## Contents

1. Tests, probabilities and quantum theory 2  
2. The Old Quantum Theory 3  
3. Bell’s inequalities 6  
4. Two-level systems 9  
5. The Schrödinger equation and wave mechanics 12  
6. Exactly solvable models 18  
7. Matrix Mechanics 35  
8. Mathematical Foundations of Quantum Mechanics 38  
9. Angular momentum and spin 49  
10. Identical particles 59  
References 60

**Note.** I created these notes for the course *ACM30210 - Foundations of Quantum Mechanics* I taught at University College Dublin in 2018. The exposition follows very closely the classical books listed at the end of the document. You can help me continue to improve these notes by emailing me with any comments or corrections you have.
1. Tests, probabilities and quantum theory

Quantum theory is a set of rules allowing the computation of probabilities for the outcomes of tests which follow specified preparations.

A preparation is an experimental procedure that is completely specified, like a recipe in a good cookbook. Preparation rules should be unambiguous, but they may involve stochastic processes, such as thermal fluctuations, provided that the statistical properties of the stochastic process are known. A test starts like a preparation, but it also includes a final step in which information, previously unknown, is supplied to an observer.

In order to develop a theory, it is helpful to establish some basic notions and terminology. A quantum system is an equivalence class of preparations. For example, there are many equivalent macroscopic procedures for producing what we call a photon, or a free hydrogen atom, etc. While quantum systems are somewhat elusive, quantum states can be given a clear operational definition, based on the notion of test. Consider a given preparation and a set of tests. If these tests are performed many times, after identical preparations, we find that the statistical distribution of outcomes of each test tends to a limit: each outcome has a definite probability. We can then define a state as follows: a quantum state is characterised by the probabilities of the various outcomes of every conceivable test.

This definition is highly redundant. We shall see that these probabilities are not independent. One can specify – in many different ways – a restricted set of tests such that, if the probabilities of the outcomes of these tests are known, it is possible to predict the probabilities of the outcomes of every other test. (A geometric analogy is the definition of a vector by its projections on every axis. These projections are not independent: it is sufficient to specify a finite number of them, on a complete, linearly independent set of axes.)

As a simple example of definition of a state, suppose that a photon is said to have right-handed polarisation. Operationally, this means that if we subject that photon to a specific test (namely, a quarter wave plate followed by a suitably oriented calcite crystal) we can predict with certainty that the photon will exit in a particular channel. For any other test, consisting of arbitrarily arranged calcite crystals and miscellaneous optically active media, we can then predict probabilities for the various exit channels. (These probabilities are computed in the same way as the classical beam intensities.) Note that the word ‘state’ does not refer to the photon by itself, but to an entire experimental setup involving macroscopic instruments.

The essence of quantum theory is to provide a mathematical representation of states, together with rules for computing the probabilities of the various outcomes of any test. This set of rules can be consistently organised assuming that
there is associated with each physical system a set of quantities, constituting a *non-commutative algebra* in the technical mathematical sense, the elements of which are the physical quantities themselves.

2. The Old Quantum Theory

There are two main groups of experimental phenomena which are inconsistent with classical physics, namely:

(i) when the internal energy of an atom changes, owing to emission or absorption of light, it does not do so evenly or continuously but in ‘quantised’ steps;

(ii) a beam of electrons exhibits interference phenomena similar to those of a light wave.

(i) *Quantum states.* A great wealth of experimental data, chiefly derived from spectroscopy, shows that an atom cannot exist in states of continuously varying energy but only in different discrete states of energy, referred to as ‘discrete energy levels’. These levels and the spacings between them are different for the various chemical elements, but are identical for all atom of the same element. The emitted energy (in the form of light) is confined to spectral lines of sharply defined frequencies, and, moreover, the frequencies are not related to one other by integral factors, as overtone, but instead show an interesting additive relation, expressed in the *Ritz-Rydberg combination principle.* The change from level $E_2$ to another $E_1$ is associate with the emission if $E_2 > E_1$ (or absorption if $E_2 < E_1$) of light, whose frequency $\nu$ is determined by the relation

$$E_2 - E_1 = h\nu,$$

where $h$ is a universal constant known as Planck’s constant. The above relation is the *Bohr frequency rule.* It is important to notice that the experimentally measurable quantities are the spacings (frequencies) and not the single energy levels! Note that Bohr’s assumption is not only in agreement with the existence of sharp spectral lines, but contains in addition the combination principle. If we order the energy levels $E_0 < E_1 < E_2 \cdots$, then in accordance to (2.1) each frequency is the difference of two terms

$$\nu(i \to j) = \frac{1}{h}(E_i - E_j) \quad (i > j).$$

Consequently, generally speaking there will occur in addition to the frequency $\nu(i \to j), \nu(j \to k)$ the frequency

$$\nu(i \to k) = \nu(i \to j) + \nu(j \to k)$$

obtained from them by addition (combination).
(ii) **Interference of waves.** It is well known that light waves when reflected or refracted from a regularly spaced grating interfere and form what is called a ‘diffraction’ pattern visible on a screen. This phenomenon was used, for instance, to prove the wave nature of X-rays. For this purpose the regular spacing is provided by the atoms of a crystal, as an ordinary grating would be too coarse. From the diffraction pattern one can determine the wavelength of the light. Davisson and Germer (1927) carried out a similar experiment with a beam of electron, all having the same velocity, or momentum \( p \), passing through a crystal, etc. The diffraction pattern so found is very similar to that produced by X-rays. This experiment shows that waves are associate with a beam of electrons. The wavelength of the electrons is then found to be inversely proportional to the momentum \( \lambda \propto \frac{1}{p} \), the slower the electrons the longer is the wavelength. The proportionality constant can be measured and it turns out to be equal to Planck’s constant

\[
\lambda = \frac{h}{p}.
\]

This relation was suggested by de Broglie in his doctoral thesis (1924), and later confirmed in the experiment by Davisson and Germer. This relation was also fundamental in the formulation of the most famous Heisenberg’s uncertainty relation

\[
\Delta x \Delta p \geq \frac{\hbar}{2}.
\]

**Example 2.1.** Hydrogen atoms in a discharge lamp emit a series of lines in the visible part of the spectrum. This series is called the Balmer Series after the Swiss teacher Johann Balmer. Here are the lines observed by Balmer (in convenient units)

\[
\begin{align*}
&5, 3, 21, 2 \\
&\frac{36}{9}, \frac{16}{9}, \frac{100}{9}, \frac{9}{9}
\end{align*}
\]

In 1885, Balmer found by trial and error a formula to describe the wavelengths of these lines. Balmer suggested that his formula may be more general and could describe spectra from other elements. Then in 1889, Johannes Robert Rydberg found several series of spectra that would fit a more general relationship, similar to Balmer’s empirical formula. A sample of lines is

\[
\begin{align*}
&11, 9, 5, 7, 6, 1, 5, 3, 21, 2, 3, 8, 15, 24, 35 \\
&\frac{900}{9}, \frac{400}{9}, \frac{144}{9}, \frac{144}{9}, \frac{225}{9}, \frac{12}{9}, \frac{36}{9}, \frac{16}{9}, \frac{100}{9}, \frac{9}{9}, \frac{4}{9}, \frac{9}{9}, \frac{16}{9}, \frac{25}{9}, \frac{36}{9} \\
&\text{infrared, visible, ultraviolet}
\end{align*}
\]

The actual spectral lines are given by the numbers above multiplied by the Rydberg constant

\[
R = \frac{2\pi^2 \mu e^2}{\hbar^2 c}
\]

where \( \mu \) is the reduced mass of the electron and the nucleus. The numerical value of the constant \( R \) agrees with values obtained from spectroscopic data. The values
for hydrogen, ionized helium, and infinite nuclear mass are

\[ R_H = 109677.759 \pm 0.05 \text{ cm}^{-1} \]
\[ R_{He} = 109722.403 \pm 0.05 \text{ cm}^{-1} \]
\[ R_{\infty} = 109737.42 \pm 0.06 \text{ cm}^{-1} . \]

