}$$

as can be verified as follows:

$$e^{-\hat{H}/(k_\mathrm{B}T)} = \sum_i \left|\psi_i\right\rangle \left\langle\psi_i\right| e^{-\hat{H}/(k_\mathrm{B}T)} \sum_j |\psi_j\rangle\langle\psi_j| = \sum_i \left|\psi_i\right\rangle e^{-\hat{E}_i/(k_\mathrm{B}T)} \left\langle\psi_i\right|. \tag{12.43}$$

Any expectation value can now in principle be evaluated. For example, consider a spin-1/2 particle connected to a heat bath of temperature T in a magnetic field **B** pointing in the z-direction. The Hamiltonian is given by

$$H = -\gamma B S_z. \tag{12.44}$$

Then the expectation value of the z-component of the spin can be calculated as

$$\langle S_z\rangle = \mathrm{Tr}\,(\rho S_z). \tag{12.45}$$

We can evaluate ρ. Using the notation $\beta = 1/(k_\mathrm{B}T)$ it reads:

$$\rho = \frac{1}{e^{\beta\gamma\hbar B/2} + e^{-\beta\gamma\hbar B/2}} \begin{pmatrix} e^{\beta\gamma\hbar B/2} & 0 \\ 0 & e^{-\beta\gamma\hbar B/2} \end{pmatrix}. \tag{12.46}$$

Now the expectation value $\langle S_z\rangle$ can immediately be found, using $S_z = \hbar\sigma_z/2$, where σ_z is the Pauli matrix:

$$\langle S_z\rangle = \mathrm{Tr}\,(\rho S_z) = \hbar/2 \tanh(\beta\gamma\hbar)B/2). \tag{12.47}$$

Considering systems of *non-interacting* particles, the density operator can be used to derive the average occupation of energy levels, leading to the well-known Fermi-Dirac distribution for fermions, and the Bose-Einstein distribution for bosons. This derivation is however beyond the scope of this course — it is treated in your statistical mechanics course.

To conclude this section, we return to the systems S, E and U (U is S and E taken together). Recalling the form of the ensemble states

$$\sqrt{p_q}\left|\eta_q^\mathrm{S}\right\rangle = \sum_j C_{jq}\left|\phi_j^\mathrm{S}\right\rangle, \tag{12.48}$$

we can formulate the density matrix on S as

$$\rho^\mathrm{S} = \sum_q p_q \left|\eta_q^\mathrm{S}\right\rangle\left\langle\eta_q^\mathrm{S}\right| = \sum_q\sum_{jk} C_{jq}C_{kr}^* \left|\phi_j^\mathrm{S}\right\rangle\left\langle\phi_k^\mathrm{S}\right|. \tag{12.49}$$

Now that we have some experience with traces, it is interesting to note that the density matrix of S can be obtained by taking the *partial trace* of the density matrix of U over the degrees of freedom of E. By this we mean the following. Remember $\left|\chi_q^\mathrm{E}\right\rangle$ is a basis of E. The density matrix of U is given by

$$\rho^\mathrm{U} = |\psi^\mathrm{U}\rangle\langle\psi^\mathrm{U}|. \tag{12.50}$$

Taking the partial trace of this over E is defined as

$$\mathrm{Tr}_\mathrm{E}\left(\rho^\mathrm{U}\right) = \sum_q \left\langle \chi_q^\mathrm{E}|\rho^\mathrm{U}|\chi_q^\mathrm{E}\right\rangle. \tag{12.51}$$

Note that the resulting object is an operator acting in S. Also note the similarity with the expression for the full trace: the difference is that we restrict ourselves to the basis states of E.

Writing ρ^{U} in terms of ψ^{U} and expanding the latter as

$$\left|\psi^{\mathrm{U}}\right\rangle = \sum_{jq} C_{jq} \left|\phi_j^{\mathrm{S}} \chi_q^{\mathrm{E}}\right\rangle, \tag{12.52}$$

and using the orthonormality of the basis on E, we obtain

$$\mathrm{Tr}_{\mathrm{E}}\left(\rho^{\mathrm{U}}\right) = \sum_q C_{jq} C_{kq}^* \left|\phi_j^{\mathrm{S}}\right\rangle \left\langle \phi_k^{\mathrm{S}}\right|. \tag{12.53}$$

We see that the resulting expression is precisely the density matrix of S:

$$\rho^{\mathrm{S}} = \mathrm{Tr}_{\mathrm{E}}\left(\rho^{\mathrm{U}}\right). \tag{12.54}$$

In the next section we come back to the use of partial traces.

Summary

We have seen that systems coupled to an environment cannot be described by a single quantum state from the Hilbert space of that system. Instead, the system is described by the *density operator* of *density matrix* $\hat{\rho}$. The density matrix has the form

$$\hat{\rho} = \sum_j p_j \left|\psi_j\right\rangle \left\langle \psi_j\right|.$$

The numbers p_j are the probabilities for finding the system in the state $\left|\psi_j\right\rangle$. They therefore add up to 1. If one $p_j = 1$ and the remaining ones are 0, the state is called *pure*. A measure for the purity of a state is called Tr $\hat{\rho}^2$.

Having $\hat{\rho}$, we can calculate statistical properties of measurements. In particular, for a general Hermitian operator $\hat{A}$, the expectation value for the measurement of that operator in a system described by a density matrix $\hat{\rho}$ is given by

$$\left\langle \hat{A} \right\rangle = \mathrm{Tr}\, \hat{\rho} \hat{A}.$$

Furthermore, the probability to find the system in a particular state $\left|\psi\right\rangle$ is given by

$$P_\psi = \left\langle \psi \middle| \hat{\rho} \middle| \psi \right\rangle.$$

A density matrix satisfies several mathematical requirements:

- It is Hermitian: $\hat{\rho} = \hat{\rho}^\dagger$.
- It is positive definite, i.e. $\left\langle \psi \middle| \hat{\rho} \middle| \psi \right\rangle \geq 0$ for all $\left|\psi\right\rangle$. This is equivalent to saying that all its eigenvalues are non-negative.
- The trace is unity: Tr $\rho = 1$.

For two coupled systems, one of the systems can be described by a *reduced density matrix*:

$$\hat{\rho}_{\mathrm{A}} = \mathrm{Tr}_{\,\mathrm{B}} \hat{\rho}$$

where $\hat{\rho}$ is the density matrix of the combined system A+B.

12.3 ENTANGLEMENT

Entanglement is a phenomenon which can occur when two or more quantum systems are coupled. We have seen in the previous section that coupling between a system S and its environment E may lead to a mixed state of S which is impossible to characterize by a single wave function. If the influence of the environment precludes the possibility of describing S by a pure state, we say that S and E are *entangled*:

Consider a system consisting of two or more subsystems. If the state of the system is such that the subsystems cannot be described by a pure state, then the subsystems are called *entangled.*

We shall focus on the simplest nontrivial system exhibiting entanglement: two particles, A and B, whose degrees of freedom span a two-dimensional Hilbert space (as usual, you may think of two spin-1/2 particles). The basis states for each particle are denoted $|0\rangle$ and $|1\rangle$. Therefore, the possible states of the two-particle system are linear combinations of the states

$$|00\rangle, \qquad |01\rangle, \qquad |10\rangle, \qquad \text{and} \qquad |11\rangle. \tag{12.55}$$

(the first number denotes the state of particle A and the second one that of particle B). We use these states (in this order) as a basis of the four-dimensional Hilbert space, that is, we may identify

$$|00\rangle \Leftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{12.56}$$

and so on.

Suppose the combined system is in the state

$$|\psi\rangle = \frac{1}{2}\left(|00\rangle + |01\rangle + |10\rangle + |11\rangle\right). \tag{12.57}$$

In vector notation, this state is represented as:

$$\psi = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \tag{12.58}$$

Note that this state is normalized.

We can find the density matrix of particle A by tracing out the degrees of freedom of particle B. This is the procedure that we followed at the very end of Section 12.2, and we can copy the result of that section simply by using $C_{jq} = 1/2$ for our example $(j, q$ both assume values 0 or 1). The result is then

$$\rho^{\mathrm{S}} = \frac{1}{2}\left(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|\right). \tag{12.59}$$

This can also be written as

$$\rho^{\mathrm{S}} = |\psi_3\rangle\langle\psi_3|, \tag{12.60}$$

with

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right). \tag{12.61}$$

We see that ρ^{S} is the density operator of a pure state – hence the two particles are not entangled.

Now we want to study the same problem from a different viewpoint. Suppose we perform measurements of the *first* spin only. More specifically, we measure the probabilities for a system to be in the states

$$|\psi_1\rangle = |0\rangle, \qquad |\psi_2\rangle = |1\rangle, \qquad |\psi_3\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \qquad \text{or} \qquad |\psi_4\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right). \tag{12.62}$$

The resulting probabilities are (check this!):

$$P_1 = P_2 = 1/2; \tag{12.63a}$$

$$P_3 = 1; \quad P_4 = 0. \tag{12.63b}$$

These are precisely the same results as those for a *single* particle in the state $\left|\psi_3\right\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, that is, if we want to predict measurements on the first particle, we can forget about the second particle. The reason for this is that we can write the state (12.57) as

$$\frac{1}{2}\left(|0\rangle_A + |1\rangle_A\right) \otimes \left(|0\rangle_B + |1\rangle_B\right). \tag{12.64}$$

The fact that (12.57) can be written as a (tensor) product of pure states of the two subsystems A and system B is responsible for the fact that the second particle does not ‘interfere’ with the first one.

We see that entanglement can be defined in three ways, which can be shown to be completely equivalent:

- Two systems A and B are entangled when the density matrix of one of them describes a *mixed state*. This is usually not the easiest definition to be used in problems!
- Two systems A and B are entangled when it is impossible to write the state as a single product of a state on A and a state on B. This definition usually is the easiest to check whether a state is entangled or not.
- Two systems A and B are entangled when the outcome of some measurement on one of them influences the probability of measurements on the other (see below).

In order to use the first definition, the use of partial traces is necessary. It is convenient to learn how a partial trace works out on a matrix representation of the density matrix. Let us first work out the density matrix of the combined system AB in the previous example:

$$\rho^{\mathrm{AB}} = \left|\psi\right\rangle\left\langle\psi\right| = \frac{1}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}(1,1,1,1) = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{12.65}$$

Taking the trace over the states of B corresponds to taking four traces over the submatrices formed by the elements in which the system A has the same values. Some inspection leads to the submatrices indicated in the following equation:

$$\mathrm{Tr}_{\mathrm{B}}\left(\begin{array}{cc|cc} a & b & p & q \\ c & d & r & s \\ \hline \alpha & \beta & \eta & \zeta \\ \gamma & \delta & \chi & \xi \end{array}\right) = \begin{pmatrix} a+d & p+s \\ \alpha+\delta & \eta+\xi \end{pmatrix}. \tag{12.66}$$

We see that we should view the full, 4 × 4 density matrix as consisting of four 2 × 2 submatrices. Of each of these submatrices, we take the trace and put this as a number in the resulting 2 × 2 matrix.

In the case where we want to take the partial trace over A rather than over B, we proceed as follows:

$$\mathrm{Tr}_{\mathrm{A}}\begin{pmatrix} a & b & p & q \\ c & d & r & s \\ \alpha & \beta & \eta & \zeta \\ \gamma & \delta & \chi & \xi \end{pmatrix} = \begin{pmatrix} a+\eta & b+\zeta \\ c+\chi & d+\xi \end{pmatrix}. \tag{12.67}$$

We see that we divide the large matrix again into four submatrices, but these submatrices have their elements two rows and/or columns apart rather than one, like in the previous case. Whether we take the trace over A or over B, the ‘reduced’ density matrix reads

$$\rho_{\mathrm{A}} = \rho_{\mathrm{B}} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{12.68}$$

This matrix has two eigenvectors:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{12.69}$$

with eigenvalues 1 and 0, respectively. This implies that the so-called *reduced density matrix* that is obtained after taking a partial trace over the full density matrix, is that of a pure state (see the previous section). Therefore, we conclude, again, that the state of the combined system AB is not and entangled state.

We now consider another example, defined by the state

$$\left|\psi_{\mathrm{E}}\right\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right). \tag{12.70}$$

In vector form, this state reads

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{12.71}$$

The density operator is (in matrix form):

$$\rho = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}(1001) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{12.72}$$

Taking the trace over either A or B gives the same result:

$$\rho_{\mathrm{A}} = \rho_{\mathrm{B}} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{12.73}$$

We immediately see two eigenvalues, 1/2 and 1/2, and we conclude that the reduced density matrix is not that of a pure state, so the state of AB we started from was entangled (see the previous section).

Now we analyze the entanglement of the state according to the other criteria discussed above. First, it is easy to see that there is no way to write the state as product state of two states of system A and B respectively. This is as usual the easiest way to determine whether the state is entangled or not (in this case, entangled).

Now consider the effect of measuring an aspect of one subsystem. If we measure for the first spin the value 0, then a measurement of the second spin will also give 0. The same holds for measuring 1 for both. We see that the measurement of one particle, influences the measurement results of the other – hence the state is entangled.

12

Let us now describe this more formally. We perform measurements on particle A and on B, checking whether these particles are found in state 1 or 0. For our entangled state (12.70) we find

$$P_{00} = P_{11} = 1/2, \tag{12.74a}$$

$$P_{10} = P_{01} = 0, \tag{12.74b}$$

where P_{01} is the probability to find particle A in state 0 and particle B in state 1 etcetera. We see that in terms of classical probabilities, the system A is strongly correlated with system B. It turns out that this correlation remains complete even when the measurement is performed with respect to another basis (see exercises):

> *Entanglement gives rise to correlation of probabilities, and this correlation cannot be lifted by a basis transformation.*

An interesting question is whether a non-entangled system may become entangled in the course of time. We therefore take a system which is in a non-entangled state — it might for example be in the state (12.57). We assume that the system evolves according to a Hamiltonian H which, in the basis $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, has the following form:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{12.75}$$

The time evolution operator is given by $T = \exp(-it\hat{H}/\hbar)$ — at $t = \pi\hbar/2$ it has the form

$$\begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix}, \tag{12.76}$$

so that we find

$$\left|\psi(t = \pi\hbar/2)\right\rangle = -\frac{i}{2}\left(|00\rangle + |01\rangle + |10\rangle - |11\rangle\right), \tag{12.77}$$

which is an entangled state (you will find no way to write it as a tensor product of two pure states of A and B). Thus we see that when a system starts in a non-entangled state, it might evolve into an entangled state in the course of time.

We have seen that non-entangled states lead to a reduced density matrix describing a pure state. The lowest value for the purity $\mathrm{Tr}\,\hat{\rho}^2$ of the reduced density matrix is obtained when the density matrix has two eigenvalues $1/2$:

$$\hat{\rho}_{\mathrm{A}} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

The two-qubit state leading to this density matrix is therefore called a 'most entangled' state. For two qubits, the space of most entangled states is spanned by the so-called *Bell states*:

$$\begin{aligned} \left|\psi_1\right\rangle &= \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right); \\ \left|\psi_2\right\rangle &= \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right); \\ \left|\psi_3\right\rangle &= \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right); \\ \left|\psi_4\right\rangle &= \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right). \end{aligned}$$

These states are generally referred to as *Bell states.*

12.4 THE EPR PARADOX AND BELL'S THEOREM

In 1935, Einstein, Podolsky and Rosen (EPR) published a thought experiment, which demonstrated that quantum mechanics is not compatible with some obvious ideas which we tacitly apply when describing phenomena. In particular the notions of a reality existing independently of experimental measurements and of locality cannot both be reconciled with quantum mechanics. Locality is used here to denote the idea that events cannot have an effect at a distance before information has travelled from that event to another place where its effect is noticed. Together, the notions of reality and locality are commonly denoted as 'local realism'. From the failure of quantum mechanics to comply with local realism, EPR concluded that quantum mechanics is not a complete theory.

The EPR paradox is quite simple to explain. At some point in space, a stationary particle with spin 0 decays into two spin-1/2 particles which fly off in opposite directions (momentum conservation does not allow the directions not to be opposite). During the decay process, angular momentum is conserved which implies that the two particles must have opposite spin: when one particle is found to have spin 'up' along some measuring axis, the other particle must have spin 'down' along the same axis. Obviously, we are dealing with an entangled state.

Suppose Alice and Bob both receive an outcoming particle from the same decay event. Alice measures the spin of the particle along the z direction, and Bob does the same with his particle. Superficially, we can say that they would both have the same probability to find either $\hbar)/2$ or $-\hbar)/2$. However, if quantum mechanics is correct, these measurements should be strongly correlated: if Alice has measured spin up, then Bob's particle must have spin down along the z-axis, so the measurement results are fully correlated. According to the 'orthodox', or 'Copenhagen' interpretation of quantum mechanics, if Alice is the first one to measure the spin, the particular value measured by her is decided at the very moment of that measurement. But this means that at the *at the same moment* the spin state of Bob's particle is determined. But Bob could be lightyears away from Alice, and perform his measurement immediately after her. According to the orthodox interpretation, his measurement would be influenced by Alice's. But this was inconceivable to Einstein, who maintained that the information about Alice's measurement could not reach the Bob's particle instantaneously, as the speed of light is a limiting factor for communication. In Einstein's view, the outcome of the measurements of the particles is determined at the moment when they leave the source, and he believed that a more complete theory could be found which would unveil the 'hidden variables' which determine the outcomes of Alice and Bob's measurements when the particles left the source. These hidden variables would then represent some "reality" which exists irrespectively of the measurement.

The EPR puzzle remained unsettled for a long time, until, in 1965, John Bell formulated a theorem which would allow to distinguish between Einstein's scenario and the orthodox quantum mechanical interpretation. We shall now derive Bell's theorem. Suppose we count in an audience the numbers of people having certain properties, such as 'red hair' or 'wearing yellow socks', 'taller than 1.70 m'. We take three such properties, called A, B and C. If we select one person from the audience, he or she will either comply to each of these properties or not. We denote this by a person being 'in the state' A^+, B^-, C^+ for example. The number of people in the state A^+, B^-, C^+ is denoted $N(A^+, B^-, C^+)$. We now write

$$N(A^+, B^-) = N(A^+, B^-, C^+) + N(A^+, B^-, C^-) \tag{12.78}$$

which is a rather obvious relation.

We use similar relations in order to rewrite this as

$$N(A^+, B^-) = N(A^+, C^-) - N(A^+, B^+, C^-) + N(B^-, C^+) - N(A^-, B^-, C^+) \le N(A^+, C^-) + N(B^-, C^+). \tag{12.79}$$

This is Bell's inequality, which can also be formulated in terms of probabilities $[P(A^+, B^-)$ instead of $N(A^+, B^-)$ etcetera]. We have used everyday-life examples in order to emphasise that there is nothing mysterious, let alone quantum mechanical, about Bell's inequality. But let us now turn to quantum mechanics, and spin determination in particular.

Consider the three axes **a**, **b** and **c** shown in the figure. A^+ is now identified with a spin-up measurement along **a** etcetera. We can now evaluate $P(A^+, C^-)$. Measuring A^+ happens with probability 1/2, but after this measurement, the particle is in the spin-up state along the **a**-axis. If the spin is then measured along the **c** direction, we have a probability $\sin^2 \pi/8$ to find C^- (see problem 16 of the exercises). The combined probability is $P(A^+, C^-)$ is therefore $\frac{1}{2}\sin^2(\pi/8)$ Similarly, $P(B^-, C^+)$ is also equal to $\frac{1}{2}\sin^2(\pi/8)$, and $P(A^+, B^-)$ is 1/4. Inserting

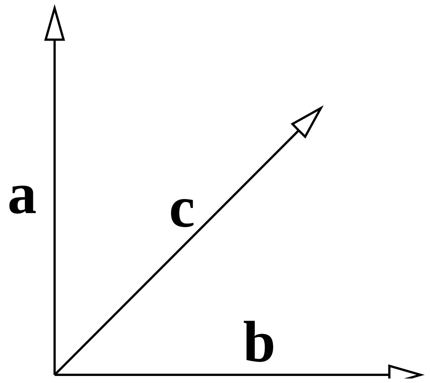


FIGURE 12.1: The measuring axis for a spin.

these numbers into Bell's inequality gives:

$$\frac{1}{4} \leq \sin^2(\pi/8) = \frac{1}{2}\left(1 - \frac{1}{2}\sqrt{2}\right), \tag{12.80}$$

which is obviously wrong. Therefore, we see that quantum mechanics does not obey Bell's inequality.

Now what does this have to do with the EPR paradox? Well, first of all, the EPR paradox allows us to measure the spin in two different directions at virtually the same moment. But, more importantly, if the particles would leave the origin with predefined probabilities, Bell's inequality would unambiguously hold. The only way to violate Bell's inequality is by accepting that Alice's measurement reduces the entangled wave function of the two-particle system, which is also noticed by Bob *instantaneously*. So, there is some 'action at a distance', in contrast to what we usually have in physics, where every action is mediated by particles such as photons, mesons, ...

In 1982, Aspect, Dalibard and Roger performed experiments with photons emerging from decaying atoms in order to check whether Bell's theorem holds or not. Since then, several other groups have redone this experiment, sometimes with different setups. The conclusion is now generally accepted that Bell's theorem does not hold for quantum mechanical probabilities. The implications of this conclusion for our view of Nature is enormous: somehow actions can be performed without intermediary particles, so that the speed of light is not a limiting factor for this kind of communication. 'Communication' is however a dangerous term to use in this context, as it suggests that information can be transmitted instantaneously. However, the 'information' which is transmitted from Alice to Bob or vice versa is purely probabilistic, since Bob nor Alice can predict the outcome of their measurements. So far, no schemes have been invented or realised which would allow us to send over a Mozart symphony at speeds faster than the speed light.

12.5 NO CLONING THEOREM

In recent years, much interest has arisen in quantum information processing. In this field, people try to exploit quantum mechanics in order to process information in a way completely different from classical methods. We have already encountered one example of these attempts: quantum cryptography, where a random encryption key can be shared between Bob and Alice without Eve being capable of eavesdropping. Another very important application, which unfortunately is still far from a realisation, is the quantum computer. When I speak of a quantum computer, you should not forget that I mean a machine which exists only on paper, not in reality. A quantum computer is a quantum machine in which *qubits* evolve in time. A qubit is a quantum system with a 2-dimensional Hilbert space. It can always be denoted

$$\left|\varphi\right\rangle = a|0\rangle + b|1\rangle, \tag{12.81}$$

where a and b are complex constants satisfying $a^2 + b^2 = 1$. The states $|0\rangle$ and $|1\rangle$ form a basis in the Hilbert space. A quantum computer manipulates several qubits in parallel. A system consisting of n qubits has a 2^n-dimensional Hilbert space. A quantum computation consists of a preparation of the qubits in some well-defined state, followed by an autonomous evolution of the qubit system, and concluded by reading out the state of the qubits. As the system is autonomous, it is described by a (Hermitian) Hamiltonian. The time-evolution operator $U = \exp(-itH/\hbar))$ is then a unitary operator, so the the quantum computation between initialisation and reading out the results can be described in terms of a sequence of unitary transformations applied to the system. In this section we shall derive a general theorem for such an evolution, the no-cloning theorem:

> *An unknown quantum state cannot be cloned.*

By cloning we mean that we can copy the state of some quantum system into some other system without losing the state of our original system.

Before proceeding with the proof of this theorem, let us assume that cloning would be possible. In that case, communication at speeds faster than light would in principle be possible. To see this, imagine Alice and Bob have a pair of entangled qubits. Alice performs a measurement on her qubit along the axis $|0\rangle$ or $|0\rangle + |1\rangle$. After this, Bob makes many clones of his qubit. As Bob has many clones, he can find out which measurement Alice performed without ambiguity (how?). So the no-cloning theorem is essential in making communication at speeds faster than the speed of light impossible.

The proof of the no-cloning theorem for qubit systems proceeds as follows. Cloning for a qubit pair means that we have a unitary evolution U with the following effect on a qubit pair:

$$U|\alpha 0\rangle = |\alpha\alpha\rangle. \tag{12.82}$$

The evolution U should work for any state α, therefore it cannot depend on α. Therefore, for some other state $\left|\beta\right\rangle$ we must have

$$U\left|\beta 0\right\rangle = \left|\beta\beta\right\rangle. \tag{12.83}$$

Now let us operate with U on the state $\left|\gamma 0\right\rangle$ with $\left|\gamma\right\rangle = \left(|\alpha\rangle + \left|\beta\right\rangle\right)/\sqrt{2}$:

$$U\left|\gamma 0\right\rangle = \left(|\alpha\alpha\rangle + \left|\beta\beta\right\rangle\right)/\sqrt{2} \neq \left|\gamma\gamma\right\rangle, \tag{12.84}$$

which completes the proof.

12.6 DENSE CODING

In this section, I describe a way of sending over more information than bits. This sounds completely impossible, but, again, quantum mechanics is *in principle* able to realise the impossible. It is however difficult to implement, as it is based on Bob and Alice having an entangled pair of qubits, in the state

$$|00\rangle + |11\rangle. \tag{12.85}$$

From now on, we shall adopt the convention in this field to omit normalisation factors in front of the wave functions. We can imagine this state to be realised by having an entangled pair generator midway between Alice and Bob, sending entangled particles in opposite directions as in the EPR setup.

Note that the following qubit operations are all unitary:

$$I\left|\phi\right\rangle = \left|\phi\right\rangle \tag{12.86a}$$
$$X|0\rangle = |1\rangle\,, \tag{12.86b}$$
$$X|1\rangle = |0\rangle \tag{12.86c}$$
$$Z|0\rangle = |0\rangle\,, \tag{12.86d}$$
$$Z|1\rangle = -|1\rangle\,. \tag{12.86e}$$
$$Y|0\rangle = |1\rangle \tag{12.86f}$$
$$Y|1\rangle = -|0\rangle \;\; Y = XZ\,. \tag{12.86g}$$

The operator I is the identity; X is called the NOT operator, We assume that Alice has a device with which she can perform any of the four transformations (I, X, Y, Z) on her member (i.e. the first) of the entangled qubit pair. The resulting perpendicular states for these four transformations are:

$$I\left(|00\rangle + |11\rangle\right) = \left(|00\rangle + |11\rangle\right) \tag{12.87a}$$
$$X\left(|00\rangle + |11\rangle\right) = \left(|10\rangle + |01\rangle\right) \tag{12.87b}$$
$$Y\left(|00\rangle + |11\rangle\right) = \left(|10\rangle - |01\rangle\right) \tag{12.87c}$$
$$Z\left(|00\rangle + |11\rangle\right) = \left(|00\rangle - |11\rangle\right) \tag{12.87d}$$

Alice does not perform any measurement — she performs one of these four transformations and then she sends her bit to Bob. Bob then measures in which of the four possible states the entangled pair is, in other words, he now knows which transformation Alice applied. This information is 'worth' two bits, but Alice had to send only one bit to Bob!

12.7 QUANTUM COMPUTING AND SHOR'S FACTORISATION ALGORITHM

A quantum computer is a device containing one or more sets of qubits (called *registers*), which can be initialised without ambiguity, and which can evolve in a controlled way under the influence of unitary transformations and which can be measured after completion of this evolution.

The most general single-qubit transformation is a four-parameter family. For more than one qubit, it can be shown that every nontrivial unitary transformation can be generated by a single-qubit transformation of the form

$$U(\theta,\phi) = \begin{pmatrix} \cos(\theta/2) & -ie^{-i\phi}\sin(\theta/2) \\ -ie^{i\phi}\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}. \tag{12.88}$$

and another unitary transformations involving more than a single qubit, the so-called 2-qubit XOR. This transformation acts on a qubit pair and has the following effect:

$$\text{XOR}\left(|00\rangle\right) = |00\rangle \tag{12.89a}$$
$$\text{XOR}\left(|01\rangle\right) = |01\rangle \tag{12.89b}$$
$$\text{XOR}\left(|10\rangle\right) = |11\rangle \tag{12.89c}$$
$$\text{XOR}\left(|11\rangle\right) = |10\rangle \tag{12.89d}$$

We see that the first qubit is left unchanged and the second one is the eXclusive OR of the two input bits. Unitary transformations are realised by hardware elements called *gates*.

Several proposals for building quantum computers exist. In the ion trap, an array of ions which can be in either the ground state ($|0\rangle$) or the excited state ($|1\rangle$), controlled by laser pulses. Coupling of neighbouring ions in order to realise an XOR-gate is realised through a controlled momentum transfer to displacement excitations (phonons) of the chain.

Here in Delft, activities focus on arrays of Josephson junctions. Josephson junctions are very thin layers of ordinary conductors separating two superconductors. Current can flow through these junctions in either the clockwise or anti-clockwise direction (interpreted as 0 and 1 respectively). Other initiatives include NMR devices and optical cavities. With this technique it has become possible recently to factorise the number 15. Realisation of a working quantum computer will take at least a few decades — if it will come at all.

A major problem in realising a working quantum computer is to ensure a unitary evolution. In practice, the system will always be coupled to the outside world. Quantum computing hinges upon the possiblity to have controlled, *coherent* superpositions. Coherent superpositions are linear combinations of quantum states into another, pure state. As we have seen in the previous section, coupling to the environment may lead to entanglement which would cause the quantum computer to be described by a density operator rather than by a pure state. In particular, any phase relation between constitutive parts of a phase-coherent superposition is destroyed by coupling to the environment. We shall now treat this phenomenon in more detail.

Consider a qubit which interacts with its environment. We denote the state of the environment by the ket $|m\rangle$. The interaction is described by the following prescription:

$$|0\rangle\,|m\rangle \to |0\rangle\,|m_0\rangle\,; \tag{12.90a}$$

$$|1\rangle\,|m\rangle \to |1\rangle\,|m_1\rangle\,. \tag{12.90b}$$

In this interaction, the qubit itself does not change — if this would be the case, our computer would be useless to start with.

12

Suppose we start with a state

$$|0\rangle + e^{i\phi}\,|1\rangle \tag{12.91}$$

which is coupled to the environment. This coupling will induce the transition

$$\left(|0\rangle + e^{i\phi}\,|1\rangle\right)|m\rangle \to |0\rangle\,|m_0\rangle + e^{i\phi}\,|1\rangle\,|m_1\rangle\,. \tag{12.92}$$

Suppose this qubit is then fed into a so-called Hademard gate, which has the effect

$$H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right); \tag{12.93a}$$

$$H|1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right). \tag{12.93b}$$

Then the outcome is

$$e^{i\phi/2}\left[|0\rangle\left(e^{-i\phi/2}\,|m_0\rangle + e^{i\phi/2}\,|m_1\rangle\right) + |1\rangle\left(e^{-i\phi/2}\,|m_0\rangle - e^{i\phi/2}\,|m_1\rangle\right)\right]. \tag{12.94}$$

If we suppose that $\langle m_0|m_1\rangle$ is real, we find for the probabilities to measure the qubit in the state $|0\rangle$ or $|1\rangle$ (after normalisation):

$$P_0 = \frac{1}{2}\left(1 + \langle m_0|m_1\rangle\cos\phi\right) \tag{12.95a}$$

$$P_1 = \frac{1}{2}\left(1 - \langle m_0|m_1\rangle\cos\phi\right) \tag{12.95b}$$

If there is no coupling, $m_0 = m_1 = m$, and we recognise the phase relation between the two states in the probabilities. On the other hand, if $\langle m_0|m_1\rangle = 0$, then we find for both probabilities 1/2, and the phase relation has disappeared completely.

It is interesting to construct a density operator for the qubit in the final state (12.94). Consider a qubit

$$\alpha\,|0\rangle| + \beta\,|1\rangle \tag{12.96}$$

which has interacted with its environment, so that we have the combined state

$$\alpha\,|0\rangle\,|m_0\rangle + \beta\,|1\rangle\,|m_1\rangle\,. \tag{12.97}$$

We can arrive at a density operator for the qubit only by performing the trace over the m-system only. Using (12.21) we find

$$\rho_{\text{qubit}} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^*\,\langle m_1|m_0\rangle \\ \alpha^*\beta\,\langle m_0|m_1\rangle & |\beta|^2 \end{pmatrix}. \tag{12.98}$$

The eigenvalues of this matrix are

$$\lambda = \frac{1}{2} \pm \frac{1}{2}\sqrt{\left(|\alpha|^2 - |\beta|^2\right)^2 + 4|\alpha|^2|\beta|^2|\,\langle m_0|m_1\rangle\,|^2} \tag{12.99}$$

and these lie between 0 and 1, where the value 1 is reached only for $\langle m_0|m_1\rangle = 1$. The terms coherence/decoherence derive from the name coherence which is often used for the matrix element $\langle m_0|m_1\rangle$.

Now let us return to the very process of quantum computing itself. The most impressive algorithm, which was developed in 1994 by Peter Shor, is that of factorising large integers, an important problem in the field of encryption and code-breaking. We shall not describe this algorithm in detail, but present a brief sketch of an important sub-step, finding the period of an integer function f. It is assumed here that all unitary transformations used can be realised with a limited number of gates.

The algorithm works with two registers, both containing n qubits. These registers are described by a 2^n-dimensional Hilbert space. As basis states we use the bit-sequences of the integers between 0 and 2^{n-1}. The basis state corresponding to such an integer x is denoted $|x\rangle_n$. Now we perform the Hademard gate (12.93) to all bits of the state $|0\rangle_n$. This yields

$$H\,|0\rangle_n \equiv |w\rangle_n = 2^{-n}\sum_{x=0}^{2^n-1} |x\rangle_n\,. \tag{12.100}$$

It is possible (but we shall not describe the method here) to construct, for any function f which maps the set of numbers 0 to 2^{n-1} onto itself, a unitary transformation U_f which has the effect

$$U_f\,|x\rangle_n\,|0\rangle_n = |x\rangle_n\,\big|f(x)\big\rangle_n \tag{12.101}$$

using a limited number of gates.

Now we are ready for the big trick in quantum computing. If we let U_f act on the state $|w\rangle_n$ then we obtain

$$U_f\,|w\rangle_n\,|0\rangle_n = 2^{-n}\sum_{x=0}^{2^n-1} |x\rangle_n\,\big|f(x)\big\rangle_n\,. \tag{12.102}$$

We see that the new state contains $f(x)$ *for all possible values of* x. In other words, applying the gates U_f to our state $|w\rangle_n|0\rangle_n$, we have evaluated the function f for 2^n different arguments. This feature is called *quantum parallelism* and it is this feature which is responsible for the (theoretical) performance of quantum computing.

Of course, if we were to read out the results of the computation for each x-value, we would have not gained much, as this would take 2^n operations. In general, however, the final result that we are after consists of only few data, so a useful problem does not consist of simply calculating f for all of its possible arguments. As an example we consider the problem of finding the periodicity of the function f, which is an important step in Shor's algorithm. This is done by reading out only one particular value of the result in the second register, $f(x) = u$, say. The first register is then the sum of all x-states for which it holds that $f(x) = u$. If f has a period r, we will find that these x-values lie a distance r apart from each other. Now we act with a (unitary) Fourier transform operator on this register, and the result will be a linear combination of the registers corresponding to the period(s) of the function f. If there is only one period, we can read this out straightforwardly.

It has been said already that finding the period of some function is an important step in the factorising algorithm. Shor's algorithm is able to factorise an n-bit integer in about $300n^3$ steps. A very rough estimate of size for the number to be factorized where a quantum computer starts outperforming a classical machine, is about 10^{130}.

12.8 PROBLEMS

1. For the operators given below, determine whether they are admissible as density matrices and, if yes, whether they describe a pure or a mixed state. Also, provide the density operators in matrix form (i.e., the density matrices). The two states $|\psi_1\rangle$ and $|\psi_2\rangle$ are linearly independent and normalised.

   (a) $\rho = \frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\psi_2\rangle\langle\psi_2|$,

   (b) $\rho = \frac{1}{2}\left[|\psi_1\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_2| + |\psi_2\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|\right]$,

   (c) $\rho = \frac{1}{2}\left[|\psi_1\rangle\langle\psi_1| + \mathrm{i}|\psi_1\rangle\langle\psi_2| - \mathrm{i}|\psi_2\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|\right]$,

   (d) $\rho = \frac{1}{2}\left[|\psi_1\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_2| - |\psi_2\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|\right]$.

2. Determine which of the following states are entangled:

   (a) $\frac{|00\rangle+|01\rangle}{\sqrt{2}}$,

   (b) $\frac{-|01\rangle+|11\rangle}{\sqrt{2}}$,

   (c) $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$,

   (d) $\frac{i|01\rangle+|00\rangle}{\sqrt{2}}$,

   (e) $\frac{|00\rangle-\mathrm{i}|01\rangle+|10\rangle-\mathrm{i}|11\rangle}{2}$,

   (f) $\frac{|00\rangle-|01\rangle+i|10\rangle-\mathrm{i}|11\rangle}{2}$,

   (g) $\frac{|000\rangle-|100\rangle+|001\rangle-|101\rangle}{2}$,

   (h) $\frac{|010\rangle+|110\rangle+|011\rangle-|111\rangle}{2}$.

3. We consider a two-qubit system prepared in the quantum state

$$|\psi(t=0)\rangle = \frac{1}{2}\left(|00\rangle + |01\rangle + |10\rangle + |11\rangle\right).$$

   We assume that this system evolves with Hamiltonian $\hat{H}$, which in the basis

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\},$$

   takes the following form:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

   (a) Is the state $|\psi(t=0)\rangle$ entangled?

   (b) Show that at $t = \pi\hbar/2$, the wave function $|\psi(t)\rangle$ is not entangled.

   (c) Does this hold for every time t? If not, can you invent a (non-trivial) Hamiltonian for which this *would* hold (change a few signs in the Hamiltonian).

   (d) Show that if the Hamiltonian can be written as the sum of a Hamiltonian acting only on the first particle, and one only acting on the second particle, the particles will never become entangled.

4. Alice and Bob have two qubits, which together are in the quantum state

$$\frac{1}{\sqrt{2}}\left(\alpha\,|00\rangle + \beta\,|01\rangle + \alpha\,|11\rangle - \beta\,|10\rangle\right),$$

where $|\alpha|^2 + |\beta|^2 = 1$. The first qubit in the ket-vector is always that of Alice.

(a) Give the density matrix for this state in the form of a 4×4 matrix. Clearly indicate which basis you choose.

(b) What is the probability for Bob to find the value 0 when he measures his qubit?

(c) Calculate the reduced density matrix for Bob in the form of a 2×2 matrix, by taking the trace over Alice's Hilbert space. Answer part (b) again based on the result.

(d) Charlie prepares this state many times for Alice and Bob. Each time, they measure their qubits in the measurement basis $|0\rangle, |1\rangle$. Finally they get together and compare their results. Can they find out the state of the system by combining their measurements? Can they characterise the state after their measurement by a density matrix? Give that density matrix.

5. Alice possesses a qubit in the quantum state

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle,$$

with $|a|^2 + |b|^2 = 1$. Bob and Charlie know this state too – that is, they know a and b.

(a) Bob measures this qubit. What are the probabities for finding $|0\rangle$ and $|1\rangle$ respectively?

(b) Following his measurement, Bob sends the qubit to Charlie, without communicating the measurement result. How would Charlie characterize the state of the qubit (i.e. is it a mixed or a pure state)? Give this state.

(c) Now both Alice and Bob have a qubit. That of Alice is in the state $|\psi\rangle$. Bob's qubit is in a different state given as

$$|\phi\rangle = c\,|0\rangle + d\,|1\rangle.$$

These qubits both traverse a controlled-NOT gate. This unitary operation has the following effect:

$$\begin{aligned} 00 &\to 00 \\ 01 &\to 01 \\ 10 &\to 11 \\ 11 &\to 10 \end{aligned}$$

After this operation, Alice's qubit is measured with respect to the basis $|0\rangle$, $|1\rangle$. Give the possible two-qubit states after this measurement.

(d) Now the qubits of Bob and of Alice are in the state

$$\frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right),$$

where the first qubit is that of Alice, and the second that of Bob.
What are now the possible states after the controlled-NOT operation, followed by a measurement of Alice's qubit? Charlie does not know the result of the measurement. Give the state for Charlie.

(e) Take the trace over Alice's Hilbert space of the density operator found in (d) to obtain the reduced density operator for Bob's qubit. Is Bob's qubit in a pure or a mixed state? Same questions for Alice's qubit after tracing over Bob's Hilbert space.

6. In this problem we analyse Bell's inequality in a different form as presented in the lecture. Alice and Bob receive particles which can each be described by a two-dimensional Hilbert space with basis $|0\rangle$ and $|1\rangle$. Alice and Bob perform measurements on these particles. Alice can perform measurements for her particle corresponding to the Pauli matrices σ_z, σ_x etc. or linear combinations of these. Bob can do the same. If Alice finds in a measurement of σ_z the value 1 (corresponding to the state $|0\rangle$), then Bob finds in a similar measurement for his particle a value -1 (corresponding to $|1\rangle$) and vice-versa. Both possibilities are equally probable.

(a) Give the most general wave function $|\psi\rangle$ with which you can describe the state of the two-particles. Also give the state when the wave function is antisymmetric under particle exchange (fermion behaviour). We assume that this antisymmetry requirement is satisfied.

Now Alice and Bob perform measurements for physical quantities corresponding to the operators σ_z, $\sigma_\theta = \cos\theta\sigma_z + \sin\theta\sigma_x$ and $\sigma_{-\theta} = \cos\theta\sigma_z - \sin\theta\sigma_x$. Alice measures either σ_z or σ_θ. Bob measures either σ_z or $\sigma_{-\theta}$. This choice is made at random with equal probabilities. In all cases they find either a value 1 or -1. We denote the measurement results by S_z^A, S_z^B, S_θ^A and $S_{-\theta}^B$. The upper index A or B indicates whether we are dealing with Alice or Bob, and the lower index indicates whether it is σ_z which is measured or $\sigma_{\pm\theta}$.

In this problem, we consider the operator

$$\hat{g} = \sigma_z^A\sigma_z^B + \sigma_\theta^A\sigma_z^B + \sigma_z^A\sigma_{-\theta}^B - \sigma_\theta^A\sigma_{-\theta}^B.$$

(b) Show that the expectation values $\langle\sigma_z^A\rangle$, $\langle\sigma_z^B\rangle$, $\langle\sigma_\theta^A\rangle$ and $\langle\sigma_{-\theta}^B\rangle$ all give the value 0.

(c) Give the matrices of the operators $\sigma_z^A\sigma_z^B$, $\sigma_z^A\sigma_x^B$, $\sigma_x^A\sigma_z^B$ en $\sigma_x^A\sigma_x^B$ with respect to the basis

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}.$$

(d) Calculate $\langle\hat{g}\rangle$, as a function of θ.

(e) Sketch $\langle\hat{g}\rangle$ as a function of θ and show that it is larger than 2 (in absolute value) for $0 < \theta < \pi/2$.

(f) Suggest a procedure for measuring $\langle\hat{g}\rangle$ in an experiment where both particles are only measured once.

Now we try to 'invent' a classical probabilistic process, which generates possible measurement outcomes S_z^A, $S_{-\theta}^B$, etcetera, and which would reproduce the values found in the experiment.

We therefore consider the number

$$g = S_z^A S_z^B + S_\theta^A S_z^B + S_z^A S_{-\theta}^B - S_\theta^A S_{-\theta}^B.$$

(g) Argue why the average of the numbers S_z^A etcetera must be 0.

(h) Show that for all possibilities for the four values S_z^A, $S_{-\theta}^B$ etcetera, $g = \pm 2$. Hint: show that the last term in the expression for g is equal to the product of the first three terms.

From the fact that $g = \pm 2$ for each pair measured by Alice and Bob it follows that:

$$|\overline{g}| \leq 2,$$

where $\overline{g}$ is the average of g is taken over many measurements. Verify this. Note that this is always satisfied for any probabilities for the combinations of the four numbers S_z^A, S_z^B, $S_{-\theta}^B$, $S_{-\theta}^A$.

(i) How does the result for $\langle g \rangle$ change if the phase relation between the two terms in the wavefunction changes? (Only the last term in the expression for g changes).

7. Consider a density matrix ρ_S of a system with a two-dimensional Hilbert space.

   (a) Show that $1/2 \le \mathrm{Tr}\left(\rho_\mathrm{S}^2\right) \le 1$.
   What is $\mathrm{Tr}\left(\rho_\mathrm{S}^2\right)$ for a pure state?

   Now consider a universe U=S+E in which the system S and environment E are each a 2-dimensional Hilbert space (i.e., they are both qubits). The state of the universe can be written as the wave function

   $$\left|\psi_\mathrm{U}\right\rangle = \alpha_{00}\left|00\right\rangle + \alpha_{01}\left|01\right\rangle + \alpha_{10}\left|10\right\rangle + \alpha_{11}\left|11\right\rangle.$$

   We would like to quantify the degree of entanglement between S and E using the *concurrence, C*:

   $$C(\left|\psi_\mathrm{U}\right\rangle) = \sqrt{2\left(1-\mathrm{Tr}\left(\rho_\mathrm{S}^2\right)\right)}.$$

   (b) Show that C is given in terms of the coefficients α_{ij} by

   $$C = 2\left|\alpha_{00}\alpha_{11} - \alpha_{10}\alpha_{01}\right|.$$

   (c) Show that $0 \le C \le 1$. What can you say about $\left|\psi_\mathrm{U}\right\rangle$ when $C=0$?

   (d) Find the concurrence for the following wave functions:
   - $\frac{1}{2}\left(\left|00\right\rangle + \left|01\right\rangle - \left|10\right\rangle - \left|11\right\rangle\right)$,
   - $\frac{1}{2}\left(\left|00\right\rangle - \left|01\right\rangle + \left|10\right\rangle + \left|11\right\rangle\right)$,
   - $\frac{1}{\sqrt{3}}\left(\left|00\right\rangle + \left|01\right\rangle + \left|11\right\rangle\right)$,
   - $\frac{1}{\sqrt{2}}\left(\left|01\right\rangle + \left|10\right\rangle\right)$.

8. The fidelity $\mathscr{F}$ of a state (pure or mixed, characterized by density operator ρ) to a target state $\left|\psi_\mathrm{target}\right\rangle$ is defined as

   $$\mathscr{F} \equiv \left\langle\psi_\mathrm{target}\right|\rho\left|\psi_\mathrm{target}\right\rangle.$$

   (a) Show that, for the case of a pure state characterized by the wave function $\left|\psi\right\rangle$,

   $$\mathscr{F} = \left|\left\langle\psi|\psi_\mathrm{target}\right\rangle\right|^2.$$

   (b) Consider a system consisting of two qubits. Show that any separable (un-entangled, product) state of the two qubits cannot have a fidelity greater than 50 % to the Bell state $\left|\psi_\mathrm{target}\right\rangle = \frac{1}{\sqrt{2}}\left(\left|01\right\rangle + \left|10\right\rangle\right)$.
   Hint: use the Bloch vector representation of the two spins, with polar angles θ, ϕ and θ', ϕ' respectively.

   (c) Show that the converse is not true. That is, $\mathscr{F} \le 0.5$ to the Bell state does not guarantee that the state $\left|\psi\right\rangle$ is un-entangled. Hint: the easiest way to show this is to give a counterexample.

9. The reduced density operator for one qubit (regardless of the size of the environment) can be expanded in the basis of Pauli operators:

   $$\rho_\mathrm{S} = \frac{v_\mathbb{1}}{2}\mathbb{1} + \frac{v_x}{2}\sigma_x + \frac{v_y}{2}\sigma_y + \frac{v_z}{2}\sigma_z.$$

   (a) Why must the coefficients v_i, v_x, v_y, and v_z be real valued?
   (b) Show $v_{\mathbb{1}} = 1$.
   (c) Express the purity $\mathrm{Tr}\left(\rho_S^2\right)$ in terms of v_x, v_y and v_z.
   (d) Show that $\mathbf{v}^2 = v_x^2 + v_y^2 + v_z^2 \leq 1$, with equality only for pure states.
   (e) Calculate $\langle\sigma_x\rangle$, $\left\langle\sigma_y\right\rangle$, $\langle\sigma_z\rangle$ in terms of the numbers v_x, v_y and v_z.

10. Consider the time-dependent state of problem 4. This described a state of a two-spin system as a function of time.

    Evaluate the trace over the Hilbert space of the *second* spin to obtain the density matrix of the first spin. Then evaluate the expectation value of the z-component of the first spin as a function of time.

13

OPEN QUANTUM SYSTEMS

In this chapter we shall discuss the time evolution of open quantum systems. In the previous chapter, we have encountered open quantum systems – in particular, we have seen that these systems are described by a density matrix rather than by a single wave function. In this chapter we focus on the time evolution of the density matrix.

So far, the time evolution that we have mostly dealt with is that of closed systems, that are fully described by a wave function $\left|\psi\right\rangle$ evolving deterministically according to the time-dependent Schrödinger equation:

$$\left|\dot{\psi}\right\rangle = \frac{1}{\mathrm{i}\hbar} H \left|\psi\right\rangle.$$

Solving this differential equation lets us relate the state at time t to that at time 0 by

$$\left|\psi(t)\right\rangle = U(t)\left|\psi(0)\right\rangle,$$

where $U(t) = e^{-\mathrm{i}Ht/\hbar}$ is the *time-evolution operator* – see chapter 2.

We have also seen that such a closed system can be equivalently described by a density operator, $\rho = \left|\psi\right\rangle\left\langle\psi\right|$. Using the time evolution of the bra- and the ket wave functions occurring in this density operator, we find that it evolves according to

$$\rho(t) = U(t)\rho(0)U^\dagger(t).$$

Taking the time derivative, we obtain the following important differential equation:

$$\mathrm{i}\hbar\dot{\rho} = \left[H, \rho\right].$$

It is interesting to note that this differential equation differs from the usual time evolution for a quantum mechanical operator as it occurs in the Heisenberg picture (see section 2.2):

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{O}(t) = \mathrm{i}\hbar\frac{\partial}{\partial t}U^\dagger(t)\hat{O}U(t) = \left[\hat{O}(t), H\right].$$

We now turn to the evolution of an open quantum system. Using the notation of section 12.2, we label the system with S and the environment with E. The environment may consist of another quantum system (often much larger) interacting with our system. It can also represent a measurement apparatus that may be either under our control or controlled by another party that does not communicate with us. The central question that concerns us is: how does the system's reduced density operator ρ_S evolve in these different scenarios? Is there an analog differential equation, or evolution operator for ρ_S? Apart from giving insight into what would really happen to the quantum systems in our labs, these questions touch upon the measurement process in quantum mechanics.

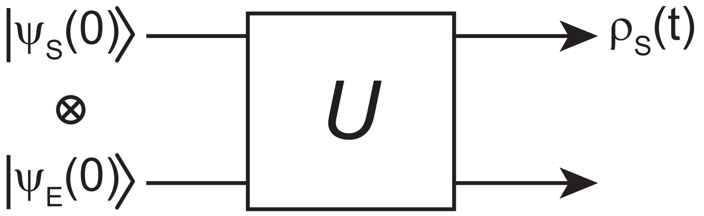


FIGURE 13.1: A quantum system and an environment, initially in a product state $|\psi_\mathrm{U}(0)\rangle = |\psi_\mathrm{S}(0)\rangle \otimes |\psi_\mathrm{E}(0)\rangle$, evolve under a globally unitary operation U. We are not interested in the state of the environment E. How is the reduced density of the system, initially in $\rho_\mathrm{S}(0) = |\psi_\mathrm{S}(0)\rangle\langle\psi_\mathrm{S}(0)|$, transformed by U?

13.1 COUPLING TO AN ENVIRONMENT

Analogously to the approach of section 12.2, let us first consider a closed universe consisting of our system and an environment evolving in a unitary fashion (see Fig. 13.1). At any time, the state of the universe is described by a state $|\psi_\mathrm{U}(t)\rangle$ and, equivalently, by a pure density matrix

$$\rho_\mathrm{U}(t) = |\psi_\mathrm{U}(t)\rangle\langle\psi_\mathrm{U}(t)|.$$

Clearly, $\rho_\mathrm{U}(t) = U(t)\rho_\mathrm{U}(0)U^\dagger(t)$, and the reduced density matrix for the system at any time is given by

$$\rho_\mathrm{S}(t) = \mathrm{Tr}_\mathrm{E}\,[\rho_\mathrm{U}(t)].$$

The answer to how the reduced density matrix of the system transforms under the unitary evolution is simple to write down formally:

$$\rho_\mathrm{S}(0) = \mathrm{Tr}_\mathrm{E}\left[\rho_\mathrm{U}(0)\right] \to \rho_\mathrm{S}(t) = \mathrm{Tr}_\mathrm{E}\left[U(t)\rho_\mathrm{U}(0)U^\dagger(t)\right].$$

But what is the form of the transformation $\mathscr{E}$ that evolves the density matrix in time:

$$\rho_\mathrm{S}(t) = \mathscr{E}(\rho_\mathrm{S}(0))?$$

To begin answering this question, let us look at a simple example. In it, and throughout this chapter, we assume that the system and environment are *initially* in an unentangled, or product state $|\psi_\mathrm{U}(0)\rangle = |\psi_\mathrm{S}(0)\rangle \otimes |\psi_\mathrm{E}(0)\rangle$.

13.1.1 EXAMPLE: THE DAMPING CHANNEL

Imagine that we have a qubit (our system) initially in a pure state $|\psi_\mathrm{S}(0)\rangle = \alpha|0\rangle + \beta|1\rangle$ and an environment initially in $|\psi_\mathrm{E}(0)\rangle$:

$$|\psi_\mathrm{U}(0)\rangle = \big(\alpha|0\rangle + \beta|1\rangle\big)|\psi_\mathrm{E}(0)\rangle.$$

The state $|0\rangle$ denotes the ground state and $|1\rangle$ the excited state. Imagine that for the system in the state $|0\rangle$, the environment evolves over a certain amount of time from the state $|\psi_\mathrm{E}\rangle$ to the (normalized) state $|\psi'_\mathrm{E}\rangle$, which is not necessarily orthonormal to $|\psi_\mathrm{E}\rangle$: For short time evolutions, these states will be almost the same.

$$|0\rangle|\psi_\mathrm{E}\rangle \to |0\rangle|\psi'_\mathrm{E}\rangle.$$

This equation expresses the fact that when the system is in its ground state, the environment will not drive it out of that state.

We allow however for the excited state $|1\rangle$ to decay to the ground state during the evolution. We assume that if there is no interaction, the system stays in $|1\rangle$ and the environment evolves in the same way as for the system in the ground state, that is, it ends up in $\left|\psi'_{\rm E}\right\rangle$. However, the excited state partly evolves into the ground state as a result of interactions with the environment, and in that case, the environment ends up in a state $\left|\psi''_{\rm E}\right\rangle$ which we take perpendicular to $\left|\psi'_{\rm E}\right\rangle$:

$$|1\rangle\otimes\left|\psi_{\rm E}\right\rangle\to\sqrt{1-p}\,|1\rangle\otimes\left|\psi'_{\rm E}\right\rangle+\sqrt{p}\,|0\rangle\otimes\left|\psi''_{\rm E}\right\rangle.$$

All environment states are normalized. The two states after the evolution are however not necessarily orthonormal to the initial environment state.

All in all, this leads to the time evolution of the state $|\psi_{\rm U}(0)\rangle$:

$$|\psi_{\rm U}(0)\rangle=\big(\alpha\,|0\rangle+\beta\,|1\rangle\big)\,|\psi_{\rm E}\rangle\to|\psi_{\rm U}(t)\rangle=\alpha\,|0\rangle\left|\psi'_{\rm E}\right\rangle+\sqrt{1-p}\beta\,|1\rangle\left|\psi'_{\rm E}\right\rangle+\sqrt{p}\beta\,|0\rangle\left|\psi''_{\rm E}\right\rangle.$$

You may check that with this choice of prefactors, the evolution preserves the norm of the state. Under this evolution, there is a probability amplitude $\sqrt{p}$ that an excited qubit will lose its excitation to the environment. The final state state of the universe can be re-expressed as a linear combination of two states with perpendicular environmental components:

$$\begin{aligned}|\psi_{\rm U}(t)\rangle&=\left(\alpha\,|0\rangle+\sqrt{1-p}\beta\,|1\rangle\right)\otimes\left|\psi'_{\rm E}\right\rangle+\sqrt{p}\beta\,|0\rangle\otimes\left|\psi''_{\rm E}\right\rangle\\&=M_0\,|\psi_{\rm S}(0)\rangle\left|\psi'_{\rm E}\right\rangle+M_1\,|\psi_{\rm S}(0)\rangle\left|\psi''_{\rm E}\right\rangle,\end{aligned}$$

where, in the $|0\rangle$, $|1\rangle$ basis, the operators M_0 and M_1 are the matrices

$$M_0=\begin{pmatrix}1&0\\0&\sqrt{1-p}\end{pmatrix}\text{ and }M_1=\begin{pmatrix}0&\sqrt{p}\\0&0\end{pmatrix}.\tag{13.1}$$

These matrices tell us what has happened to the system state if we find the environment in either the state $\left|\psi'_{\rm E}\right\rangle$ or $\left|\psi''_{\rm E}\right\rangle$.

The final reduced density matrix for the system is found by tracing the density matrix of the universe over the environment:

$$\begin{aligned}\rho_{\rm S}(t)&=&&{\rm Tr}_{\rm E}\left(\left|\psi'_{\rm U}\right\rangle\left\langle\psi'_{\rm U}\right|\right)\\&=&&M_0\,|\psi_{\rm S}(0)\rangle\,\langle\psi_{\rm S}(0)|\,M_0^\dagger+M_1\,|\psi_{\rm S}(0)\rangle\,\langle\psi_{\rm S}(0)|\,M_1^\dagger,\end{aligned}$$

where we have made use of the orthogonality between $\left|\psi'_{\rm E}\right\rangle$ and $\left|\psi''_{\rm E}\right\rangle$. Substituting

$$\rho_{\rm S}(0)=|\psi_{\rm S}(0)\rangle\,\langle\psi_{\rm S}(0)|$$

above, we find

$$\rho_{\rm S}(t)=\sum_i M_i\rho_{\rm S}(0)M_i^\dagger.$$

So, once we have found the matrices M_i, this expression gives our sought-after connection between the initial and final reduced density matrices for the system, expressed as a transformation only in the Hilbert space of S. This specific form is known as the *operator-sum representation* of the transformation $\mathscr{E}$. The M_i are called *operation elements.* Note that they act solely on the qubit Hilbert space, and that they are not unitary. You can easily check this for M_0 and M_1 above. We have found the following important result:

> The state of a system that is initially in a pure state, and coupled to an environment, will generally evolve into a mixed state. The evolution of the density matrix is encoded in the so-called operator elements M_i which, for each final state i of the environment, map the initial state of the system to a final state.
> Viewing the M_i as matrices, their dimension is that of the Hilbert space of the system, but we have a number of them equal to the number of possible final environment states.

But how general is the operator-sum representation? Can it describe the evolution of the reduced density matrix for *any* system in interaction with an environment in a closed universe? We shall address this general question in the following section.

Before concluding this example, we note that, for a single qubit, $\mathcal{E}$ is easy to visualize on the Bloch sphere. Using

$$\rho_{\mathrm{S}} = \frac{1}{2}\begin{pmatrix} 1+\langle Z\rangle & \langle X\rangle - \mathrm{i}\langle Y\rangle \\ \langle X\rangle + \mathrm{i}\langle Y\rangle & 1-\langle Z\rangle \end{pmatrix}, \tag{13.2}$$

and we can relate the final and initial Bloch vectors:

$$\begin{pmatrix} \langle X\rangle(t) \\ \langle Y\rangle(t) \\ \langle Z\rangle(t) \end{pmatrix} = \begin{pmatrix} \sqrt{1-p} & 0 & 0 \\ 0 & \sqrt{1-p} & 0 \\ 0 & 0 & 1-p \end{pmatrix} \begin{pmatrix} \langle X\rangle(0) \\ \langle Y\rangle(0) \\ \langle Z\rangle(0) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ p \end{pmatrix}. \tag{13.3}$$

This represents a motion of the point towards $Z=1$ and a compression within the XY plane. Anticipating that p would grow in time, the state will therefore end up in its ground state $Z=1$.

13.1.2 THE OPERATOR-SUM REPRESENTATION

In this section we consider the general case of a unitary system-environment interaction, starting from an unentangled state

$$|\psi_{\mathrm{U}}(0)\rangle = |\psi_{\mathrm{S}}(0)\rangle \otimes |\psi_{\mathrm{E}}(0)\rangle\,.$$

Following the interaction, the universe state becomes

$$|\psi_{\mathrm{U}}(t)\rangle = U(t)\left(|\psi_{\mathrm{S}}(0)\rangle \otimes |\psi_{\mathrm{E}}(0)\rangle\right).$$

We can expand the unitary operator $U(t)$ in terms of a basis of the universe. We use orthonormal basis vectors $\left|j\right\rangle$ for the system S, and $|i\rangle$ for the environment E, so $\left|j,i\right\rangle$ forms an (orthonormal) basis of the universe, and this is the one we shall use:

$$U(t) = \sum_{j}^{N_{\mathrm{S}}}\sum_{i}^{N_{\mathrm{E}}}\sum_{j'}^{N_{\mathrm{S}}}\sum_{i'}^{N_{\mathrm{E}}} U_{jij'i'}(t)\left|j,i\right\rangle\left\langle j',i'\right|.$$

Thus,

$$\begin{aligned}
|\psi_{\mathrm{U}}(t)\rangle &= \sum_{j}^{N_{\mathrm{S}}}\sum_{i}^{N_{\mathrm{E}}}\sum_{j'}^{N_{\mathrm{S}}}\sum_{i'}^{N_{\mathrm{E}}} U_{jij'i'}(t)\left|j,i\right\rangle\left\langle j',i'\right|\left(|\psi_{\mathrm{S}}(0)\rangle \otimes |\psi_{\mathrm{E}}(0)\rangle\right) \\
&= \sum_{j}^{N_{\mathrm{S}}}\sum_{i}^{N_{\mathrm{E}}}\sum_{j'}^{N_{\mathrm{S}}}\sum_{i'}^{N_{\mathrm{E}}} U_{jij'i'}(t)\left\langle i'\right|\psi_{\mathrm{E}}(0)\rangle\left|j\right\rangle\left\langle j'\right|\psi_{\mathrm{S}}(0)\rangle \otimes |i\rangle \\
&= \sum_{i}^{N_{\mathrm{E}}}\left(\sum_{j}^{N_{\mathrm{S}}}\sum_{j'}^{N_{\mathrm{S}}}\sum_{i'}^{N_{\mathrm{E}}} c_{i'}U_{jij'i'}(t)\left|j\right\rangle\left\langle j'\right|\right)|\psi_{\mathrm{S}}(0)\rangle \otimes |i\rangle \\
&= \sum_{i}^{N_{\mathrm{E}}} M_i\,|\psi_{\mathrm{S}}(0)\rangle \otimes |i\rangle\,,
\end{aligned} \tag{13.4}$$

where $c_{i'} = \left\langle i'\right\rangle\psi_{\mathrm{E}}(0)$ and $M_i \equiv \sum_{j}^{N_{\mathrm{S}}}\sum_{j'}^{N_{\mathrm{S}}}\sum_{i'}^{N_{\mathrm{E}}} c_{i'}U_{jij'i'}(t)\left|j\right\rangle\left\langle j'\right|$. We can write M_i succinctly as

$$M_i = \langle i|\,U(t)\,|\psi_{\mathrm{E}}(0)\rangle\,.$$

At this point it is very important to realize that the M_i's are operators acting in the *system* Hilbert space – this may not be obvious from the notation. We can write the reduced density operator for the system following the interaction:

$$\begin{aligned}
\rho_S(t) &= \mathrm{Tr}_E\left(|\psi_U(t)\rangle\langle\psi_U(t)|\right) \\
&= \mathrm{Tr}_E\left(\left(\sum_i^{N_E} M_i|\psi_S(0)\rangle\otimes|i\rangle\right)\left(\sum_{i'}^{N_E} M'_i\langle\psi_S(0)|\otimes\left\langle i'\right|\right)\right) \\
&= \sum_i M_i|\psi_S(0)\rangle\langle\psi_S(0)|M_i^\dagger \\
&= \sum_i M_i\rho_S(0)M_i^\dagger. \qquad (13.5)
\end{aligned}$$

We have arrived, without loss of generality, at an operator-sum representation for the transformation $\rho_S(0)\to\rho_S(t)$ that involves "sandwiching" the initial system reduced density operator by operation elements M_i. The detailed form of these operation elements depends on the initial state of the environment, the global evolution operator U, and our choice of orthonormal basis for the environment. Crucially, however, they are independent of the initial system state $|\psi_S(0)\rangle$. It is important to highlight that the number of nonzero operation elements can equal the dimensionality N_E of the environment. Thus, their number can largely exceed the dimensionality N_S of the system Hilbert space, because most often the environment is much larger than the system. Our second example will illustrate this point. Before moving to it, we note a property of the operation elements M_i arising from the unitarity of the global evolution operator.

$$\begin{aligned}
1 &= \left\langle\psi_U(t)\right\rangle\psi_U(t) \\
&= \left(\sum_i\langle i|\langle\psi_S(0)|M_i^\dagger\right)\left(\sum_{i'}M_{i'}|\psi_S(0)\rangle\left|i'\right\rangle\right) \\
&= \sum_i\langle\psi_S(0)|M_i^\dagger M_i|\psi_S(0)\rangle \\
&= \langle\psi_S(0)|\left(\sum_i M_i^\dagger M_i\right)|\psi_S(0)\rangle.
\end{aligned}$$

As this must hold for any $|\psi_S(0)\rangle$, it follows that

$$\sum_i M_i^\dagger M_i = I_S,$$

where I_S is the identity operator in the system Hilbert space. We say that the operation elements form a *decomposition of unity.*

13.1.3 EXAMPLE: QUBIT DEPOLARIZATION

We consider a qubit-environment interaction that transforms the initially unentangled state $|\psi_U(0)\rangle=|\psi_S(0)\rangle|\psi_E(0)\rangle$ to

$$\begin{aligned}
|\psi_U(t)\rangle = &\sqrt{1-3p/4}\,|\psi_S(0)\rangle\left|\psi'_E\right\rangle+\sqrt{p/4}X|\psi_S(0)\rangle\left|\psi''_E\right\rangle \\
&+\sqrt{p/4}Y|\psi_S(0)\rangle\left|\psi'''_E\right\rangle+\sqrt{p/4}Z|\psi_S(0)\rangle\left|\psi''''_E\right\rangle,
\end{aligned}$$

where the four environment states on the right-hand side are orthonormal. Under this interaction, the qubit undergoes a bit flip (X operation), phase flip (Z operation) and bit-phase flip (Y operation), each with a probability amplitude $\sqrt{p/4}$. The final reduced density matrix for the system is given by

$$\rho_S(t)=\sum_{i=1}^4 M_i\rho_S(0)M_i^\dagger,$$

with operation elements

$$M_1=\sqrt{1-3p/4}I,\ M_2=\sqrt{p/4}X,\ M_3=\sqrt{p/4}Y,\ \text{and}\ M_4=\sqrt{p/4}Z. \qquad (13.6)$$

The dimension 2 of the system's Hilbert space is reflected in the fact that the matrices M_i are 2×2. We have four independent 2×2 matrices in this example. By noting that the square of every Pauli operator is the identity, it is easy to show that these operation elements indeed form a decomposition of unity.

The effect of this process is easy to visualize in the Bloch sphere:

$$\begin{pmatrix} \langle X \rangle (t) \\ \langle Y \rangle (t) \\ \langle Z \rangle (t) \end{pmatrix} = (1-p) \begin{pmatrix} \langle X \rangle (0) \\ \langle Y \rangle (0) \\ \langle Z \rangle (0) \end{pmatrix}.$$

The interaction with the environment reduces the magnitude of the Bloch vector without changing its direction. For $p = 1$, every Bloch vector collapses to the origin. Quite generally, a process under which the purity of the reduced density matrix decreases is said to cause *decoherence* of the system.

13.2 DIRECT QUANTUM MEASUREMENTS

We now turn our attention to the evolution of a system that undergoes measurement. We first focus on *direct quantum measurements* (Fig. 13.2). Despite their perhaps unfamiliar name, these measurements are exactly the type considered in standard introductory courses. Such measurements follow Born's rule: *With every such measurement is asssociated an operator* $\hat{\mathcal{O}}$ *in the system Hilbert space. The act of measurement collapses the system state* $|\psi_S\rangle$ *to one of the eigenstates* $|i\rangle$ *of* $\hat{\mathcal{O}}$*, and the measurement result, which we denote by* m*, is the corresponding eigenvalue* λ_i*. The probability of measuring* λ_i *and collapsing onto eigenstate* $|i\rangle$ *is given by the square of the overlap between* $|\psi_\mathrm{S}\rangle$ *and* $|i\rangle$*:*

$$p_i = |\langle i \rangle \, \psi_\mathrm{S}|^2.$$

13.2.1 SYSTEM EVOLUTION CONDITIONED ON THE RESULT OF DIRECT MEASUREMENT

Suppose we prepare the system in a pure state $|\psi\rangle$. According to the Born rule, the post-measurement state of the system is simply the eigenstate $|i\rangle$ of $\hat{\mathcal{O}}$ corresponding to the measurement result λ_i. Note that $\hat{\mathcal{O}}$ can be written in terms of its eigenvalues and eigenvectors as The evolution under a direct quantum measurement with result $m = \lambda_i$ is

$$\rho_\mathrm{S} = |\psi_\mathrm{S}\rangle \langle \psi_\mathrm{S}| \rightarrow \rho_{\mathrm{S}|m=\lambda_i} = |i\rangle \langle i|.$$

Another, admittedly more complicated, way to write the post-measurement density matrix conditioned on $m = \lambda_i$ is

$$\rho_{\mathrm{S}|m=\lambda_i} = \frac{1}{p_i} \Pi_i \rho_\mathrm{S} \Pi_i, \tag{13.7}$$

where $\Pi_i \equiv |i\rangle \langle i|$ is a projection operator. This last expression puts the transformation in operator-sum representation. In this case, the only operator element is $\Pi_i / \sqrt{p_i}$, which is obviously not a decomposition of unity.

13.2.2 UNCONDITIONED SYSTEM EVOLUTION UNDER DIRECT MEASUREMENT

In the previous section we have considered the evolution of the system when the measurement result is known. We call this the *conditional evolution.* Now we consider the same measurement, but for the case that the result is not known or discarded. For example, this may arise when the system is measured by someone else (a lab partner perhaps) who does not communicate the result to us. We can ask ourselves whether the system density matrix undergoes a nontrivial evolution in this situation. Indeed, this turns out to be the case. We call the evolution of the density matrix in this process *unconditional.*

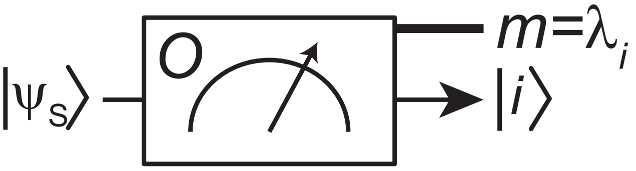


FIGURE 13.2: A direct quantum measurement is associated with an Hermitian operator. The measurement result m is an eigenvalue λ_i of $\hat{\mathcal{O}}$. The post-measurement state of the system is the corresponding eigevenvector $|i\rangle$. The probability of measuring $m = \lambda_i$ is given by $p_i = |\langle i\rangle \psi_S|^2$.

Following the measurement, the system is left in a statistical mixture of the eigenstates of $\hat{\mathcal{O}}$, and is described by the density matrix:

$$\begin{aligned}
\rho'_S &= \sum_i p_i \,|i\rangle \langle i| \\
&= \sum_i |\langle i\rangle \psi_S|^2 \,|i\rangle \langle i| \\
&= \sum_i |i\rangle \langle i\rangle \psi_S \left\langle \psi_S \right\rangle i \,\langle i| \\
&= \sum_i \Pi_i \,|\psi_S\rangle \langle \psi_S| \,\Pi_i \\
&= \sum_i \Pi_i \rho_S \Pi_i .
\end{aligned}$$

In the final line, we have arrived at an expression for ρ'_S in operator-sum representation, where the operation elements M_i are the projectors Π_i onto the eigenstates of $\hat{\mathcal{O}}$. Because these eigenstates form an orthonormal basis for the system Hilbert space, the operation elements clearly form a decomposition of unity.

In summary, we have seen that the evolution of the reduced density matrix of a system that undergoes quantum measurement can be written in operator-sum representation, wherein all operator elements are projection operators.

13.2.3 MEASUREMENT STATISTICS

Before turning to an example, we consider the statistics of the measurement. The expectation value of the measurement is

$$\begin{aligned}
\langle m\rangle &= \sum_i \lambda_i p_i \\
&= \sum_i \lambda_i |\langle i\rangle \psi_S|^2 \\
&= \sum_i \lambda_i \left\langle \psi_S \right\rangle i \,\langle i\rangle \psi_S \\
&= \langle \psi_S| \left(\sum_i \lambda_i \,|i\rangle \langle i| \right) |\psi_S\rangle \\
&= \langle \psi_S| \hat{\mathcal{O}} \,|\psi_S\rangle ,
\end{aligned}$$

a result that should be familiar to us.

13.2.4 EXAMPLE: PROJECTIVE MEASUREMENT OF A QUANTUM BIT

We consider the direct measurement of Z for a qubit. As illustrated in Fig. 13.3, it is easy to visualize the effect of the operation elements $M_+ = \Pi_0 = |0\rangle \langle 0|$ and $M_- = \Pi_1 = |1\rangle \langle 1|$ on the Bloch vector. For a measurement result $m = +1$, the Bloch vector is collapsed onto the north pole. For $m = -1$, it is collapsed onto the south pole. When we ignore the measurement

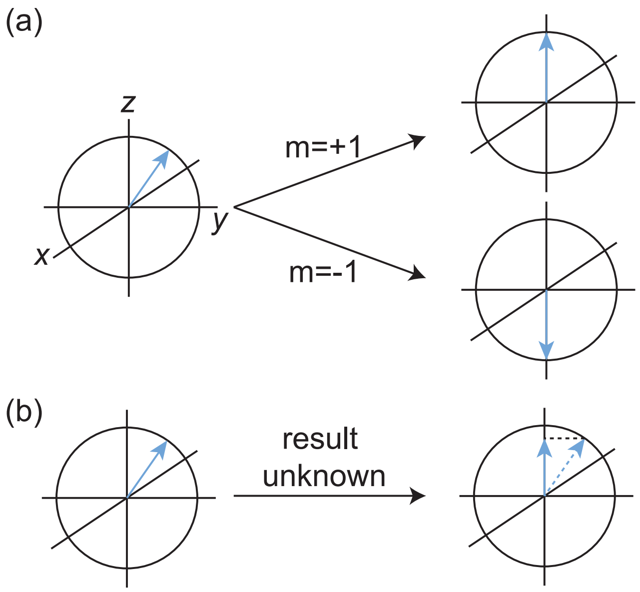


FIGURE 13.3: Illustration of the effect of the projective operation elements in the direct quantum measurement of Z for a qubit. Depending on the measurement result, the qubit Bloch vector is collapsed to the north ($m = +1$) or south pole ($m = -1$). Under unconditioned evolution, the Bloch vector x and y components vanish, while the z component remains unchanged.

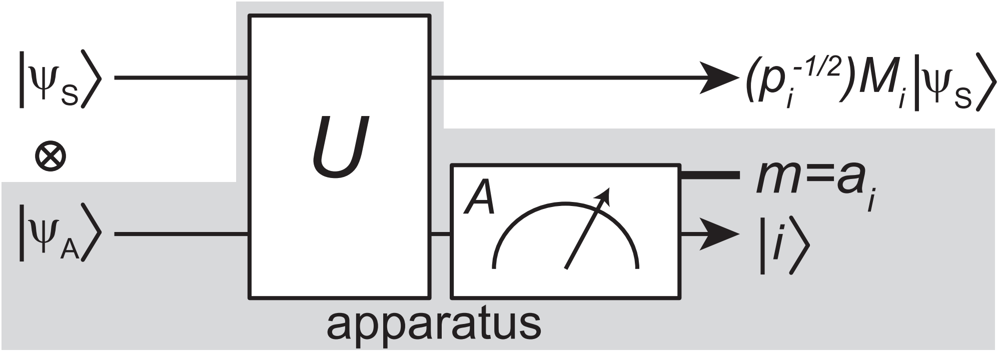


FIGURE 13.4: Anatomy of an indirect quantum measurement.

result, the post-measurement density matrix evolves as

$$\rho_S = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \to \rho_S' = M_+ \rho_S M_+^\dagger + M_- \rho_S M_-^\dagger = \begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix}.$$

Thus, we see that the unconditioned evolution preserves the z component of the Bloch vector, and zeroes out its x and y components.

13

13.3 INDIRECT QUANTUM MEASUREMENTS

We now turn to the evolution of a quantum system that is subject to a different kind of measurement, known as *indirect quantum measurement*. An indirect measurement consists of three steps, as illustrated in Fig. 13.4. First, a measurement apparatus (itself a quantum system with a Hilbert space of dimension N_A) is initialized in a pure state $|\psi_A\rangle$. The system and apparatus then undergo a known joint unitary evolution U, which in general entangles the two. Finally, a direct measurement of the apparatus observable $\hat{A}$ is performed.

Let us take a moment to consider the wealth of choices available within indirect measurements:

- dimensionality N_{A} of the measurement apparatus.
- initial state $|\psi_{\mathrm{A}}\rangle$ of the measurement apparatus.
- unitary time evolution U corresponding to Hamiltonian that contains an interaction between system + apparatus.
- operator $\hat{A}$ for direct measurement of the apparatus.

These choices make for a wide variety of indirect measurements. Indeed, there is a whole zoology of such measurements. The interested reader will find a thorough presentation in Ch. 1 of Wiseman and Milburn.