**Exercise 2.1.** Try to find Balmer’s empirical formula. If you succeed, try to find Rydberg’s generalisation to explain the above sequence. Hint: rearrange the sequence as a table

\[
\begin{array}{cccc}
3 & 4 \\
8 & 9 \\
15 & 16 & 17 & 18 \\
24 & 21 & 16 & 9 \\
35 & 2 & 1 & 5 & 11 & 900 \\
& \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}
\]

**Table 1.** Relations between experimental interpretation and theoretical inferences.

<table>
<thead>
<tr>
<th>Diffraction (Young 1803, Laue 1912)</th>
<th>Electromagnetic waves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackbody radiation (Planck 1900)</td>
<td>Electromagnetic quanta</td>
</tr>
<tr>
<td>Photoelectric effect (Einstein 1904)</td>
<td></td>
</tr>
<tr>
<td>Combination principle (Ritz-Rydberg 1908)</td>
<td>Discrete values</td>
</tr>
<tr>
<td>Franck-Hertz experiment (1913)</td>
<td>for physical quantities</td>
</tr>
<tr>
<td>Stern and Gerlach (1922)</td>
<td>Matter waves</td>
</tr>
<tr>
<td>Davisson and Germer (1927)</td>
<td></td>
</tr>
</tbody>
</table>

**2.1. Quantisation rules.** In 1915 W. Wilson and A. Sommerfeld discovered independently a simple method of quantisation which was soon applied in the discussion of several atomic phenomena with good success.

The first step of the method consists in solving the classical equations of motion defined by the classical Hamiltonian \( H = H(p, q) \). The assumption is then introduced that only those classical orbits are allowed as stationary states for which the following conditions are satisfied

\[
(2.6) \quad \oint p_k dq_k = n_k h, \quad n_k \text{ an integer},
\]
for all \( k = 1, \ldots, N \) (\( N \) being the number of degrees of freedom of the system). These integrals, which are called action integrals, can be calculated only for integrable systems or, in the language of old quantum theory, 'conditionally periodic systems'.

**EXAMPLE 2.2 (Quantisation of the harmonic oscillator).** A classical harmonic oscillator is defined by the Hamiltonian (total energy)

\[
H(p, q) = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2, \quad (p, q) \in \mathbb{R}^2.
\]

The solution of the Hamilton equations

\[
\begin{align*}
\frac{dq}{dt} &= \frac{\partial H}{\partial p} = \frac{1}{m}p \\
\frac{dp}{dt} &= -\frac{\partial H}{\partial q} = -m\omega^2 q
\end{align*}
\]

are oscillations about the point \((p, q) = (0, 0)\)

\[
\begin{align*}
q(t) &= A \sin(\omega t + \phi) \\
p(t) &= m\omega A \cos(\omega t + \phi)
\end{align*}
\]

where \( A \) and \( \phi \) are arbitrary constants. It is easy to see that the \((p, q)\)-plane is partitioned into ellipses of equation

\[
\frac{q^2}{A^2} + \frac{p^2}{m^2\omega^2 A^2} = 1.
\]

The action integral can be evaluated explicitly

\[
\oint pdq = \pi m\omega A^2 = hn.
\]

Applying the quantisation rule, we see that the amplitude \( A \) is restricted to the quantised values \( A_n = (hn/\pi m\omega)^{1/2} \). The corresponding energy values are

\[
E_n = \frac{1}{2}m\omega^2 A_n^2 = (h/2\pi)\omega n = \hbar \omega n.
\]

The energy levels predicted by the old quantum theory are integral multiples of \( \hbar \omega \).

3. Bell’s inequalities

**3.1. An experiment with polarizers.** In a beam of monochromatic (fixed color) light we put a pair of polarizing filters, each of which can be rotated around the axis formed by the beam. As is well known, the light emerging from both filters changes in intensity when the filters are rotated relative to each other. Starting from the orientation where the resulting intensity is maximal, and rotating one of the filters through an angle \( \alpha \), the light intensity decreases with \( \alpha \), vanishing for \( \alpha = \pi/2 \). If we call the intensity of the beam before the filters \( I_0 \), after the first
$I_1$, and after the second $I_2$, then $I_1 = \frac{1}{2} I_0$ (we assume the original beam to be unpolarized), and

(3.1) \[ I_2 = I_1 \cos^2 \alpha. \]

So far the phenomenon is described well by classical physics. During the last century, however, it has been observed that for very low intensities (monochromatic) light comes in small packages, which were called photons, whose energy depends on the color, but not on the total intensity. So the intensity must be proportional to the number of these photons, and formula (3.1) must be given a statistical meaning: a photon coming from the first filter has probability $\cos^2 \alpha$ to pass through the second. Thinking along the lines of classical probability, we may associate to a polarization filter in the direction $\alpha$ a random variable $P_\alpha$ taking value 0 if the photon is absorbed by the filter and value 1 if the photon passes through. Then, for two filters in the direction $\alpha$ and $\beta$ these random variables should be correlated as follows

(3.2) \[ \mathbb{E} P_\alpha = \mathbb{E} P_\beta = \frac{1}{2} \quad \text{and} \quad \mathbb{E} P_\alpha P_\beta = \mathbb{P}(P_\alpha = 1 \text{ and } P_\beta = 1) = \frac{1}{2} \cos^2(\alpha - \beta). \]

Here we hit on a difficulty: the function on the right hand side is not a possible correlation function! To see this, consider three polarizing filters having polarization directions $\alpha_1$, $\alpha_2$, and $\alpha_3$ respectively. They should give rise to random variables $P_1$, $P_2$ and $P_3$ satisfying

\[ \mathbb{E} P_i P_j = \frac{1}{2} \cos^2(\alpha_i - \alpha_j). \]

**Proposition 3.1** (Bell’s three variable inequality). *For any three 0-1-valued random variables $P_1$, $P_2$, and $P_3$ on the same probability space, the following inequality holds:*

\[ \mathbb{P}(P_1 = 1, P_3 = 0) \leq \mathbb{P}(P_1 = 1, P_2 = 0) + \mathbb{P}(P_2 = 1, P_3 = 0) \]

**Proof.**

\[
\begin{align*}
\mathbb{P}(P_1 = 1, P_3 = 0) &= \mathbb{P}(P_1 = 1, P_2 = 0, P_3 = 0) + \mathbb{P}(P_1 = 1, P_2 = 1, P_3 = 0) \\
&\leq \mathbb{P}(P_1 = 1, P_2 = 0) + \mathbb{P}(P_2 = 1, P_3 = 0).
\end{align*}
\]

In our example,

\[ \mathbb{P}(P_i = 1, P_j = 0) = \mathbb{P}(P_i = 1) - \mathbb{P}(P_i = 1, P_j = 1) = \frac{1}{2} \sin^2(\alpha_i - \alpha_j). \]

Bell’s inequality thus reads

\[ \frac{1}{2} \sin^2(\alpha_1 - \alpha_3) \leq \frac{1}{2} \sin^2(\alpha_1 - \alpha_2) + \frac{1}{2} \sin^2(\alpha_2 - \alpha_3) \]

which is clearly violated for the choices $\alpha_1 = 0$, $\alpha_2 = \pi/6$ and $\alpha_3 = \pi/3$. 
The above calculation could be summarized as follows: we are looking for a family of 0-1-valued random variables \((P_\alpha)_{0 \leq \alpha < \pi}\) on the same probability space with \(\mathbb{P}(P_\alpha = 1) = 1/2\), satisfying
\[
\mathbb{P}(P_\alpha \neq P_\beta) = \sin^2(\alpha - \beta).
\]
Now, on the space of 0-1-valued random variables on a probability space, the function \((X,Y) \mapsto \mathbb{P}(X \neq Y)\) equals the \(L^1\)-distance of \(X\) and \(Y\). On the other hand, the function \((\alpha, \beta) \mapsto \sin^2(\alpha - \beta)\) does not satisfy the triangle inequality. Therefore no family \((P_\alpha)_{0 \leq \alpha < \pi}\) exists which meets the above requirement.