Let us now consider the evolution of the system under such measurements. Like for direct measurements, we consider two scenarios: one in which we are aware of the measurement result (because we perform the measurement), and another in which we are ignorant of it (because another party performs the measurement and does not tell us the result).

13.3.1 SYSTEM EVOLUTION CONDITIONED ON THE RESULT OF INDIRECT MEASUREMENT

We first analyze what happens to the state of the system in an indirect measurement. The system and apparatus are initially in a product state $|\psi_{\mathrm{U}}(0)\rangle = |\psi_{\mathrm{S}}\rangle\,|\psi_{\mathrm{A}}\rangle$. Following the unitary evolution, the joint state is

$$|\psi_{\mathrm{U}}\rangle = \sum_{i}^{N_{\mathrm{A}}} M_i\,|\psi_{\mathrm{S}}\rangle\,|i\rangle\,,$$

where $M_i = \langle i|\,U\,|\psi_{\mathrm{A}}\rangle$ – see Eq. (13.4). We again emphasize that, although this form of M_i seems to give a number, it is actually an operator in S, as U acts on the combined Hilbert space of the system and the apparatus, and $\langle i|$ and $|\psi_{\mathrm{A}}\rangle$ only act within the Hilbert space of the apparatus – see the discussion following Eq. (13.1.2). We choose for orthonormal basis $|i\rangle$ of the apparatus Hilbert space the eigenstates of the operator $\hat{A}$, $\hat{A}\,|i\rangle = a_i\,|i\rangle$. This choice simplifies the next step. Upon a measurement giving result $m = a_i$, the joint system collapses to the product state

$$\left|\psi_{\mathrm{U}|m=a_i}\right\rangle = \mathscr{N}\left[M_i\,|\psi_{\mathrm{S}}\rangle\right]\otimes|i\rangle\,.$$

Here, $\mathscr{N} = 1/\sqrt{\langle\psi_{\mathrm{S}}|\,M_i^\dagger M_i\,|\psi_{\mathrm{S}}\rangle}$ is a normalization factor. Note that, by Born's rule, $\langle\psi_{\mathrm{S}}|\,M_i^\dagger M_i\,|\psi_{\mathrm{S}}\rangle$ is precisely the probability p_i of getting measurement result $m = a_i$. Thus, we can write the post-measurement reduced density matrix conditioned on $m = a_i$ as

$$\rho_{\mathrm{S}|m=a_i} = \frac{1}{p_i} M_i \rho_{\mathrm{S}} M_i^\dagger.$$

Note that this form is the same as Eq. (13.7), except for a crucial distinction: M_i is not necessarily a projection operator!

13.3.2 UNCONDITIONED SYSTEM EVOLUTION UNDER INDIRECT MEASUREMENT

If the indirect measurement is performed by another party that does not reveal the measurement result, the post-measurement reduced density matrix of the system is mixed. It is given by a weighted sum of the conditioned density matrices,

$$\begin{aligned}
\rho'_{\mathrm{S}} &= \sum_i p_i \rho_{s|m=a_i} && (13.8)\\
&= \sum_i p_i \left(\frac{1}{p_i} M_i\,|\psi_{\mathrm{S}}\rangle\,\langle\psi_{\mathrm{S}}|\,M_i^\dagger\right) && (13.9)\\
&= \sum_i M_i\,|\psi_{\mathrm{S}}\rangle\,\langle\psi_{\mathrm{S}}|\,M_i^\dagger && (13.10)\\
&= \sum_i M_i \rho_{\mathrm{S}} M_i^\dagger. && (13.11)
\end{aligned}$$

Thus, the transformation $\mathscr{E}(\rho_\mathrm{S})$ of the reduced density matrix of a system undergoing an indirect quantum measurement can also be written in operator-sum representation. When the measurement result is known, there is only one operation element. When the result is unknown (or ignored) the sum includes N_A operator elements.

13.3.3 WHAT DOES AN INDIRECT QUANTUM MEASUREMENT ACTUALLY MEASURE?

Before considering our first example, we briefly discuss the statistics of the measurement result. The average value of the measurement is

$$\begin{aligned}
\langle m\rangle &= \sum_{i=1}^{N_\mathrm{A}} a_i p_i \\
&= \sum_{i=1}^{N_\mathrm{A}} a_i \langle\psi_\mathrm{S}| M_i^\dagger M_i |\psi_\mathrm{S}\rangle \\
&= \langle\psi_\mathrm{S}| \left(\sum_{i=1}^{N_\mathrm{A}} a_i M_i^\dagger M_i\right) |\psi_\mathrm{S}\rangle \\
&= \langle\psi_\mathrm{S}|\hat{\mathscr{O}}|\psi_\mathrm{S}\rangle ,
\end{aligned}$$

where $\hat{\mathscr{O}} = \sum_{i=1}^{N_\mathrm{A}} a_i M_i^\dagger M_i$ is an Hermitian operator acting on the system Hilbert space. Let us take a moment to appreciate how this system operator is determined. It is given by weighted sum of N_A Hermitian operators built from the operation elements M_i (themselves determined by the initial apparatus state and the evolution U). The weighing coefficients are given by the eigenvalues of the direct measurement of the apparatus operator A.

Evidently, like the observable in a direct measurement, $\hat{\mathscr{O}}$ is an Hermitian operator. However, note the crucial distinction. The measurement results a_i and corresponding post-measurement states $\left|\psi_{\mathrm{S}|m=a_i}\right\rangle$ are categorically not eigenvalues and eigenvectors of $\hat{\mathscr{O}}$!

13.3.4 POVMS

We have placed a lot of emphasis on the post-measurement state of the system. However, it is often the case in experiment that a measurement is performed only as the final step, and the post-measurement state of the system is really not of interest. Rather, only the properties of the pre-measurement state revealed by the measurement are of interest. In such a case, it suffices to know the operator as $\hat{\mathscr{O}} = \sum_{i=1}^{N_\mathrm{A}} a_i E_i$. The E_i are called *effects*, and are related to the operation elements by $E_i = M_i^\dagger M_i$. Note that knowing the effects E_i lets us calculate all the statistics of the measurement result, but precludes us from saying anything about the post-measurement state. The effects add up to the identity operator in the system Hilbert space,

$$\sum_i E_i = I,$$

and are positive operators:

$$\langle\psi_\mathrm{S}| E_i |\psi_\mathrm{S}\rangle \geq 0 \;\text{for any}\; |\psi_\mathrm{S}\rangle .$$

Measurements that are specified by their effects, rather than their operation elements, are known as *positive-operator-valued measures*, or *POVMs* in short.

13.3.5 EXAMPLE: WEAK QUANTUM MEASUREMENT OF A QUBIT

One of the simplest and most beautiful examples of an indirect quantum measurement is the so-called *weak* measurement of a qubit. This process is illustrated in Fig. 13.5. The apparatus consists of another qubit (often called the *ancilla*[1]) initialized in state $|0\rangle$. The interaction between the qubit S and the ancilla A leads to the unitary evolution

$$U = e^{-\mathrm{i} Z_\mathrm{S} Y_\mathrm{A} \theta},$$

---

[1] Ancilla is the latin word for a slave or servant.

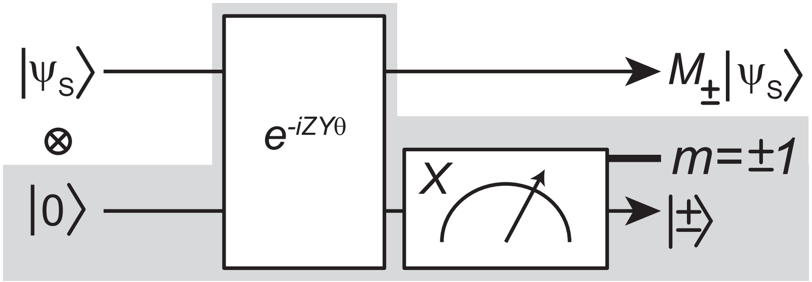


FIGURE 13.5: Weak measurement of a qubit by indirect quantum measurement. The apparatus consists contains an ancilla qubit initialized in $|0\rangle$. The ancilla and qubit undergo an entangling unitary $U = e^{-iZY\theta}$, where Z acts on the qubit and Y on the ancilla. Following this interaction, a direct measurement of X is performed on the ancilla. The strength of the measurement is controlled by θ. The choice $\theta = \pi/4$ realizes a fully projective measurement of the qubit in the qubit Z basis. The choice $\theta = 0$ implies no measurement at all. For in-between values of θ, a weak measurement is performed where the operation elements only rotate the qubit toward the north and south poles of the Bloch sphere.

where Z_S acts on the qubit and Y_A on the ancilla and $\theta \in [0, \pi/4]$. Finally, a direct measurement of the operator X_A is performed on the ancilla. By choosing θ small, the act of measurement distorts the initial state only slightly - in this sense, it is (or can be) a weak measurement.

The unitary evolution is easy to visualize: the ancilla undergoes a rotation by $\pm 2\theta$ around the y axis of the Bloch sphere, with the sign depending on the state of the qubit:

$$\begin{aligned}
U\left(\alpha\,|0\rangle_\mathrm{S} + \beta\,|1\rangle_\mathrm{S}\right)|0\rangle_\mathrm{A} &= \alpha\,|0\rangle_\mathrm{S}\, e^{-\mathrm{i}Y\theta}\,|0\rangle_\mathrm{A} + \beta\,|1\rangle_\mathrm{S}\, e^{\mathrm{i}Y\theta}\,|0\rangle_\mathrm{A} \\
&= \alpha\,|0\rangle_\mathrm{S}\,(\cos\theta\,|0\rangle_\mathrm{A} + \sin\theta\,|1\rangle_\mathrm{A}) + \beta\,|1\rangle_\mathrm{S}\,(\cos\theta\,|0\rangle_\mathrm{A} - \sin\theta\,|1\rangle_\mathrm{A}) \\
&= \cos\theta\left(\alpha\,|0\rangle_\mathrm{S} + \beta\,|1\rangle_\mathrm{S}\right)|0\rangle_\mathrm{A} + \sin\theta\left(\alpha\,|0\rangle_\mathrm{S} - \beta\,|1\rangle_\mathrm{S}\right)|1\rangle_\mathrm{A}.
\end{aligned}$$

Writing this state in the $|\pm\rangle_\mathrm{A} = (|0\rangle\,\_A \pm |1\rangle_\mathrm{A})/\sqrt{2}$ basis for the ancilla, dropping the subscripts and using $C_\theta = \cos(\theta)$ and $S_\theta = \sin(\theta)$ for conciseness, we obtain

$$\begin{aligned}
U\left(\alpha\,|0\rangle + \beta\,|1\rangle\right)|+\rangle &= C_\theta\left(\alpha\,|0\rangle + \beta\,|1\rangle\right)\frac{|+\rangle + |-\rangle}{\sqrt{2}} + S_\theta\left(\alpha\,|0\rangle - \beta\,|1\rangle\right)\frac{|+\rangle - |-\rangle}{\sqrt{2}} \\
&= \frac{\left((C_\theta + S_\theta)\,\alpha\,|0\rangle + (C_\theta - S_\theta)\,\beta\,|1\rangle\right)}{\sqrt{2}}\,|+\rangle + \frac{\left((C_\theta - S_\theta)\,\alpha\,|0\rangle + (C_\theta + S_\theta)\,\beta\,|1\rangle\right)}{\sqrt{2}}\,|-\rangle.
\end{aligned}$$

From this, the operation elements are seen to be

$$M_+ = \frac{C_\theta}{\sqrt{2}} I + \frac{S_\theta}{\sqrt{2}} Z \text{ and } M_- = \frac{C_\theta}{\sqrt{2}} I - \frac{S_\theta}{\sqrt{2}} Z.$$

In the qubit $|0\rangle, |1\rangle$ basis, they are represented by the matrices

$$M_+ = \begin{pmatrix} \frac{C_\theta + S_\theta}{\sqrt{2}} & 0 \\ 0 & \frac{C_\theta - S_\theta}{\sqrt{2}} \end{pmatrix} \text{ and } M_- = \begin{pmatrix} \frac{C_\theta - S_\theta}{\sqrt{2}} & 0 \\ 0 & \frac{C_\theta + S_\theta}{\sqrt{2}} \end{pmatrix}.$$

Let us take a moment to visualize the effect of these operation elements on the qubit state. For $\theta = \pi/4$, the operation elements become projection operators onto $|0\rangle$ and $|1\rangle$. This is identical to a direct measurement of Z. For $\theta < \pi/4$, the measurement is no longer projective, and in this sense is called *weak*. Rather than collapse the Bloch vector onto the poles, a measurement result $m = +1$ ($m = -1$) only rotates the Bloch vector *toward* the north (south) pole. This rotation preserves the azimuthal angle of the Bloch vector. The rotation angle magnitude depends on θ, on the latitude of the Bloch vector, and on the measurement result.

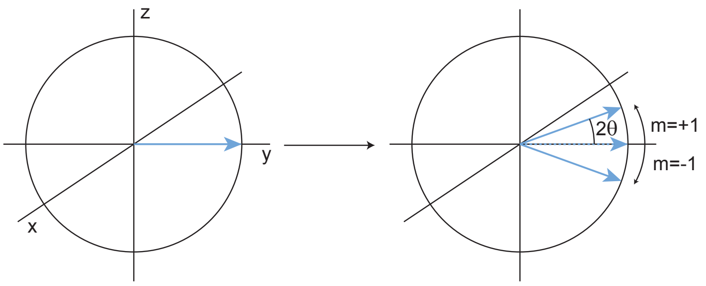


FIGURE 13.6: Illustration of the operation elements for weak indirect measurement of a qubit. (a) For a measurement result $m = +1$ ($m = -1$), a Bloch vector initially on the equator is rotated toward the north (south) pole by a change of polar angle of $\mp 2\theta$. The azymuthal angle of the Bloch vector is preserved. (b) In general, the change of polar angle is a function of θ, the latitude of the Bloch vector, and the measurement result.

When the Bloch vector lies on the equator, the rotation angle has magnitude (2θ) for both measurement results, as illustrated in Fig. 13.6.

The unconditioned post-measurement density matrix

$$\rho'_S = M_+\rho_S M_+ + M_-\rho_S M_-$$

is easily visualized on the Bloch sphere. It is straightforward to show that

$$\begin{pmatrix} \langle X\rangle' \\ \langle Y\rangle' \\ \langle Z\rangle' \end{pmatrix} = \begin{pmatrix} C_{2\theta} & 0 & 0 \\ 0 & C_{2\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \langle X\rangle \\ \langle Y\rangle \\ \langle Z\rangle \end{pmatrix}.$$

The effects are

$$E_+ = M_+^\dagger M_+ = \frac{1}{2}I + \frac{S_{2\theta}}{2}Z \text{ and } E_- = M_-^\dagger M_- = \frac{1}{2}I - \frac{S_{2\theta}}{2}Z.$$

The measurement operator $\hat{\mathcal{O}}$ is

$$\begin{aligned} \hat{\mathcal{O}} &= +1M_+^\dagger M_+ - 1M_-^\dagger M_- \\ &= +1\left(\frac{1}{2}I + \frac{S_{2\theta}}{2}Z\right) - 1\left(\frac{1}{2}I - \frac{S_{2\theta}}{2}Z\right) \\ &= S_{2\theta}Z. \end{aligned}$$

13.4 REPEATED MEASUREMENTS

It is natural to ask, both for direct and indirect quantum measurements, how ρ_S evolves under repeated measurements. After N measurements resulting in a measurement record $\{m_1, m_2, m_3, ...m_N\} = \{i_1, i_2, i_3, ..., i_N\}$, ρ_S transforms into

$$\rho'_S = \frac{1}{p}\left(\Pi_{n=1}^N M_{i_n}\right)\rho_S(0)\left(\Pi_{n=1}^N M_{i_n}\right)^\dagger,$$

where $p = \left(\Pi_{n=1}^N M_{i_n}\right)\rho_S(0)\left(\Pi_{n=1}^N M_{i_n}\right)^\dagger$ is the probability of getting the measurement record.

For direct measurements, the evolution of the ρ_S is not particularly rich: it is frozen after the first measurement! This is because the operation elements M_i are orthogonal projectors, and thus all subsequent measurements equal the first, and $\Pi_{n=1}^N M_{i_n} = M_{i_1}^N = M_{i_1}$. The two possible *trajectories* for the Bloch vector of a qubit undergoing a series of direct measurements of Z are shown in Fig. 13.7(a,b). Remember that the conditional evolution is subject to a known measurement outcome. The evolution can be much richer for weak indirect quantum measurements. Here, we imagine that the apparatus state is re-initialized to a fixed $|\psi_A\rangle$

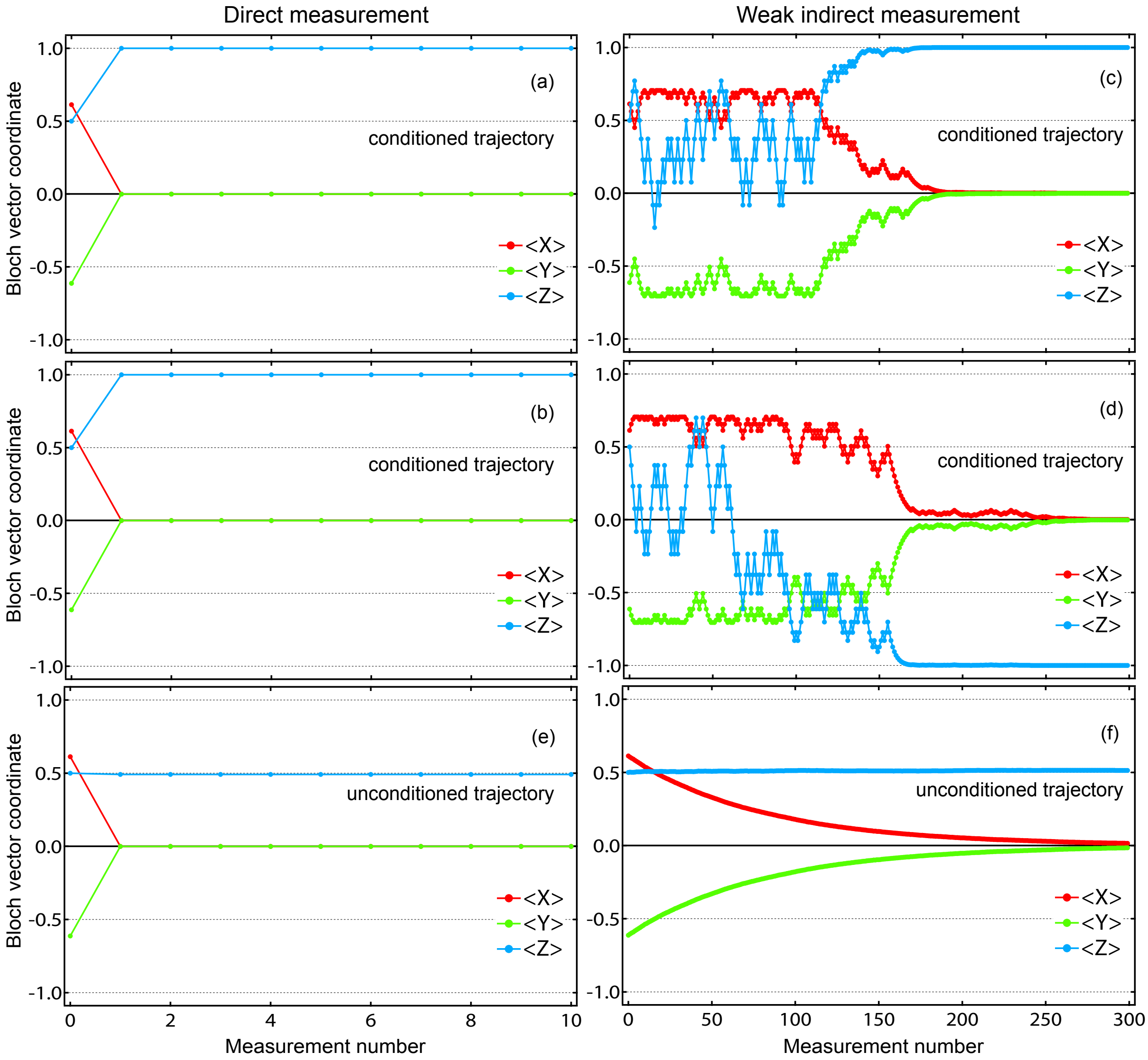


FIGURE 13.7: Sample quantum trajectories of a qubit undergoing repeated direct measurements of Z (a,b) and repeated weak indirect measurements with $\theta = \pi/40$ (c,d). (e) The unconditioned trajectory of the Bloch vector under repeated direct measurement. (f) The unconditioned trajectory under repeated weak indirect measurements. The initial Bloch vector is in all cases $(\langle X\rangle, \langle Y\rangle, \langle Z\rangle) = \left(\sqrt{3/8}, -\sqrt{3/8}, 1/2\right)$.

following each measurement. Note that otherwise the operation elements M_i would change from measurement to measurement! Example trajectories for the Bloch vector of a qubit undergoing weak indirect measurements with $\theta = \pi/40$ are shown in Fig. 13.7(c,d).

If the measurement results are concealed (or disregarded), the reduced density matrix of the system becomes mixed. After N measurements,

$$\rho'_S = \sum_{i_i=1}^{N_A} \sum_{i_2=1}^{N_A} \cdots \sum_{i_N=1}^{N_A} \left(\Pi_{n=1}^{N} M_{i_n}\right) \rho_S \left(\Pi_{n=1}^{N} M_{i_n}\right)^{\dagger}.$$

Note that for direct measurements, the projective character of the M_i also stops the evolution of the unconditioned density matrix after the first measurement (Fig. 13.7(e)). The evolution of the unconditioned density matrix for a qubit undergoing weak indirect measurements with $\theta = \pi/40$ is shown in Fig. 13.7(f).

13.5 LINDBLAD REPRESENTATION

So far we have described transformations $\rho_S \to \rho'_S = \mathscr{E}(\rho_S)$ using the operator-sum, or *Kraus*, representation. In all of our examples in this chapter - coupling to environment, direct and indirect quantum measurements - the final reduced density matrix is given by "sandwiching" the initial matrix ρ_S by operation elements[2] M_i. In this section, we will derive a general expression for the *change* in ρ_S, that is, for $\mathscr{E}(\rho_S) - \rho_S$. We will show that this change can always be represented in a compact form known as *Lindblad* form.

Our starting point is to expand all operation elements M_i using an orthonormal basis $F_1, F_2, ..., F_{N_S^2}$ for operators in S. By orthonormal, we mean

$$\mathrm{Tr}_S \left[F_i^\dagger F_j\right] = \delta_{ij}. \tag{13.12}$$

For convenience, we always define $F_{N_S^2}$ as the normalized identity operator,

$$F_{N_S^2} \equiv \frac{1}{\sqrt{N_S}} I_S.$$

Note that this choice of $F_{N_S^2}$ combined with the orthonormality condition make $F_1, ..., F_{N_S^2-1}$ traceless,

$$\mathrm{Tr}_S\ [F_i] = 0 \text{ for } i = 1, .., N_S^2 - 1.$$

Expanding the operation elements in this basis, we have

$$M_i = \sum_{j=1}^{N_S^2} \alpha_{ij} F_j,$$

with coefficients $\alpha_{ij} = \mathrm{Tr}_S \left[F_j^\dagger M_i\right]$.

Example: Orthonormal bases of operators.

For a two-dimensional S, a basis of orthonormal operators is

$$\left\{\frac{1}{\sqrt{2}} Z, \sigma_-, \sigma_+, \frac{1}{\sqrt{2}} I\right\}, \tag{13.13}$$

where

$$\sigma_+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } \sigma_- = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Another possibility is

$$\left\{\frac{1}{\sqrt{2}} X, \frac{1}{\sqrt{2}} Y, \frac{1}{\sqrt{2}} Z, \frac{1}{\sqrt{2}} I\right\}. \tag{13.14}$$

You should check that each of these bases satisfies the orthonormality condition (13.12). Using Eq. (13.5),

$$\begin{aligned} \mathscr{E}(\rho_S) &= \sum_i M_i \rho_S M_i^\dagger \\ &= \sum_i \left(\sum_j \alpha_{ij} F_j\right) \rho_S \left(\sum_k \alpha_{ik}^* F_k^\dagger\right) \\ &= \sum_{j,k} c_{jk} F_j \rho_S F_k^\dagger, \end{aligned}$$

where

$$c_{jk} \equiv \sum_i \alpha_{ij} \alpha_{ik}^*.$$

---

[2] Operation elements are also called *Kraus* operators in the literature

Separating the terms involving $F_{N_S^2}$, we arrive at

$$\mathscr{E}(\rho_S) = \frac{c_{N_S^2 N_S^2}}{N_S}\rho_S + \sum_{j=1}^{N_S^2-1}\left(\frac{c_{jN_S^2}}{\sqrt{N_S}}F_j\rho_S + \frac{c_{N_S^2 j}}{\sqrt{N_S}}\rho_S F_j^\dagger\right) + \sum_{j,k=1}^{N_S^2-1} c_{jk}F_j\rho_S F_k^\dagger$$

Defining

$$F \quad = \quad \tfrac{1}{\sqrt{N_S}}\textstyle\sum_{i=2}^{N_S^2} c_{iN_S^2}F_i,$$

and from it the Hermitian operators

$$\begin{aligned} H &= \tfrac{1}{2\mathrm{i}}\left(F^\dagger - F\right), \\ G &= \tfrac{c_{N_S^2 N_S^2} - N_S}{2N_S} I + \tfrac{1}{2}\left(F^\dagger + F\right), \end{aligned}$$

the equation can be written as

$$\mathscr{E}(\rho_S) - \rho_S = -\mathrm{i}\left[H, \rho_S\right] + \left\{G, \rho_S\right\} + \sum_{j,k=1}^{N_S^2-1} c_{jk}F_j\rho_S F_k^\dagger.$$

The transformation must conserve the (unity) trace of any ρ_S. The condition

$$\mathrm{Tr}_S\,[\mathscr{E}(\rho_S) - \rho_S] = 0$$

implies

$$G = -\frac{1}{2}\sum_{j,k=1}^{N_S^2-1} c_{jk}F_k^\dagger F_j,$$

and thus

$$\mathscr{E}(\rho_S) - \rho_S = -\mathrm{i}\left[H, \rho_S\right] + \sum_{j,k=1}^{N_S^2-1} c_{jk}\left(F_j\rho_S F_k^\dagger - \frac{1}{2}\left\{F_k^\dagger F_j, \rho_S\right\}\right). \tag{13.15}$$

This compact expression is said to be in *first standard form.*

The $(N_S^2 - 1) \times (N_S^2 - 1)$ matrix

$$C \equiv \begin{pmatrix} c_{1,1} & \cdots & c_{2,N_S^2-1} \\ \vdots & \ddots & \vdots \\ c_{N_S^2-1,1} & \cdots & c_{N_S^2-1,N_S^2-1} \end{pmatrix}$$

is Hermitian and positive semi-definite. It may be diagonalized by a unitary transformation U:

$$UCU^\dagger = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \gamma_{N_S^2-1} \end{pmatrix},$$

with diagonal elements $\gamma_i \geq 0$.

Introducing a new set of operators A_k:

$$F_j = \sum_{k=1}^{N_S^2-1} u_{kj}A_k,$$

and plugging into Eq. (13.15), we arrive finally at

$$\mathscr{E}(\rho_S) - \rho_S = -\mathrm{i}\left[H, \rho_S\right] + \sum_{i=1}^{N_S^2-1} \gamma_i\left(A_i\rho_S A_i^\dagger - \frac{1}{2}\left\{A_i^\dagger A_i, \rho_S\right\}\right),$$

which can be written very compactly as

$$\mathscr{E}(\rho_{\mathrm{S}}) - \rho_{\mathrm{S}} = -\mathrm{i}\left[H, \rho_{\mathrm{S}}\right] + \sum_{i=1}^{N_{\mathrm{S}}^2-1} \gamma_i \mathscr{D}[A_i]\rho_{\mathrm{S}},$$

with

$$\mathscr{D}[A]\rho_{\mathrm{S}} \equiv A\rho_{\mathrm{S}}A^\dagger - \frac{1}{2}A^\dagger A\rho_{\mathrm{S}} - \frac{1}{2}\rho_{\mathrm{S}}A^\dagger A. \tag{13.16}$$

This final expression is in so-called *Lindblad form.* $\mathscr{D}$ bears the name of *dissipation superoperator.* It warrants being called a *super*-operator because it transforms operators to operators. The operator A associated with $\mathscr{D}$ is called a *Lindblad operator.* It is important to remark that, in general, the Hermitian operator H is not the free Hamiltonian H_{S} of S in the full Hamiltonian $H_{\mathrm{U}} = H_{\mathrm{S}} + H_{\mathrm{E}} + H_{\mathrm{I}}$.

Example: The damping channel.
Let us write the damping transformation in Lindblad form. In the basis (13.13), the two operation elements are

$$M_0 = \alpha_{01}\frac{1}{\sqrt{2}}Z + \alpha_{04}\frac{1}{\sqrt{2}}I \text{ and } M_1 = \alpha_{13}\sigma_-,$$

with $\alpha_{04} = \frac{1+\sqrt{1-p}}{\sqrt{2}}$, $\alpha_{01} = \frac{1-\sqrt{1-p}}{\sqrt{2}}$, and $\alpha_{13} = \sqrt{p}$ (all other $\alpha_{ij} = 0$). These give $c_{11} = \left(1-\sqrt{1-p}\right)^2/2$, $c_{33} = p$ and $c_{jk} = 0$ for all other j, k. Since $F \propto Z$ and $Z = Z^\dagger$, it follows that $H = 0$. The matrix C is already diagonal, and we thus arrive at

$$\mathscr{E}(\rho_{\mathrm{S}}) - \rho_{\mathrm{S}} = c_{33}\mathscr{D}[\sigma_-]\rho_{\mathrm{S}} + \frac{c_{11}}{2}\mathscr{D}[Z]\rho_{\mathrm{S}}.$$

Example: The depolarization channel.
We can proceed similarly to write the depolarization transformation in Lindblad form. The four operation elements in (13.6) are already expanded in the operator basis (13.14). It is easy to see that $F = 0$ and thus also $H = 0$. The matrix C is already diagonal. We leave it as an exercise to arrive at

$$\mathscr{E}(\rho_{\mathrm{S}}) - \rho_{\mathrm{S}} = \frac{p}{4}\mathscr{D}[X]\rho_{\mathrm{S}} + \frac{p}{4}\mathscr{D}[Y]\rho_{\mathrm{S}} + \frac{p}{4}\mathscr{D}[Z]\rho_{\mathrm{S}}.$$

13.5.1 UNCONDITIONED WEAK MEASUREMENTS IN LINDBLAD FORM

Let us now consider a weak indirect quantum measurement in which the interaction between system and apparatus is given by the Hamiltonian

$$H_{\mathrm{U}} = H_{\mathrm{I}} = B_{\mathrm{S}} \otimes A_{\mathrm{A}},$$

with Hermitian system operator B_{S} and apparatus operator A_{A}. The time-evolution operator for an interaction time τ is given by

$$\rho'_{\mathrm{U}} = U(\tau)\rho_{\mathrm{U}}U(\tau)^\dagger,$$

where

$$U = e^{-\mathrm{i}H_{\mathrm{U}}\tau/\hbar}.$$

Because we assume the interaction to be weak, we can approximate U by an expansion up to second order in τ:

$$U = 1 + \frac{\tau}{\mathrm{i}\hbar}B_{\mathrm{S}}A_{\mathrm{A}} - \frac{\tau^2}{\hbar^2}B_{\mathrm{S}}^2A_{\mathrm{A}}^2 + \ldots$$

The final universe density matrix is thus

$$\begin{aligned} U\rho_{\mathrm{U}}U^\dagger &= \left(I + \tfrac{\tau}{\mathrm{i}\hbar}BA - \tfrac{\tau^2}{2\hbar^2}B^2A^2\right)\rho_{\mathrm{U}}\left(I - \tfrac{\tau}{\mathrm{i}\hbar}BA - \tfrac{\tau^2}{2\hbar^2}B^2A^2\right) \\ &= \rho_{\mathrm{U}} + \tfrac{\tau}{\mathrm{i}\hbar}\left(BA\rho_{\mathrm{U}} - \rho_{\mathrm{U}}BA\right) + \tfrac{\tau^2}{\hbar^2}\left(BA\rho_{\mathrm{U}}BA - \tfrac{1}{2}B^2A^2\rho_{\mathrm{U}} - \tfrac{1}{2}\rho_{\mathrm{U}}B^2A^2\right) + O(\tau^3). \end{aligned}$$

Tracing over the apparatus and taking into consideration that initially the system is unentangled with the apparatus,

$$\rho_{\mathrm{U}} = \rho_{\mathrm{S}} \otimes |\psi_{\mathrm{A}}\rangle\langle\psi_{\mathrm{A}}|,$$

we find

$$\begin{aligned} \mathscr{E}(\rho_{\mathrm{S}}) &= \mathrm{Tr}_{\mathrm{A}}\,[U\rho_{\mathrm{U}}U^{\dagger}] \\ &= \rho_{\mathrm{S}} + \tfrac{\tau}{\mathrm{i}\hbar}\langle A\rangle[B,\rho_{\mathrm{S}}] + \langle A^2\rangle\tfrac{\tau^2}{\hbar^2}\left(B\rho_{\mathrm{S}}B - \tfrac{1}{2}\{B^2,\rho_{\mathrm{S}}\}\right) + O(\tau^3), \end{aligned}$$

with $\langle A\rangle \equiv \langle\psi_{\mathrm{A}}|\,A\,|\psi_{\mathrm{A}}\rangle$ and $\langle A^2\rangle \equiv \langle\psi_{\mathrm{A}}|\,A^2\,|\psi_{\mathrm{A}}\rangle$.

We have arrived at a Lindblad-form expression for the change in ρ_{S}:

$$\mathscr{E}(\rho_{\mathrm{S}}) - \rho_{\mathrm{S}} = \frac{\tau}{\mathrm{i}\hbar}\langle A\rangle[B,\rho_{\mathrm{S}}] + \langle A^2\rangle\frac{\tau^2}{\hbar^2}\mathscr{D}[B]\rho_{\mathrm{S}} + O(\tau^3). \tag{13.17}$$

This expression has only one dissipator with associated Lindblad operator B_{S}, and an effective Hamiltonian term proportional to B_{S}. The latter induces systematic backaction on the system. Notice that the strength of these two different terms depends on the initial state of the apparatus through $\langle A\rangle$ and $\langle A^2\rangle$.

Example: Weak indirect measurement of a qubit.

Let us take a moment to relate this result to the example of weak measurement of a qubit from above. Linking the notation, we have $B_{\mathrm{S}} = Z$ and $A_{\mathrm{A}} = Y$, $\theta = \tau/\hbar$, and $|\psi_{\mathrm{A}}\rangle = |0_{\mathrm{A}}\rangle$. Thus, we have $\langle A_{\mathrm{A}}\rangle = 0$ and $\langle A_{\mathrm{A}}^2\rangle = 1$. The first tells us that there is no systematic backaction on the measured qubit, and that all dynamics of ρ_{S} arises from the single dissipator.

An interesting prediction of Eq. (13.17) is that initializing the ancilla qubit in a state $|\psi_{\mathrm{A}}\rangle$ with non-zero expectation value for Y_{A} ($\langle\psi_{\mathrm{A}}|\,Y_{\mathrm{A}}\,|\psi_{\mathrm{A}}\rangle \neq 0$)) will produce systematic backaction of the measured qubit. Note that $A_{\mathrm{A}}^2 = Y^2 = I$, so the dissipative term is independent of the initial ancilla state. We explore this in Fig. 13.8 by initializing the ancilla qubit on the y-z plane with polar angle ranging from 0 to $\pi/2$.

13.6 PROBLEMS

1. *Measurement-induced qubit dephasing*

   In lecture, we discussed an indirect quantum measurement of a qubit using a second, ancillary qubit. This two-step process consists of an interaction

   $$U = e^{-\mathrm{i}Z_{\mathrm{S}}Y_{\mathrm{A}}\theta},$$

   followed by direct measurement of the ancilla with operator X_{A}. Here, we will explore the consequences of projecting the ancilla onto different bases. We adopt the usual convention where X, Y and Z stand for the Pauli matrices – they are used instead of σ_x, σ_y and σ_z.

   (a) Consider an ancilla measurement in Y_{A}. Derive the two operation elements in this case. What does each of these operation elements do to the qubit, i.e., what is their backaction on the qubit? Check that these operation elements form a decomposition of unity.

   (b) Consider an ancilla measurement in Z_{A}. Derive the two operation elements in this case. What does each of these operation elements do to the qubit? Check that these operation elements form a decomposition of unity.

   (c) Show that in both cases, the unconditioned post-measurement density matrix is the same as that derived in class for measurement in X_{A}.

   Note: This result is quite general. When an indirect measurement is performed on a system, but the measurement result ignored, the post-measurement system density matrix is independent of the projective measurement chosen for the probe.

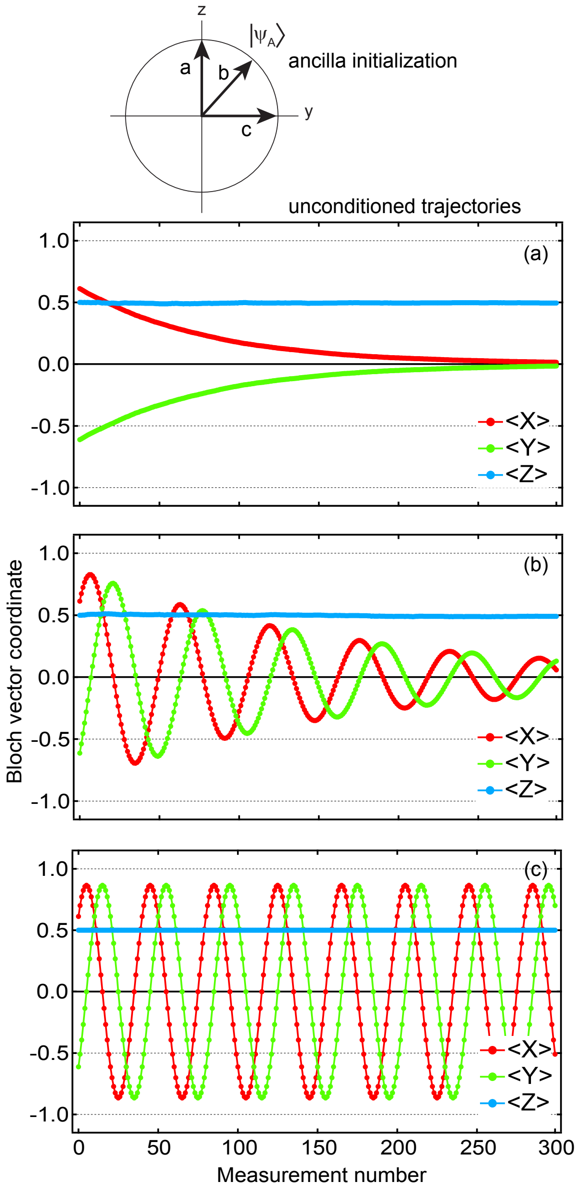


FIGURE 13.8: Unconditioned evolution of a qubit undergoing weak quantum measurement with the ancilla initialized in three different states on the y-z plane of the Bloch sphere: (a) $|\psi_\mathrm{A}\rangle = |0\rangle$, (b) $|\psi_\mathrm{A}\rangle = C_{\pi/8}\,|0\rangle + iS_{\pi/8}\,|1\rangle$, (c) and $|\psi_\mathrm{A}\rangle = \frac{1}{\sqrt{2}}\,(|0\rangle + i\,|1\rangle)$. The initial condition of the ancilla affects the systematic backaction on the qubit. Consistent with Eq. (13.17), the systematic backaction in this example is a rotation about the z axis. In (a), there is no systematic backaction, and all dynamics arises from the dissipator. In (c), the systematic backaction term dominates the evolution.

2. Consider the problem of discriminating between non-orthogonal polarization states of a single photon using projective measurements. Imagine a source of single photons that prepares them in either of the following two non-orthogonal states:

$$\begin{aligned} \left|\psi_0\right\rangle &= \cos(\alpha/2)\,|H\rangle + \sin(\alpha/2)\,|V\rangle\,, \\ \left|\psi_1\right\rangle &= \cos(\alpha/2)\,|H\rangle - \sin(\alpha/2)\,|V\rangle\,, \end{aligned}$$

where $|H\rangle$ and $|V\rangle$ correspond to horizontal and vertical polarization. We denote the probability that the source sends $\left|\psi_0\right\rangle$ as p_0 and the probability that it sends $\left|\psi_1\right\rangle$ as $p_1 = 1 - p_0$.

(a) Calculate the overlap $\left\langle\psi_0|\psi_1\right\rangle$.

(b) If $|\psi_0\rangle$ and $|\psi_1\rangle$ were orthogonal, we would discriminate perfectly by performing projective measurements in the basis $|\psi_0\rangle$, $|\psi_1\rangle$. For non-orthogonal $|\psi_0\rangle$ and $|\psi_1\rangle$, what is the measurement basis which minimizes the probability for a mistake in deciding which state we received? Hint: Parametrize the measurement basis by an angle β:

$$\begin{aligned} |\Phi_0\rangle &= \cos(\beta/2)\,|H\rangle + \sin(\beta/2)\,|V\rangle\,, \\ |\Phi_1\rangle &= \sin(\beta/2)\,|H\rangle - \cos(\beta/2)\,|V\rangle\,, \end{aligned}$$

and write the probability of error as a function of p_0, p_1, α and β. Minimize with respect to β.

(c) Show that the minimum error probability is attained for

$$\tan(\beta^{\text{opt}}) = \tan(\alpha)/(p_0 - p_1).$$

(d) Find an expression for the minimum error probability, p_{min}. This minimum error probability is known as the *Helstrom lower bound.*

(e) Make a plot of p_{min} as a function of α for $p_0 = p_1 = 1/2$.

3. In this problem, we consider a different strategy altogether for discriminating between the two non-orthogonal states. Here, we will devise a measurement scheme which yields the possible results:

- We know for sure that the state is $|\psi_0\rangle$.
- We know for sure that the state is $|\psi_1\rangle$.
- Inconclusive: we do not know which state was sent to us.

Consider a POVM with the three effects:

$$\begin{aligned} E_0 &= M_0^\dagger M_0 = \gamma_0\,|\Psi_2\rangle\,\langle\Psi_2|\,, \\ E_1 &= M_1^\dagger M_1 = \gamma_1\,|\Psi_3\rangle\,\langle\Psi_3|\,, \\ E_2 &= M_2^\dagger M_2 = 1 - E_0 - E_1, \end{aligned}$$

where

$$\begin{aligned} |\psi_2\rangle &= \sin(\alpha/2)\,|H\rangle - \cos(\alpha/2)\,|V\rangle\,, \\ |\psi_3\rangle &= \sin(\alpha/2)\,|H\rangle + \cos(\alpha/2)\,|V\rangle\,, \end{aligned}$$

and the coefficients γ_i are real and satisfy $\gamma_i > 0$. Note that $\langle\psi_2|\psi_0\rangle = \langle\psi_3|\psi_1\rangle = 0$, with $|\psi_0\rangle$ and $|\psi_1\rangle$ as defined in the previous problem.

(a) Determine the conditions for this POVM to be valid (the effects are positive operators, and form a decomposition of unity). That is, what constraint does this impose on γ_0 and γ_1?

(b) Show that the probability of getting measurement result 0 given that the photon sent is in $|\psi_0\rangle$ is 0.

(c) Calculate the probability of getting measurement result 1 given that the photon sent is in $|\psi_0\rangle$.

(d) Calculate the total probability of getting measurement result 2. Express this as a function of γ_0, γ_1, p_0, and p_1.

(e) The strategy for discrimination proposed is: If we measure 0, declare that the state sent was $|\psi_1\rangle$. If we measure 1, declare that state sent was $|\psi_0\rangle$. If we measure 2, report Inconclusive. Optimize over the γ to minimize the probability that our result will be inconclusive.

4. (10 pts) Consider the indirect weak measurement of a quantum bit S using another ancillary qubit A, as covered in class and lecture notes. As before, the ancilla is initialized in $|0\rangle_A$. However, on this occasion the chosen interaction between S and A is

$$H = Z_S X_A,$$

which is on for a time $t = \hbar\theta$.

   (a) Show that the operation elements corresponding to ancilla measurement along the Y_A axis with results $m = +1$ and $m = -1$ are (in the $|0\rangle_S, |1\rangle_S$ basis):

$$M_{+1} = \begin{pmatrix} \frac{C_\theta - S_\theta}{\sqrt{2}} & 0 \\ 0 & \frac{C_\theta + S_\theta}{\sqrt{2}} \end{pmatrix} \text{ and } M_{-1} = \begin{pmatrix} \frac{C_\theta + S_\theta}{\sqrt{2}} & 0 \\ 0 & \frac{C_\theta - S_\theta}{\sqrt{2}} \end{pmatrix}$$

   where the usual notation $C_\theta = \cos(\theta)$, $S_\theta = \sin(\theta)$ is used.

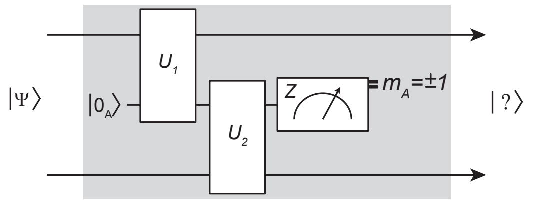
   (b) Consider now performing two such weak measurements (measurement a followed by measurement b). It is understood that the ancilla qubit is re-initialized to $|0\rangle_A$ between measurements. What is the operation element corresponding to the measurement record $\{m_a, m_b\} = \{+1, +1\}$.

   (c) Repeat for $\{m_a, m_b\} = \{+1, -1\}$.

   (d) For what measurement records can you say that the post-measurements' state of the qubit is the same as the initial state?

   (e) Generalize to the case of a series of N measurements. Consider the cases where N is odd or even. For what measurement records does the qubit return to the initial state after the measurements are done?

5. (10 pts) *Entanglement by measurement*
In this problem we consider the possibility of entangling two qubits (Q_1 and Q_2) by an indirect measurement using a third, ancillary qubit A. Consider the scheme below.



Q_1 and Q_2 are initially in a maximal superposition state

$$|\Psi\rangle = \frac{1}{2}\left(|0_1 0_2\rangle + |0_1 1_2\rangle + |1_1 0_2\rangle + |1_1 1_2\rangle\right).$$

The interaction between Q_i and A results in the unitary evolution:

$$U_i = |0_i\rangle\langle 0_i| e^{-i\theta Y_A} + |1_i\rangle\langle 1_i| e^{+i\theta Y_A},$$

where as usual Y_A is the Pauli-Y_A operator, and θ is a fixed angle, somewhere between 0 and $\pi/2$, equal for the two interactions.

   (a) Demonstrate that the initial state $|\Psi\rangle$ has no entanglement between Q_1 and Q_2.

Be careful: Do not try to solve the following parts using elaborate calculations. Try to visualize the action of the two U_i operators on the ancilla.

(a) Work out the combined action of U_1 and U_2 on the state $|0_1 1_2\rangle|0_A\rangle$. Do the same for the other possible initial states of the $Q1$-$Q2$ system.

(b) What is the probability of obtaining the measurement result $m_A = +1$? What is the post-measurement state in that case? Does this state represent entangled qubits Q_1 and Q_2? If yes, are the qubits *maximally* entangled?

(c) What is the post-measurement state of the Q_1-Q_2 system when the measurement result is $m_A = -1$? Are the two qubits entangled in this case? If so, are they *maximally* entangled?

6. **H** Consider a qubit whose state is described by a known reduced density matrix ρ_s. We now hand this qubit over to Jos, who performs a projective measurement in the $|0\rangle, |1\rangle$ basis. Jos then returns the qubit to us, without telling us the result of his measurement.

(a) How is ρ_s transformed by this process? Write down this process using operator-sum representation:

$$\rho'_s = \sum_i M_i \rho_s M_i^\dagger.$$

Specify the two operation elements M_0 and M_1 (each one as a 2×2 matrix in the $|0\rangle, |1\rangle$ basis).

(b) How does this process transform the Bloch vector representing the qubit state? Hint: $(\langle X\rangle, \langle Y\rangle, \langle Z\rangle) \rightarrow (?, ?, ?)$.

(c) Now consider another process, known as the phase-flip channel. With probability 1/2, nothing happens. With probability 1/2, the qubit acquires a phase shift:

$$\alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|0\rangle - \beta|1\rangle.$$

What are the two operation elements M'_0 and M'_1 for this process? How does the Bloch vector transform in this case?

(d) Compare results for the transformation of the Bloch vector obtained in (b) and in (c). These transformations should be the same! This simple example shows that different processes can lead to the same quantum operation on a quantum system. Show that the operation elements M_0 and M_1 are related to the operation elements M'_0 and M'_1 by a unitary transformation

$$\begin{pmatrix} M'_0 \\ M'_1 \end{pmatrix} = S \begin{pmatrix} M_0 \\ M_1 \end{pmatrix},$$

where S is a 2×2 matrix.

14

TIME EVOLUTION OF THE DENSITY OPERATOR

In this chapter, we revisit the material covered in the previous chapter from a different viewpont: we calculate the time evolution of the density operator for a system coupled to a bath. We shall discuss the different assumptions necessary to arrive at the final result in detail: these boil down to requiring that the system under consideration is coupled sufficiently weakly to the bath that the latter can always be considered to be in equilibrium.

14.1 THE BORN-MARKOV MASTER EQUATION

In this section, we develop the formal equation describing the evolution of the density operator of a system coupled to an environment or bath – the terms 'bath' and 'environment' have the same meaning in this section. The equation we search for is the analogue for an open system of the Schrödinger equation for a closed quantum system. Of course, different environments give rise to different evolutions of the density matrix and the same holds for different couplings to the same environment. Therefore, we should anticipate that our equations will depend on the model(s) used for both the environment and the coupling. On the other hand, one may ask the question what is the most general form of this equation. We shall address this issue towards the end of this chapter.

The material discussed in this and the following sections can be found in many books. The discussion here is based on *Quantum Measurement and Control*, by H. M. Wiseman and G. J. Milburn, Cambridge University Press, 2010. Another useful text is *The Theory of Open Quantum Systems*, by H. P. Breuer and F. Petruccione, Oxford University Press, 2002.

To find the equation of motion for the system's density operator is straightforward: we start by writing up the time evolution of the 'Universe' and then we trace out the environment. However, in order to arrive at a useful and convenient result, we must make two important approximations as we shall see.

The Hamiltonian of the universe can be written as

$$H = H_\mathrm{S} + H_\mathrm{E} + V_\mathrm{SE}$$

where H_S and H_E are the Hamiltonians of the system and the environment without interaction, and V_SE is the coupling between the two.

Now we step back. In section 2.2, we have introduced the Schrödinger and Heisenberg pictures of quantum mechanics. Let us summarise them here:

- **Schrödinger picture:**

  The *states* evolve in time according to the time evolution operator $U(t) = \exp(-\mathrm{i}tH/\hbar)$:

  $$\left|\psi(t)\right\rangle = U(t)\left|\psi(0)\right\rangle.$$

  Equivalently, the time evolution of the states is given by the time-dependent Schrödinger equation:

  $$\mathrm{i}\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = H\left|\psi(t)\right\rangle.$$

  The density operator evolves in time as

  $$\rho(t) = U(t)\rho(0)U^\dagger(t).$$

  The operators do not evolve in time.

- **Heisenberg picture:**

  The *operators* $\hat{\mathscr{O}}$ evolve in time according to

  $$\hat{\mathscr{O}}_\mathrm{H}(t) = e^{\mathrm{i}Ht/\hbar}\hat{\mathscr{O}}e^{-\mathrm{i}Ht/\hbar} = U^\dagger(t)\hat{\mathscr{O}}U(t).$$

  Equivalently, they evolve according to the Heisenberg equation of motion:

  $$\frac{d}{dt}\mathscr{O}_\mathrm{H}(t) = \frac{\mathrm{i}}{\hbar}\left[H,\mathscr{O}_\mathrm{H}(t)\right].$$

  The states do not depend on time in this picture and the same holds for the density operator ρ_H.

It should be clear from this overview that the density operator evolves in a way different from a 'usual' operator: it is time dependent in the Schrödinger picture, where usual operators are time-independent, and in the Heisenberg picture we have the opposite situation.

For our problem where we have a non-interacting system described by the Hamiltonian $H_0 = H_\mathrm{S} + H_\mathrm{E}$ and an interaction V_SE, it turns out convenient to introduce a third picture: the 'interaction picture', in which the states and the operators are time dependent. The states are defined by

$$\left|\psi_\mathrm{I}(t)\right\rangle = e^{\mathrm{i}tH_0/\hbar}\left|\psi(t)\right\rangle = e^{\mathrm{i}tH_0/\hbar}e^{-\mathrm{i}t(H_0+V_\mathrm{SE})}\left|\psi_0\right\rangle.$$

Note that the two factors $\mathrm{i}tH_0$ in the exponent do not cancel due to the fact that H_0 does not commute with V_SE. The operators also evolve in the interaction picture – they are defined as

$$\hat{\mathscr{O}}_\mathrm{I}(t) = e^{\mathrm{i}H_0t/\hbar}\hat{\mathscr{O}}e^{-\mathrm{i}H_0t/\hbar}$$

and the density matrix is in this picture defined as:

$$\rho_\mathrm{I}(t) = e^{\mathrm{i}H_0t/\hbar}\rho(t)e^{-\mathrm{i}H_0t/\hbar},$$

where $\rho(t)$ is the density matrix in the Schrödinger picture. This directly gives the inverse relation

$$\rho(t) = e^{-\mathrm{i}H_0t/\hbar}\rho_\mathrm{I}(t)e^{\mathrm{i}H_0t/\hbar}. \tag{14.1}$$

Armed with this new picture, the time evolution of the density matrix can now be cast in a compact form. We now set

$$\hbar \equiv 1$$

for convenience, and start with the general relation (in the Schrödinger picture):

$$\dot{\rho} = -\mathrm{i}\left[H,\rho\right].$$

Rewriting this equation in the interaction picture gives after some calculation:

$$\dot{\rho}_{\mathrm{I}}(t) = -\mathrm{i}\left[V_{\mathrm{SE,\,I}}(t), \rho_{\mathrm{I}}(t)\right]. \tag{14.2}$$

This equation can be solved implicitly:

$$\rho_{\mathrm{I}}(t) = \rho_{\mathrm{I}}(0) - \mathrm{i}\int_0^t \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{I}}(t')\right] dt'.$$

This equation is then put back into (14.2), which then leads to

$$\dot{\rho}_{\mathrm{I}}(t) = -\mathrm{i}\left[V_{\mathrm{SE,\,I}}(t), \rho_{\mathrm{I}}(0)\right] - \int_0^t \left[V_{\mathrm{SE,\,I}}(t), \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{I}}(t')\right]\right] dt'.$$

We may iterate further but this does not turn out to be necessary for our purposes. Such an equation requires an initial condition. Without fully specifying the initial density matrix, we assume that it is of the form:

$$\rho_{\mathrm{I}}(0) = \rho_{\mathrm{S}}(0) \otimes \rho_{\mathrm{E}}(0),$$

i.e. at $t = 0$, there are no correlations between the system and the environment. This may be realised in practice by preparing a system and an equilibrated environment, and bringing them into contact at $t = 0$.

We want to take the trace of the density matrix over the environment in order to arrive at a density matrix for the system only. Doing this for the uncorrelated density operator at $t = 0$, we usually find

$$\mathrm{Tr}_{\,\mathrm{E}}\left[V_{\mathrm{SE,\,I}}(0)\rho_{\mathrm{I}}(0)\right] = 0.$$

If this is not the case, we simply add a constant to H_0:

$$H_0 \to H_0 + a I_{\mathrm{S}} \otimes I_{\mathrm{E}}$$

where $I_{\mathrm{S,E}}$ are the unit operators within the respective Hilbert spaces and a is chosen such as to make the above trace zero. We then have

$$\mathrm{Tr}_{\,\mathrm{E}}\dot{\rho}_{\mathrm{I}}(t) = -\int_0^t \mathrm{Tr}_{\,\mathrm{E}}\left(\left[V_{\mathrm{SE,\,I}}(t), \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{I}}(t')\right]\right]\right) dt'.$$

We can make this equation manageable if we realise that the coupling V_{SE} between the environment and the system is weak. Subsequent iterations lead to higher and higher powers of V_{SE} in the integral. So we may replace the ρ_{I} in that integral by a simpler form without spoiling the $\dot{\rho}_{\mathrm{I}}$ on the left hand side too much. The approximation we make for ρ_{I} is based on the physical assumption that the environment contains so many degrees of freedom that it is hardly perturbed by the system, and that it does not build up extensive correlations with the latter. This means that we assume the environment to be at equilibrium at all times. In that case we may put

$$\rho_{\mathrm{I}}(t') = \rho_{\mathrm{S,I}}(t') \otimes \rho_{\mathrm{E,I}},$$

so that the evolution of the density operator is determined by the equation

$$\dot{\rho}_{\mathrm{S,I}} = -\int_0^t \mathrm{Tr}_{\,\mathrm{E}}\left(\left[V_{\mathrm{SE,\,I}}(t), \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{S,I}}(t') \otimes \rho_{\mathrm{E,I}}(0)\right]\right]\right) dt'.$$

Note that this approximation is useful if we are mainly interested in the effect of the bath on the system – there is also a reverse effect of the system on the bath, which becomes important when the environment carries away information concerning the system, as happens during a measurement. The approximation that the environment is not influenced by the system is called the *Born approximation*. It is strongly related to the Born approximation in scattering, where we replace the full wave function, rather than the density matrix, by an unperturbed one. Finally, note that $\rho_{\mathrm{S,I}}$ and $\rho_{\mathrm{E,I}}$ are both calculated in the interaction picture.

The evolution equation for the system density operator still has a major disadvantage: the left hand side is a function of t, whereas the right hand side contains an integral over t' for which we need all the density matrices at times between 0 and t. Noting that

$$V_{\mathrm{SE,\,I}}(t) = e^{\mathrm{i}t(H_\mathrm{S}+H_\mathrm{E})} V_{\mathrm{SE}} e^{-\mathrm{i}t(H_\mathrm{S}+H_\mathrm{E})}$$

we can assume that, for a large bath, the terms with phase factor $e^{\mathrm{i}(t-t')H_\mathrm{E}}$ oscillate so quickly in the integral that their contribution averages out to zero except for $t \approx t'$. This is the *Markov condition*, which can be formulated as the requirement that

$$\mathrm{Tr}_\mathrm{E}\left(\rho_\mathrm{E} V_{\mathrm{SE,\,I}}(t) V_{\mathrm{SE,\,I}}(t')\right) \equiv \Gamma(t-t'),$$

which is recognized as the autocorrelation function of the environment, is sharply peaked around $t = t'$. In physical terms, it says that the environment decorrelates so quickly that all contributions in which t and t' are separated more than a minimum time τ, average out to zero; τ is furthermore supposed to be much smaller than the time scale at which the system changes. The evolution of the system is now governed by the equation

$$\dot\rho_\mathrm{S}(t) = -\int_0^t \mathrm{Tr}_\mathrm{E}\left(\left[V_{\mathrm{SE,\,I}}(t), \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{S,I}}(t)\otimes\rho_\mathrm{E}(0)\right]\right]\right) dt', \tag{14.3}$$

where we have omitted the subscript 'I' with $\rho_\mathrm{E}(0)$ as the different pictures are identical at $t = 0$. Given the fact that the integrand of the original integral is peaked around $t \approx t'$, we can set the lower bound of the integral to $-\infty$, which leads to

$$\dot\rho_{\mathrm{S,I}}(t) = -\int_{-\infty}^t \mathrm{Tr}_\mathrm{E}\left(\left[V_{\mathrm{SE,\,I}}(t), \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{S,\,I}}(t)\otimes\rho_\mathrm{E}(0)\right]\right]\right) dt'.$$

This is the so-called *Born-Markov* equation for the density matrix.

Summary so far We have seen that a convenient picture for describing the evolution of a system coupled to a bath is the *interaction picture*, defined by

$$\begin{aligned} \left|\psi_\mathrm{I}(t)\right\rangle &= e^{\mathrm{i}tH_0}\left|\psi(t)\right\rangle; \\ \rho_\mathrm{I}(t) &= e^{\mathrm{i}tH_0}\rho(t)e^{-\mathrm{i}H_0 t}; \\ \mathscr{O}_\mathrm{I}(t) &= e^{\mathrm{i}tH_0}\mathscr{O}e^{-\mathrm{i}tH_0}. \end{aligned}$$

In this picture, the evolution of a system can be derived if we make the following two assumptions:

- The environment is considered to remain in the initial state and does not evolve significantly from its state at $t = 0$ (Born approximation).
- The interactions with the environment decorrelate quickly in time (Markov approximation).

The result is therefore called the *Born-Markov equation*:

$$\dot\rho_{\mathrm{S,\,I}}(t) = -\int_{-\infty}^t \mathrm{Tr}_\mathrm{E}\left(\left[V_{\mathrm{SE,\,I}}(t), \left[V_{\mathrm{SE,\,I}}(t'), \rho_{\mathrm{S,\,I}}(t)\otimes\rho_\mathrm{E}(0)\right]\right]\right) dt'. \tag{14.4}$$

The name Redfield or Bloch-Redfield equation is often given to this or to a very similar equation.

We shall now consider some concrete examples to illustrate the ideas and to derive some physically relevant results.

14.2 EXAMPLES

14.2.1 THE DAMPED HARMONIC OSCILLATOR

We consider a harmonic oscillator S with frequency ω_0 coupled to a bath E of other harmonic oscillators with frequencies $\omega_{\mathbf{k}}$, one for each $\mathbf{k}$-vector in a cavity. The Hamiltonian is

$$H = H_{\rm S} + H_{\rm E} + V_{\rm SE},$$

where

$$H_{\rm S} = \omega_0 \left(a^\dagger a + 1/2 \right); \qquad H_{\rm E} = \sum_{\mathbf{k}} \left(b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + 1/2 \right);$$

$$V_{\rm SE} = \sum_{\mathbf{k}} g_{\mathbf{k}} \left(a^\dagger + a \right) \left(b^\dagger_{\mathbf{k}} + b_{\mathbf{k}} \right).$$

The environment, described by the creation and annihilation operators $b, b^\dagger$'s, could for example be a 1D chain of N harmonic oscillators like the one studied in section 8.2 (a quick revision of the harmonic oscillator is strongly recommended!). Note that the last term represents a coupling which, for a 1D system, would be of the form $x_a x_0$, as $x_a \propto a + a^\dagger$, and the position x_0 of the zeroeth oscillator of the environment would be

$$x_0 = \frac{1}{N} \sum_{\mathbf{k}} x_{\mathbf{k}} \propto \sum_{\mathbf{k}} \left(b^\dagger_{\mathbf{k}} + b_{\mathbf{k}} \right).$$

The bath has a density matrix

$$\rho_{\rm E} = \frac{1}{Z} \exp(-\beta H_{\rm E}),$$

where $\beta = 1/(k_{\rm B} T)$ and. Note that this can also be written as

$$\rho_{\rm E} = \frac{1}{Z} \prod_{\mathbf{k}} |n_{\mathbf{k}}\rangle \langle n_{\mathbf{k}}| \, e^{-\beta\omega_{\mathbf{k}}(n_{\mathbf{k}}+1/2)},$$

with

$$Z = \prod_{\mathbf{k}} \left[\sum) n_{\mathbf{k}} = 0^{\infty} e^{-\beta\omega_{\mathbf{k}}(n_{\mathbf{k}}+1/2)} \right].$$

In the interaction picture, the a and $b_{\mathbf{k}}$ become time-dependent – their time dependence is caused by the unperturbed Hamiltonian and we obtain:

$$V_{\rm SE,\,I}(t) = \sum_{\mathbf{k}} g_{\mathbf{k}} \left(a^\dagger e^{{\rm i}\omega_0 t} + a e^{-{\rm i}\omega_0 t} \right) \left(b_{\mathbf{k}} e^{-{\rm i}\omega_{\mathbf{k}} t} + b^\dagger_{\mathbf{k}} e^{{\rm i}\omega_{\mathbf{k}} t} \right)$$

where the time dependence has been made explicit in the form of the phase factors. All ω's are positive. The time-dependent phase factors rotate in the complex plane, and their *average* effect is expected to decay rapidly with time. An exception occurs when $\omega_{\mathbf{k}} \approx \omega_0$, and we usually only keep the terms which are expected to give a non-negligible transition rate:

$$V_{\rm SE,\,I}(t) = \sum_{\mathbf{k}} g_{\mathbf{k}} \left(a^\dagger b_{\mathbf{k}} e^{{\rm i}(\omega_0 - \omega_{\mathbf{k}}) t} + b^\dagger_{\mathbf{k}} a e^{-{\rm i}(\omega_0 - \omega_{\mathbf{k}}) t} \right).$$

This expression will give significant contributions when $\omega_{\mathbf{k}}$ is close to ω_0. This approximation which consists of eliminating the rapidly oscillating phase factors is called the *rotating wave approximation.*

Expanding the double commutator occurring in the Born-Markov equation seems cumbersome: it gives rise to no less than 16 terms. However, half of these vanish since the trace over the terms $b_{\mathbf{k}} b_{\mathbf{k}'}$ and $b^\dagger_{\mathbf{k}} b^\dagger_{\mathbf{k}'}$ vanishes (why?). The other two possible combinations give (see problem 1)

$$\mathrm{Tr}_{\rm E} \left(\rho_{\rm E} b^\dagger_{\mathbf{k}} b_{\mathbf{k}'} \right) = \delta(\mathbf{k} - \mathbf{k}') \overline{n}_{\mathbf{k}}$$

and

$$\mathrm{Tr}_{\,\mathrm{E}}\left(\rho_{\mathrm{E}} b_{\mathbf{k}} b^{\dagger}_{\mathbf{k}'}\right) = \delta(\mathbf{k}-\mathbf{k}')\left(\overline{n}_{\mathbf{k}}+1\right).$$

Apart from pre-factors, sums over **k** and the integral over t, but after having taken the trace over the environment Hilbert space, we are then left with the following:

$$\begin{aligned}
&aa^{\dagger}\rho_{\mathrm{S,I}}\overline{n}_{\mathbf{k}}e^{-\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} + a^{\dagger}a\rho_{\mathrm{S,I}}(\overline{n}_{\mathbf{k}}+1)e^{\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} - \\
&a\rho_{\mathrm{S,I}}a^{\dagger}\overline{n}_{\mathbf{k}}e^{-\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} - a^{\dagger}\rho_{\mathrm{S,I}}a(\overline{n}_{\mathbf{k}}+1)e^{\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} - \\
&a^{\dagger}\rho_{\mathrm{S,I}}a^{\dagger}\overline{n}_{\mathbf{k}}e^{-\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} - a\rho_{\mathrm{S,I}}a^{\dagger}(\overline{n}_{\mathbf{k}}+1)e^{\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} + \\
&\rho_{\mathrm{S,I}}aa^{\dagger}\overline{n}_{\mathbf{k}}e^{\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')} + \rho_{\mathrm{S,I}}a^{\dagger}a(\overline{n}_{\mathbf{k}}+1)e^{-\mathrm{i}(\omega_0-\omega_{\mathbf{k}})(t-t')}.
\end{aligned}$$

Two operations remain to be performed: the sum over the modes **k** and the integral over the time t. It can be shown that

$$\int_{-\infty}^{0} e^{\pm\mathrm{i}\omega t} = \int_{0}^{\infty} e^{\mp\mathrm{i}\omega t} = \pi\delta(\omega) \mp \mathrm{i}\mathscr{P}\left(\frac{1}{\omega}\right).$$

It turns out that the imaginary contribution shifts the frequency over a small amount – we neglect this effect. Replacing the sum over **k** by an integral over the energies ω:

$$\sum_{\mathbf{k}} \to \int \eta(\omega)\,d\omega$$

where $\eta(\omega)$ represents the density of states, we finally obtain, putting $\gamma = 2\pi\eta(\omega)\left|g_{\mathbf{k}}\right|^2$:

$$\begin{aligned}
\dot{\rho}_{\mathrm{S,\,I}} = \frac{\gamma}{2}\overline{n}(\omega_0)\left(2a^{\dagger}\rho_{\mathrm{S,I}}a - aa^{\dagger}\rho_{\mathrm{S,I}} - \rho_{\mathrm{S,I}}aa^{\dagger}\right) + \frac{\gamma}{2}\left(\overline{n}(\omega_0)+1\right)\left(2a\rho_{\mathrm{S,I}}a^{\dagger} - a^{\dagger}a\rho_{\mathrm{S,I}} - \rho_{\mathrm{S,I}}a^{\dagger}a\right)& \\
\equiv \mathscr{L}\rho_{\mathrm{S,I}},& \quad (14.5)
\end{aligned}$$

where we have introduced the Lindblad operator $\mathscr{L}$, acting on ρ_{S}.

We now want to evaluate the time evolution of different physical quantities (represented by Hermitian operators). We first calculate the time dependence of the expectation value of a:

$$\frac{d}{dt}\langle a\rangle(t) = \frac{d}{dt}\mathrm{Tr}_{\,\mathrm{S}}\left(a_{\mathrm{I}}(t)\rho_{\mathrm{S,\,I}}(t)\right) = -\mathrm{i}\langle\omega_0 a_{\mathrm{I}}\rangle + \mathrm{Tr}_{\,\mathrm{S}}\left(\mathscr{L}\rho_{\mathrm{S,I}}a_{\mathrm{I}}\right),$$

where the first term on the right hand side derives from $\dot{a}_{\mathrm{I}}(t)$. The rightmost term gives six contributions (see the form of $\mathscr{L}$). The first of these has the form:

$$\frac{\gamma}{2}\overline{n}(\omega_0)\mathrm{Tr}_{\,\mathrm{S}}\left(2a^{\dagger}\rho_{\mathrm{S,I}}aa_{\mathrm{I}}(t)\right) = \frac{\gamma}{2}\overline{n}(\omega_0)\mathrm{Tr}_{\,\mathrm{S}}\left(2a^{\dagger}\rho_{\mathrm{S,I}}aae^{-\mathrm{i}\omega t}\right) = \frac{\gamma}{2}\overline{n}(\omega_0)\mathrm{Tr}_{\,\mathrm{S}}\left(2\rho_{\mathrm{S,I}}a^2a^{\dagger}e^{-\mathrm{i}\omega t}\right).$$

Working out the next two terms in a similar fashion and adding them, yields the term proportional to $\overline{n}(\omega_0)$:

$$\frac{\gamma}{2}\overline{n}(\omega_0)\mathrm{Tr}_{\,\mathrm{S}}\left(\rho_{\mathrm{S,I}}a_{\mathrm{I}}(t)\right) = \frac{\gamma}{2}\overline{n}(\omega_0)\langle a\rangle.$$

The next three terms (proportional to $\overline{n}(\omega_0)+1$) combine into:

$$-\frac{\gamma}{2}\left[\overline{n}(\omega_0)+1\right]\langle a\rangle,$$

so that, collecting all terms, we obtain

$$\frac{d}{dt}\langle a\rangle(t) = -\mathrm{i}\omega\langle a\rangle(t) - \frac{\gamma}{2}\langle a\rangle(t).$$

The first term on the right hand side arises from the commutator with the (unperturbed) Hamiltonian; the Lindblad operator yields the damping term γ. We see that after a long time, the expectation value of a reduces to a simple oscillation.

We can also work out the relaxation of the energy. This is addressed in problem 2.

14.2.2 SPONTANEOUS EMISSION FROM AN ELECTRONIC EXCITATION IN AN ATOM

The environment is in this case composed of the photons of the field. The Hamiltonian of the environment can be written as

$$H_{\mathrm{E}} = \sum_k \omega_k a_k^\dagger a_k$$

where we use k as a general index containing information about the wave vector $\mathbf{k}$ and the polarization (there are two transverse polarization modes in vacuum) – the sum is therefore over *all* independent modes. We have furthermore put $\hbar \equiv 1$. The creation and annihilation operators $a_k^\dagger$ and a_k satisfy the boson commutation relation

$$[a_k, a_l^\dagger] = \delta_{kl}.$$

The interaction between the electron and the electromagnetic field follows from the Hamiltonian of a charged particle in an electromagnetic field:

$$H = \frac{1}{2m}\left(\mathbf{p} + e\mathbf{A}(\mathbf{r}, t)\right)^2 - e\varphi(\mathbf{r}, t).$$

We use the *dipole approximation* which takes the wavelength of electromagnetic waves to be much larger than the size of the atom. This is realistic, as the size of the atom is of the order of Angstroms, whereas the wavelength for light inducing an atomic transition is typically three orders of magnitude larger.

We thus take

$$\mathbf{A}(\mathbf{r}, t) \to \mathbf{A}(t).$$

The term proportional to $\mathbf{A}(t)^2$ is thus a constant oscillating field. We neglect the influence of this term, which is typically very small. We are then left with the interaction term

$$V_{\mathrm{SE}}(t) = \frac{e}{m}\mathbf{A}(t)\cdot\mathbf{p}.$$

and the atomic Hamiltonian is

$$H_{\mathrm{S}} = \frac{p^2}{2m} - e\phi(\mathbf{r}).$$

Expressed in terms of the ground state $|g\rangle$ and the excited state $|e\rangle$ of the atom, this Hamiltonian takes the form

$$H_{\mathrm{S}} = \frac{\omega_a}{2}\sigma_z,$$

where σ_z is the Pauli-matrix operator which works in the two-dimensional Hilbert space spanned by $|g\rangle$ and $|e\rangle$. It is a diagonal operator with eigenvalues $\langle e|\sigma_z|e\rangle = 1$ and $\langle g|\sigma_z|g\rangle = -1$. The energy difference between ground and excited state is ω_a.

For the vector potential occurring in V_{SE} we have, putting the atom at $\mathbf{r} = 0$:

$$\mathbf{A}(t) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\epsilon_0\omega_{\mathbf{k},\alpha}}}\hat{\boldsymbol{\epsilon}}_\alpha\left[a_{\mathbf{k},\alpha}(t) + a^\dagger_{-\mathbf{k},\alpha}(t)\right],$$

with $a_{\mathbf{k},\alpha}(t) = e^{-\mathrm{i}\omega_{\mathbf{k}}t}a(0)$ (see section 8.3.2). Note that we have replaced the integral over the $\mathbf{k}$-modes by a sum, including the factor $1/\sqrt{V}$.

The other part of the interaction Hamiltonian V_{SE} is the momentum. We find a suitable form for this via a trick. We note that

$$\mathbf{p} = \mathrm{i}m\,[H_{\mathrm{S}}, \mathbf{r}]$$

as can easily be verified. We therefore find, expressed in the basis $|g\rangle, |e\rangle$

$$\langle g|\mathbf{p}|e\rangle = \mathrm{i}m\langle g|[H_{\mathrm{S}}, \mathbf{r}]|e\rangle = -\mathrm{i}m\omega_a\langle g|\mathbf{r}|e\rangle$$

with $\omega_a = E_e - E_g$. The diagonal matrix elements vanish due to anti-symmetry of the integral: $\int \mathbf{r}\left|\psi(\mathbf{r})\right|^2 d^3r = 0$ due to the fact that $\psi(\mathbf{r}) = \pm\psi(-\mathbf{r})$. All in all we see that the operator $\mathbf{p}$ can be written as

$$\mathbf{p} = \mathrm{i}m\omega_a\,(\sigma_- - \sigma_+)\left\langle g\middle|\mathbf{r}\middle|e\right\rangle,$$

where

$$\sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

A trivial basis transformation:

$$\begin{aligned} |g\rangle &\to |g\rangle \\ |e\rangle &\to \mathrm{i}\,|e\rangle \end{aligned}$$

transforms the expression for $\mathbf{p}$ into

$$\mathbf{p} = m\omega_a\,(\sigma_- + \sigma_+)\left\langle g\middle|\mathbf{r}\middle|e\right\rangle.$$

Note that we have not changed the *number* $\left\langle g\middle|\mathbf{r}\middle|e\right\rangle$; we have only changed the operator representation. We envisage that the relevant modes which excite or de-excite the atom have frequencies close to the transition energy: $\omega_{\mathbf{k}} \approx \omega_a$.

Note that, so far, we have taken the $\mathbf{A}$-field to be time dependent, whereas the $\sigma_\pm$ were taken time-independent. In the interaction picture, they vary with time as

$$\begin{aligned} \sigma_+(t) &= \sigma_+ e^{\mathrm{i}\omega_a t}; \\ \sigma_-(t) &= \sigma_- e^{-\mathrm{i}\omega_a t}. \end{aligned}$$

All in all, we find that the interaction Hamiltonian takes the form (with the factors $\hbar$ restored):

$$V_{\mathrm{SE,\,I}}(t) = \sum_{\mathbf{k},\alpha}\sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 V}}\,\hat{\boldsymbol{\epsilon}}_\alpha\cdot\left\langle g\middle|\mathbf{r}\middle|e\right\rangle\left(\sigma_- e^{-\mathrm{i}\omega_a t} + \sigma_+ e^{\mathrm{i}\omega_a t}\right)\left(a_{\mathbf{k}} e^{-\mathrm{i}\omega_{\mathbf{k}} t} + a^\dagger_{\mathbf{k}} e^{\mathrm{i}\omega_{\mathbf{k}} t}\right).$$

Lumping all prefactors into a coupling constant $g_{\mathbf{k}}$, we obtain:

$$V_{\mathrm{SE,I}} = \sum_{\mathbf{k}} g_{\mathbf{k}}(a_{\mathbf{k}} e^{-\mathrm{i}\omega_{\mathbf{k}} t} + a^\dagger_k e^{\mathrm{i}\omega_k t})(\sigma_+ e^{\mathrm{i}\omega_a t} + \sigma_- e^{-\mathrm{i}\omega_a t});$$

with

$$g_{\mathbf{k}} = \sum_\alpha \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 V}}\,\hat{\boldsymbol{\epsilon}}_\alpha\cdot\mathbf{d}, \tag{14.6}$$

where $\mathbf{d}$ is the matrix element of the dipole moment between the ground and excited state.