3.2. An improved experiment. On closer inspection the above example is not very convincing. Indeed, when two polarizers are arranged on the optical bench, why should not the random variable for the second polarizer depend on the angle of the first?

In fact we can do a better experiments using a clever technique from quantum optics. It is possible to build a device that produces pairs of photons, such that the members of each pair move in opposite directions and show opposite behaviour towards parallel polarization filters: if one passes the filter, then the other is surely absorbed. With these photon pairs, the very same experiment can be performed, but this time the polarizers are far apart, each one acting on its own photon. (This is the optical version of the Gedankenexperiment was proposed and discussed by Einstein, Podolski, Rosen and Bohm.) The same outcomes are found, violating Bell’s three variable inequality.

3.3. The decisive experiment. Advocates of classical probability could still find serious fault with the argument given so far. So the argument has to be tightened still further. This brings us to the experiment (conceptually devised by Bell in the sixties) which was actually performed by A. Aspect and his team at Orsay, Paris, in 1982. In this experiment a random choice out of two different polarization measurements was performed on each side of the pair-producing device, say in the direction \(\alpha_1\) or \(\alpha_2\) on the left and in the direction \(\beta_1\) or \(\beta_2\) on the right, giving rise to four random variables \(P_1 = P(\alpha_1), P_2 = P(\alpha_2)\) and \(Q_1 = Q(\beta_1), Q_2 = Q(\beta_2)\), two of which are measured and compared at each trial.

**Proposition 3.2 (Bell’s four variable inequality).** For any four 0-1-valued random variables \(P_1, P_2, Q_1,\) and \(Q_2\) on the same probability space, the following inequality holds:
\[
\mathbb{P}(P_1 = Q_1) \leq \mathbb{P}(P_1 = Q_2) + \mathbb{P}(P_2 = Q_1) + \mathbb{P}(P_2 = Q_2).
\]

Quantum mechanics predicts, and the experiment of Aspect showed that
\[
\mathbb{P}(P(\alpha) = Q(\beta) = 1) = \mathbb{P}(P(\alpha) = Q(\beta) = 0) = \frac{1}{2} \sin^2(\alpha - \beta).
\]
Hence
\[ \mathbb{P}(P(\alpha) = Q(\beta)) = \sin^2(\alpha - \beta). \]

So Bell’s four variable inequality reads
\[ \sin^2(\alpha_1 - \beta_1) \leq \sin^2(\alpha_1 - \beta_2) + \sin^2(\alpha_2 - \beta_1) + \sin^2(\alpha_2 - \beta_2), \]
which is clearly violated for the choices \( \alpha_1 = 0, \alpha_2 = \pi/3, \beta_1 = \pi/2 \) and \( \beta_2 = \pi/6 \).

There does not exist, on any classical probability space, a quadruple \( P_1, P_2, Q_1, \) and \( Q_2 \) of random variables with the correlations measured in this experiment.

4. Two-level systems

Consider the historic Stern-Gerlach experiment whose purpose was to determine the magnetic moment of atoms, by measuring the deflection of a neutral atomic beam by a magnetic field. If the mean magnetic field is oriented along the \( e_1 \)-axis, i.e. \( \vec{B} = B e_1 \), then the classical equations of motion predict a deflection of the beam along the \( e_1 \)-axis proportional to \( \mu_1 = \mu \cdot e_1 \). The surprising result found by Gerlach and Stern was that \( \mu_1 \) could take only two values, \( \pm \mu \).

This result is extremely surprising from the point of view of classical physics, because Gerlach and Stern could have chosen different orientations for their magnet, for example \( e_2 \) and \( e_3 \), making angles of \( \pm 120^\circ \) with \( e_1 \) to measure \( \mu_2 = \mu \cdot e_2 \) or \( \mu_2 = \mu \cdot e_2 \), respectively. As the outcome of the experiment cannot be affected by merely rotating the magnet, the would have found, likewise, \( \mu_2 = \pm \mu \) or \( \mu_3 = \pm \mu \). This creates, however an apparent contradiction because
\[ \mu_1 + \mu_2 + \mu_3 = \mu \cdot (e_1 + e_2 + e_2) = 0. \]

Obviously, \( \mu_1, \mu_2 \) and \( \mu_3 \) cannot all be equal to \( \pm \mu \), and also sum up to zero.

Of course, it is impossible to measure in this way the values of \( \mu_1, \mu_2 \) and \( \mu_3 \) of the same atom (this is purely classical impossibility). Quantum theory tells that this is not a defect of this particular experimental method for measuring a magnetic moment: no experiment whatsoever can determine \( \mu_1, \mu_2 \) and \( \mu_3 \) simultaneously.

Example 4.1. Consider an experiment with sequential Stern-Gerlach (SG) apparatus. We measure the magnetic moment (spin) \( S_z \) of a neutral beam in the direction \( e_z \). Two beams of equal intensity emerge from the apparatus.

Suppose that we block the \( S_z = -1/2 \) beam (the bottom beam say) and we let only the \( 1/2 \) beam go into a second identical SG apparatus. This apparatus lets those out the top output and nothing comes out in the bottom output. We say that the states with spin \( +1/2 \) have no component along \( S_z = -1/2 \).

Now, suppose that we measure the \( S_x \) component of the \( S_z = +1/2 \) beam. Classically an object with angular momentum along the z axis has no component
of angular momentum along the x axis, these are orthogonal directions. Here however, about half of the beam exits through the top \( S_x = +1/2 \), and the other half through the bottom \( S_x = -1/2 \).

Now we block the top beam and we only have an \( S_x = -1/2 \) output. The beam is fed into a SG along \( e_z \). We find half of the particles make it out of the third SG with \( S_z = +1/2 \) and the other half with \( S_z = -1/2 \).

### 4.1. Spin and Pauli matrices.

One can consider experiments with a few Stern-Gerlach apparata in series measuring the magnetic moment (the spin) in the directions \( e_x \), \( e_y \), \( e_z \), and so on. A mathematical framework consistent with the outcomes of the experiments and their probabilities is as follows. The experiment suggests that the spin state can be described using normalised vectors in \( \mathbb{C}^2 \). In particular we can describe a state using two basis vectors, for instance

\[
|\uparrow_z\rangle \quad \text{and} \quad |\downarrow_z\rangle.
\]

The first corresponds to a spin \( +1/2 \) in the z-direction (‘spin up along z’). The second corresponds to the state of spin \( -1/2 \) ‘spin down along z’. A magnet measuring the spin in the z direction is an operator \( S_z \) for which \( |\uparrow_z\rangle \) and \( |\downarrow_z\rangle \) are eigenvectors with eigenvalue \( \pm 1/2 \), respectively:

\[
S_z |\uparrow_z\rangle = +\frac{1}{2} |\uparrow_z\rangle \quad \text{and} \quad S_z |\downarrow_z\rangle = -\frac{1}{2} |\downarrow_z\rangle.
\]

It is customary to represent the two vectors in the computational basis

\[
|\uparrow_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow_z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\]