Working out the product in the sum gives terms with $\exp\left[\pm\mathrm{i}(\omega_k + \omega_a)t\right]$ and terms with $\exp\left[\pm\mathrm{i}(\omega_k - \omega_a)t\right]$. Just as in the previous section, we neglect the terms of the first form as they give a negligible contribution (rotating wave approximation). In this approximation, the interaction reads:

$$V_{\mathrm{SE,I}} = \sum_k g_k\left[a_k\sigma_+ e^{\mathrm{i}(\omega_a - \omega_k)t} + a^\dagger_k\sigma_- e^{-\mathrm{i}(\omega_a - \omega_k)t}\right].$$

The Hamiltonian describing the system with this interaction is called the *Jaynes-Cummings* Hamiltonian. It is omnipresent in systems where bosons interact with fermions.

We now have all the ingredients for formulating the Born-Markov equation for this case. We do this for very low temperatures, in which the electromagnetic field is in its ground state (all occupations 0). In addition to terms with two creation or two boson annihilation operators, the expectation value

$$\left\langle 0\middle| b^\dagger_{\mathbf{k},\alpha} b_{\mathbf{k},\alpha}\middle|0\right\rangle = 0.$$

Working out all the terms in the Born-Markov equation leads to (we leave out the subscripts 'I'):

$$\dot{\rho}_S = -\int_{-\infty}^{t} \Gamma(t-t')\left[\sigma_+\sigma_-\rho_S(t') - \sigma_-\rho_S(t')\sigma_+\right] dt' + \text{h.c.} \tag{14.7}$$

Here h.c. denotes Hermitian conjugate and

$$\Gamma(\tau) = \sum_k g_k^2 e^{-i(\omega_k - \omega_a)\tau}.$$

You are strongly advised to verify this (see problem 1).

In analogy with the harmonic oscillator problem of the previous section, we have

$$\int_0^\infty \Gamma(\tau) d\tau = \frac{\gamma}{2} - \Delta\omega_a.$$

We then obtain the Born-Markov equation:

$$\dot{\rho}_S(t) = -i\frac{\Delta\omega_a}{2}\left[\sigma_z, \rho_S(t)\right] + \gamma\sigma_-\rho_S(t)\sigma_+ - \frac{\gamma}{2}\left(\sigma_+\sigma_-\rho_S(t) + \rho_S(t)\sigma_+\sigma_-\right).$$

We shall neglect the term with $\Delta\omega$ – this is called the *Lamb shift*. From the definition of $\Gamma(\tau)$ it can be seen that the Lamb shift only occurs when g is not symmetrically distributed around ω_k: if there are for example more states in the bath with frequency $\omega_k > \omega_a$ than the other way round, the effective frequency of the atom is slightly shifted upward. Calculating the Lamb shift is quite difficult; it is one of the major exercises of quantum electrodynamics. We refrain from going into details here.

The term on the right hand side turns out to generate a non-unitary time evolution. For a *closed* quantum system, we expect only unitary time evolutions. The fact that this is not the case here reflects the leak of information to the environment. Let us make this more explicit by calculating the time evolution of the state of the atom. This is found from

$$\frac{d}{dt}\langle\sigma_z\rangle = \text{Tr}_S\left(\dot{\rho}_S\sigma_z\right).$$

Using the above equation for $\dot{\rho}_S$ we obtain, with $z \equiv \langle\sigma_z\rangle$:

$$\dot{z} = -\gamma(z+1),$$

which decays from a starting value in the excited state $z = 1$ to -1 with a decay time $1/\gamma$.

The value of the emission rate can be calculated: it is given as $\left|g_{\mathbf{k}}\right|^2 \eta(\omega)$, with η the density of states. We can calculate the latter straightforwardly:

$$\sum_{\mathbf{k}} = \frac{V}{(2\pi)^3}\int d^3k = \frac{V}{2\pi^2}\int \frac{\omega^2}{c^3} d\omega,$$

so that

$$\eta(\omega) = \frac{V}{2\pi^2}\frac{\omega^2}{c^3}.$$

We furthermore need to calculate $\left|\langle g|\mathbf{d}|e\rangle \cdot \hat{\boldsymbol{\epsilon}}_\alpha\right|^2$ which occurs in the expression for $g_{\mathbf{k}}$ see eq. 14.6. Now we use a symmetry argument to evaluate this expression:

$$\sum_\alpha \left|\langle g|\mathbf{d}|e\rangle\hat{\boldsymbol{\epsilon}}_\alpha\right|^2 = \frac{2}{3}\left|\langle g|\mathbf{d}|e\rangle\right|^2.$$

The factor 2/3 arises because there are 2 polarization directions, and for each direction we consider the component of the vector $\left|\langle g|\mathbf{d}|e\rangle\right|^2$ along that polarization direction. Averaging over all polarization directions of the field then leads to 1/3 because isotropy. All in all we obtain, restoring the factor $\hbar$:

$$\gamma = 2\pi\eta \frac{2}{3}\frac{\omega_a}{2\epsilon_0} \left|\langle g | \mathbf{d} | e \rangle\right|^2 = \frac{\omega_a^3 |\mathbf{d}|^2}{3\pi\hbar\epsilon_0 c^3}.$$

This is the spontaneous emission rate.

It is instructive to work out the Lindblad operator in the space of the states $|g\rangle$ and $|e\rangle$. In that space, the density operator is a 2×2 matrix. Using the forms of σ_+ and σ_- we directly see that

$$\dot\rho_{\mathrm{S}} = \frac{1}{dt}\begin{pmatrix} \rho_{++} & \rho_{+-} \\ \rho_{+-} & \rho_{--} \end{pmatrix} = \gamma \begin{pmatrix} \rho_{--} & -\rho_{+-}/2 \\ -\rho_{-+}/2 & -\rho_{--} \end{pmatrix}.$$

From this we can easily calculate the time evolution of the three components of the Bloch vector. We call these components X, Y and Z:

$$\begin{aligned} \frac{d}{dt}\langle X\rangle &= -\frac{\gamma}{2}\langle X\rangle, \\ \frac{d}{dt}\langle Y\rangle &= -\frac{\gamma}{2}\langle Y\rangle, \\ \frac{d}{dt}\langle Z\rangle &= -\gamma\langle Z\rangle + \gamma. \end{aligned}$$

Fortunately, these equations are independent and we can solve them at once, finding

$$\begin{aligned} \langle X\rangle(t) &= e^{-\gamma t/2}\langle X\rangle(0), \\ \langle Y\rangle(t) &= e^{-\gamma t/2}\langle Y\rangle(0), \\ \langle Z\rangle(t) &= e^{-\gamma t}\langle Z\rangle(0) + 1 - e^{-\gamma t}. \end{aligned}$$

14.3 PROBLEMS

1. Consider a boson bath described by a density matrix ρ_{B}. The expectation value for the number of bosons in mode k is given by

$$\langle n_k\rangle = \mathrm{Tr}\left(\rho_{\mathrm{B}} b_k^\dagger b_k\right).$$

   (a) Show that $\mathrm{Tr}\left(b_k \rho_{\mathrm{B}} b_k^\dagger\right) = \langle n_k\rangle$.
   (b) Show that $\mathrm{Tr}\left(b_k^\dagger \rho_{\mathrm{B}} b_k\right) = \langle n_k\rangle + 1$.
   (c) Derive Eq. (14.7) of the lecture notes. This equation is derived for the vacuum state $\langle n_k\rangle = 0$.
   (d) Derive the same equation if the system is not in the vacuum state.

2. *Energy damping in the damped harmonic oscillator* In section 14.2.1, we have derived the evolution equation for the density matrix of the harmonic oscillator coupled to a bath of other oscillators. From the result, Eq. 14.5 calculate the time evolution of the expectation value of the operator $a^\dagger a$.

3. In honour of the Physics Nobel prize 2012 (Serge Haroche and David Wineland), we consider a cavity quantum electrodynamics system. We consider two atoms (A and B) and one single-mode cavity, all with matching transition frequencies. We realise an indirect quantum measurement of the two-atom system by interacting each atom in sequence (first A, then B) with the cavity and then measuring the photon number n in the cavity. We work in the interaction picture that makes the non-interacting Hamiltonian terms disappear, leaving us only with the interaction term (in the rotating-wave approximation):

$$H_I = \gamma_A(t)\left(a^\dagger\sigma_{-,A} + a\sigma_{+,A}\right) + \gamma_B(t)\left(a^\dagger\sigma_{-,B} + a\sigma_{+,B}\right).$$

Initially, the two-atom system is in state $|\psi\rangle$, and the cavity in the $n = 0$ Fock state. To make atom A interact with the cavity, we turn on γ_A to γ at $t = 0$, for a time τ – γ_B is kept at zero. At $t = \tau$, we turn γ_A off and turn γ_B to γ for another time τ. At $t = 2\tau$, we turn γ_B off. After these interactions, we perform a photon-number measurement.

(a) We choose τ so that an excitation in A is fully transferred to the cavity. Please express τ in terms of γ.

(b) Consider $|\psi\rangle = |g_A g_B\rangle$ (both atoms in ground state). What is the state of the atoms+cavity universe after the two interaction steps?

(c) Repeat for $|\psi\rangle = |g_A e_B\rangle$ (atom B excited).

(d) Repeat for $|\psi\rangle = |e_A g_B\rangle$ (atom A excited).

(e) Finally, repeat for $|\psi\rangle = |e_A e_B\rangle$ (both atoms excited).

(f) Write the three operation elements M_n corresponding to measuring $n = 0$, 1 or 2 photons in the cavity. Verify that these operation elements form a decomposition of unity.

(g) What is the measurement operator acting in the two-atom Hilbert space giving the expected value of the measurement result $\overline{n}$? Please write this operator as a matrix expressed in the basis $\{|g_A g_B\rangle, |g_A e_B\rangle, |e_A g_B\rangle, |e_A e_B\rangle\}$.

4. *Collective coupling in cavity QED*
In this problem we return to the Nobel-prize winning field of Cavity QED. We consider the problem of N identical two-level atoms coupling with equal strength γ to a single mode of a cavity. The cavity mode frequency is resonant with the atomic transition. Working in the interaction picture, the effective Hamiltonian in the rotating-wave approximation is given by

$$H_I = \gamma \sum_{i=1}^{N} \left(a^\dagger \sigma_{-,i} + a\sigma_{+,i} \right),$$

where as usual $a^\dagger$ and a are creation and annihilation operators for the cavity mode, and $\sigma_{+,i}$ and $\sigma_{-,i}$ are raising and lowering operators for atom i. We denote the ground state of atom i by $|g_i\rangle$ and the excited state by $|e_i\rangle$. Written in the basis $\{|g_i\rangle, |e_i\rangle\}$,

$$\sigma_{+,i} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } \sigma_{-,i} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The initial state ($t = 0$) of the system is a 1-photon Fock state of the cavity, with all atoms in the ground state:

$$|\Psi(t=0)\rangle = |n=1\rangle \otimes |g_1 ... g_N\rangle.$$

(a) Calculate the action of H_I^2 on $|\psi(t=0)\rangle$.

(b) Give an explicit expression for $|\psi(t)\rangle$ by acting with the time evolution operator $\exp(-\mathrm{i}tH_I/\hbar)$ on the initial state. Show that the state of the atoms+cavity oscillates coherently between the initial state and another (normalized) state. Specify this other state and the oscillation frequency as a function of γ and N.

(c) Imagine instead that the system is initially in the state

$$|\Psi(t=0)\rangle = |n=0\rangle \otimes \sum_{i=1}^{N} (-1)^i |g_1 ... g_{i-1} e_i g_{i+1} ... g_N\rangle$$

and suppose N is even. Describe the time evolution in this case.

5. In this problem we consider the interaction between one particular mode of an electromagnetic field with an atom that can be in the ground state $|g\rangle$ or in the excited state $|e\rangle$. The electric field mode can in practice be realized using a cavity, whose dimensions precisely accommodate the single mode, which we assume to have a frequency ω_0. The set of orthonormal states $|g,n\rangle$ and $|e,n\rangle$ forms a basis of the Hilbert space of the cavity + atom. Here $|n\rangle$ denotes a cavity state containing n photons. In this problem, we assume that the cavity is tuned to the frequency corresponding to the energy difference between the two atomic states $|g\rangle$ and $|e\rangle$. The Hamiltonian of the atom is

$$H_{\mathrm{A}} = -\frac{\hbar\omega_A}{2}\sigma_z,$$

where σ_z is the Pauli matrix which operates in the Hilbert space spanned by

$$|g\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |e\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and $\omega_0 = \omega_A$. The interaction between the field and the cavity is given by

$$W = \gamma\left(\sigma_+ a + \sigma_- a^\dagger\right),$$

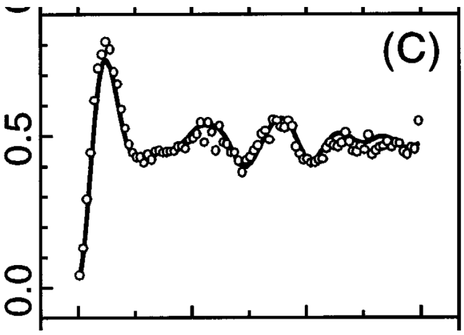
where coefficient γ sets the coupling strength, a and $a^\dagger$ are the creation and annihilation operators for photons in the cavity, and $\sigma_\pm$ are the operators which move the atom from ground to excited state (σ_+) and viceversa (σ_-).

(a) Explain the two processes described by the Hamiltonian W.

(b) Let H_0 be the Hamiltonian of the atom plus the field, without the interaction. Give the eigenstates and eigenvalues of H_0. Give the degeneracies of the eigenstates.

(c) Determine the eigenstates of $H = H_0 + W$ and the corresponding energies. Show that this problem reduces to the diagonalisation of a set of 2×2 matrices. You should find the values $\hbar\omega_0 n \pm \gamma\sqrt{n}$ for the energy eigenvalues.

(d) Now we assume that at $t = 0$, the cavity is in a coherent state α, and the atom is in its excited state $|e\rangle$, i.e., cavity plus atom are in the state $|\alpha\rangle \otimes |e\rangle$. Calculate the probability of finding, at time T, the atom in the ground state $|g\rangle$. The result is a series expansion in n. Plot the series, cutting it off for different values of n.



15

(MORE THAN A) SURVIVAL GUIDE TO SPECIAL RELATIVITY

This chapter is more extensive due to the fact that I had these lecture notes lying around from about ten years ago. It contains more than you need on special relativity. What you need can be found in chapter 31 of Desai.

15.1 HISTORY AND EINSTEIN'S POSTULATES

The theory of classical mechanics based on Newton's laws gives an excellent description of everyday life systems. At speeds of the order of the speed of light, relativistic effects come into play and necessitate a different formulation of mechanics. It was Einstein who in 1905 gave a definite physical interpretation to the mathematics which was already known in essence. This is the *special theory of relativity*.

The theory of special relativity is based on two fundamental postulates, one of which holds for classical mechanics as well, and which is known as *Galilei invariance*. This postulate pertains to *inertial frames*. We shall frequently use the name *reference frame* for the same thing. An inertial frame is a system which moves at constant velocity, i.e. bodies standing still in this frame are not subject to any force (that is, the forces acting on it add up to zero). Obviously the same holds for uniformly moving objects in that frame. Galilei invariance for a physical theory can be formulated as follows:

- It is impossible to determine the absolute velocity of an inertial frame. Only the relative velocity of two inertial frames can be determined. Physical laws are the same in *all* inertial frames.

Galilei invariance holds for classical mechanics, but not if classical mechanics is combined with Maxwell's equations, which describe the phenomena of electricity and magnetism. In fact, Maxwell's equations predict electromagnetic waves to move (in vacuum) at the speed of light. It is not clear from Maxwell's equations how waves emitted by a source moving with velocity $\mathbf{v}$ can be described, as in the context of classical mechanics, these waves should move at a speed $c\hat{\mathbf{k}}+\mathbf{v}$ (the unit vector $\hat{\mathbf{k}}$ defines the direction of the radiation) which would mean that we obtain a speed different from that of light, but this behaviour is not obtained by Maxwell's equations. The solution to this problem comes from the notion (or rather the experimental fact) that the speed of light is independent of the velocity of the source. This is formulated in the second postulate:

- The speed of light is the same for observers in different inertial frames.

This postulate is a very counter-intuitive one. If you drive a very fast car, the headlights emit EM radiation at the speed of light. But by an observer standing still at the road that radiation is also perceived as moving the speed of light! The validity of this postulate has been firmly established through experiments, of which the *Michaelson-Morley experiment* stands out as one of the landmarks of experimental physics. From Einstein's two postulates the full theory of special relativity can be derived.

15.2 THE LORENTZ TRANSFORMATION

The Lorentz transformation directly follows from the fact that the speed of light is independent of the reference frame you are in. The Lorentz transformation for systems with one space and one time dimension relates the space coordinates, x and x', and the time coordinates, t and t', where the 'primed' system is moving with a velocity $v = \beta c$ with respect to the 'unprimed' system. The first Lorentz equation is

$$x' = x/\gamma - \beta t', \tag{15.1}$$

or

$$x = \gamma(x' + \beta t'). \tag{15.2}$$

Here we have introduced $\gamma = (1 - v^2/c^2)^{-1/2}$. Another notational convention is $\beta = v/c$. Finally, c is usually taken equal to 1. From now on, we shall conform to these units and this notation, unless stated otherwise:

$$\beta = v/c; \qquad c \equiv 1 \tag{15.3a}$$

$$\gamma = 1/\sqrt{1-(v/c)^2} = 1/\sqrt{1-\beta^2}. \tag{15.3b}$$

A similar formula which relates t to x' and t' is given by

$$t' = \gamma(t - \beta x) \tag{15.4}$$

which, together with (15.2) gives

$$t = \gamma(t' + \beta x'). \tag{15.5}$$

The conclusion is that for a fixed time t' in the observer's frame, moving clocks which are synchronised in their rest frame, indicate different times! In other words: synchronous time has a meaning in one and the same reference frame, but it is not an invariant condition for different reference frames.

These are the *Lorentz equations*, which give the relation between space-time points x, t in the primed and unprimed frame:

$$t = \gamma(t' + \beta x'); \tag{15.6a}$$

$$x = \gamma(x' + \beta t'). \tag{15.6b}$$

It is clear that the inverse is obtained by swapping primed and unprimed quantities and setting $\beta \to -\beta$:

$$t' = \gamma(t - \beta x); x' = \gamma(x - \beta t). \tag{15.7a}$$

Restoring the factors c in these equations, they read:

$$t' = \gamma(t - \beta x/c); \tag{15.8a}$$
$$x' = \gamma(x - \beta ct). \tag{15.8b}$$

Having these transformation equations, we can derive a transformation equation for the velocity $\mathbf{u} = d\mathbf{r}/dt$. Consider two reference frames moving with respect to each other. The relative velocity is oriented along the x-axis. We now calculate the velocity components along the x and y axis of a particle which moves in the primed system with a velocity $\mathbf{u}'$.[1] It is easy to see that

$$u_x = \frac{dx}{dt} = \frac{d\gamma(x' + \beta t')}{d\gamma(t' + \beta x')} = \frac{u'_x + \beta}{1 + \beta u'_x}. \tag{15.9}$$

In 'full units', i.e. without putting $c = 1$, this reads:

$$u_x = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}}. \tag{15.10}$$

Now consider the y-component. This transforms according to

$$u_y = \frac{dy'}{d\gamma(t' + \beta x')} = \frac{1}{\gamma}\frac{u'_y}{1 + \beta u'_x} \tag{15.11}$$

which can be written out as

$$u_y = \sqrt{1 - (v/c)^2}\,\frac{u'_y}{1 + \frac{vu'_x}{c^2}}. \tag{15.12}$$

We see that the velocity does not transform according to the Lorentz transformation.

15.3 MORE ABOUT THE LORENTZ TRANSFORMATION

The second postulate of special relativity directly leads to the Lorentz transformation. It can easily be shown that for any two points in space time separated by a distance Δt in the time direction and by Δx in the space direction, Lorentz transformation leaves the quantity

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 \tag{15.13}$$

invariant. For three spatial dimensions plus one temporal dimension, the invariant quantity Δs is

$$\Delta s^2 = c^2 \Delta t^2 - \Delta r^2, \qquad \Delta r = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}. \tag{15.14}$$

We have only given the Lorentz transformation for the case where one frame moves with respect to the other with a relative speed which is directed along the x-direction. The general Lorentz transformation is represented by a 4 × 4 matrix. It not only describes relative motions in different directions, but also rotations and reflections in the spatial part, as this also preserves the quantity Δs^2. We shall not give the full matrix expression here.

The Lorentz transformation can be represented in a graphical way – see figure 15.1. The left picture shows the transformation starting from a Cartesian (x, t) space-time, and the right hand shows the inverse, i.e., starting from the (x', t') system. In the right picture, a heavy line is shown at some time t in the un-primed system. The dashed lines show the previous position in the un-primed system if the line does not move in the unprimed system. In the moving (primed) system, the line moves in the $-x$ direction and it is *contracted*: the heavy dash-dotted line represents this line at some time in the unprimed system.

The fact that $\Delta s^2 = c^2\Delta t^2 - c^2\Delta r^2$ is invariant under the Lorentz transformation enables us to write physical equations in space-time in a very elegant form. This form involves scalars,

---

[1] Particle velocities will be denoted by **u** from now on. For relative velocities between inertial frames we use **v**.

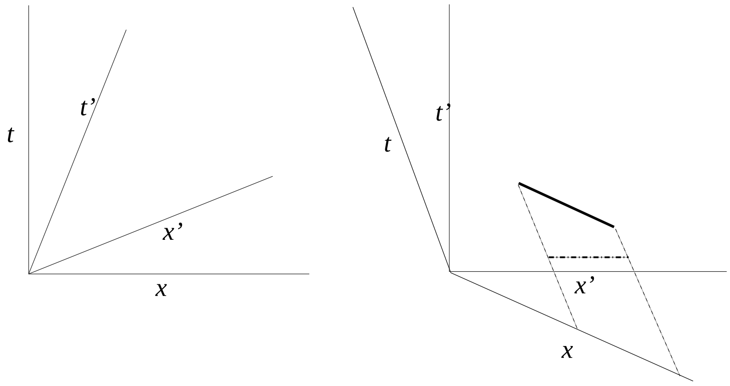


FIGURE 15.1: Graphical representation of the Lorentz transformation.

vectors and tensors, objects which should not be considered as arrays of numbers only, but as objects with certain transformation properties. As an example, we introduce the *four vector*

$$x^\mu = (x^0, x^1, x^2, x^3) = (t, x, y, z) \tag{15.15}$$

of which we know that it transforms according to the Lorentz transformation represented by a matrix L^μ_ν:

$$x'^\mu = \sum_{\nu=0}^{3} L^\mu{}_\nu x^\nu \equiv L^\mu{}_\nu x^\nu. \tag{15.16}$$

In this equation, we have introduced the notational convention that repeated upper and lower indices are summed over. This is the *Einstein summation convention.*

The quantity Δs^2 defined in (15.14) looks like a norm in 4-dimensional space-time. The only difference with the well known norm from Euclidean vector spaces is the minus-sign in front of the space-components. This minus-sign makes it useful to introduce the *metric tensor*:

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{15.17}$$

Let us now consider the quantity $s^2 = c^2t^2 - x^2$, which is the same as Δs^2 when one of the two points is at the origin of space-time. Using the definition of $g_{\mu\nu}$, we can write this as

$$s^2 = x^\mu g_{\mu\nu} x^\nu. \tag{15.18}$$

Note that the Einstein summation convention has been used here. We define this inner product as the *interval* in space-time.

For a Lorentz transformation represented by the linear operator $L^\mu{}_\nu$, invariance of the interval under Lorentz transformation gives:

$$s'^2 = x'^\mu g_{\mu\nu} x'^\nu = L^\mu{}_\nu x^\nu g_{\mu\rho} L^\rho{}_\sigma x^\sigma = x^\mu g_{\mu\nu} x^\nu. \tag{15.19}$$

It follows that $L^\mu{}_\nu$ must satisfy

$$L^\rho{}_\mu g_{\rho\sigma} L^\sigma{}_\nu = g_{\mu\nu}. \tag{15.20}$$

This can be taken as the definition of a Lorentz transformation. Note that the operator can also be written as $\left(L^T\right)_\mu{}^\rho$ (the superscript T denotes the transpose). The equation can therefore be written as

$$L^T g L = g. \tag{15.21}$$

In order to avoid having to put $g_{\mu\nu}$ and the like in all the equations, we define

$$x_\mu = g_{\mu\nu} x^\nu = (x^0, -x^1, -x^2, -x^3), \tag{15.22}$$

so that the interval can now simply be written as

$$s^2 = x^\mu x_\mu. \tag{15.23}$$

From now on we shall use this notation.

Now suppose we have a vector a_μ which transforms according to

$$a'_\mu = L_\mu{}^\nu a^\nu \tag{15.24}$$

then a^μ transforms as

$$a'^\mu = g^{\mu\nu} a'_\nu = g^{\mu\nu} L_\nu{}^\rho a_\rho = g^{\mu\nu} L_\nu{}^\rho g_{\rho\sigma} a^\sigma \equiv M^\mu{}_\sigma a^\sigma. \tag{15.25}$$

A vector like a_μ with a *lower* index is called *covariant.* We see that a covariant vector transforms according to M which is related to the inverse Lorentz transformation for a covariant vector as

$$M = gLg. \tag{15.26}$$

From a covariant vector, we can construct a *contravariant* vector.ÂăThis is a vector with an upper index – it is related to a covariant vector by

$$a^\mu = g^{\mu\nu} a_\nu.$$

Below we shall see that for a covariant vector x_μ, transforming according to the Lorentz transformation, $\partial/\partial x_\mu$ transforms as a contravariant vector. Therefore we can write $\partial/\partial x_\mu \equiv \partial^\mu$, expressing the fact that this is a contravariant vector.

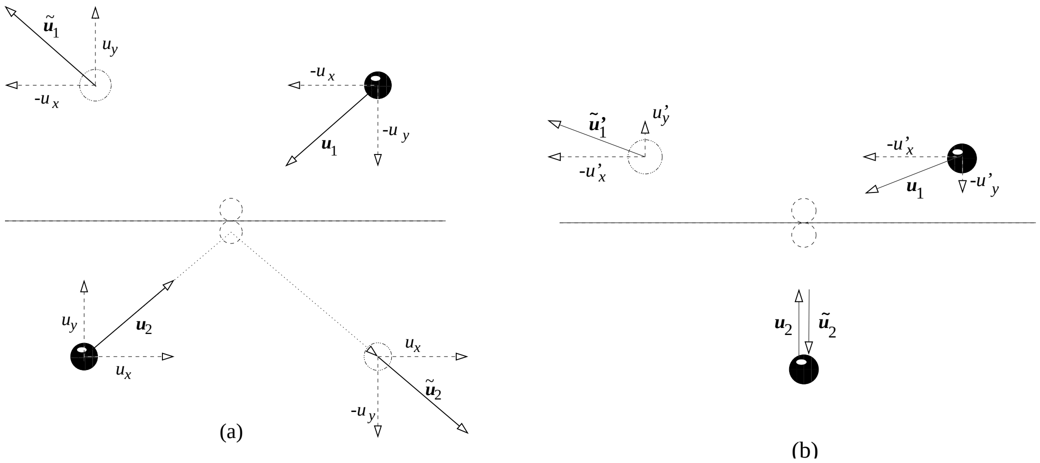
Finally, a few remarks about naming conventions. The components 1, 2 and 3 of a four-vector form its *spatial part*, the component 0 is the *temporal* part. These names are also used when the four-vector is not $(t, \mathbf{x})$ (we shall encounter other examples of four-vectors below). A four-vector s_μ is called *time-like* when $(s^0)^2 > (s^1)^2 + (s^2)^2 + (s^3)^2$, and *space-like* when this is not the case.

15.4 ENERGY AND MOMENTUM

In nonrelativistic classical mechanics, physical laws are usually phrased in terms of expressions involving scalar quantities (such as mass or energy) or vector quantities (position, momentum, angular momentum). Vector quantities are usually derived from the position vector, and they transform accordingly. For example, a rotation of the three dimensional Euclidean space, acts exactly the same on position vectors as on the momentum. In relativistic mechanics we want physical laws to be invariant under relative displacements at uniform speed, which means that we require the equations to assume the same form in any inertial frame. This can be done in two ways: we can formulate the equations as equalities between scalars, which should be invariant, or as vector equalities. As the physical vector quantities we are interested in are assumed to be derived from the positions of the particles involved, it is natural to assume that these quantities should transform according to the Lorentz transformations, just as the rotational transformations in nonrelativistic mechanics – i.e., they should be four-vectors.

In the light of this, when looking for physical laws, it is natural to find these laws as expressions in terms of four-vectors or scalars. Scalars can be constructed by taking the inner product of two four-vectors, analogous to Eqs. (15.19) and (15.23). This leads to some scalar quantity q^2:

$$q^2 = a_\mu g^{\mu\nu} b_\nu = a_\mu b^\mu = a^0 b^0 - \mathbf{a} \cdot \mathbf{b}. \tag{15.27}$$



(a)

(b)

FIGURE 15.2: Collision of two particles, (a) in the CM frame, and (b) in the rest frame of particle 1.

Let us now consider conservation of momentum and energy in more detail. The procedure is to start with the nonrelativistic definition of momentum of a point particle of mass m moving at velocity $\mathbf{u}$:

$$\mathbf{p} = m\mathbf{u}. \tag{15.28}$$

In this expression, the velocity $\mathbf{u}$ is defined as the time-derivative of position, and we have already seen how this quantity transforms [see Eq. (15.10) and (15.12)]:

$$u_x = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}} \tag{15.29a}$$

$$u_y = \sqrt{1-(v/c)^2}\,\frac{u'_y}{1 + \frac{vu'_x}{c^2}} \tag{15.29b}$$

where the relative velocity v of the frames is directed along the x-axis.

The aim is now to construct a four-vector starting from the definition of nonrelativistic momentum. We know that in the absence of external forces, momentum is conserved in classical nonrelativistic mechanics. We now want to find a relativistic analogue of momentum which is also conserved. As we do not yet know how to treat forces in relativistic mechanics, we consider elastic collisions, where the forces only act at the moment of the collision, and where energy and momentum conservation forms a general framework for describing the physics.

We consider the collision shown in figure 15.2. The collision involves two particles, 1 and 2, of equal mass m. In the left hand part, the collision is shown in the frame in which the centre of mass is at rest (CM frame), whereas the right hand part shows the same collision in the frame in which particle 2 moves along the y-axis. Note that the right hand frame moves at velocity $u_x\hat{\mathbf{x}}$ with respect to the CM frame. The particle velocities in the CM frame are equal in magnitude (but with opposite directions). We can now evaluate the velocities of the particles in the right hand frame by the velocity transformation law. A tilde denotes the velocity *after* the collision.

$$\tilde{u}'_{1x} = u'_{1x} = \frac{-2u_x}{1+u_x^2}; \tag{15.30a}$$

$$\tilde{u}'_{2x} = u'_{2x} = 0; \tag{15.30b}$$

$$-\tilde{u}'_{1y} = u'_{1y} = \frac{1}{\gamma}\frac{-u_y}{1+u_x^2}; \tag{15.30c}$$

$$-\tilde{u}'_{2y} = u'_{2y} = \frac{1}{\gamma}\frac{u_y}{1-u_x^2}. \tag{15.30d}$$

It is clear that in the CM frame, the total nonrelativistic momentum is conserved (it vanishes!). This is however no longer the case in the right hand frame, where the y-component of the total momentum is different before and after the collision. We search for a modified definition of momentum, which is conserved in relativistic mechanics. The simplest possibility is to add an extra velocity-dependent factor to the momentum:

$$\mathbf{p} = f(u) m\mathbf{u}. \tag{15.31}$$

If the momentum is conserved we have in the frame in which particle 2 moves along the y-axis:

$$f(u'_1) m\Delta u'_{1y} = f(u'_2) m\Delta u'_{2y}. \tag{15.32}$$

where $\Delta u'_{1y}$ is the change in velocity due to the collision. It is easy to find expressions for these changes, using the transformed velocities [Eq. (15.30)]:

$$f(u'_1) m \frac{2u_y}{\gamma(1+u_x^2)} = f(u'_2) m \frac{2u_y}{\gamma(1-u_x^2)} \tag{15.33}$$

so that

$$\frac{f(u'_1)}{f(u'_2)} = \frac{1+u_x^2}{1-u_x^2}. \tag{15.34}$$

We now express the right hand side in terms of u'_1 and u'_2, which are given by

$$u'^2_1 = \frac{4u_x^2 + u_y^2(1-u_x^2)}{(1+u_x^2)^2}; \tag{15.35a}$$

$$u'^2_2 = \frac{u_y^2}{1-u_x^2}. \tag{15.35b}$$

Eliminating u_y, this can be cast in the form

$$\frac{1-u'^2_1}{1-u'^2_2} = \frac{(1-u_x^2)^2}{(1+u_x^2)^2}. \tag{15.36}$$

The conclusion is therefore that

$$f(u) = \mathrm{Const}/\sqrt{1-u^2} = \mathrm{Const}\cdot\gamma_u \tag{15.37}$$

where we have introduced γ_u to distinguish it from γ_v which is determined by the relative velocity of the two frames rather than the speed of a particle in one reference frame or another. As we want the momentum to coincide with the nonrelativistic momentum for small speeds ($\gamma \approx 1$), we have

$$\mathbf{p}_{\mathrm{rel}} = \gamma_u m\mathbf{u}. \tag{15.38}$$

Often, the factor γ is included in the mass:

$$m(u) = \gamma_u m. \tag{15.39}$$

At $u = 0$, this relativistic mass assumes the value m, which we call the *rest mass.*

We now know the expression for the three-dimensional momentum. If the momentum is a physical quantity involved in acceptable physical laws, it should be part of a four-vector. Let us assume then that this is the case. What could the zeroth element of this four-vector be? Consider a particle moving in some (primed) reference frame in the x-direction with momentum $p'_x = m\gamma u'_x$. Consider the frame in which this particle is at rest. This rest frame moves at speed $-u'_x$ with respect to our reference frame. In the rest frame, the momentum p_x is zero, so according to the Lorentz transformation we must have:[2]

$$p_x = 0 = \gamma(p'_x - u'_x p'_0), \tag{15.40}$$

---

[2] note that as $\mathbf{v} = -u_x\hat{\mathbf{x}}$, $\gamma_v = \gamma_u$, so their distinction is not made in the following.

that is, $p'_0 = \gamma m$. So we take

$$p^\mu = (\mathbf{p}, p^4) = \gamma(m\mathbf{u}, mc). \tag{15.41}$$

What is the meaning of the first term? Let's expand this term for low velocities:

$$\gamma mc = mc\left(1 + \frac{(u/c)^2}{2}\right) = \frac{mc^2 + \frac{1}{2}mu^2}{c}. \tag{15.42}$$

The second term on the right hand side of this expression is the kinetic energy, and the first term is a constant. In classical mechanics, a constant added to the energy does not change the physics. Therefore, the first component is the relativistic expression for the energy. For a particle at rest, we have

$$E = mc^2. \tag{15.43}$$

This is the famous equation which is usually presented as Einstein's main achievement. Note that this expression for the energy only holds for particles at rest. For moving particles, the energy is

$$E = \gamma mc^2 = \sqrt{m^2c^4 + \mathbf{p}^2c^2}. \tag{15.44}$$

The last expression can easily be derived from $E = \gamma mc^2$.

It is useful to explicitly show that

$$p^\mu = (\mathbf{p}, p^4) = \left(\mathbf{p}, \frac{E}{c}\right) \tag{15.45}$$

is indeed a four vector. This is left as an exercise to the reader.

It is possible to formulate the foregoing in a different way, using the concept of *proper time*. The proper time is the time as measured in the particle's rest frame. Because of time dilatation, we know that when we observe a particle move at speed u, its proper time τ is related to our time t by

$$dt = \gamma d\tau. \tag{15.46}$$

Observers in different reference frames have non-synchronised clocks. However, they can both relate their time to the proper time which itself is independent of the speed of an observer's reference frame, and is therefore an invariant. In fact, for a point particle not subjected to external forces, the only two relativistic invariants we know are the mass and the proper time. It now is immediately clear that the vector quantity

$$u^\mu = (dt/d\tau, d\mathbf{x}/d\tau) = (\gamma, \gamma d\mathbf{x}/dt) = \gamma(1, \mathbf{v}) \tag{15.47}$$

transforms as a covariant four-vector. Multiplying this by the rest mass, we obtain the four-momentum:

$$p^\mu = (\mathbf{p}, E/c) = mu^\mu = (\gamma m\mathbf{v}, \gamma m). \tag{15.48}$$

In the previous section we have shown that the inner product of any two four-vectors is an invariant quantity. For the four-momentum we find that the inner product of this vector with itself is:

$$p^\mu p_\mu = p_0^2 - \mathbf{p}^2 = \gamma^2 m^2(1 - v^2) = m^2, \tag{15.49}$$

and the rest mass m is obviously invariant.

15.5 MATHEMATICAL STRUCTURE OF SPACE-TIME

We have encountered two different examples of vectors which transform according to a Lorentz transformation: x_μ and p_μ. We also have seen that for such a vector a_μ, the product:

$$s^2 = a^\mu a_\mu \equiv +\left(a^0\right)^2 - \left(a^1\right)^2 - \left(a^3\right)^2 - \left(a^3\right)^2 \tag{15.50}$$

is invariant under the Lorentz transformation. Recall that the components of a^μ are:

$$a_0 = a^0; \quad a^1 = -a_1; \quad a^2 = -a_2; \quad a^3 = -a_3. \tag{15.51}$$

The vectors a_μ are called *covariant* and their counterparts a^μ are *contravariant.* Going from co- to contravariant means swapping the sign of the space-like components of the vectors.

The invariance of the quantity s^2 fixes the form of the possible Lorentz transformation up to a sign. The Lorentz transformation corresponding to a relative velocity $v = \beta c$ along the x-axis has the form

$$L_\mu{}^\nu = \begin{pmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \tag{15.52}$$

The inverse transformation is obtained by moving in the opposite direction: $\beta \to -\beta$. This last result is immediately clear on physical grounds, and it can be checked explicitly for the matrix form.

Note that in the form given, the Lorenz transform describes what happens to a covariant vector:

$$x'_\mu = L_\mu{}^\nu x_\nu. \tag{15.53}$$

We can find the form of the Lorentz transform of a contravariant vector by swapping upper- and lower indices, using the metric tensor $g^{\mu\nu} = g_{\mu\nu}$, as we have seen above:

$$a'^\mu = M^\mu{}_\nu a^\nu; \tag{15.54a}$$

$$M^\mu{}_\nu = g^{\mu\rho} L_\rho{}^\sigma g_{\rho\sigma}. \tag{15.54b}$$

In fact, as L for contravariant vectors always occurs with subscripts $L_\mu{}^\nu$ and M with $M^\mu{}_\nu$, we can also replace $M^\mu{}_\nu$ by $L^\mu{}_\nu$ without risking confusion. Moreover, this convention allows us to use g to turn lower indices into upper ones and vice versa, as Eq. 15.54b) shows.

In physics, we often deal with space-time derivatives. An example is the continuity equation:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \tag{15.55}$$

which involves the spatial gradient ∇ and the time derivative $\partial/\partial t$. It turns out that the derivatives with respect to covariant vectors are contravariant and vice versa. We shall show this now explicitly.

Let us investigate this for the covariant vector x_μ. Suppose we have a function f depending on $x^2 = x^\mu x_\mu$:

$$f = f(s^2) = f(x_\mu x^\mu). \tag{15.56}$$

We calculate the gradient of the function f with respect to the vector x_μ, applying the chain rule:

$$\frac{\partial}{\partial x_\mu} f(s^2) = f'(s^2) \frac{\partial}{\partial x_\mu} (x^\nu x_\nu) = f'(s^2) 2x^\mu. \tag{15.57}$$

From the fact that $f(s^2)$ is invariant, we conclude that $\partial/\partial x_\mu$ transforms as a *contravariant* vector.

We have obtained a few important results:

- We can consider two types of relativistically transforming four-vectors: covariant and contravariant. They are related by a sign change of their spatial components.
- The Lorentz transformations for co- and contravariant vectors are each other's inverse.
- The gradient with respect to a covariant vector transforms as a contravariant vector and vice versa.

The last remark leads to the notational convention:

$$\frac{\partial}{\partial x^\mu} \equiv \partial_\mu; \qquad \frac{\partial}{\partial x_\mu} \equiv \partial^\mu. \tag{15.58}$$

15.6 ELECTROMAGNETIC FIELDS AND RELATIVITY

In section 8.3.1, we have seen that the Maxwell equations in the absence of sources can be formulated in terms of the vector and scalar potential. In the presence of sources (charge and current), the equations for the potentials read:

$$-\nabla^2 \mathbf{A}(\mathbf{r},t) + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}(\mathbf{r},t)}{\partial t^2} = \mu_0 \mathbf{j}(\mathbf{r},t); \tag{15.59a}$$

$$-\nabla^2 \phi(\mathbf{r},t) + \frac{1}{c^2}\frac{\partial^2 \phi(\mathbf{r},t)}{\partial t^2} = \rho(\mathbf{r},t)/\epsilon_0. \tag{15.59b}$$

The electric and magnetic fields can be obtained from these potentials:

$$\mathbf{B} = \nabla \times \mathbf{A};$$

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

where we have left out the position and time dependences for sake of brevity.

The electric and magnetic fields transform in a complicated way under a Lorentz transformation. The potentials are convenient vehicles for expressing the behaviour of electromagnetic fields under a Lorentz transformation. In section 15.5, we have seen that the gradient with respect to the contravariant space-time transforms as a covariant vector and vice versa. Hence, we can write

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

which is therefore invariant.

We can write the potential formulation of the Maxwell equations, Eqs. 15.59a in the form

$$\partial_\mu \partial^\mu \mathbf{A} = \mu_0 \mathbf{j}; \tag{15.60a}$$

$$CHECKSIGN!!!!\partial_\mu \partial^\mu \phi = \rho/\epsilon_0. \tag{15.60b}$$

It is clear that we can turn this into a relativistically invariant expression if we combine $\phi, \mathbf{A}$ and $\rho, \mathbf{j}$ into four-vectors. In order to get an idea to do this in a sensible way, note that ρ and $\mathbf{j}$ should satisfy the *continuity equation*:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0,$$

which, after introducing the *four current* J_μ:

$$J^\mu = (c\rho, \mathbf{j}),$$

can also be written as

$$\partial_\mu J^\mu = 0.$$

If we now write $A^\mu = (\phi/c, \mathbf{A})$, then we can write

$$\partial_\mu \partial^\mu A^\nu = \mu_0 J^\nu CHECK!!!$$

which clearly exhibits the relativistically invariant form of the Maxwell equations. In the next chapter, we shall comment further on the relativistically invariant formulation of Maxwell's equations.

15.7 RELATIVISTIC DYNAMICS

In classical mechanics, we start from Newton's law of motion and find the appropriate expression for the force. We could follow a similar approach in relativistic mechanics by replacing Newton's law by some relativistically invariant analogue. Newton's law is a vector equality, so we want to replace it by a relation between four-vectors. It is now important to realise that although p^μ is a four-vector, dp_μ/dt is not. This is because the time t is not a relativistic invariant. Therefore, we use the four vector $dp_\mu/d\tau$, which is an invariant. Newtons' law now reads:

$$K_\mu = \frac{dp_\mu}{d\tau}. \tag{15.61}$$

K_μ is a generalisation of the force – in the nonrrelativistic limit, its spatial part should reduce to the classical force.

The approach described here is used in many textbooks. A problem with it however is that obtaining the correct expression for K_μ requires guessing and hand-waving. Therefore, we shall follow here a more consistent approach which ties in closely with the first part of this course as it is based on constructing a relativistically invariant action. Again, we want to find the dynamic trajectory of a particle as the one which minimises some functional. The action is a scalar, and if we want to derive a general principle, we should construct it as a relativistic invariant. This invariant should reduce to the classical action in the norelativistic limit. In the following derivation we leave factors c in the expressions for clarity.

We start with a particle not subject to a force. We have only two relativistic invariants: the proper time and the mass. As the classical action is written as an integral over time, we write it now as an integral over proper time:

$$S = \int_A^B d\tau \lambda, \tag{15.62}$$

where λ may depend on the rest mass m. Using (15.46), we can write S as

$$S = \int_A^B \frac{dt}{\gamma} \lambda. \tag{15.63}$$

A Taylor expansion of the integrand gives

$$\frac{\lambda}{\gamma} = \lambda \sqrt{1 - \frac{v^2}{c^2}} \approx \lambda \left(1 - \frac{v^2}{2c^2}\right). \tag{15.64}$$

This is seen to equal the kinetic energy plus a constant (which does not affect the stationary solutions of S) if

$$\lambda = -mc^2. \tag{15.65}$$

So we have found

$$S = -mc^2 \int_A^B \frac{dt}{\gamma}. \tag{15.66}$$

as the relativistically invariant expression for the action. In analogy to the classical expression for the action, we can say that the Lagrangian is given as

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}. \tag{15.67}$$

The canonical momentum, defined in section 2.7, is given by

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{15.68}$$

which corresponds precisely to the expression for the relativistic momentum, found earlier.

The Euler-Lagrange equation for a particle not subject to external forces therefore reads

$$\frac{d\mathbf{p}}{dt} = 0. \tag{15.69}$$

The Hamiltonian is found as

$$H = \mathbf{p}\cdot\mathbf{v} - L = \gamma mc^2, \tag{15.70}$$

which, as we have seen above, is indeed the energy.

Now we want to include a force. As an example we consider the force experienced by a particle in an electromagnetic field. Strictly speaking, we cannot derive this because we have not investigated the behaviour of electromagnetic fields in the relativistic limit, which is a topic beyond the scope of this course. Therefore we shall quote the main result of this analysis:

> The vector $(\varphi, \mathbf{A})$ formed by the electric potential φ and the vector potential $\mathbf{A}$, transforms as a four vector. This four-vector is usually denoted as A_μ.

So our task is to construct a relativistically invariant analogue to the classical Lagrangian of a charged particle in an electromagnetic field:

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 + q\dot{\mathbf{r}}\cdot\mathbf{A}(\mathbf{r},t) - q\phi(\mathbf{r},t). \tag{15.71}$$

From the classical expression we immediately guess

$$L = -mc^2\sqrt{1-\frac{v^2}{c^2}} - \frac{q}{\gamma m}p_\mu A^\mu. \tag{15.72}$$

In that case, the action reads:

$$S = \int_A^B d\tau\,\left(-mc^2 - \frac{q}{m}p_\mu A^\mu\right), \tag{15.73}$$

which is immediately seen to be relativistically invariant.

From the Lagrangian, we derive the momentum as

$$\mathbf{p} = \gamma m\mathbf{v} - q\mathbf{A}. \tag{15.74}$$

The Euler-Lagrange equations read:

$$\frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-v^2/c^2}} - q\mathbf{A}\right) = -q\nabla\varphi + q\nabla(\mathbf{v}\cdot\mathbf{A}). \tag{15.75}$$

Except for the first term on the left hand side, everything is exactly the same as in the classical derivation, and we immediately find:

$$\frac{d}{dt}(\gamma m\mathbf{v}) = q\mathbf{E} + q(\mathbf{v}\times\mathbf{B}). \tag{15.76}$$

It is interesting to study the case of a constant, homogeneous electric field in the x-direction. Then

$$\frac{d}{dt}(\gamma m v_x) = qE_x. \tag{15.77}$$

We find

$$\gamma m v_x = qE_x t, \tag{15.78}$$

from which it follows that

$$v_x = \frac{qcE_x t}{\sqrt{m^2c^2 + q^2E_x^2t^2}}. \tag{15.79}$$

The acceleration cannot be constant, as the velocity saturates at the value c. In fact, the acceleration is given as

$$a = \frac{dv}{dt} = \frac{qE_x}{m}\left[1 + \left(\frac{qE_x t}{mc}\right)^2\right]^{-3/2}. \tag{15.80}$$

15.8 SUMMARY

At the end of this chapter, it seems useful to summarize those results that are necessary for the remainder of these lecture notes.

We have introduced several four-vectors:

- Space-time: $x^\mu = (\mathbf{r}, ct)$;
- Energy-momentum: $p^\mu = (\mathbf{p}, E/c)$;
- Four-current: $j^\mu = (\mathbf{j}, c\rho)$;
- Four-vector potential: $A^\mu = (\mathbf{A}, \phi)$.

These vectors transform under a Lorentz transformation as follows. For a vector a^μ:

$$a'^\mu = L^\mu{}_\nu a^\nu,$$

where a'^μ is the transformed vector and where we have used the *Einstein summation convention*:

> In the Einstein summation convention, we sum over indices that occur twice in an expression, once as an upper index and once as a lower index.

The Lorentz transformation is defined by the requirement that it leaves

$$\Delta\mathbf{r}^2 - c^2\Delta t^2$$

invariant. It includes rotations of the space vectors and 'boosts', resulting from changing the reference frame from the original, 'unprimed' one to a second, 'primed' frame which is moving with velocity v with respect to the unprimed one. If the relative velocity of the primed frame with respect to the unprimed frame is v along the x-axis, the Lorentz transformation is

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

where $\beta = v/c$ and $\gamma = 1/\sqrt{1 - v^2/c^2}$.

All the vectors mentioned so far in this summary had an upper index. We define for such a vector a^μ another vector with a lower index, which has its space part reversed. For the space-time vector

$$x^\mu = (\mathbf{r}, ct), \qquad x_\mu = (-\mathbf{r}, ct).$$

> A vector with an upper index is called *covariant*. A vector with a lower index is called *contravariant*.

A covariant vector transforms into a contravariant vector through the so-called *metric tensor*, g:

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We have, for an arbitrary vector a^μ:

$$a_\mu = g_{\mu\nu} a^\nu; \qquad a^\mu = g^{\mu\nu} a_\nu.$$

Here, the Einstein summation convention is obviously used.

A transformation L of four-vectors is a Lorentz transformation if and only of it satisfies

$$L^{\mathrm{T}} g L = g,$$

which should be read as an equation of 4 × 4 matrices.

The beauty of the Lorentz transformation is that it leaves *any* product of the form $a^{\mu} b_{\mu}$ invariant. Such a quantity is an invariant scalar. Also, any four-vector equality $a_{\mu} = b_{\mu}$ or $a^{\mu} = b^{\mu}$ is left invariant under a Lorentz transformation.

Finally, we introduce the four-derivative

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}},$$

which has the form $(\nabla, 1/c\partial/\partial t)$. This transforms as a contravariant vector – hence the notation ∂_{μ}. Likewise,

$$\partial^{\mu} = \frac{\partial}{\partial x_{\mu}},$$

which has the form $(-\nabla, 1/c\partial/\partial t)$, transforms as a covariant vector.

15.9 PROBLEMS

1. An equation which holds in any reference frame is relativistically invariant. We also use the term *covariant* for this property, which is not to be confused with 'covariant' as in covariant/contravariant. Which of the following are covariant equations (give brief explanations?)?

   (a)
   $$\frac{\partial \phi}{\partial x^{\mu}} = A^{\mu}.$$

   (b)
   $$\frac{\partial \phi}{\partial x_{\mu}} = a(c^2 t^2 - r^2).$$

   (c)
   $$T^{\mu\nu} A^{\nu} = B^{\mu}.$$

   (d)
   $$A^{\mu} = B^{\mu\nu} C_{\mu}.$$

   (e)
   $$A^{\mu} = B^{\mu\nu} C_{\nu}.$$

   (f)
   $$\partial_{\mu} A^{\mu} = C_{\mu}.$$

   (g)
   $$\partial_{\mu} T^{\rho\sigma\mu} = 1.$$

   Now prove the following:

   (h) For a tensor T^{μ}_{ν}, show that $\mathrm{Tr}\left(T^{\mu}_{\nu}\right) = T^{\mu}_{m} u$ is an invariant scalar.

   (i) Knowing that the number $K = A_{\mu\nu} C^{\mu\nu}$ is an invariant scalar for *any* tensor $C^{\mu\nu}$, show that $A_{\mu\nu}$ transforms as a tensor.

2. In this problem we consider the relativistically invariant formulation of the Maxwell equations. Physical systems are described by a Lagrangian, and an advanced example of this idea is the electromagnetic field. The field is described by a four vector which encapsulates the vector and scalar potential as explained in the lecture notes. The four vector is $A_\mu = (\mathbf{A}, \phi/c)$.

   We introduce the *electromagnetic tensor* $F_{\mu\nu}$:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

   Note that this tensor is anti-symmetric: $F_{\mu\nu} = -F_{\nu\mu}$. We construct the Lagrangian directly from this tensor:

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - J^\mu A_\mu.$$

   The field equations which minimize this Lagrangian are obtained through the generalized Euler-Lagrange equations:

$$\partial_\beta \frac{\partial \mathscr{L}}{\partial \left(\partial_\beta A_\alpha\right)} = \frac{\partial \mathscr{L}}{\partial A_\alpha}.$$

   (a) Show that

$$\frac{\partial F_{\mu\nu}}{\partial \left(\partial_\beta A_\alpha\right)} = \delta_{\mu\beta}\delta_{\nu\alpha} - \delta_{\nu\beta}\delta_{\mu\alpha},$$

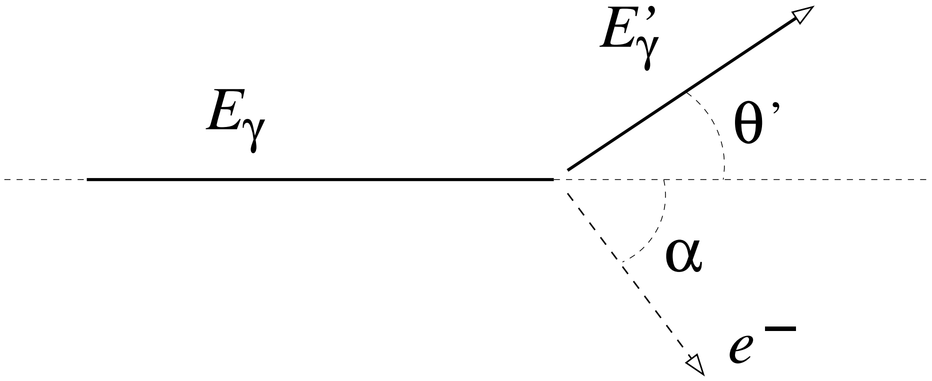
   (b) Show that the generalized Euler-Lagrange equation leads to

$$\partial_\beta F^{\beta\alpha} = J^\alpha.$$

   (c) Show that this leads directly to the Maxwell equations, formulated in terms of the potentials:

$$\partial^\mu \partial_\mu A^\nu = J^\nu.$$

3. This problem is about Compton scattering.
   Individual photons behave like particles of rest mass equal to zero. When a photon scatters off a free electron, the electron can recoil, taking up both energy and momentum. There is a unique relationship between the energy of the scattered photon, E'_γ and the scattering angle from the incident photon, θ'. This relation also depends on the electron rest mass m_e and on the incident photon energy, E_γ, see figure. If the four-momentum of the incident and the scattered photons are denoted by k^μ, k'^μ respectively, the statement of energy and momentum conservation is $k^\mu + p_e^\mu = k'^\mu + p_e'^\mu$, where p_e^μ and $p_e'^\mu$ are the initial and final electron energy-momentum four vectors of the electron.



(a) Prove, using $k^\mu k_\mu = k'^\mu k'_\mu = 0$, that the conservation laws of energy and momentum require that

$$E'_\gamma(\theta') = \frac{E_\gamma}{1 + \frac{E_\gamma(1-\cos\theta')}{m_e c^2}}.$$

(b) Using the above, prove that the maximum wavelenth shift of the photon is given by

$$\lambda' - \lambda = 2\frac{h}{m_e c},$$

where $h/(m_e c)$ is the Compton wavelength of the electron.

4. (a) Derive, starting from the Maxwell equations as given in section 8.3.1 a second order equation for the electric and magnetic fields separately. For example, take the curl of the equation involving $\nabla \times \mathbf{E}$ and use the identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E}$. You should find

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{E} = -4\pi\left[\frac{1}{c}\frac{\partial}{\partial t}\left(\frac{\mathbf{j}}{c}\right) + \nabla\rho\right].$$

(b) The operator on the left-hand side of the last equation is relativistically invariant. Show that on the right-hand side, we have an expression proportional to $\partial^\mu J^\nu - \partial^\nu J^\mu$ with $\mu = 0$ and $\nu = i = 1,2,3$. From this it follows that **E** must be three components of a tensor $F_{\mu\nu}$.

(c) The tensor $F^{\mu\nu}$ has the form

$$\begin{pmatrix} 0 & -B_z & B_y & E_x \\ B_z & 0 & -B_x & E_y \\ -B_y & B_x & 0 & E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix}.$$

By deriving an equation for **B** similar to that for **E**, verify that

$$\partial^\rho \partial_\rho F^{\mu\nu} = C\left(\partial^\mu j^\nu - \partial^\nu j^\mu\right),$$

where C is a constant.

16

THE KLEIN-GORDON AND MAXWELL EQUATIONS

16.1 THE KLEIN-GORDON EQUATION

In an attempt to generalise the Schrödinger equation to relativistic problems, one may use the standard representation of the momentum operator in quantum mechanics

$$\mathbf{p} = -\mathrm{i}\hbar\nabla,$$

together with the 'energy operator', to construct a four-vector, analogous to what we have in special relativity. What would be a suitable form of the energy operator? We can use two ideas for guiding us: first, from the fact that **p** involves a spatial derivative, we see that the fourth component must be a time-derivative because the gradient with respect to a covariant space–time vector four vector is a contravariant vector. All in all, this leads to the four vector

$$p_\mu = \mathrm{i}\hbar\partial_\mu$$

where, as usual, $\partial_\mu = (\nabla, \frac{1}{c}\frac{\partial}{\partial t})$.

Another idea may be the fact that for an eigenstate of the Schrödinger equation with definite energy E, we know that

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = E\left|\psi(t)\right\rangle$$

which suggests the identification $\mathrm{i}\hbar\frac{\partial}{\partial t}$ with the fourth component of the energy-momentum four vector.

Now we use the general result

$$p^\mu p_\mu = m^2c^2$$

of special relativity to write down the so-called *Klein-Gordon equation*:

$$\left[-\hat{\mathbf{p}}^2 + \left(\frac{\hat{E}}{c}\right)^2\right]\left|\psi\right\rangle = m^2c^2\left|\psi\right\rangle.$$

This is an analogue of the Schrödinger equation for free particle at relativistic speed.

In section 7.3, we have analyzed the mass conservation problem and concluded that for a density given by

$$\rho(\mathbf{r}, t) = \left|\psi(\mathbf{r}, t)\right|^2$$

and a current

$$\mathbf{j} = \frac{\hbar}{2\mathrm{i}m}\left[\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\nabla\psi^*(\mathbf{r})\right],$$

the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0$$

holds.

We now would like to play the same game using the Klein-Gordon equation. It turns out that this fails as this equation does not involve the *first* time derivative like the Schrödinger equation, but the *second* time derivative. The best option seems to preserve the expression for the flux and see what we then obtain for the material density (which was $\left|\psi(\mathbf{r})\right|^2$ in the case of Schrödinger quantum mechanics). This can be done by taking the Klein-Gordon equation and its conjugate, and subtracting the two;

$$\nabla \cdot \mathbf{j} = \frac{\hbar^2}{2\mathrm{i}m}\nabla \cdot \left[\psi^* \nabla \psi - \psi \nabla \psi^*\right] = \frac{\hbar^2}{2\mathrm{i}m}\left[\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*\right] = \frac{\hbar^2}{2\mathrm{i}mc^2}\left[\psi^* \partial_t^2 \psi - \psi \partial_t^2 \psi^*\right] = -\partial_t \rho,$$

which then leads to

$$\rho = -\frac{\hbar}{2\mathrm{i}mc^2}\left[\psi^*(\mathbf{r},t)\frac{\partial \psi(\mathbf{r},t)}{\partial t} - \frac{\partial \psi^*(\mathbf{r},t)}{\partial t}\psi(\mathbf{r},t)\right].$$

For a free particle state $\psi(\mathbf{r},t) = C \exp\left[\mathrm{i}\left(\mathbf{k}\cdot\mathbf{r} - \omega t\right)\right]$, this leads to

$$\rho = \frac{E}{mc^2}|C|^2.$$

This forms nicely a four-vector with the current

$$\mathbf{j} = \frac{\mathbf{p}}{m}|C|^2.$$

This four-vector is then

$$j^\mu = \frac{p^\mu}{m}|C|^2.$$

The expression for the current density suggests, by analogy to the nonrelativistic case, that an ‘inner product’ can be defined as follows:

$$\langle \psi_1 | \psi_2 \rangle = \mathrm{i}\int \psi_1^*(x)\overleftrightarrow{\partial}_t \psi_2(x) d^3 r.$$

In this expression, x denotes the space-time four vector, and

$$\psi_1^*(x)\overleftrightarrow{\partial}_t \psi_2(x) = \psi_1^*\left(\frac{\partial \psi_2}{\partial t}\right) - \left(\frac{\partial \psi_1^*}{\partial t}\right)\psi_2.$$

From this inner product, we can determine the normalisation of a plane wave state. This turns out to be $1/\sqrt{2\omega_k}$, with $\hbar\omega_k = E$. With proper units we have

$$\psi(x) = \frac{1}{\sqrt{2\omega_k}}\frac{1}{(2\pi\hbar)^{3/2}}e^{-\mathrm{i}p\cdot x}. \tag{16.1}$$

It is important to appreciate the consequences of the fact that the energy operator occurs with a square in the Klein-Gordon equation, whereas it occurs with power 1 in the Schrödinger equation. The difference is clear from the dispersion relation:

$$E = \pm\sqrt{p^2c^2 + m^2c^4},$$

showing that E occurs with two signs: for every positive energy, there exists also a negative energy. This differs quite significantly from the situation in the Schrödinger equation: as the momentum p^2 is not bounded from above, the energy can become infinitely negative. This means that there is no well-defined ground state, and that we have an ‘energy hole’, into which energy can flow ad inifinitum. This is obviously not physical.

Another problem follows from the fact that the density is proportional to the energy. Therefore, negative densities are possible, which does not seem sensible either.

16.2 ANALOGY WITH THE MAXWELL EQUATIONS

The structure of the Maxwell equations is quite analogous to that of the Klein-Gordon equation as we have seen in the previous chapter. Here we shall discuss the relativistic formulation of Maxwell's equations more in-depth. We start from the relativistic formulation of the Maxwell equations. Note that we can express the electric and magnetic field completely in terms of the scalar and vector potential, which together form a four vector:

$$A^\mu = (\mathbf{A}, \varphi/c).$$

The fields are found from

$$\begin{aligned}\mathbf{B} &= \nabla \times \mathbf{A};\\ \mathbf{E} &= -\nabla\varphi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}.\end{aligned}$$

The Maxwell equations in the absence of electric charges and currents can conveniently be written in terms of the tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{16.2}$$

Note that this expression does not change under a *gauge transformation*

$$A_\mu \to A_\mu - g\partial_\mu \chi(x).$$

The Maxwell equations read:

$$\partial_\mu F^{\mu\nu} = 0$$

as we shall now show.

Working our this formula in terms of the four-vector potential gives

$$\partial_\mu \partial^\mu A^\nu + \partial^\nu \left(\partial_\mu A^\mu\right) = 0.$$

An appropriate gauge transformation (called the *Lorenz gauge*), ensuring $\partial_\nu A^\nu$ to vanish, eliminates the second term in this formulation of Maxwell's equations in the absence of currents and charges, leaving

$$\partial_\nu \partial^\nu A^\mu = 0.$$

This now looks remarkably similar to the Klein-Gordon equation, if this is written in the form

$$\partial_\nu \partial^\nu \phi = m^2 \phi.$$

The fact that the right hand side vanishes in the case of the Maxwell equations shows that the particles described by these equations have zero mass. These particles are the photons.

The operator $\partial_\nu \partial^\nu = \frac{1}{c^2}\partial_t^2 - \nabla^2$ is usually written as $\Box$.

Note that the Maxwell equations are treated here completely on the classical level! The Laplace operator in the Klein-Gordon equation comes about as a result of the quantisation of the momentum operator $p = -i\hbar\nabla$. The classical Maxwell equations have this operator on board without quantization. This difference is also clear from the fact that the Klein-Gordon equation contains Planck's constant, whereas the Maxwell equations do not.

Another important remark is that a right hand side containing current and charge sources does not introduce a mass in the Maxwell equations – a mass term would have the form mA_ν, which is essentially different from an external source term.

16.3 SOURCE TERMS

The Maxwell equations in the presence of charges and sources read

$$\Box A^\mu(x) = J^\mu(x).$$

Such an inhomogeneous equation can be solved in terms of the Green's function D_F which is defined as

$$\Box D_F = \delta(ct)\delta^{(3)}(\mathbf{r}) = \delta^{(4)}(x).$$

With this definition, it can be verified that

$$A^\mu = A_0^\mu(x) + \int D_F(x-x')J^\mu(x')d^4x'$$

where $A_0^\mu(x)$ is a solution of the Maxwell equations without sources, solves this equation with sources.

The Green's function (or *free propagator*) D_F can be found by Fourier-transforming:

$$D_F(x) = -\frac{1}{2\pi^4}\int D_F(p)e^{-\mathrm{i}px}d^4p$$

where $p^\mu = (\mathbf{p}, E/c)$ is the momentum four vector. Transformations from p to x and vice versa always hinge upon the expression for the delta-function:

$$\delta^{(4)}(x) = \frac{1}{(2\pi)^4}\int e^{-\mathrm{i}px}d^4p.$$

Considering the Fourier transform $\tilde{\phi}(p)$ of a function $\phi(x)$, where these two are related through:

$$\phi(x) = -\frac{1}{2\pi^4}\int \tilde{\phi}(p)e^{-\mathrm{i}px}d^4p,$$

we see that $\partial_\mu\phi(x)$, gives for the Fourier transform the result $\mathrm{i}\hbar p_\mu\tilde{\phi}(p)$. Extending this result to the present case, we see that the free propagator for space-time can be found as

$$D_F(p) = \frac{1}{p^2}.$$

Note that we do not use a tilde ˜for denoting the Fourier transform of D_F – the argument tells us whether we deal with the function in direct or in reciprocal space. This expression is singular at the origin, which we usually get rid off by writing

$$D_F(p) = \frac{1}{p^2 + \mathrm{i}\epsilon},$$

which shifts the singularity slightly off the real axis. In that case we can use the tricks from chapter 1 in order to calculate the Green's function:

$$D_F(x) = -\frac{1}{2\pi^4}\int \frac{e^{-\mathrm{i}px}}{p^2+\mathrm{i}\epsilon}d^4p.$$

We can also add a source term to the Klein-Gordon equation. This does not immediately have a physical interpretation, but the propagator can play a useful role in solving the problem of a Klein-Gordon particle in a (weak) potential.

For the Klein-Gordon equation, an analysis similar to that of the Maxwell equations leads to

$$\Delta_F(x) = -\frac{1}{2\pi^4}\int \frac{e^{-\mathrm{i}px}}{p^2 - m^2 + \mathrm{i}\epsilon}d^4p.$$

We see that the Maxwell propagator is a special case of the Klein-Gordon propagator with $m = 0$.

The Klein-Gordon propagator can be reworked to

$$\Delta_F(x - x') = \frac{\mathrm{i}}{(2\pi)^3}\left[\theta(t - t')\int \frac{e^{-\mathrm{i}p\cdot(x-x')}}{2\omega_p} d^3p + \theta(-t + t')\int \frac{e^{\mathrm{i}p\cdot(x-x')}}{2\omega_p} d^3p\right].$$

where $\omega_p = \sqrt{p^2 + m^2}$. We leave this form for future reference.

16.4 STATIC SOLUTIONS FOR THE PROPAGATORS

It is useful at this point to recall the forms of the propagators in the static limit (no t-dependence). The equation for the Maxwell propagator:

$$\nabla^2\phi_e(\mathbf{r}) = \delta^{(3)}(\mathbf{r})$$

is the Poisson equation for a point charge (up to a factor of 4π). We have used the notation ϕ_e to distinguish the static propagator from the dynamic one D_F. The solution is given by

$$\phi_e(\mathbf{k}) = \frac{1}{k^2}.$$

In real space, this becomes

$$\phi_e(\mathbf{r}) = \frac{1}{4\pi}\frac{1}{r}.$$

In 1D, this would be $\phi_e(x) = |x|$ and in 2D, it is $\phi_e(\mathbf{r}) = 2\pi\ln(r)$.

Now we turn to the static Klein-Gordon equation. This has the form

$$\left(\nabla^2 - m^2\right)\phi(\mathbf{r}) = \delta^{(3)}(\mathbf{r}).$$

This is a *Helmholtz equation*. The solution is given by

$$\phi(\mathbf{k}) = \frac{1}{k^2 + m^2}.$$

In real space, this becomes

$$\phi(\mathbf{r})\frac{1}{4\pi}\frac{e^{-mr}}{r},$$

which is consistent with the solution of the Poisson equation for $m \to 0$.

We see that zero-mass gives a power-law decay of the potential, whereas the effect of the mass is to set a finite range beyond which the potential decays exponentially to zero.

16.5 SCATTERING AS AN EXCHANGE OF VIRTUAL PARTICLES

The Green's function formulation of scattering results in a nice interpretation of scattering processes as we shall explain in this section. Remember that in the Born approximation, the scattering amplitude is given as

$$f(\theta) = -\frac{m}{2\pi\hbar^2}\int e^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}V(\mathbf{r}')d^3r'.$$

Here, $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ where $\mathbf{k}$ is the incoming, and $\mathbf{k}'$ the outgoing momentum. Note that in the elastic case, $|\mathbf{k}| = |\mathbf{k}'|$; therefore we have

$$q = 2k\sin(\theta/2),$$

where θ is the angle between $\mathbf{k}$ and $\mathbf{k}'$.

In the case of Rutherford scattering ($V(r) = -Ze^2/r$), the scattering amplitude becoms

$$f(\theta) = \left(\frac{2mZe^2}{\hbar^2}\right)^2 \frac{1}{q^2},$$

which is reminiscent of the propagator for the zero-mass Klein-Gordon equation (or for the Maxwell equations), which makes sense as the scattering amplitude in the Born approximation is just the Fourier transform of the interaction potential. The fact that this interaction is described as a Klein-Gordon solution with zero mass, suggests an interpretation of the term $D(\mathbf{q}) = 1/q^2$ in the expression as a massless particle, the photon.

In the case of a Yukawa potential

$$V(\mathbf{r}) = -\frac{Ze^2}{r}\frac{e^{-mr}}{\quad,}$$

the factor $1/q^2$ is replaced by $1/(q^2 + m^2)$ which derives from a Klein-Gordon equation with mass m. In this case the 'virtual particles' responsible for the effect of the interaction carry a mass m. An example are *mesons*, elementary particles that are responsible for the interactions between protons and neutrons.

16.6 PROBLEMS

1. *The nonrelativistic limit of the Klein-Gordon equation*
The Klein-Gordon equation is compatible with an energy-momentum relation

$$E^2 = p^2c^2 + m^2c^4.$$

(a) Show that, in the non-relativistic limit, this reduces to the sum of the rest energy and the classical expression for the kinetic energy.

In order to study the non-relativistic limit of the Klein-Gordon equation, we write the solution to that equation in the form $\psi e^{-\mathrm{i}mc^2t/\hbar}$.

(b) Show that this leads to the Schrödinger equation to be obeyed by ψ.

2. *Klein-Gordon equation and Coulomb potential*
It is interesting te study the 'Klein-Gordon hydrogen atom', where we use the Coulomb potential in this relativistic wave equation:

$$\left[\mathrm{i}\hbar\frac{\partial}{\partial t} + \frac{Ze^2}{r}\right]^2 \psi(\mathbf{r},t) = -\hbar^2c^2\nabla^2\psi(\mathbf{r},t) + m^2c^4\psi(\mathbf{r},t).$$

(a) Write $\psi(\mathbf{r},t) = R(\mathbf{r})Y_l^m(\theta,\phi)e^{-\mathrm{i}Et/\hbar}$ and show that this leads to the following equation for $R(r)$:

$$\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) + \frac{l(l+1)}{r^2}\right]R(r) = \frac{\left(E + \frac{Ze^2}{r}\right)^2 - m^2c^4}{\hbar^2c^2}.$$

Here, you may look into the analysis of the Schrödinger equation for a spherically-symmetric potential (e.g. Griffiths Ch. 4) where the analysis of the Laplace operator via a separation of variables using Y_l^m is done in a similar way.

(b) Writing

$$\gamma = \frac{Ze^2}{\hbar c}; \quad \alpha^2 = \frac{4\left(m^2c^4 - E^2\right)}{\hbar^2c^2}; \quad \lambda = \frac{2E\gamma}{\hbar c\alpha},$$

and $\rho = \alpha r$, show that the radial Klein Gordon equation of part (a) can be written as

$$\left\{\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{d}{d\rho}\right) + \left[\frac{\lambda}{\rho} - \frac{1}{4} - \frac{(l(l+1) - \gamma^2}{\rho^2}\right]\right\}R(\rho) = 0.$$

(c) Now write R in the form $R(\rho) = f(\rho)g(\rho)v(\rho)$, where f and g are the correct solutions for R for small and large ρ respectively:

$$R(\rho) \to \begin{cases} f(\rho) = \rho^s & \text{for } \rho \to 0; \\ g(\rho) = g(\rho) = e^{-\rho/2}. \end{cases}$$

Prove this and give an equation for s in terms of l and γ. We now have

$$R(\rho) = \rho^s e^{-\rho/2} v(\rho).$$

(d) As usual, we expand $v(\rho)$ as a power series $\sum_{k=0}^{\infty} a_k \rho^k$. By putting this into the radial Klein-Gordon equation, derive a recurion relation for the a_k. Requiring that for $k \geq N$, $a_k = 0$, derive that

$$E_n = mc^2 \left(1 + \frac{\gamma^2}{n^2}\right)^{-1/2},$$

where $n = N + s + 1$.

(e) Expand the energy found in (d) to fourth order in γ^2 and show that the zeroth order gives the rest energy, and the second order the Rydberg series. Calculate a formula for the lowest relativistic correction to the Rydberg formula. You should find

$$-\gamma^4 \frac{mc^2}{n^4} \left(\frac{n}{l + 1/2} - 3/2\right).$$

Hint: First show that $s = l - \gamma^2/(2l + 1)$.

17

THE DIRAC EQUATION

17.1 IMPROVING ON THE KLEIN-GORDON EQUATION

In the previous chapter we have seen that the problem with generalising the Schrödinger equation to the relativistic case is two-fold:

- For every positive energy, there is also a negative emergy of the same modulus, so that the spectrum is not bounded from below.
- The density may become negative.

Dirac tried to solve these issues by postulating a *linear* form of the Hamiltonian operator in terms of the momentum and energy:

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = \left(\frac{\hbar}{\mathrm{i}}\boldsymbol{\alpha}\cdot\nabla + \beta m\right)\psi = H\psi.$$

The α_i and β are fixed by satisfying the following physical requirements.

- The free-particle wave functions ψ should be a plane wave with a wave vector satisfying $E^2 = \mathbf{p}^2c^2 + m^2c^4$.
- There exists a four-vector current density j^μ whose fourth component should be a positive density.

Before we continue we adopt the usual convention in relativistic quantum mechanics to put $\hbar = c = 1$. This defines particular units in which we work. Any physically meaningful result which involves a physical dimension like energy, time etcetera, can be transformed into conventional units by multiplying with suitable powers of c and $\hbar$.

Let us work out the first condition. We work this out, interpreting the Hamiltonian as the energy, and operating with H^2 on ψ. We then find conditions for α_i and β from the requirement that the following equation must hold:

$$p^2 + m^2 = \sum_i(\alpha_i p_i)\sum_j(\alpha_j p_j) + \beta^2 m^2 + \sum_i(\alpha_i p_i)\beta + \beta\sum_j(\alpha_j p_j), \qquad (17.1)$$

where the indices i, j run from 1 to 3. I have deliberately respected the order in which the different terms occur in the products as it turns out that *commuting* objects α_i and β do not lead to the desired result. This can directly be seen from the following conditions that should be obeyed in order to satisfy the equation (17.1) (note that $\{a, b\}$ denotes an anti-commutator $ab + ba$):

$$\begin{aligned}\{\alpha_i, \alpha_j\} &= 0 \text{ for } i \neq j;\\ \{\alpha_i, \beta\} &= 0;\\ \alpha_i^2 = \beta^2 &= 1.\end{aligned}$$

The anti-commutation rules immediately block the possibility to use scalars for the α_i and β. Moreover, they suggest to involve the Pauli matrices, which also anti-commute. However, together with the unit matrix, the Pauli matrices form an independent set of four objects, but the unit matrix fails to anti-commute with the Pauli matrices.

The simplest option to satisfy the conditions, is to choose four-dimensional matrices. Then, several acceptable forms of these matrices can be found – for example:

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix};$$
$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$$

where $\mathbb{1}$ denotes the 2×2 unit matrix.