A state can be represented as a superposition (i.e. normalised linear combination)

\[
|\psi\rangle = a |\uparrow_z\rangle + b |\downarrow_z\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \quad a, b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1.
\]

The probability to measure a state \( \phi \) given that the system is prepared in the state \( \psi \) is \( |\langle \phi | \psi \rangle|^2 \). The state \( |\uparrow_z\rangle \) entering the second Stern-Gerlach must have zero overlap with \( |\downarrow_z\rangle \) since no such down spins emerge. Moreover the overlap of \( |\uparrow_z\rangle \) with itself must be one, as all states emerge from the second Stern-Gerlach top output. Indeed,

\[
\langle \uparrow_z | \downarrow_z \rangle = (1 \ 0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0, \quad \langle \uparrow_z | \uparrow_z \rangle = (1 \ 0) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1, \quad \text{etc.}
\]

The operator \( S_z \) can be represented as the \( 2 \times 2 \) matrix:

\[
S_z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]
When we measure the $z$-component of a state $|\uparrow_x\rangle$ we get $|\uparrow_z\rangle$ or $|\downarrow_z\rangle$ with equal probability $1/2$; hence

\[ |\langle \uparrow_z | \uparrow_x \rangle| = |\langle \downarrow_z | \uparrow_x \rangle| = \frac{1}{\sqrt{2}}. \]

Therefore we can write

\[ |\uparrow_x\rangle = \frac{1}{\sqrt{2}} |\uparrow_z\rangle + \frac{1}{\sqrt{2}} e^{i\delta_1} |\downarrow_z\rangle. \]

The state $|\downarrow_x\rangle$ must be orthogonal to $|\uparrow_x\rangle$; this leads to

\[ |\downarrow_x\rangle = \frac{1}{\sqrt{2}} |\uparrow_z\rangle - \frac{1}{\sqrt{2}} e^{i\delta_2} |\downarrow_z\rangle. \]

We can now give a matrix representation of $S_x$:

\[ S_x = \frac{1}{2} \begin{pmatrix} 0 & e^{-i\delta_1} \\ e^{i\delta_1} & 0 \end{pmatrix}. \]

A similar argument with $S_x$ replaced by $S_y$ leads to

\[ S_y = \frac{1}{2} \begin{pmatrix} 0 & e^{-i\delta_2} \\ e^{i\delta_2} & 0 \end{pmatrix}. \]

Is there any way of determining $\delta_1$ and $\delta_2$? We can consider a sequential Stern-Gerlach experiment with a measure along $e_x$ followed by a measure along $e_y$, leading to

\[ |\langle \uparrow_y | \uparrow_x \rangle| = |\langle \downarrow_y | \uparrow_x \rangle| = \frac{1}{\sqrt{2}}. \]

Thus we obtain

\[ \frac{1}{2} \left| 1 \pm e^{i(\delta_1-\delta_2)} \right| = \frac{1}{\sqrt{2}} \]

which is satisfied only if

\[ \delta_2 - \delta_1 = \pm \pi/2. \]

We see that the matrix elements of $S_z$, $S_x$ and $S_y$ cannot all be real. The standard representation is $\delta_1 = 0$ and $\delta_2 = \pi/2$ so that $S_i = \frac{1}{2} \sigma_i$, $i = x, y, z$, with

\[ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \]

**Exercise 4.1.**

1. Verify the following commutation relations of the Pauli matrices $\sigma_x$, $\sigma_y$ and $\sigma_z$.

\[ [\sigma_x, \sigma_y] = i\sigma_z, \quad [\sigma_y, \sigma_z] = i\sigma_x, \quad [\sigma_z, \sigma_x] = i\sigma_y. \]

2. Compute the anticommutators $\{\sigma_x, \sigma_y\}$, $\{\sigma_y, \sigma_z\}$, $\{\sigma_z, \sigma_x\}$.

3. Compute $\sigma_x^2$, $\sigma_y^2$ and $\sigma_z^2$.

4. Show that $S^2 = S_x^2 + S_y^2 + S_z^2$ is a multiple of the identity operator.
The description and analysis of spin states that point in arbitrary directions, as specified by a unit vector $\mathbf{n}$ goes as follows. Let

$$\mathbf{n} = (n_x, n_y, n_z) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad 0 \leq \theta \leq \pi, 0 \leq \phi < 2\pi,$$

be a unit vector. Here $\theta$ and $\phi$ are the familiar polar and azimuthal angles. We can consider $\mathbf{S} = (S_x, S_y, S_z)$, and define the spin operator $S_n$ along $\mathbf{n}$ as

$$S_n = \mathbf{n} \cdot \mathbf{S} = n_x S_x + n_y S_y + n_z S_z.$$

**EXERCISE 4.2.**

1. Check that, just like $S_x$, $S_y$, and $S_z$, the eigenvalues of $S_n$ are $\pm 1/2$.

2. Show that the eigenvectors $|\pm; \mathbf{n}\rangle$ of $S_n$ can be written as

$$|+; \mathbf{n}\rangle = \cos \frac{\theta}{2} |\uparrow_z\rangle + \sin \frac{\theta}{2} e^{i\phi} |\downarrow_z\rangle$$

(4.15)

$$|--; \mathbf{n}\rangle = \sin \frac{\theta}{2} |\uparrow_z\rangle - \cos \frac{\theta}{2} e^{i\phi} |\downarrow_z\rangle.$$  

(4.16)

The unit vector $\mathbf{n}$ denotes a location on the unit sphere, or ‘Bloch sphere’. Each point on the surface of the sphere represents a (pure) state, i.e. a normalised vector $|\psi\rangle \in \mathbb{C}^2$. For example, $|\uparrow_z\rangle$ corresponds to the north pole ($\theta = 0$), etc.

### 5. The Schrödinger equation and wave mechanics

In the first half of 1926, E. Schrödinger showed that the rules of quantisations can be replaced by another postulate, in which there occurs no mention of whole numbers. Instead, the introduction of integers arises in the same way as, for example, in a vibrating string, for which the number of nodes is integer.

Schrödinger postulated a sort of wave equation and applied it to a number of problems, including the hydrogen atom and the harmonic oscillator.

Schrödinger’s approach to the dynamics differs from that of classical mechanics in its aims as well in its method. Instead of attempting to find equations, such as Newton’s equations, which enable a prediction to be made of the exact positions and momenta of the particles of a system at any time, he devised a method of calculating a function $\Psi$ of the coordinates and time (and not the momenta), with the aid of which, probabilities for the coordinates or other quantities can be predicted.

The Schrödinger equation enables to determine a certain function $\Psi$ (of the coordinates and time) called wavefunction or probability amplitude. The square of the modulus of a wavefunction is interpreted as a probability measure for the coordinates of the system in the state represented by this wavefunction.

Besides yielding the probability amplitude $\Psi$, the Schrödinger equation provides a method of calculating values of the energy of the stationary states of the system. No arbitrary postulates concerning quantum numbers are required in the calculations; instead, integers enter automatically in the process of finding satisfactory solutions of the wave equation.
For our purposes, the Schrödinger equation, the auxiliary restrictions upon the wavefunction and its probabilistic interpretation are conveniently taken as fundamental postulates.

5.1. The Schrödinger equation. Consider a system with one degree of freedom, consisting of a particle of mass \( m \) that can move along a straight line, and let us assume that the system is further described by a potential energy function \( V(x) \) throughout the region \(-\infty < x < +\infty\). The classical Hamiltonian of the system is

\[
H(p, x) = \frac{p^2}{2m} + V(x)
\]

and by the conservation of energy the Hamiltonian is constant on solutions of the Hamilton equations:

\[
H(p, x) = W.
\]

Schrödinger’s idea was to replace \( p \) by the differential operator \( \hbar \frac{\partial}{\partial x} \) and \( W \) by \( -\frac{\hbar}{i} \frac{\partial}{\partial t} \), and introduce the function \( \Psi(x, t) \) on which these operators can operate

\[
H \left( \frac{\hbar}{i} \frac{\partial}{\partial x}, x \right) \Psi(x, t) = W \Psi(x, t).
\]

For this system the Schrödinger equation is

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

**Remark.** The analogy between the terms in the wave equation and the classical energy equation has only a formal significance.