This form of the α_i and the β has as a consequence that the wave function becomes four-dimensional. Note that this four-dimensionality has no direct relation to space-time being four-dimensional. The free-particle wave function ψ can now be written as

$$\psi(\mathbf{r}) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} e^{-\mathrm{i}p\cdot x},$$

where u_1 and u_2 are *two-dimensional* vectors. They are chosen because of the block structure of the 4×4 matrices α_i and the β. Writing the Dirac equation for the components of ψ, we obtain for u_1 and u_2:

$$(E - m)\, u_1 - \boldsymbol{\sigma} \cdot \mathbf{p}\, u_2 = 0; \tag{17.2}$$
$$-\boldsymbol{\sigma} \cdot \mathbf{p}\, u_1 + (E + m) u_2 = 0. \tag{17.3}$$

It is important to realise that the relativistic equation for the energy, $E^2 = p^2 + m^2$ still allows for both positive and negative energies. Let us first consider the energy to be positive, and the momentum to be small (this is the nonrelativistic limit). In that case, we see that the size of u_2 is *much smaller* than that of u_1. So the positive energy solution causes – in the weakly relativistic limit – one component of the four-spinor ψ to be much smaller than the other. For negative energies, the roles are reversed and u_2 is the large component and u_1 the small one. We now change notation and speak henceforth of the *large component* u_L and the *small component* u_S.

A second important note is that we can combine the two equations for u_L and u_S into a single one for u_L. Noting that $(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p}) = p^2$ (verify this!) we find

$$\left(|E| - m - \frac{p^2}{|E| + m} \right) u_L = 0.$$

In the nonrelativistic limit and positive energy, $E, m \gg |E - m|$, we see that

$$\frac{p^2}{2m} u_L \approx (E - m) u_L,$$

so that, as required, we recover the Schrödinger equation with energy $E - m$.

As for both u_L and u_S, their nonrelativistic Hamiltonian commutes with the Pauli matrix σ_z, we can assign a spin-value to u_L and to u_S. The two components of u_L are identified with the two spin states along the z-axis, corresponding to the eigenstates of σ_z:

$$u_L^+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad u_L^- = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and similar for $u_S^{\pm}$. We use the Dirac equation in the form (17.2) to find the full (non-normalized) four-vectors

$$u^+(p) = C\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{|E|+m} \\ \frac{p_x+\mathrm{i}p_y}{|E|+m} \end{pmatrix} \tag{17.4a}$$

and

$$u^-(p) = C\begin{pmatrix} 0 \\ 1 \\ \frac{p_x-\mathrm{i}p_y}{|E|+m} \\ \frac{-p_z}{|E|+m} \end{pmatrix}. \tag{17.4b}$$

For the negative energy solutions we can play the same game to arrive at

$$u^+(p) = C\begin{pmatrix} \frac{-p_z}{|E|+m} \\ -\frac{p_x+\mathrm{i}p_y}{|E|+m} \\ 1 \\ 0 \end{pmatrix} \tag{17.4c}$$

and

$$u^-(p) = C\begin{pmatrix} -\frac{p_x-\mathrm{i}p_y}{|E|+m} \\ \frac{p_z}{|E|+m} \\ 0 \\ 1 \end{pmatrix}. \tag{17.4d}$$

17.2 THE PROBABILITY DENSITY

One problem with the Klein-Gordon equation was the non-positiveness of the density. We now construct a density for the Dirac equation by requiring again that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0,$$

where $\rho = \psi^\dagger \psi$ (which is always positive). The wave functions satisfy the Dirac equation

$$\mathrm{i}\frac{\partial \psi}{\partial t} = H\psi = -\mathrm{i}\boldsymbol{\alpha} \cdot \nabla\psi + \beta m \psi$$

and its Hermitian conjugate is

$$-\mathrm{i}\frac{\partial \psi^\dagger}{\partial t} = \mathrm{i}\left(\boldsymbol{\nabla}\psi^\dagger\right) \cdot \boldsymbol{\alpha} + \beta m \psi^\dagger,$$

which directly leads to

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \left(\psi^\dagger \boldsymbol{\alpha} \psi\right),$$

so that we identify $\mathbf{j} = \psi^\dagger \boldsymbol{\alpha} \psi$.

It is very important that we have succeeded in finding a four-current with a density ρ which is always positive. This means that we have solved the problem of the sometimes negative densities of the Klein-Gordon equation! The problem of the negative energy states still remains. We shall come back to this later in this course.

17.3 A NEW FORM FOR THE DIRAC EQUATION

The Dirac equation can quite trivially be reformulated by introducing the so-called *gamma matrices*, γ^μ. These are 4×4 matrices that are expressed in terms of the α_j and β matrices:

$$\gamma^4 = \beta; \quad \gamma^j = \beta\alpha_j,$$

from which we easily find the explicit form of these matrices:

$$\gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$\gamma^4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Note that the first three of these have the form

$$\begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}.$$

The Dirac equation now takes the simple form

$$\left(\gamma^\mu p_\mu - m\right)\psi(\mathbf{r}, t) = 0.$$

In this equation, p_μ is the operator $\mathrm{i}\hbar\partial_\mu$ (time-dependent form) or $(E, -\mathrm{i}\hbar\partial_j)$, $j = 1,2,3$ in the stationary case (when ψ is a solution at energy E).

We now make a very important remark: *Although this equation looks relativistically invariant, this is not immediately clear from its form! It should be emphasised that the γ^μ do NOT transform under a Lorentz transformation, only the p_μ and the ψ do.* The transformation of the four vector ψ is necessary to obtain a relativistically invariant equation. We shall not go into details here (see section 35.7 of Desai's book) but it is necessary for guaranteeing covariance, that under a Lorentz transformation

$$x'^\mu = L^\mu_\nu x^\nu,$$

ψ transforms according to a 4×4 matrix S:

$$\psi'(x') = S\psi(x)$$

where S obeys the requirement that

$$S^{-1}\gamma^\mu S = L^\mu{}_\nu \gamma^\nu.$$

We list some properties of the gamma matrices:

$$\left(\gamma^4\right)^2 = 1;$$

$$\left\{\gamma^\mu, \gamma^\nu\right\} = 2g^{\mu\nu}; \tag{17.5}$$

and

$$\left(\gamma^\mu\right)^\dagger = \gamma^4\gamma^\mu\gamma^4.$$

17.4 SPIN-1/2 FOR THE DIRAC PARTICLE

The angular momentum operator can be viewed as the 'generator of rotations'. With this statement, we mean that, for a rotation around the z-axis for example, the rotation operator about an angle α can be written as

$$R = \exp(i\alpha L_z/\hbar).$$

Note that we have restored the $\hbar$ in the equation – this will turn out useful in this section. This means that any component of the angular momentum operator commutes with any spherically symmetric operator. Note that angular momentum comes in two 'flavors': the orbital angular momentum is defined as

$$L = \mathbf{r} \times \mathbf{p}$$

and an additional type of angular momentum which has no analogue in classical mechanics: the *spin*.

We consider a spherically symmetric problem, defined by a potential $V(r)$ and evaluate the commutator between the orbital angular momentum component L_j and the Hamiltonian

$$H = -\mathrm{i}\boldsymbol{\alpha}\cdot\nabla + \beta m + V(r).$$

Due to the spherical symmetry, we immediately verify that the only non-vanishing commutator is

$$[L_j, H] = \boldsymbol{\alpha}\cdot[L_j, \mathbf{p}] = \sum_{k=1}^{3} \alpha_k [L_j, p_k].$$

The expression on the right hand side can be evaluated from and from $[p_j, r_k] = -\mathrm{i}\hbar\delta_{jk}$, and leads to the result

$$[\mathbf{L}, H] = \mathrm{i}\hbar(\boldsymbol{\alpha}\times\mathbf{p}).$$

We see that the orbital angular momentum does not commute with the Hamiltonian, which is quite strange in view of the rotational symmetry the Hamiltonian should have. The solution to this apparent paradox is the notion that we have neglected the spin, which is an essential part of the total angular momentum. A natural form of the spin is the vector operator with components

$$\sigma'_j = \begin{pmatrix} \sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix}.$$

Here $j = x, y$ or z and σ'_j is a 4×4 matrix containing two Pauli spin matrices on the diagonal. Note that the prime with σ indicates that we are dealing with a 4×4 matrix containing two Pauli spin matrices. Note the difference with the boldface $\boldsymbol{\sigma}$ which denotes a vector operator with the three 2×2 matrices σ_x, σ_y and σ_z as its components. The vector operator $\boldsymbol{\sigma}'$ has the three 4×4 matrices σ'_j, $j = x, y, z$ as its components. Using the relation $[\sigma_j, \sigma_k] = 2\mathrm{i}\epsilon_{jkl}\sigma_l$, the commutator of this operator with H can be worked out straightforwardly and the result is

$$[\boldsymbol{\sigma}', H] = 2\mathrm{i}(\mathbf{p}\times\boldsymbol{\alpha}).$$

Combining this with the commutation relation for the orbital angular momentum operator, we obtain

$$\left[\left(\mathbf{L} + \frac{\hbar}{2}\boldsymbol{\sigma}'\right), H\right] = 0,$$

i.e. the Hamiltonian *does* commute with the *total* angular momentum, provided we assume that the particles have spin-1/2, described by the operator $\hbar\boldsymbol{\sigma}'/2$. The total angular momentum

$$\mathbf{J} = \mathbf{L} + \frac{\hbar}{2}\boldsymbol{\sigma}'$$

therefore commutes with H.

17.5 THE HYDROGEN ATOM

In section 17.1 we have found the solutions of the Dirac equation with a constant potential: they are plane waves with a four-dimensional amplitude vector. The constant potential is also a standard problem that we can solve for in the Schrödinger equation case. Other problems for which the Schrödinger equation can be solved exactly are the harmonic oscillator and the hydrogen atom. It turns out that the hydrogen atom can also be solved for in the case of the Dirac equation. This turns out to be a very useful exercise as the Dirac equation contains the spin, which should give us important parts of the solution such as the spin-orbit coupling. For the full solution, we refer to Desai, sections 34.3-6. Here we shall restrict ourselves to considering the weakly relativistic limit, and show that all the fine structure terms of the hydrogen atom naturally emerge from the Dirac equation.

Similar to the standard approach one usually follows to obtain a stationary Schrödinger equation, we construct a stationary Dirac equation by writing

$$\psi(\mathbf{r}, t) = \exp(-\mathrm{i}Et)\phi(\mathbf{r})$$

which, after plugging this into the Dirac equation leads to the following stationary equation:

$$\left[E - \boldsymbol{\alpha}\cdot\mathbf{p} - \beta m - V(r)\right]\phi(\mathbf{r}) = 0,$$

where we have replaced E by $E - V(r)$ to account for the (spherically symmetric) potential. Note that the solution to this equation is a four vector. Writing this solution as

$$\phi = \begin{pmatrix} \phi_{\mathrm{L}} \\ \phi_{\mathrm{S}} \end{pmatrix},$$

where S and L stand for small and large respectively, and plugging this into the Dirac equation, we obtain

$$(E - m - V)\,\phi_{\mathrm{L}} - \boldsymbol{\sigma}\cdot\mathbf{p}\phi_{\mathrm{S}} = 0;$$

$$-\boldsymbol{\sigma}\cdot\mathbf{p}\phi_{\mathrm{L}} + (E + m - V)\,\phi_{\mathrm{S}} = 0.$$

We now show that the in the weakly relativistic limit the fine structure terms emerge from this equation.

To this end, we take $E - m \ll m$. We define $E_{\mathrm{T}} = E - m$ to rewrite the two equations we just derived in the form

$$(E_{\mathrm{T}} - V)\,\phi_{\mathrm{L}} = \boldsymbol{\sigma}\cdot\mathbf{p}\phi_{\mathrm{S}};$$

$$(2m + E_{\mathrm{T}} - V)\,\phi_{\mathrm{S}} = \boldsymbol{\sigma}\cdot\mathbf{p}\phi_{\mathrm{L}}.$$

Now we use the second of these equations to eliminate ϕ_{S} in the first:

$$(E_{\mathrm{T}} - V)\,\phi_{\mathrm{L}} = \left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\frac{1}{2m + E_{\mathrm{T}} - V}\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\phi_{\mathrm{L}}.$$

It would seem that we have a weakly relativistic Dirac equation for the spinor represented by the large component which has the form similar to the stationary Schrödinger equation. However, there is a subtle problem with this equation. Usually when we solve an equation like this, we search for a solution which is normalized to 1. To see this, imagine that we minimize the expectation value of the energy:

$$\langle E\rangle = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$$

with respect to ψ. This is most easily done by varying $\psi \to \psi + \delta\psi$ and then requiring that the first-order Taylor terms of vanish. Some calculation leads to two equations. The first is

$$H|\psi\rangle = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}|\psi\rangle$$

and the second is the Hermitian conjugate of this equation (which does not yield new information). Replacing then $\langle\psi|H|\psi\rangle / \langle\psi|\psi\rangle$ by E, the standard stationary Schrödinger equation $H\left|\psi\right\rangle = E\left|\psi\right\rangle$ is recovered. Obviously, the problem can also be solved by finding the minimum of the numerator under the condition that $\langle\psi|\psi\rangle$ is normalized to 1. However, in the present case, it is ϕ, and not $\phi_{\rm L}$ which is normalized to 1. Using

$$\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right) = \sigma_j p_j \sigma_k p_k = \mathbf{p}^2,$$

we can write, using $m \gg E, V$:

$$\langle\phi_{\rm S}|\phi_{\rm S}\rangle = \left\langle\phi_{\rm L}\left|\frac{p^2}{4m^2}\right|\phi_{\rm L}\right\rangle,$$

so that

$$\int \phi^\dagger\phi d^3r = 1 = \int \left[|\phi_{\rm L}|^2 + \phi_{\rm L}^\dagger\frac{p^2}{4m^2}\phi_{\rm L}\right] d^3r.$$

Therefore we now introduce a two-spinor function ψ in terms of $\phi_{\rm L}$:

$$\psi \approx \left(1 + \frac{p^2}{8m^2}\right)\phi_{\rm L}$$

which is properly normalized to 1 for $\left|p/m\right| \ll 1$. Inverting this relation we have

$$\phi_{\rm L} \approx \left(1 - \frac{p^2}{8m^2}\right)\psi.$$

We shall consistently neglect corrections of the order p^4/m^4. Rephrasing our weakly relativistic Dirac equation as an equation for the properly normalized function ψ, we obtain

$$(E_{\rm T} - V)\left(1 - \frac{p^2}{8m^2}\right)\psi = \left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\frac{1}{2m + E_{\rm T} - V}\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\left(1 - \frac{p^2}{8m^2}\right)\psi.$$

Approximating the term on the right hand side within our consistent approach:

$$\frac{1}{2m + E_{\rm T} - V} \approx \frac{1}{2m} - \frac{E_{\rm T} - V}{4m^2},$$

we obtain after some calculation

$$(E_{\rm T} - V)\left(1 - \frac{p^2}{8m^2}\right)\psi = \left[\frac{p^2}{2m} - \frac{p^4}{16m^3} - E_{\rm T}\frac{p^2}{4m^2} + \frac{\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)V\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)}{4m^2}\right]\psi.$$

Reorganising this equation further and multiplying left and right hand side by $1 - p^2/(8m^2)$ we obtain:

$$E_{\rm T}\psi = \left[\frac{p^2}{2m} - \frac{p^4}{8m^3} + V - \frac{p^2V + Vp^2}{8m^2} + \frac{\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)V\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)}{4m^2}\right]\psi.$$

This equation contains the somewhat intricate term $\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)V\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)$. This can be reworked using

$$(\boldsymbol{\sigma}\cdot\mathbf{a})\,(\boldsymbol{\sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + \mathrm{i}\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})$$

to

$$\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)V\left(\boldsymbol{\sigma}\cdot\mathbf{p}\right)\psi = \left(\mathbf{p}V\right)\cdot\left(\mathbf{p}\psi\right) + \mathrm{i}\boldsymbol{\sigma}\cdot\left[\left(\mathbf{p}V\right)\times\left(\mathbf{p}\psi\right)\right].$$

Working out all the terms, one obtains, after some algebra:

$$H\psi = E_{\rm T}\psi,$$

where

$$H = \frac{p^2}{2m} - \frac{p^4}{8m^3} + V + \frac{\nabla^2 V}{8m^2} + \frac{\boldsymbol{\sigma}\left(\nabla V \times \mathbf{p}\right)}{4m^2}.$$

In addition to the nonrelativistic Hamiltonian

$$H_{\text{nonrel}} = \frac{p^2}{2m} + V,$$

we see that this equation contains corrections which we shall now discuss.

The term $p^4/(8m^3)$ is a correction due to the kinetic energy and does not depend on the potential V. Two more terms do depend on V. The term $\nabla^2 V/(8m^2)$ is called the *Darwin term*. Using the fact that $\nabla V = 4\pi\delta(\mathbf{r})$ for the specific case of the Coulomb potential of a charge located at the origin, it can be formulated as a correction term located only at the origin.

The last term can also be reformulated. Using the fact that V is spherically symmetric, we have

$$\nabla V = \frac{\mathbf{r}}{r}\frac{dV}{dr},$$

so that we can write

$$\boldsymbol{\sigma} \cdot \left(\nabla V \times \mathbf{p}\right) = \frac{1}{r}\frac{dV}{dr}\boldsymbol{\sigma} \cdot \left(\mathbf{r} \times \mathbf{p}\right) = \frac{1}{r}\frac{dV}{dr}\boldsymbol{\sigma} \cdot (\mathbf{r} \times \mathbf{p}) = \frac{1}{r}\frac{dV}{dr}\boldsymbol{\sigma} \cdot \mathbf{L}.$$

This is the famous *spin-orbit coupling*.

17.6 INTERACTION WITH AN ELECTROMAGNETIC FIELD

The Dirac equation can easily be formulated for a particle moving in an electromagnetic field. Such a field contains an electric potential ϕ and a vector potential $\mathbf{A}$. We follow the same rules as in the ordinary Schrödinger equation:

$$\mathbf{p} \to \mathbf{p} - e\mathbf{A}$$

and

$$E \to E - e\phi.$$

This is a relativistically correct procedure as we can write this transformation as

$$p^\mu \to p^\mu - eA^\mu.$$

The energy-momentum relation in this case is replaced by the condition

$$(E - e\phi)^2 = \left(\mathbf{p} - e\mathbf{A}\right)^2 + m^2,$$

which, in the nonrelativistic limit, yields (by taking the square root of the last equation):

$$E - e\phi \approx m + \frac{\left(\mathbf{p} - e\mathbf{A}\right)^2}{2m}.$$

The Dirac equation now reads:

$$\left[\gamma^\mu (p_\mu - eA_\mu) - m\right]\psi = 0.$$

Multiplying this from the left by $\gamma^\nu(p_\nu - eA_\nu) + m$, yields

$$\left[\gamma^\nu (p_\nu - eA_\nu)\gamma^\mu (p_\mu - eA_\mu) - m^2\right]\psi = 0.$$

We now consider the factor $\gamma^\mu\gamma^\nu$, which we write in the form

$$\gamma^\mu\gamma^\nu = \frac{1}{2}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right) + \frac{1}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right).$$

The first term is an anti-commutator which is equal to $g^{\mu\nu}$ (see Eq. (17.5)). The second term is defined as

$$\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu = -2\mathrm{i}\Sigma^{\mu\nu}.$$

As $\Sigma^{\mu\nu}$ is a commutator, the diagonal elements ($\mu = \nu$) are zero. The elements Σ^{4i} are equal to

$$\Sigma^{4i} = \mathrm{i}\begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \qquad i = x, y, z.$$

The remaining nonvanishing terms can be shown to be

$$\sigma^{ij} = 2\epsilon_{ijk}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \qquad i, j, k = x, y, z.$$

All in all, we have

$$\left[(p^\mu - eA^\mu)(p_\mu - eA_\mu) - \mathrm{i}\Sigma^{\mu\nu}(p_\mu - eA_\mu)(p_\nu - eA_\nu) - m^2\right]\psi(x) = 0.$$

Using the anti-symmetry of Σ (i.e. $\Sigma^{\mu\nu} = -\Sigma^{\nu\mu}$), we can write the second term as

$$-\mathrm{i}\Sigma^{\mu\nu}(p_\mu - eA_\mu)(p_\nu - eA_\nu)\psi = -\frac{e}{2}\Sigma^{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi$$

where in the last form, the partial derivatives only act on the components of A and not on ψ as can be seen by carefully writing out all terms. The term in parenthesis is recognised as the Maxwell tensor (see Eq. (16.2)):

$$-\frac{e}{2}\Sigma^{\mu\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu) = -\frac{e}{2}\Sigma^{\mu\nu}F_{\mu\nu}.$$

Using the fact that the electric field components are $E_i = F_{0i}$ and that the magnetic field is given as $B_i = -\frac{1}{2}\epsilon_{ijk}F_{jk}$, we have

$$\frac{e}{2}\Sigma_{\mu\nu}F_{\mu\nu} = e\left(\boldsymbol{\sigma}'\cdot\mathbf{B} - \mathrm{i}\boldsymbol{\alpha}\cdot\mathbf{E}\right).$$

All in all, the Dirac equation for an electron in an electromagnetic field becomes

$$\left[(E - e\phi)^2 - (\mathbf{p} - e\mathbf{A})^2 + e\boldsymbol{\sigma}'\cdot\mathbf{B} + \mathrm{i}e\boldsymbol{\alpha}\cdot\mathbf{E} - m^2\right]\psi = 0.$$

It can be shown that, in the weakly relativistic limit, the term involving the electric field becomes much smaller than the other ones. The term involving the magnetic field is the natural coupling between the spin and the magnetic field. If we restore the prefactors and give it a dimension of energy, it reads

$$\frac{e\hbar}{2mc}\boldsymbol{\sigma}';$$

the prefactor is called the *Bohr magneton.* Restricting ourselves to the large component of the spinor in the nonrelativistic limit, we obtain

$$\mathrm{i}\hbar\frac{\partial\psi_L}{\partial t} = \left[\frac{1}{2m}(\mathrm{i}\hbar\nabla + e\mathbf{A})^2 - \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B} + e\phi\right]\psi_L.$$

This equation can also be derived in a more direct way for a plane wave solution. Starting from the Dirac equation for this plane wave we have

$$\begin{pmatrix} E - e\phi - m & -\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A}) \\ -\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A}) & E - e\phi + m \end{pmatrix}\begin{pmatrix} u_L \\ u_S \end{pmatrix} = 0$$

which yields the two equations

$$(E - e\phi - m)u_L - \boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})u_S = 0,$$

and

$$(E - e\phi + m)u_S - \boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})u_L = 0.$$

In the nonrelativistic approximation, $|E - m| \ll m$ and $e\phi \ll m$, and we can write the second equation as

$$u_S = \frac{\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})}{E - e\phi + m}u_L \approx \frac{\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})}{2m}u_L.$$

Inserting this into the first equation gives

$$(E - e\phi - m)u_L = \boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})\frac{\boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})}{2m}u_L.$$

This then yields, after some algebra

$$\left(E - m - e\phi - \frac{1}{2m}(\mathbf{p} - e\mathbf{A})^2 + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B}\right)u_L = 0.$$

This equation is in accordance with the result of our more general calculation, but the derivation is not so general as we have assumed a plane wave form for the solution.

17.7 PROBLEMS

1. Establish the following properties of the γ matrices:
   (a) $\gamma_\nu^2 = g_{\nu\nu}\mathbb{1}$.
   (b) $g^{\mu\nu}\gamma_\mu\gamma_\nu = 4\mathbb{1}$.
   (c) $\gamma_\lambda\gamma_\nu\gamma_\lambda = -\gamma_\nu$.
   (d) $\gamma_\lambda\gamma_\nu\gamma_\kappa\gamma_\lambda = 4\delta_{\kappa\nu}$.
   (e) $\gamma_\lambda\gamma_\nu\gamma_\kappa\gamma_\mu\gamma_\lambda = -2\gamma_\mu\gamma_\lambda\gamma_\nu$.
   (f) $\mathrm{Tr}(\gamma_\kappa\gamma_\nu) = 4\delta_{\kappa\nu}$.

2. *The Dirac equation with a central potential, Part 1.*
   We consider a particle in a radial potential. We define the operator

$$K = \beta\left(\frac{\boldsymbol{\sigma}'\cdot\mathbf{L}}{\hbar} + 1\right),$$

and the operator

$$p_r = -\mathrm{i}\hbar\frac{1}{r}\frac{\partial}{\partial r}r = \frac{1}{r}\left(\mathbf{r}\cdot\mathbf{p} - \mathrm{i}\hbar\right).$$

This is the radial momentum operator. Also, we define

$$\alpha_r = \frac{1}{r}\left(\boldsymbol{\alpha}\cdot\mathbf{r}\right).$$

(a) Show that

$$\boldsymbol{\sigma}\times\boldsymbol{\sigma} = \mathrm{i}\boldsymbol{\sigma}.$$

for the vector $\boldsymbol{\sigma}$ which has the Pauli matrices as its components. (This result can be generalized to an arbitrary angular momentum operator **J**.)

(b) Show that

$$(\boldsymbol{\alpha}\cdot\mathbf{r})\left(\boldsymbol{\alpha}\cdot\hat{\mathbf{p}}\right) = \mathbf{r}\cdot\mathbf{p} + \mathrm{i}\boldsymbol{\sigma}'\mathbf{L} = rp_r + \mathrm{i}\left(\boldsymbol{\sigma}\cdot\mathbf{L} + \hbar\right).$$

(c) Multiplying this from the left by α_r/r, show that

$$\boldsymbol{\alpha}\cdot\mathbf{p} = \alpha_r\left[p_r + \frac{\mathrm{i}}{r}\beta K\right].$$

This leads to the following form of the Dirac Hamiltonian:

$$H = c\alpha_r p_r + \mathrm{i}\hbar\frac{c}{r}\alpha_r\beta K + \beta mc^2 + V(r).$$

3. *Dirac equation with a spherical potential, Part 2*

We proceed with the Dirac Hamiltonian of problem 2.

(a) Show that $[K, H] = 0$.

(b) The matrix operator K is block diagonal; hence it operates separately on two spinor wave functions ϕ_L and ϕ_S. Take as a basis in the space of spinor wave functions, $Y_l^m \chi_\pm$, where

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \qquad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are the spin-up and down spinors, respectively, and Y_l^m is the usual spherical harmonic function.

Show these basis functions are all eigenstates of $\hbar K$ and that the eigenvalues are $\hbar\kappa$, where $\kappa = \pm(j + 1/2)$.

4. **H** We consider a system described by the Dirac Hamiltonian

$$H = -\mathrm{i}\boldsymbol{\alpha} \cdot \nabla + \beta m.$$

We take $\hbar = c = 1$.

(a) Show that the Heisenberg equation

$$-\mathrm{i}\frac{\partial Q}{\partial t}(t) = (H, Q),$$

applied to the position operator $x(t)$ gives

$$\frac{\partial x}{\partial t}(t) = \alpha_x.$$

So, α_x can be seen as a velocity (measured in units of c) in the x-direction.

(b) Show that α_x(t) by itself satisfies the equation of motion

$$\frac{\partial \alpha_x}{\partial t}(t) = 2\mathrm{i}\left[p_x - \alpha_x H\right].$$

(c) Solve the two equations of motion for α_x and x to obtain

$$\alpha_x(t) = Ae^{-2\mathrm{i}Ht} + p_x H^{-1}$$

and express A in terms of $\alpha_x(0)$. Hint: use the time independence of H and of p_x.

(d) Verify that

$$x(t) = x(0) + p_x H^{-1} t + \frac{1}{2}\mathrm{i}H^{-1}\left[\alpha_x(0) - p_x H^{-1}\right]\left(e^{-2\mathrm{i}Ht} - 1\right).$$

We see that the second term contains an oscillatory contribution. It can be shown that it arises from an interference between positive- and negative-energy waves. This phenomenon is known as the *Zitterbewegung*.

5. **H** In 2010, the Nobel prize was awarded to A. Geim and K. Novoselov for their research into the physical properties of electrons in graphene, which is a single layer of carbon atoms, ordered in a hexagonal grid. One of the striking properties of these electrons is that they behave as massless spin-1/2 particles in two dimensions. In 1929, Klein had analyzed the properties of Dirac particles when they scatter off a potential barrier, and

in 2006, Katsnelson, Geim and Novoselov published a paper in Nature Physics on an experimental verification in graphene of this process. The so-called Klein paradox is addressed in this problem.

We consider electrons incident on a potential barrier

$$V(z) = \begin{cases} V_0 & \text{for } z > 0; \\ 0 & \text{for } z < 0. \end{cases}$$

For $z < 0$, the Dirac equation reads

$$\left(\alpha p + \beta m\right)\psi = E\psi.$$

For $z > 0$ we have in turn:

$$\left(\alpha p + \beta m\right)\psi = (E - V_0)\psi.$$

Note that α is the matrix $\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}$ with σ_z the Pauli matrix. Also, $p = p_z$. From the Dirac theory, we know that an incident spin-up particle is described by

$$\psi_{\text{inc}} = \begin{pmatrix} 1 \\ 0 \\ p/(E+m) \\ 0 \end{pmatrix} e^{ipz}.$$

As the potential step is not expected to change the spin of the particle, we expect the reflected and transmitted wave to only have spin-up character. The reflected wave has the form

$$\psi_{\text{refl}} = a \begin{pmatrix} 1 \\ 0 \\ -p/(E+m) \\ 0 \end{pmatrix} e^{-ipz}$$

and the transmitted wave is

$$\psi_{\text{trans}} = b \begin{pmatrix} 1 \\ 0 \\ p'/(E - V_0 + m) \\ 0 \end{pmatrix} e^{ip'z}.$$

Show that matching the solution at $z = 0$ gives the following equations:

$$\begin{aligned} 1 + a &= b; \\ 1 - a &= \frac{p'}{p}\frac{E+m}{E - V_0 + m} b \equiv rb. \end{aligned}$$

We distinguish between three cases for the value of V_0:

$$\begin{cases} V_0 < E - m & \text{case I;} \\ E - m \le V_0 < E + m & \text{case II;} \\ V_0 \ge E + m & \text{case III.} \end{cases}$$

Show that in case I,

$$r = \sqrt{\frac{(V_0 - E + m)(E+m)}{(V_0 - E - m)(E - m)}}.$$

whereas for case II,

$$r = i\sqrt{\frac{(V_0 - E + m)(E + m)}{(E - V_0 + m)(E - m)}}.$$

and, in case III:

$$r = -\sqrt{\frac{(V_0 - E + m)(E + m)}{(V_0 - E - m)(E - m)}}.$$

Also show that the reflected current versus the incident current is given by

$$\frac{j_{\text{refl}}}{j_{\text{inc}}} = |a|^2 = \left|\frac{1 - r}{1 + r}\right|^2 .$$

Show that in region I, the reflected flux is less than 1, that in region 2 the reflected flux is 1, and in region 3 it exceeds 1!

Find an explanation for this unexpected result.

18

SECOND QUANTIZATION FOR RELATIVISTIC PARTICLES

18.1 INTRODUCTION

In this chapter, we apply the second quantization formalism to relativistic particles. As we have seen in the previous chapters, we need to describe the particles with the appropriate relativistic quantum equation: the Klein-Gordon equation for relativistic bosons and the Dirac equation for relativistic fermions. We start this chapter, by considering these two cases in some detail. Then we analyze an important example of fermion systems in some detail: electrons in graphene.

18.2 SECOND QUANTIZATION FOR KLEIN-GORDON PARTICLES

We first formulate the procedure of second quantization for the Klein-Gordon field. Let us first recall this equation:

$$\left(-\nabla^2+\partial_t^2+m^2\right)\phi(\mathbf{r},t)=0.$$

Looking at the solutions to this equation, we concluded in chapter 16, that there are two differences in comparison with the the Schrödinger particles considered in chapter 9:

- We have particles with positive and with negative energies
- We have a different expression for the density, which is given by

$$\rho(\mathbf{r},t)=\mathrm{i}\phi^\dagger(\mathbf{r},t)\overset{\leftrightarrow}{\partial}_t\phi(\mathbf{r},t)=\mathrm{i}\left[\phi^\dagger(\mathbf{r},t)\partial_t\phi(\mathbf{r},t)-\partial_t\phi^\dagger(\mathbf{r},t)\partial_t\phi(\mathbf{r},t)\right].$$

The first point leads to the conclusion that the wavefunctions have the form

$$Ce^{-\mathrm{i}(\omega t-\mathbf{p}\cdot\mathbf{r})}\text{ or }Ce^{\mathrm{i}(\omega t+\mathbf{p}\cdot\mathbf{r})},$$

where $\omega^2=\mathbf{p}^2+m^2$ is *always* positive, therefore $E=\omega$ (first case) or $E=-\omega$ (second case).

The normalization condition gives us the constant C. Evaluating the inner product of two plane waves with positive energies, and momenta $\mathbf{p}$ and $\mathbf{p}'$, we obtain, using the Klein-Gordon density:

$$\int d^3r\,\left[\mathrm{i}e^{\mathrm{i}(\omega_{\mathbf{p}}t-\mathbf{p}\cdot\mathbf{r})}\left(-\mathrm{i}\omega_{\mathbf{p}'}\right)e^{-\mathrm{i}(\omega_{\mathbf{p}'}t-\mathbf{p}'\cdot\mathbf{r})}-\mathrm{i}\left(\mathrm{i}\omega_{\mathbf{p}}\right)e^{\mathrm{i}(\omega t-\mathbf{p}\cdot\mathbf{r})}e^{-\mathrm{i}(\omega_{\mathbf{p}'}t-\mathbf{p}'\cdot\mathbf{r})}\right]=2\omega_{\mathbf{p}}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}').$$

Also, performing a similar calculation for a positive and a negative energy state, we obtain

$$\int d^3r\,\left[\mathrm{i}e^{-\mathrm{i}(\omega_{\mathbf{p}}t-\mathbf{p}\cdot\mathbf{r})}\left(-\mathrm{i}\omega_{\mathbf{p}'}\right)e^{-\mathrm{i}(\omega_{\mathbf{p}'}t-\mathbf{p}'\cdot\mathbf{r})}-\mathrm{i}\left(-\mathrm{i}\omega_{\mathbf{p}}\right)e^{-\mathrm{i}(\omega_{\mathbf{p}}t-\mathbf{p}\cdot\mathbf{r})}e^{-\mathrm{i}(\omega_{\mathbf{p}'}t-\mathbf{p}'\cdot\mathbf{r})}\right]=0.$$

So we see that properly normalized positive-energy wavefunctions have the form

$$\phi^{(+)}(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_{\mathbf{p}}}} e^{-\mathrm{i}(\omega_{\mathbf{p}} t - \mathbf{p}\cdot\mathbf{r})}.$$

For negative-energy wave functions, the wavefunction

$$\phi^{(-)}(\mathbf{r},t) = \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_{\mathbf{p}}}} e^{\mathrm{i}(\omega t + \mathbf{p}\cdot\mathbf{r})}$$

is normalized to -1 as can be checked by a calculation similar to that for $\phi^{(+)}$. Furthermore, $\phi^{(+)}$ and $\phi^{(-)}$ are orthogonal.

It is now possible to formulate field operators by a similar expansion as in chapter 9:

$$\hat{\phi}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3p}{\sqrt{2\omega_{\mathbf{p}}}} \left[e^{-\mathrm{i}(\omega t - \mathbf{p}\cdot\mathbf{r})} a(\mathbf{p}) + e^{\mathrm{i}(\omega t + \mathbf{p}\cdot\mathbf{r})} b(\mathbf{p}) \right].$$

First we realize that within the second integral we can replace $\mathbf{p}$ by $-\mathbf{p}$, and, using $p\cdot x = \omega t - \mathbf{p}\cdot\mathbf{r}$, we obtain

$$\hat{\phi}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3p}{\sqrt{2\omega_{\mathbf{p}}}} \left[e^{-\mathrm{i}p\cdot x} a(\mathbf{p}) + e^{\mathrm{i}p\cdot x} b(-\mathbf{p}) \right].$$

If we now require that the field operators are Hermitian fields ($\hat{\phi}(\mathbf{r},t) = \hat{\phi}^\dagger(\mathbf{r},t)$), we have $a^\dagger(\mathbf{p}) = b(-\mathbf{p})$. We then have

$$\hat{\phi}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3p}{\sqrt{2\omega_{\mathbf{p}}}} \left[e^{-\mathrm{i}p\cdot x} a(\mathbf{p}) + e^{\mathrm{i}p\cdot x} a^\dagger(\mathbf{p}) \right].$$

Let us consider this form. We see that the field operators are composed of creation operators of particles with momentum $\mathbf{p}$ and of annihilation operators with momentum $\mathbf{p}$, which are equivalent to creation operators for momentum states $-\mathbf{p}$. Both types of particles are assumed to have positive energies – therefore the Hamiltonian will have the form

$$H = \frac{1}{2}\int d^3p\, \hbar\omega \left[a^\dagger(\mathbf{p}) a(\mathbf{p}) + a(\mathbf{p}) a^\dagger(\mathbf{p}) \right].$$

Of course you will now question how we can write this form from the information we had so far – to be honest, we cannot. A more sophisticated field-theoretical description however confirms that this is the right form (see for example Bjorken and Drell, *Relativistic Quantum Fields*).

Using the Boson commutation relations for the $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$:

$$\left[a(\mathbf{p}), a^\dagger(\mathbf{p}) \right] = 1,$$

we can rewrite the Hamiltonian as

$$H = \int d^3p \hbar\omega \left[a^\dagger(\mathbf{p}) a(\mathbf{p}) + 1/2 \right].$$

We see that the ground state energy is infinite! This is another example of an infinity occurring in field theory – previously we encountered this when discussing the Casimir effect.

In order to resolve the 'inifinite energy problem', we should realize that the *values* of the energies do not matter – it is *energy differences* that drive physical processes. Therefore, the physics does not change if we subtract the ground state energy from *all* energies. This leads to the 'renormalized' Hamiltonian

$$H = \int d^3p\, \hbar\omega a^\dagger(\mathbf{p}) a(\mathbf{p}).$$

In addition to this, we can count the number of particles as

$$N = \int d^3p\, a^\dagger(\mathbf{p}) a(\mathbf{p}).$$

These definitions lead to a field theory with positive energies and particle numbers (or densities). The fact that we have not propely shown that our expressions for the Hamiltonian and particle number are correct may be somewhat unsatisfactory – for Dirac particles the situation is better so we quickly move on to that case.