5.2. The time-independent Schrödinger equation. In order to solve (5.2) let us (as usual in the analysis of partial differential equations) first study the non-trivial solutions \( \Psi \) (if any exist) which can be expressed as product of two functions, one involving time alone and the other the coordinate alone:

\[
\Psi(x, t) = \psi(x) \varphi(t).
\]

On introducing this in Eq. (5.2) we get

\[
\frac{1}{\psi(x)} \left( \frac{-\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} + V(x) \psi(x) \right) = \frac{-\hbar}{i} \frac{1}{\varphi(t)} \frac{\partial \varphi(t)}{\partial t}
\]

The right-hand side of this equation is a function of time \( t \) alone and the left-hand side a function of \( x \) alone. It is consequently necessary that the quantity to which each side is equal to be dependent on neither \( x \) nor \( t \); that is, that it be a constant.
Let us call it $E$. Then,

\begin{equation}
\begin{cases}
\frac{-\hbar \partial \varphi(t)}{i \partial t} = E \varphi(t) \\
\frac{-\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial^2 x} + V(x) \psi(x) = E \psi(x).
\end{cases}
\end{equation}

(5.4)

The second equation is customarily written as a spectral problem for the \textbf{Schrödinger operator} $\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$, i.e.,

\begin{equation}
\left( \frac{-\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi(x) = E \psi(x).
\end{equation}

(5.5)

Eq. (5.5) is often called time independent Schrödinger equation, or amplitude equation, inasmuch $\psi(x)$ determines the amplitude of $\Psi(x,t)$.

It is found that typically the equation possesses various solutions, corresponding to various values of the constant $E$. Let us denote these values of $E$ by attaching a subscript $n$, and similarly the amplitude corresponding to $E_n$ as $\psi_n(x)$. The corresponding equation for $\varphi(t)$ can be integrated at once to give

\begin{equation}
\varphi_n(t) = e^{-i \frac{E_n t}{\hbar}}.
\end{equation}

(5.6)

The general solution of the Schrödinger equation is a linear combination of all the particular solutions with arbitrary coefficients

\begin{equation}
\Psi(x,t) = \sum_n a_n \Psi_n(x,t) = \sum_n a_n \psi_n(x) e^{-i \frac{E_n t}{\hbar}}.
\end{equation}

(5.7)

5.3. Wave functions: discrete and continuous sets of energy values. It is found that satisfactory solutions $\psi_n(x)$ exist only for certain values of the parameter $E_n$. The values $E_n$ are called eigenvalues of the Schrödinger operator. It will be shown later that the physical interpretation of the wavefunction requires that the values $E_n$ represent the energy of the system in its various stationary states.

In addition to being a solution of the Schrödinger equation, the wavefunction must be single-valued, continuous and square integrable.

For a given system, the energy levels $E_n$ may occur only as a set of discrete values, or as a set of values covering a continuous range, or as both. From analogy with spectroscopy it is often said that in these three cases the energy values comprise a discrete spectrum, a continuous spectrum, or both. The way in which the postulates regarding the wave equation and its acceptable solutions lead to the selection of definite energy values may be understood by the qualitative consideration of a simple example. Let us consider, for a system of one degree of freedom, that the potential-energy function satisfies $V(x) \to \infty$ as $|x| \to \infty$. For a given value of the energy parameter $E$, the wave equation is

\begin{equation}
\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} = (V(x) - E) \psi(x).
\end{equation}

(5.8)
For sufficiently large $|x|$, the quantity $V(x) - E$ will be positive. Hence in this region the curvature $\frac{\partial^2 \psi}{\partial x^2}$ will be positive if $\psi$ is positive, and negative if $\psi$ is negative. Suppose that at an arbitrary point $x = c$ the function $\psi$ has a certain value and a certain slope $\frac{\partial \psi}{\partial x}$. The behaviour of the function, as it is continued both to the right and to the left, is completely determined by the values assigned to two quantities. We see that, for a given value of $E$ only by a very careful selection of the slope of the function at the point $x = c$ can the function be made to behave properly for large values of $x$. Similarly for large negative values of $x$. This selection is such as to cause the wave function to approach the value zero asymptotically with increasing $|x|$. In view of the sensitiveness of the curve to the parameter $E$, an infinitesimal change from this satisfactory value will cause the function to behave improperly. We conclude that the parameter $E$ and the slope at the point $x = c$ (for a given value of the function itself at this point) can have only certain values if $\psi$ is to be an acceptable wave function. For each satisfactory value of $E$ there is one (or, in certain cases, more than one) satisfactory value of the slope, by the use of which the corresponding wave function can be built up. For this system the characteristic values $E_n$ of the energy form a discrete set, and only a discrete set, inasmuch as for every value of $E$, no matter how large, $V(x) - E$ is positive for sufficiently large values of $|x|$.

It is customary to number the energy levels for such a system as $E_0$ (the lowest), $E_1$ (the next), and so on, corresponding to wavefunctions $\psi_0(x)$, $\psi_1(x)$, etc. The integer $n$, which is written as a subscript in $E_n$ and $\psi_n(x)$, is called quantum number.

Let us now consider a system in which the potential-energy function remains finite as $x \to \infty$ or as $x \to -\infty$ or at both limits. For a value of $E$ smaller than both $V(+\infty)$ or $V(-\infty)$ the argument presented above is valid. Consequently the energy levels will form a discrete set for this region. If $E$ is greater than $V(+\infty)$, however, a similar argument shows that the curvature will be such as always to return the wave function to the $x$ axis, about which it will oscillate. Hence any value of $E$ greater than $V(+\infty)$ or $V(-\infty)$ will be an allowed value, corresponding to an acceptable wave function, and the system will have a continuous spectrum of energy values in this region.

### 5.4. Probability measures.

Let $\Psi(x, t)$ be a normalised solution of the Schrödinger equation; that is

$$\int_{-\infty}^{+\infty} \Psi^*(x, t)\Psi(x, t)dx = 1.$$  

For any $t$, the quantity $\Psi^*(x, t)\Psi(x, t)dx = |\Psi(x, t)|^2dx$ is a probability measure. It can be interpreted as the probability that the system represented by the state $\Psi(x, t)$ is in the region between $x$ and $x + dx$ at time $t$. 

It is also convenient to normalise the individual amplitude function $\psi_n(x)$,

$$\int_{-\infty}^{+\infty} \psi_n^*(x) \psi_n(x) dx = 1.$$  

Moreover, the independent solutions of the amplitude equation can always be chosen in such a way that they form an orthonormal system; that is

$$\int_{-\infty}^{+\infty} \psi_m^*(x) \psi_n(x) dx = \delta_{mn}.$$  

Using this relation, it is found that a wavefunction $\Psi(x, t) = \sum_n a_n \psi_n(x, t)$ is normalised when the coefficients $a_n \in \mathbb{C}$ satisfy the condition

$$\sum_n a_n^* a_n = 1.$$  

**5.5. Stationary states.** Let us consider the probability measure $\Psi^* \Psi$ for a system in the state $\Psi(x, t) = \sum_n a_n \psi_n(x) e^{-iE_n t / \hbar}$ and its conjugate $\Psi^*(x, t) = \sum_n a_n^* \psi_n^*(x) e^{iE_n t / \hbar}$. Then,

$$\Psi^*(x, t) \Psi(x, t) = \sum_n a_n^* a_n \psi_n^*(x) \psi_n(x) + \sum_{m \neq n} a_m^* a_n \psi_m^*(x) \psi_n(x) e^{-i(E_m - E_n) t / \hbar}.$$  

In general, then, the probability function and hence the properties of the system depend on the time, inasmuch as the time enters in the exponential factors of the double sum. Only if the coefficients are zero for all except one value of $E_n$ is $\Psi^* \Psi$ independent of $t$. In such a case the wavefunction will contain only a single term $\Psi_n(x, t) = \psi_n(x) e^{-iE_n t / \hbar}$; For such a state the probability measure $\Psi^* \Psi$ is independent of time, and the state is called **stationary state**.

**5.6. Average values.** If we inquire as to what average value would be expected on measurement at a given time $t$ of the coordinate $x$ of the system in a physical situation represented by the wavefunction $\Psi(x, t)$, the above interpretation of $\Psi^* \Psi$ leads to the answer

$$\langle x \rangle = \int_{-\infty}^{+\infty} x \Psi(x, t)^* \Psi(x, t) dx.$$  

A similar integral gives the average value predicted for $x^2$, $x^3$, or any function $f(x)$ of the coordinate $x$:

$$\langle f(x) \rangle = \int_{-\infty}^{+\infty} f(x) \Psi(x, t)^* \Psi(x, t) dx.$$  

In order that the same question can be answered for a more general dynamical function $f(p, x)$ involving the momentum $p$ as well as the coordinate $x$, we assume
that the average value of \( f(p,x) \) predicted for a system in the physical situation represented by the wavefunction \( \Psi \) is given by the integral

\[
\langle f(p,x) \rangle = \int_{-\infty}^{+\infty} \Psi(x,t)^* f \left( \frac{\hbar}{i} \frac{\partial}{\partial x}, x \right) \Psi(x,t) dx.
\]

in which the operator \( f \) obtained from \( f(p,x) \) by replacing \( p \) by \( \frac{\hbar}{i} \frac{\partial}{\partial x} \), acts on the function \( \Psi(x,t) \).

**Remark.** In some cases further considerations are necessary in order to determine the exact form of the operator \( f \left( \frac{\hbar}{i} \frac{\partial}{\partial x}, x \right) \), but we shall not encounter such difficulties.

In general, the result of a measurement of \( f \) will not be given by the average value \( \langle f \rangle \); it is instead described by a probability distribution.

Even if the system is in a stationary state \( \Psi_n(x,t) = \psi_n(x)e^{-i\frac{E_n}{\hbar}t} \), the variance \( \text{Var}(f) = \langle f^2 \rangle - \langle f \rangle^2 \) of an arbitrary observable \( f \) is, in general, not zero. The energy of the system, corresponding to the Hamiltonian function \( H(p,x) \), has, however, a definite value for a stationary state \( \Psi_n \) of the system, equal to \( E_n \). To prove this, we evaluate \( \langle H \rangle \) and \( \text{Var}(H) \). The average value of \( H \) for the state \( \Psi_n \) is given by

\[
\langle H \rangle = \int_{-\infty}^{+\infty} \psi_n(x)^* \left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi_n(x) dx = E_n.
\]

By a similar computation \( \langle H^r \rangle = E^r_n, r = 0, 1, 2, \ldots \). In particular, this shows that \( \text{Var}(H) = 0 \); thus, the energy of the system in the state \( \Psi_n \) has a definite value \( E_n \).

**Remark.** Note that the probabilistic interpretation of \( \Psi(x,t) \) is unchanged if we multiply the wavefunction by a complex number \( \alpha \) of modulus \( |\alpha| = 1 \). Thus, a state is an equivalence class of the form \( e^{i\theta} \Psi(x,t) \), \( \theta \in \mathbb{R} \). (This is especially relevant when discussing symmetries of a quantum system. More on this later.)

**Exercise 5.1 (Probability conservation law).** Show that, if \( \Psi(x,t) \) is a solution of the Schrödinger equation

\[
i\hbar \frac{\partial \Psi(x,t)}{\partial t} = H \Psi(x,t), \quad x \in \mathbb{R}^n,
\]

\((H = -\frac{\hbar^2}{2m} \Delta + V(x))\), then the probability density \( |\Psi(x,t)|^2 \) satisfies the continuity equation (conservation of probability):

\[
\frac{\partial}{\partial t} |\Psi|^2 + \vec{\nabla} \cdot \vec{J} = 0,
\]

where the probability current is

\[
\vec{J} = \frac{i\hbar}{2m} \left( \Psi \nabla \Psi^* - \Psi^* \nabla \Psi \right).
\]
Exercise 5.2 (Correspondence principle). Prove the Ehrenfest theorem: if \( H(p, q) \) is the classical Hamiltonian of the systems, then the expectation values \( \langle q \rangle \) and \( \langle p \rangle \) of the quantised variables satisfy the classical equation of motion

\[
\begin{align*}
\frac{d\langle q \rangle}{dt} &= +\langle \frac{\partial H}{\partial p} \rangle \\
\frac{d\langle p \rangle}{dt} &= -\langle \frac{\partial H}{\partial q} \rangle.
\end{align*}
\]

(This is an instance of the correspondence principle.)

6. Exactly solvable models

6.1. The quantum harmonic oscillator. You may be familiar with several examples of harmonic oscillators form classical mechanics, such as particles on a spring or the pendulum for small deviation from equilibrium, etc.

The classical Hamiltonian of a particle in a potential \( V(x) = \frac{1}{2}m\omega^2x \) is

\[
H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2x.
\]

The corresponding quantum Hamiltonian is

\[
H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2.
\]

Using dimensionless coordinates, the corresponding Schrödinger equation is

\[
\left( -\frac{1}{2} \frac{d^2}{dx^2} + \frac{1}{2}x^2 \right) \psi(x) = E\psi(x).
\]

There is a very useful procedure to solve (6.3) based on the determination of the form of \( \psi \) in the regions of large \( |x| \), and the subsequent discussion, by the introduction of a factor in the form of a power series (which later reduces to a polynomial), of the value of \( \psi \) for generic \( x \in \mathbb{R} \). This procedure may be called the polynomial method.

6.1.1. Eigenvalues and eigenfunctions of the harmonic oscillator. The first steps is the asymptotics of the solution of (6.3) when \( |x| \gg 1 \). For large \( |x| \), \( E \) is negligibly small relative to \( x^2/2 \), and the asymptotic form of the Schrödinger equation becomes

\[
\frac{d^2\psi}{dx^2} = x^2\psi.
\]

This equation is asymptotically satisfied by \( e^{-x^2/2} \) and \( e^{x^2/2} \). Of the two asymptotic solutions, the second tends rapidly to infinity with increasing values of \( |x| \); the first, however, leads to a satisfactory treatment of the problem.

We now proceed to obtain a solution of (6.3) throughout the whole configuration space \( x \in \mathbb{R} \), based upon the asymptotic solution, by introducing as factor a power series in \( x \) and determining its coefficients by substitution and recursion.
Let $\psi(x) = H(x)e^{-x^2/2}$. Then

$$\frac{d^2 \psi}{dx^2} = e^{-x^2/2} \left((x^2 - 1)H(x) - 2xH'(x) + H''(x)\right).$$

If $\psi(x)$ solves the Schrödinger equation, then we get a differential equation for $H(x)$:

(6.4) \hspace{1cm} H''(x) - 2xH'(x) + (2E - 1)H(x) = 0.

We now represent $H(x)$ as a power series

(6.5) \hspace{1cm} H(x) = \sum_{k\geq0} a_k x^k, \quad H'(x) = \sum_{k \geq 0} k a_k x^{k-1}, \quad H''(x) = \sum_{k \geq 0} k(k-1)a_k x^{k-2}.