18.3 SECOND QUANTIZATION FOR DIRAC PARTICLES

When finding plane wave solutions for the Dirac equation, we distinguish between the solutions with positive and those with negative energy. We write spinors with positive energy as

$$\langle x|\psi_s\rangle = u_s e^{-\mathrm{i}p\cdot x}$$

where

$$\left(\gamma\cdot p - m\right) u_s(\mathbf{p}) = 0,$$

where the fourth component of p in the left hand side is $E_p = \sqrt{\mathbf{p}^2 + m^2}$, $s = \pm$ denotes the spin. Just as in the case of the Klein-Gordon equation, we also have negative energy solutions of the form:

$$\psi(\mathbf{r}, t) = v_s e^{\mathrm{i}p\cdot x}.$$

Putting this into the Dirac equation leads to

$$\left(\gamma\cdot p + m\right) v_s(\mathbf{p}) = 0.$$

The question is now what the spinors u_s and v_s look like. We have already encountered them in chapter 17. They have the form

$$u_+(\mathbf{p}) = C \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{|E|+m} \\ \frac{p_x + \mathrm{i}p_y}{|E|+m} \end{pmatrix}$$

and

$$u_-(\mathbf{p}) = C \begin{pmatrix} 0 \\ 1 \\ \frac{p_x - \mathrm{i}p_y}{|E|+m} \\ -\frac{p_z}{|E|+m} \end{pmatrix};$$

$$v_+(\mathbf{p}) = C \begin{pmatrix} \frac{p_z}{|E|+m} \\ \frac{p_x + \mathrm{i}p_y}{|E|+m} \\ 1 \\ 0 \end{pmatrix}$$

and

$$v_-(\mathbf{p}) = C \begin{pmatrix} \frac{p_x - \mathrm{i}p_y}{|E|+m} \\ -\frac{p_z}{|E|+m} \\ 0 \\ 1 \end{pmatrix}.$$

The important question now is what the C's are in order for u and v to be normalized. We have seen in section 17.2 that the correct four-verctor form of the current is

$$(\mathbf{j}, \rho) = (\psi^\dagger \boldsymbol{\alpha}\psi, \psi^\dagger\psi).$$

This can also be written in the more elegant form

$$j^\mu = \psi^\dagger \gamma^4 \gamma^\mu \psi.$$

The quantity

$$\psi^\dagger(x)\gamma^4$$

occurs so often that a special notation is used for it:

$$\bar{\psi} = \psi^\dagger(x)\gamma^4.$$

Now there is a subtle problem. The norm of the wavefunction should *always* be 1, i.e. is should be a relativistically invariant scalar. But the density transforms as the 4-th component of a four vector, and is therefore not an invariant scalar itself! It turns out that the quantity $\bar{\psi}\psi$ is invariant (see Bjorken and Drell; this result is also used in Desai, sec. 35.7).

For $\mathbf{p} = 0$, the forms of u_s and v_s are taken to be

$$u_+ = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \quad u_- = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

and

$$v_+ = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}; \quad v_- = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

For $\mathbf{p} \neq 0$ our norm $\bar{u}_s u_{s'}$ remains constant. So we use the relativistically invariant normalization conditions:

$$\bar{u}_s u_{s'} = -\bar{v}_s v_{s'} = \delta_{ss'}, \qquad \bar{u}_s v_{s'} = \bar{v}_s u_{s'} = 0.$$

Note the minus-sign in the normalization for the v's: it is due to the γ^4 appearing in the definition of $\bar{\psi}$. These normalisation conditions fix C for u_s to

$$|C|^2 \left(1 - \frac{p^2}{(\omega+m)^2}\right) = 1$$

(note that the minus-sign is caused by the γ^4 in the definition of the norm). Therefore, we see that

$$u_s^\dagger u_{s'} = |C|^2 \left(1 + \frac{p^2}{(\omega+m)^2}\right) \delta_{ss'}$$

which, after some calculus, can be seen to yield

$$u_s^\dagger u_{s'} = \frac{\omega}{m} \delta_{ss'}.$$

The same holds for the v_s:

$$v_s^\dagger v_{s'} = \frac{\omega}{m} \delta_{ss'}.$$

Interestingly, there is a 'skewed' orthogonality condition between the u's and the v's:

$$u_s^\dagger(\mathbf{p}) v_{s'}(-\mathbf{p}) = v_s^\dagger(\mathbf{p}) u_{s'}(-\mathbf{p}) = 0. \tag{18.1}$$

This is natural as can be seen by writing out the normalization conditions including the plane wave part:

$$\int d^3 r\, u_s^\dagger(\mathbf{p}) e^{\mathrm{i}p\cdot x} v_s(\mathbf{p}') e^{\mathrm{i}p'\cdot x}.$$

The integral over $\mathbf{r}$ forces $\mathbf{p} = -\mathbf{p}'$, and for this situation, the above orthogonality condition works.

We see that a properly orthonormalized set of wavefunctions is given by

$$\sqrt{\frac{m}{\omega_{\mathbf{p}}}} e^{-\mathrm{i}p\cdot x} u_s(\mathbf{p}); \qquad \sqrt{\frac{m}{\omega_{\mathbf{p}}}} e^{\mathrm{i}p\cdot x} v_s(\mathbf{p}).$$

From this, we can immediately write down the field operator

$$\boldsymbol{\psi}(x) = \sum_{s=\pm} \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p \sqrt{\frac{m}{E_p}} \left[u_s(\mathbf{p}) a_s(\mathbf{p}) e^{-\mathrm{i}p\cdot x} + v_s(\mathbf{p}) b_s^\dagger(\mathbf{p}) e^{\mathrm{i}p\cdot x} \right].$$

where the $a_s(\mathbf{p})$, $a_s^\dagger(\mathbf{p})$ annihilate and create positive and negative energy particles with positive energy, and the operators $b_s(\mathbf{p})$, $b_s^\dagger(\mathbf{p})$ create or annihilate a particle with negative energy.

Obviously, these operators satisfy the anti-commutator relation:

$$\left\{a_s(\mathbf{p}), a_{s'}^\dagger(\mathbf{q})\right\} = \delta_{ss'}\delta^{(3)}(\mathbf{p}-\mathbf{q}); \qquad \left\{a_s(\mathbf{p}), a_{s'}(\mathbf{q})\right\} = \left\{a_s^\dagger(\mathbf{p}), a_{s'}^\dagger(\mathbf{q})\right\} = 0;$$

$$\left\{b_s(\mathbf{p}), b_{s'}^\dagger(\mathbf{q})\right\} = \delta_{ss'}\delta^{(3)}(\mathbf{p}-\mathbf{q}); \qquad \left\{b_s(\mathbf{p}), b_{s'}(\mathbf{q})\right\} = \left\{b_s^\dagger(\mathbf{p}), b_{s'}^\dagger(\mathbf{q})\right\} = 0.$$

All anti-commutators involving a *a and* a *b* vanish.

The Hermitian conjugate of the field operator is

$$\boldsymbol{\psi}^\dagger(x) \sum_{s=\pm} \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p \sqrt{\frac{m}{E_p}} \left[u_s^\dagger(\mathbf{p}) a_s^\dagger(\mathbf{p}) e^{\mathrm{i}p\cdot x} + v_s^\dagger(\mathbf{p}) b_s(\mathbf{p}) e^{-\mathrm{i}p\cdot x} \right].$$

Let us again calculate the particle number:

$$N(t) = \int \boldsymbol{\psi}^\dagger(\mathbf{r},t)\boldsymbol{\psi}(\mathbf{r},t) d^3r.$$

Using the normalization conditions for the basis functions, we see that we get a number of positive-energy particles

$$N^+ = \sum_s \int d^3p\, a_s^\dagger(\mathbf{p}) a_s(\mathbf{p}).$$

For the negative energies, we have

$$N^- = \sum_s \int d^3p\, b_s(\mathbf{p}) b_s^\dagger(\mathbf{p}) = \sum_s \int d^3p \left[1 - b_s^\dagger(\mathbf{p}) b_s(\mathbf{p})\right].$$

It is a convention to neglect the inifinite constant $\sum_s \int d^3p$ from this number, so that we have a negative density left. This is common practice in quantum field theory: for quantities like energy and particle number, we use the *normal order* procedure: we move all creation operators to the left, and the annihilation operators to the right. The notation for this procedure is two colons: $:\ldots:$. For example

$$\sum_s : \int d^3p\, b_s(\mathbf{p}) b_s^\dagger(\mathbf{p}) := -\sum_s \int d^3p\, b_s^\dagger(\mathbf{p}) b_s(\mathbf{p}).$$

As you can see, the re-ordering has consequences for the sign in front of the expression. Finally, cross-terms such as

$$\sum_{ss'} \frac{1}{(2\pi\hbar)^3} \int d^3r \int d^3p \sqrt{\frac{m}{E_p}} a^\dagger(\mathbf{p}) \int d^3q \sqrt{\frac{m}{E_q}} u_s^\dagger(\mathbf{p}) v_{s'}(\mathbf{q}) b^\dagger(\mathbf{q}) e^{\mathrm{i}p\cdot x} e^{\mathrm{i}q\cdot x}$$

vanish as a result of the orthogonality condition (18.1).

We see that the negative-energy particles contribute with a minus-sign to the particle number. This can be related back to the negative density problem which always seems to pop up in relativistic quantum mechanics. A solution may be to interpret this quantity not so much as the number density, but rather as the charge density.

We may also formulate the Hamiltonian of the system, which proceeds in a way similar to the particle density:

$$H = \int d^3r \boldsymbol{\psi}^\dagger(x) \left(-\mathrm{i}\boldsymbol{\alpha}\cdot\nabla + \beta m\right) \boldsymbol{\psi}(x).$$

We can rewrite this after insertion of γ^4, and using $\bar{\boldsymbol{\psi}} = \boldsymbol{\psi}^\dagger\gamma^4$, as

$$H = \int d^3r \bar{\boldsymbol{\psi}}(x) \left(-\mathrm{i}\boldsymbol{\gamma}\cdot\nabla - m\right) \boldsymbol{\psi}(x).$$

Inserting the expansion of the field operators in terms of the $a(\mathbf{p})$ and $b(\mathbf{p})$ operators and their Hermitian conjugates gives

$$H = \int d^3p E_p \sum_s \left[a_s^\dagger(\mathbf{p}) a_s(\mathbf{p}) - b_s(\mathbf{p}) b_s^\dagger(\mathbf{p})\right].$$

The right hand side again contains an infinite energy as can be seen by writing $bb^\dagger = 1 - b^\dagger b$. Removing this infinite offset then gives

$$H = \int d^3p E_p \sum_s \left[a_s^\dagger(\mathbf{p}) a_s(\mathbf{p}) + b_s^\dagger(\mathbf{p}) b_s(\mathbf{p})\right].$$

which leads to the surprising conclusion that the negative energy particles contribute a positive amount to the total energy!

At this stage it is useful to enter the interpretation of the Dirac theory. We have an equation which describes the behaviour of electrons extremely successfully (remember our analysis of the hydrogen atom). However, it allows for negative energies which by themselves may not be a problem, unless we allow the electron to interact with an electromagnetic field. If this is done, then the electron may lower its energy by emitting a photon. If it continues to do so, it can lower its energy ad infinitum. An electron moving at constant momentum through space will not emit any radiation according to electrodynamics, but if it is orbiting in a hydrogen atom in a nucleus, the notorious process of losing energy will continue to happen. This is not in agreement with observations – therefore these transitions should somehow be ‘forbidden’: Dirac realized this by invoking Pauli’s principle: he stated that all the negative energy states were already filled with negative-energy electrons. Moving a positive energy electron to a negative energy state would then be impossible. Outrageous the suggestion may seem, it actually works very well. In fact, in condensed matter we use this picture all the time: the state below the Fermi energies are filled with electrons, and this prevents electrons with energies above the Fermi energies to ‘fall down’ and occupy the states below.

We also know from condensed matter physics that electrons can be excited from below the Fermi energy to states above, when they absorb a photon. Suppose that the negative electrons in a ‘vacuum’ could do the same. Then we would observe the excited electron as a particle with negative charge and positive energy. However that is not all: we also see a *hole* in the Fermi sea, which, due to the absence of an electron with negative charge there, acts as a positively charged particle. In condensed matter physics, we call such a particle indeed a hole. Furthermore, as we are ‘missing’ an electron in the sea of negative energy states, the hole carries a *positive* energy.

In observations, there is no distinction between a hole and a positively charged particle. As the suggestion that the vacuum consists of an infinite amount of negative-energy electrons, which should somehow be compensated for by a positive background charge, is not the most elegant option, we usually choose for the interpretation that the a-operators create and annihilate negatively charged particles, whereas the b-s create and annihilate positively charged particles. These particles are called *positrons*. The energy to create an electron-positron pair (or, in the Dirac sea language, to excite a negative energy electron to a positive energy state) requires an energy of at least 2 electron masses (one electron mass corresponds to 0.51 MeV).

18.4 A PHYSICAL REALIZATION OF A DIRAC FIELD THEORY: GRAPHENE

In 2010, A. Geim and K. Novoselov were awarded the Nobel prize for the isolation of and research on graphene: a single layer of carbon atoms, arranged in an hexagonal lattice – see figure 18.1. The nearest-neighbour distance between two carbon atoms in graphene is about $a = 1.42$ Å. The hexagonal lattice is spanned by two unit vectors:

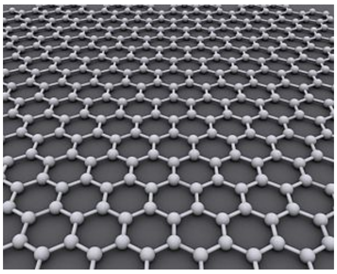


FIGURE 18.1: The hexagonal graphene lattice.

$$\mathbf{a}_1 = \frac{a}{2}(3, \sqrt{3}); \quad \mathbf{a}_2 = \frac{a}{2}(3, -\sqrt{3}).$$

Each unit cell contains *two* atoms: we say that the hexagonal lattice is a lattice *with a basis*, where the basis is the set of (in this case) two atoms in the unit cell. The basis vectors of the reciprocal lattice are

$$\mathbf{b}_1 = \frac{2\pi}{3a}(1, \sqrt{3}); \quad \mathbf{b}_2 = \frac{2\pi}{3a}(1, -\sqrt{3}).$$

In this lattice, we can define the *Brillouin zone* as usual: it is the set of points which are closer to the origin than to any reciprocal lattice points. We see that the Brillouin zone is again a hexagon. Two of the corner points of the hexagon lie in the first Brillouin zone; we call those points **K** and **K**′. They are given as

$$\mathbf{K} = \frac{2\pi}{3a}\left(1, \frac{1}{3}\sqrt{3}\right); \qquad \mathbf{K}' = \frac{2\pi}{3a}\left(1, -\frac{1}{3}\sqrt{3}\right)$$

The lattice and the Brillouin zone are graphically illustrated in figure 18.2.

As you can see in the figure, we can divide up the lattice into two different kinds of points: *A* points and *B* points. Each *A* point has three *B* points as its neighbours and vice versa. The vectors connecting an *A* point to its three neighbours are

$$\mathbf{d}_1 = \frac{a}{2}(1, \sqrt{3}), \qquad \mathbf{d}_2 = \frac{a}{2}(1, -\sqrt{3}), \qquad \mathbf{d}_3 = -a(1, 0).$$

The Hamiltonian for electrons moving on this lattice contains as usual a potential and a kinetic energy term. We use wave functions localized near the nuclei as basis states.[1] The operators $a^\dagger_{\sigma,n_A}$ and a_{σ,n_A} are the creation and annihilation for an electron with spin $\sigma = \pm 1/2$ occupying the orbital of the *A* atom at site n_A. Similarly, we use $b^\dagger_{\sigma,m_B}$ and b_{σ,m_B} for the electrons in the orbitals of the *B*-atoms. The potential energy of an electron when it is in such a localized state is state is ϵ:

$$H_{\text{pot}} = \epsilon \sum_{\sigma=\pm 1/2} \left(\sum_{n_A} a^\dagger_{\sigma,n_A} a_{\sigma,n_A} + \sum_{m_B} b^\dagger_{\sigma,m_B} b_{\sigma,m_B} \right).$$

---

[1] These orbitals are the p_z orbitals which stick out from two sides of the graphene sheet.

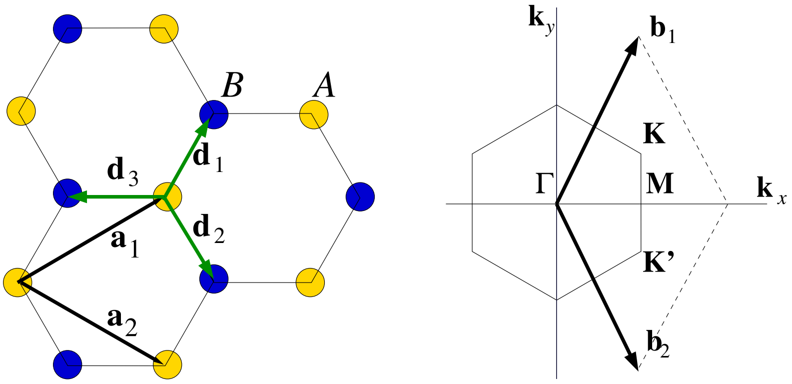


FIGURE 18.2: Hexagonal lattice with its unit cell. The vectors $\mathbf{d}_j$, $j = 1, 2, 3$ connect the A-sites of the lattice to its three nearest B-neighbours. On the right the reciprocal lattice is shown with the hexagonal Brillouin zone with some of the points with special symmetry indicated.

Note that this can be written as

$$H_{\text{pot}} = \epsilon \sum_{\sigma=\pm 1/2} \left(\sum_{n_A} n_{\sigma,n_A} + \sum_{m_B} n_{\sigma,m_B} \right).$$

where n_{σ,n_A} and n_{σ,m_B} are the number operators.

The electrons may hop to a localized orbital of another atom. This hopping process is most prominent for nearest neighbour orbitals. In fact, we cut off possible hops to farther than nearest neighbours and apply a hopping parameter t to arrive at a neighbouring point, and obtain the following kinetic energy:

$$H_{\text{kin}} = -t \sum_{\sigma=\pm 1/2} \sum_{\langle n_A, m_B \rangle} \left(a^\dagger_{\sigma,n_A} b_{\sigma,m_B} + b^\dagger_{\sigma,m_B} a_{\sigma,n_A} \right).$$

The angular brackets $\langle n_A, m_B \rangle$ indicate that the sum is only over nearest neighbours. The particles do not change their spin in the hopping process. Note that this Hamiltonian is Hermitian.

The velocity of electrons near the Fermi energy is 10^6 m/s in graphene. This means that the electrons can safely be described as nonrelativistic (their velocity is much smaller than the speed of light). Therefore, and as the two spin directions do not occur explicitly in the Hamiltonian, we can neglect the electron spin: spin-up electrons behave the same as spin-down electrons, and there is no coupling between the two.

In order to find the energies, we make use of Bloch's theorem, which tells us that the eigenstates of a periodic Hamiltonian can be written as

$$\langle \mathbf{r} | \psi_{\mathbf{k}} \rangle = u(\mathbf{r}) e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}},$$

where u is a function which has the periodicity of the lattice, and $\mathbf{k}$ can be chosen in the first Brillouin zone. This theorem is formulated for continuum space. Here we are however dealing with a discrete space. In that case, Bloch's theorem reads:

$$\langle n_c | \psi \rangle = u(n_c) e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{n_c}},$$

where n_c denotes a unit cell of the lattice, located at $\mathbf{r}_{n_c}$. As we have two orbitals in each unit cell, one on an A atom and another one on a B atom, all wave functions can be written as two-vectors:

$$\langle n_c | \psi \rangle = \begin{pmatrix} \phi_A(n_c) \\ \phi_B(n_c) \end{pmatrix}. \tag{18.2}$$

The periodic function u also has this form:

$$u = \begin{pmatrix} u_A \\ u_B \end{pmatrix}.$$

The phase factor $\exp(i\mathbf{k}\cdot\mathbf{r})$, however, is obviously just a scalar. Note that we have left out the argument n_c which is justified since u is a *periodic* function: it has the same value in each cell.

The Schrödinger equation for such a wave function with Bloch wave-vector $\mathbf{k}$, yields an equation for u_A and u_B:

$$E_{\mathbf{k}} u_A = \epsilon u_A - t u_B \left(e^{-i\mathbf{k}\cdot(\mathbf{a}_1+\mathbf{a}_2)} + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2}\right).$$

Similarly,

$$E_{\mathbf{k}} u_B = \epsilon u_B - t u_A \left(e^{i\mathbf{k}\cdot(\mathbf{a}_1+\mathbf{a}_2)} + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2}\right).$$

These are two homogeneous, linear equations with two unknowns. A nontrivial solution to these only exists when the determinant of this system of equations vanishes:

$$\begin{vmatrix} E-\epsilon & t\left(e^{-i\mathbf{k}\cdot(\mathbf{a}_1+\mathbf{a}_2)} + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2}\right) \\ t\left(e^{i\mathbf{k}\cdot(\mathbf{a}_1+\mathbf{a}_2)} + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2}\right) & E-\epsilon \end{vmatrix} = 0$$

leads to

$$E-\epsilon = \pm t\sqrt{3+2\cos(\mathbf{k}\cdot\mathbf{a}_1)+2\cos(\mathbf{k}\cdot\mathbf{a}_2)+2\cos\left[\mathbf{k}\cdot(\mathbf{a}_2-\mathbf{a}_1)\right]}.$$

The basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ have been given above, so we easily find

$$E-\epsilon = \pm t\sqrt{3+2\cos\left(\frac{3a}{2}k_x+\frac{a\sqrt{3}}{2}k_y\right)+2\cos\left(\frac{3a}{2}k_x-\frac{a\sqrt{3}}{2}k_y\right)+2\cos\left(\sqrt{3}k_y a\right)},$$

which can be rewritten as

$$E-\epsilon = \pm t\sqrt{3+2\cos\left(\sqrt{3}k_y a\right)+4\cos\left(\frac{3a}{2}k_x\right)\cos\left(\frac{a\sqrt{3}}{2}k_y\right)}. \tag{18.3}$$

A picture of this dispersion relation is given in figure 18.3. We see that the band structure has positive and negative values, as is immediately clear from the ± sign in Eq. (18.3). Furthermore we observe that the positive and negative values 'touch' each other at special points in the Brillouim zone with a cone-like shape. These cones are described by the mathematical expression

$$E \propto \pm|\mathbf{q}|$$

where $\mathbf{q}$ is the point in reciprocal space relative to the cone position. Comparing this with the relativistic expression for the energy

$$E = \pm\sqrt{p^2+m^2},$$

and identifying p with q, we see that these cones describe particles with zero mass. In fact, the relation $\omega = c|\mathbf{k}|$ is well known for photons, the carriers of light, which have zero mass and travel at the speed of light. The $\mathbf{k}$ points where the cones touch, are $\mathbf{K}$ and $\mathbf{K}'$ which are shown in figure 18.2. Writing $\mathbf{k} = \mathbf{K}+\mathbf{q}$ or $\mathbf{k} = \mathbf{K}'+\mathbf{q}$, it is easy to verify that for small vectors $\mathbf{q}$, the dispersion relation reads:

$$E(\mathbf{q}) = v_{\mathrm{F}}|\mathbf{q}| + \mathcal{O}(q^2).$$

The points $\mathbf{K}$ and $\mathbf{K}'$ are called 'Dirac points' for reasons that will become clear later on.

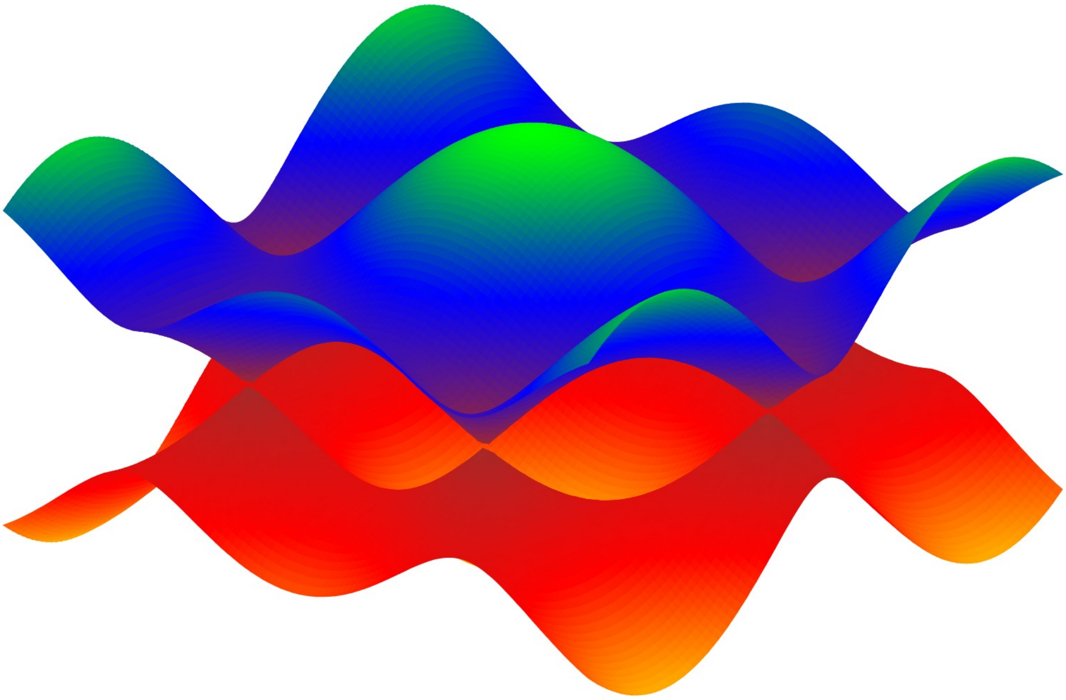


FIGURE 18.3: Band structure of graphene. Shown are the energies as a function of the two-dimensional vectors **k**.

Now we need some additional piece of information which comes from chemistry. Carbon has a nuclear charge $Z = 6$ and therefore it has 6 electrons. Two of these are in the low-lying 1s orbital, which has spherical symmtry and which strongly localized near the nucleus. This leaves four electrons, three of which are used to bind to the three neighbouring atoms. This leaves *one* electron in each atomic p_z orbital, and as we have two atoms per unit cell in the hexagonal lattice, we have two electrons per cell. The band structure which we have calculated is only for these p_z electrons, and for each k point we should fill the energy values with two electrons. At zero temperature, we fill the lowest possible energy values. As we can put two electrons with opposite spin in the negative energy states, we conclude that these states are filled at zero temperature, whereas the positive energy states are empty. The picture we now have obtained is that of a filled Dirac sea of negative energy states, and for energies close to zero (i.e. close to the Fermi energy) we have a dispersion relation

$$E = v_{\mathrm{F}}|\mathbf{q}| = c|\mathbf{p}|,$$

where in the second expression we have emphasised the similarity to massless particles in relativity.

We want to show that the electrons in graphene satisfy a Dirac equation in two dimensions. Let us therefore first analyze such a Dirac theory. In section 17.1, Eq. (17.1), we have seen that we need ‘objects’ α_i, $i = x, y$ and β, satisfying

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij},$$

$$\{\alpha_i, \beta\} = 0,$$

and

$$\alpha_i^2 = \beta^2 = 1.$$

The fact that we now only need *two* α_i's makes it much easier to find a solution than in the four-dimensional case we studied before: we choose $\alpha_i = \sigma_i$, and $\beta = \sigma_z$! The Dirac Hamiltonian should then read

$$H = -\mathrm{i}\boldsymbol{\sigma}\cdot\nabla - \beta m,$$

with $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$ and $\beta = \sigma_z$, and where the dot product is of course two dimensional.

In order to show that in graphene the electrons are indeed described by such a Hamiltonian, recall that the wavefunction could be written in the form of a two-spinor, see Eq. (18.2). Let us return to the Hamiltonian in the Fock space which describes particles hopping to nearest neighbour positions:

$$H = -t \sum_{\langle n_A m_B\rangle} \left(a^\dagger_{n_A} b_{m_B} + b^\dagger_{m_B} a_{n_A}\right),$$

where $\langle n_A m_B\rangle$ denotes nearest neighbour pairs on the lattice as usual. Note that the points n_A and m_B both belong to some unit cell of the hexagonal lattice. These cells are indicated by a position vector $\mathbf{R}_n$, which may be the position of the A-atom in that cell – other conventions, such as the point midway of a 'horizontal' point pair in figure 18.2 are possible. We know that the dispersion relation has Dirac cones located near the points $\mathbf{K}$ and $\mathbf{K}'$ in that figure. We want to describe the electrons with wave vectors close to these points. To this end, we consider the Fourier transforms of the operators a_{n_A} etcetera:

$$a_{\mathbf{k}} = \frac{1}{\sqrt{N_c}} \sum_n e^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}_n} a_{n_A},$$

where the sum is over the N_c cells, and n_A indicates the A-point within the cell located at $\mathbf{R}_n$, and $\mathbf{k}$ is a vector inside the Brillouin zone. The inverse transform is

$$a_{n_A} = \frac{1}{\sqrt{N_c}} \sum_{\mathbf{k}\epsilon \mathrm{BZ}} a(\mathbf{k}) e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{R}_n}.$$

Furthermore

$$a^\dagger_{\mathbf{k}} = \frac{1}{\sqrt{N_c}} \sum_n e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{R}_n} a^\dagger_{n_A}$$

and we have similar transforms for the b_{n_B}.

Now we make an important step: as we want to describe the behaviour near $\mathbf{K}$ and $\mathbf{K}'$, we write

$$a_{n_A} = \frac{1}{\sqrt{N_c}} \sum_{\mathbf{q}\text{ small}} \left[a(\mathbf{K}+\mathbf{q}) e^{-\mathrm{i}(\mathbf{K}+\mathbf{q})\cdot\mathbf{R}_n} + a(\mathbf{K}'+\mathbf{q}) e^{-\mathrm{i}(\mathbf{K}'+\mathbf{q})\cdot\mathbf{R}_n} \right],$$

i.e., we focus on the regions near the two Dirac points. We now define

$$a_{1,n_A} = \frac{1}{\sqrt{N_c}} \sum_{\mathbf{q}} a(\mathbf{K}+\mathbf{q}) e^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_n}$$

and

$$a_{2,n_A} = \frac{1}{\sqrt{N_c}} \sum_{\mathbf{q}} a(\mathbf{K}'+\mathbf{q}) e^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_n},$$

so that

$$a_{n_A} = e^{-\mathrm{i}\mathbf{K}\cdot\mathbf{R}_n} a_{1,n_A} + e^{-\mathrm{i}\mathbf{K}'\cdot\mathbf{R}_n} a_{2,n_A}.$$

A similar approach for the B particles gives

$$b_{n_B} = e^{-\mathrm{i}\mathbf{K}\cdot\mathbf{R}_n} b_{1,n_B} + e^{-\mathrm{i}\mathbf{K}'\cdot\mathbf{R}_n} b_{2,n_B}.$$

Note that the Fourier-expansions of the a and b operators both contain an $\mathbf{R}_n$: the a-operators automatically pertain to the A-point of that cell, and the b-operators to the B-point. As the a_{i,n_A} and the b_{i,n_B} are expanded in terms of Fourier components with small $\mathbf{q}$, their spatial variation is small.

The Hamiltonian now reads

$$H = -t \sum_{\langle n_A m_B\rangle} \left(e^{\mathrm{i}\mathbf{K}\cdot\mathbf{R}_n} a^\dagger_{1,n_A} + e^{\mathrm{i}\mathbf{K}'\cdot\mathbf{R}_n} a^\dagger_{2,n_A}\right)\left(e^{-\mathrm{i}\mathbf{K}\cdot\mathbf{R}_m} b_{1,m_B} + e^{-\mathrm{i}\mathbf{K}'\cdot\mathbf{R}_m} b_{2,m_B}\right) + \text{h.c.},$$

where ‘h.c.’ denotes the Hermitian conjugate of the term written down.

Let us analyse the term

$$H_{11} = -t \sum_{\langle n_A m_B\rangle} e^{\mathrm{i}\mathbf{K}\cdot\mathbf{R}_n} a^\dagger_{1,n_A} e^{-\mathrm{i}\mathbf{K}\cdot\mathbf{R}_m} b_{1,m_B} + \text{h.c.} = -t \sum_{\langle n_A m_B\rangle} e^{\mathrm{i}\mathbf{K}\cdot(\mathbf{R}_n - \mathbf{R}_m)} a^\dagger_{1,n_A} b_{1,m_B} + \text{h.c.}$$

Careful inspection of Fig. 18.2 leads to the conclusion that for an A-point in cell $\mathbf{R}_n$, the three neighbouring B-points are in the cells with $\mathbf{R}_m = \mathbf{R}_n - \mathbf{a}_1$, $\mathbf{R}_m = \mathbf{R}_n - \mathbf{a}_2$ and $\mathbf{R}_m = \mathbf{R}_n - \mathbf{a}_1 - \mathbf{a}_2$. This implies that if we act on a wave function $|\psi\rangle$ with components

$$\langle n_c|\psi\rangle = \begin{pmatrix} \phi_A(n_c) \\ \phi_B(n_c) \end{pmatrix},$$

where n_c denotes, as usual, a cell, we see that, after acting with the Hamiltonian on this wavefunction, the new upper component, corresponding to $\phi_A(n)$, becomes

$$-t\left(e^{\mathrm{i}\mathbf{K}\cdot\mathbf{a}_1}\phi_B(\mathbf{R}_n - \mathbf{a}_1) + e^{\mathrm{i}\mathbf{K}\cdot\mathbf{a}_2}\phi_B(\mathbf{R}_n - \mathbf{a}_2) + e^{\mathrm{i}\mathbf{K}\cdot(\mathbf{a}_1+\mathbf{a}_2)}\phi_B(\mathbf{R}_n - \mathbf{a}_1 - \mathbf{a}_2)\right).$$

We work this and Taylor-expand the ϕ_B around $\mathbf{R}_n$ (this is possible because all our ϕ’s are slowly varying functions of the position) to obtain as a new upper component

$$\frac{3ta}{2}\left(-\partial_x + \mathrm{i}\partial_y\right)\phi_B(\mathbf{R}_n). \tag{18.4}$$

Now we consider the Hermitial conjugate of this term in H_{11}:

$$-t \sum_{\langle n_B m_A\rangle} e^{\mathrm{i}\mathbf{K}\cdot(\mathbf{R}_n - \mathbf{R}_m)} b^\dagger_{1,n_B} a_{1,m_A}.$$

This term acts on the wavefunction on the neighbouring sites m_A of n_B and it uses the values on those sites to fill the new value of the wavefunction on the site n_B. These A-neighbours are located in cells at a relative distance $-\mathbf{a}_1$, $-\mathbf{a}_2$ and $\mathbf{0}$ with respect to $\mathbf{R}_n$. Writing this out we obtain a value of the B-component of $H|\psi\rangle$ at n_B:

$$\frac{3ta}{2}\left(\partial_x + \mathrm{i}\partial_y\right)\phi_A(\mathbf{R}_n). \tag{18.5}$$

The two equations (18.4) and (18.5) can be written as

$$\langle n_c|H|\phi\rangle = \frac{3ta}{2}\begin{pmatrix} 0 & \left(-\partial_x + \mathrm{i}\partial_y\right) \\ \partial_x + \mathrm{i}\partial_y & 0 \end{pmatrix}\begin{pmatrix} \phi_A(\mathbf{R}_n) \\ \phi_B(\mathbf{R}_n) \end{pmatrix}.$$

This can also be written as

$$\langle n_c|H|\phi\rangle = -\frac{3ta}{2}\mathrm{i}\left(-\partial_x\sigma_y + \partial_y\sigma_x\right)\langle n_c|\psi\rangle,$$

where σ_x and σ_y are Pauli matrices. These matrices have a fixed form, but the choice of the x and y axis are of course arbitrary. Swapping these two ($x \to -y$, $y \to x$) we can write the Hamiltonian in the form

$$H = -\frac{3\mathrm{i}ta}{2}\boldsymbol{\sigma}\cdot\nabla,$$

which is recognised as the Dirac Hamiltonian in two dimensions for massless particles.

The terms coupling the a_2 and b_2 operators gives the same Hamiltonian, whereas the coupling terms involving a_1 and b_2 contain quickly oscillating terms which vanish in the sum over the lattice. We see that we have two independent systems of Dirac Hamiltonians for massless particles. The two systems correspond to the labels 1 and 2, which can ultimately be related to the two Dirac points $\mathbf{K}$ and $\mathbf{K}'$. The two components of the wavefunctions correspond to A and B and are therefore related to the fact that graphene has a hexagonal lattice which has *two* points in the unit cell.

18.5 PROBLEMS

1. (20 pts) In this problem, we study electrons in graphene subject to a magnetic field perpendicular to the graphene sheet. In the lectures, the Hamiltonian for electrons (without a magnetic field) was shown to be

$$-\mathrm{i} v_\mathrm{F} \boldsymbol{\sigma} \cdot \nabla \psi(\mathbf{r}) = E\psi(\mathbf{r}).$$

Here, $\boldsymbol{\sigma}$ is a vector whose components are the 2×2 σ_x and σ_y Pauli matrices, $\psi(\mathbf{r})$ is a two-spinor and $\mathbf{r}$ is a two-dimensional position vector on the sheet, which is taken to be in the xy plane.

The effect of the magnetic field is accounted for by replacing

$$-\mathrm{i}\nabla \to -\mathrm{i}\nabla + e\mathbf{A}/c.$$

We use the 'Landau gauge':

$$\mathbf{A} = B(-y, 0, 0).$$

(a) Show that this leads to a magnetic field of magnitude B along the positive z-axis.

(b) Write out the two components of the Dirac equation for electrons in graphene in the presence of a magnetic field. Write $\psi(\mathbf{r}) = e^{\mathrm{i}kx}\phi(y)$, ($\phi(y)$ is then a two-spinor too) and show that the Dirac equation can be cast into the form

$$\omega_c \begin{pmatrix} 0 & \mathcal{O} \\ \mathcal{O}^\dagger & 0 \end{pmatrix} \phi(\xi) = E\phi(\xi),$$

where

$$\xi = \frac{y}{l_\mathrm{B}} - l_\mathrm{B} k, \qquad l_\mathrm{B} = \sqrt{\frac{c}{eB}}.$$

The parameter l_B is called the 'magnetic length'. Here

$$\mathcal{O} = \frac{1}{\sqrt{2}}\left(\partial_\xi + \xi\right).$$

Give ω_c and the form of $\mathcal{O}^\dagger$.

This Dirac equation can also be written as

$$\left(\mathcal{O}\sigma^+ + \mathcal{O}^\dagger \sigma^-\right)\phi = \frac{2E}{\omega_c}\phi.$$

(c) Show that $[\mathcal{O}, \mathcal{O}^\dagger] = 1$, i.e., the operators $\mathcal{O}^\dagger$ and $\mathcal{O}$ are boson creation and annihilation operators, respectively. Show that, given a ground state wave function

$$\mathcal{O}\chi_0(\xi) = 0,$$

all solutions can be found as

$$\phi_{N,\pm} = \begin{pmatrix} \chi_{N-1}(\xi) \\ \pm\chi_N(\xi) \end{pmatrix},$$

with $\chi_{-1}(\xi) = 0$ energies

$$E_\pm = \pm\omega_c\sqrt{N}.$$

Oscillations with this particular structure have been observed for graphene.

(d) The existence of a zero-energy state in graphene, observed in experiments as a resistance peak (see figure), is called the 'anomalous integer quantum Hall effect'. Here, n on the horizontal axis denotes the electron density, i.e. $n = 0$ means charge-neutral graphene. What can you say about the Fermi energy in that case in view of the band structure of graphene?

Explain why the observation in the graph is called 'anomalous' in view of the density of states of graphene at zero magnetic field.