On substitution of these expressions, Eq. (6.4) assumes the form

$$\sum_{k \geq 0} [(k + 2)(k + 1)a_{k+2} - 2k a_k + (2E - 1)a_k] x^k = 0.$$

In order for this series to vanish for all values of $x$ (i.e., for $H(x)$ to be a solution of (6.4)), the coefficients of each power of $x$ must vanish

$$(k + 2)(k + 1)a_{k+2} + (2E - 2k - 1)a_k = 0$$

or

(6.6) \hspace{1cm} a_{k+2} = -\frac{2(E - k - 1/2)}{(k + 2)(k + 1)} a_k.

This recursion formula enables us to calculate successively the coefficients $a_2, a_3, a_4, \ldots$ in terms of $a_0$ and $a_1$, which are arbitrary. Note that if $a_0 = 0$ (resp. $a_1 = 0$) only odd (resp. even) powers appear.

For arbitrary values of the parameter $E$, the above series consists of an infinite number of terms and does not correspond to a satisfactory wavefunction. To see this, note that for large enough values of $k$, we have

$$a_{k+2} \simeq \frac{2}{k} a_k.$$

But this is the same recursion of the series coefficients of $e^{x^2}$,

$$e^{x^2} = 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \cdots + \frac{x^k}{(k/2)!} + \frac{x^{k+2}}{(k+2)/2)!} + \cdots$$

so that the higher terms of the series for $H(x)$ differ from those of $e^{x^2}$ only by a multiplicative constant. Therefore, for large values of $|x|$, $H(x) \sim e^{x^2}$, and the product $e^{-x^2/2}H(x)$ will diverge like $e^{x^2}$, thus making unacceptable as a wavefunction.
We must therefore restrict $E$ to values which will cause the series for $H(x)$ to terminate, leaving a polynomial. The value of $E$ which causes the series to break off after the $n$th term is seen from (6.6) to be

(6.7) \[ E = n + \frac{1}{2}. \]

It is, moreover, also necessary that the value of either $a_0$ or of $a_1$ be put equal to zero, according as $n$ is odd or even. The solutions are thus either odd or even functions of $x$. The normalised solutions of the Schrödinger equation may be written in the form

(6.8) \[ \psi_n(x) = N_n e^{-\frac{x^2}{2}} H_n(x). \]

$H_n(x)$ is the polynomial of degree $n$th determined by the recursion (6.6) with $E = n + 1/2$, and $N_n$ is a constant which is adjusted so that $\int |\psi_n(x)|^2 dx = 1$ (this condition fixes the initial value $a_0$ or $a_1$ of the recursion). The polynomials $H_n(x)$, called Hermite polynomials, did not originate with Schrödinger’s work but were well known to mathematicians in connection with other problems.

They can be defined using the formula

(6.9) \[ H_n(x) = e^{x^2} \left( -\frac{d}{dx} \right)^n e^{-x^2}. \]

Another definition makes use of the generating function

(6.10) \[ S(x, y) = \sum_{n=0}^\infty \frac{H_n(x)}{n!} y^n = e^{x^2-(y-x)^2}. \]

**Exercise 6.1.** Show the equivalence of the two definitions (6.9)-(6.10) of $H_n(x)$.

It is easy to show that $S(x, y)$ satisfies the following identities

(6.11) \[ \frac{\partial S}{\partial y} = -2(y - x) S \]

(6.12) \[ \frac{\partial S}{\partial x} = 2y S. \]

**Exercise 6.2.** Using (6.11) and (6.12) show that the Hermite polynomials satisfy the three-term recursion

(6.13) \[ H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0 \]

(with initial conditions $H_0(x) = 1$, $H_1(x) = 2x$), and the differential equation

(6.14) \[ H''_n(x) - 2xH'_n(x) + 2nH_n(x) = 0. \]

In particular, the latter is just (6.4) if we put $E = n + 1/2$. 

The functions
\[ \psi_n(x) = N_n e^{-\frac{x^2}{2}} H_n(x), \quad \text{with} \quad N_n = \frac{1}{\sqrt{n!2^n\sqrt{\pi}}} \]
are called Hermite functions. They satisfy the orthogonality relation (why?)
\[ \int_{-\infty}^{+\infty} \psi_n(x) \psi_m(x) dx = \delta_{nm}. \]

**Example 6.1.** By using the generating function \( S \) we can evaluate certain integrals involving \( \psi_n \) which are of importance. For example, we may study the integral which determines the probability of transition from the state \( n \) to the state \( m \). This is
\[ \langle n | x | m \rangle = \int_{-\infty}^{+\infty} \psi_n^*(x) x \psi_m(x) dx = N_n N_m \int_{-\infty}^{+\infty} x H_n(x) H_m(x) e^{-x^2} dx. \]
Using the generating functions \( S(x,y) \) and \( S(x,z) \) we obtain the relation
\[ \int_{-\infty}^{+\infty} x S(x,y) S(x,z) dx = \sum_{n,m} \frac{1}{n!m!} y^n z^m \int_{-\infty}^{+\infty} x H_n(x) H_m(x) e^{-x^2} dx \]
\[ = e^{2yz} \int_{-\infty}^{+\infty} x e^{-(x-y-z)^2} dx \]
\[ = e^{2yz} (y+z) \sqrt{\pi} \]
\[ = \sqrt{\pi} \left( y + 2y^2z + \frac{2^2 y^3 z^2}{2!} + \ldots + \frac{2^n y^{n+1} z^n}{n!} + \ldots \right) \]
\[ + z + 2yz^2 + \frac{2^2 y^2 z^3}{2!} + \ldots + \frac{2^n y^n z^{n+1}}{n!} + \ldots \right). \]
Hence, comparing the coefficients of \( y^n z^m \), we see that \( \langle n | x | m \rangle \) is zero except for \( m = n \pm 1 \), and
\[ \langle n | x | n+1 \rangle = \sqrt{\frac{n+1}{2}}, \quad \langle n | x | n-1 \rangle = \sqrt{\frac{n}{2}}. \]

6.1.2. Creation and annihilation operators. The harmonic oscillator is of importance for general theory, because it forms a cornerstone in the theory of radiation. An alternative method of solution is based on the creation and annihilation operators. They are defined as
\[ a^* = \frac{1}{\sqrt{2}} (x - ip) = \frac{1}{\sqrt{2}} (x - \partial_x) \]
\[ a = \frac{1}{\sqrt{2}} (x + ip) = \frac{1}{\sqrt{2}} (x + \partial_x) \]
Note that \( a \) and \( a^* \) do not commute
\[ [a, a^*] = 1. \]
Using the properties of the Hermite functions it is easy to verify that

\begin{align}
    a \psi_n(x) &= \sqrt{n} \psi_{n-1}(x) \\
    a^* \psi_n(x) &= \sqrt{n+1} \psi_{n+1}(x)
\end{align}

We can also show that the number operator \( N = a^* a \) satisfies

\begin{equation}
    N \psi_n(x) = n \psi_n(x).
\end{equation}

We have the following interpretations: \( a^* \) creates a ‘quantum’ of energy, \( a \) annihilates a ‘quantum’ of energy, \( N \) counts the number of quanta of energy. We write

\begin{align}
    a^* |n\rangle &= \sqrt{n+1} |n+1\rangle \\
    a |n\rangle &= \sqrt{n} |n-1\rangle \\
    N |n\rangle &= n |n\rangle.
\end{align}

Note that \( H = a^* a + \frac{1}{2} = N + \frac{1}{2} \), and \( H |n\rangle = (n + 1/2) |n\rangle \). The eigenstates \( |n\rangle \) are obtained by the repeated action of the creation operator on the ground state

\begin{equation}
    |n\rangle = \frac{1}{\sqrt{n!}} (a^*)^n |0\rangle.
\end{equation}

**EXERCISE 6.3.** Prove by induction that

\[ [a, (a^*)^n] = n (a^*)^{n-1}, \quad \text{for all } n \in \mathbb{N}. \]

**6.2. A quantum particle in a box.** The classical Hamiltonian (energy) of a particle in a box \([0, L]\) is

\begin{equation}
    H = \frac{1}{2m} p^2.
\end{equation}

The corresponding quantum Hamiltonian is formally

\begin{equation}
    H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}.
\end{equation}

This corresponds to a force-free \((V = 0)\) situation in the interior of the box. We have to take into account the walls of the box, i.e. the boundary conditions (b.c.). We may assume that the wavefunctions vanishes at the walls \( \psi(0) = 0 \) and \( \psi(L) = 0 \) (so-called Dirichlet b.c.). The eigenfunction of the Schrödinger operator \( H \) with Dirichlet b.c. satisfy the problem

\begin{equation}
    \begin{cases}
        -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E \psi(x) & \text{for } 0 < x < L \\
        \psi(x) = 0 & \text{at } x = 0 \text{ and } x = L.
    \end{cases}
\end{equation}

We see that a generic solution \( \psi \) is of the form

\[ \psi(x) = Ae^{i\alpha x} + Be^{-i\alpha x} \]
with $\alpha^2 = \frac{2mE}{\hbar^2}$ and constants $A$ and $B$. Imposing the boundary conditions

$$\psi(0) = 0 \Rightarrow A + B = 0$$
$$\psi(L) = 0 \Rightarrow Ae^{i\alpha L} + Be^{-i\alpha L} = 0.$$ 

This imposes $A = -B$ and $\alpha L = \pi k$, $k = 1, 2, \ldots$. The value of $A$ is fixed by the normalisation condition $\int_0^L |\psi(x)|^2 dx = 1$. Therefore we find that the Schrödinger equation in a box with Dirichlet b.c. has eigenfunctions

$$\psi_k = \sqrt{\frac{2}{L}} \sin \left( \frac{\pi k}{L} x \right)$$  

with eigenvalues $E_k = \frac{\pi^2 \hbar^2}{2mL^2} k^2$, $k = 1, 2, \ldots$. Note that a particle in a box has a minimum energy

$$E_1 = \frac{\pi^2 \hbar^2}{2L^2m} > 0.$$  

This is a manifestation of the uncertainty principle. If we know that a particle is localised in $[0, L]$, then the momentum $p_x$ has uncertainty $\Delta p_x \geq \frac{2\pi \hbar}{L}$. This amounts to not knowing whether the particle moves with a momentum $|p_x| = \frac{\pi \hbar}{L}$ in the positive or negative sense. The energy corresponding to this momentum is $E = p_x^2/2m = \frac{\pi^2 \hbar^2}{2L^2m} = E_1$. This is a simple method to obtain a rough estimate of the minimum energy of a system. (It can be used to estimate the energy of electrons in an atom or of nucleons in a nucleus if the atomic or nuclear radius, i.e., the size of the box, is known.)

**Example 6.2**. Consider the Hamiltonian $H$ of a free particle in a box of length $L$ with Dirichlet b.c.. Suppose that a particle (of mass $m$) is in the state $\psi(x)$, where

$$\psi(x) = Ax(L - x)$$

if $0 \leq x \leq L$ and zero otherwise. The normalisation constant is $A = \sqrt{30L^5}$. Suppose that we measure $H$ in the state $\psi(x)$. The mean value of the energy is

$$\langle H \rangle = \int \psi^*(x)H\psi(x)dx = \frac{Ah^2}{m} \int_0^L \psi_n^*(x)dx = \frac{5h}{mL^2}.$$ 

To find the most probable outcome of the measurement we expand the wavefunction $\psi(x)$ in the eigenfunctions of the Hamiltonian $\psi(x) = \sum_{k \geq 1} c_n \psi_n(x)$, where

$$c_n = \int \psi_n^*(x)\psi(x)dx.$$ 

A simple calculation reveals that the most probable value is $E_1 = \frac{\pi^2 \hbar^2}{2L^2m}$ with probability $|c_1|^2 = 960/\pi^6 \simeq 0.998$.

**Example 6.3 (Quantum fractals)**. Let $H_D$ be the Hamiltonian of a free particle in the box $[0, L]$ with Dirichlet b.c.. The simplicity of the spectrum and eigenfunctions is deceptive and could lead us to think that the dynamics, which is the
solution to the Schrödinger equation, is very simple. In fact, this belief is false, as showed by M. V. Berry: the dynamics is instead very intricate. Consider the evolution \(-i\hbar \partial_t \psi = H_D \psi\) with initial condition \(\psi(x,0) = \sqrt{1/L}\) (uniform initial state). The wavefunction \(\psi(x,t)\) at time \(t \in \mathbb{R}\) is given by \(\psi(x,t) = \theta(\xi, \tau)\) where 

\[
\xi = \frac{x}{L}, \tau = \frac{2\pi \hbar}{mL^2}, \quad \theta(\xi, \tau) = \text{const} \cdot \sum_{k \in \mathbb{Z}} \frac{1}{(k + \frac{1}{2})^2} e^{2\pi i (k + \frac{1}{2}) - i\pi (k + \frac{1}{2})^2}.
\]

(This is related to a Jacobi theta function.) It is easy to check that the wavefunction is quasiperiodic \(\theta(\xi, \tau + 1) = e^{-i\pi/4} \theta(\xi, \tau)\). Thus, at integer times \(\tau\) the wave function comes back (up to a phase) to its initial flat form: these are the quantum revivals. More generally, at rational values of \(\tau\) the graph of \(|\theta(\xi, \tau)|^2\) is piecewise constant and there is a partial reconstruction of the initial wavefunction. On the other hand, it can be proved that at irrational times, the wavefunction is a fractal in space and time and form a beautifully intricate quantum carpet.

**EXERCISE 6.4.** Different boundary conditions correspond to different Schrödinger operators, i.e., different physical systems. Show that the stationary states \(\psi_k(x)\) and energy levels \(E_k\) of a particle in a box with Neumann b.c. \((\psi'(0) = 0 \text{ and } \psi'(L) = 0)\) and are

\[
\psi_0(x) = \sqrt{\frac{1}{L}}, \quad \psi_k(x) = \sqrt{\frac{2}{L}} \cos \left(\frac{\pi k}{L} x\right), \quad k = 1, 2, \ldots
\]

with \(E_k = \frac{\pi^2 \hbar^2}{2L^2 m} k^2\), \(k = 0, 1, 2, \ldots\).

**EXERCISE 6.5.** Diagonalise the Hamiltonian \(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2}\), \(0 \leq x \leq L\) with periodic b.c. \(\psi'(0) = \psi(L)\) (quantum particle in a ring).

**REMARK.** The strictly positivity of the energy is a property of the Dirichlet boundary conditions. Indeed, with the Neumann ones the energy of the ground state is 0, and, more surprisingly, it is even negative with Robin’s boundary conditions.

**6.3. Free evolution of a quantum particle in \(\mathbb{R}^n\).** The evolution of the state of a particle of mass \(m > 0\) in \(n\) dimensions in the absence of a potential (i.e., free evolution, \(V(x) = 0\)) is governed by the Hamiltonian

\[
(6.32) \quad H = -\frac{\hbar^2}{2m} \Delta,
\]

The operator \(H\) is essentially self-adjoint, and we will identify it with its unique self-adjoint extension. If \(\psi_0\) is the state at time \(t = 0\), the state \(\psi(x,t)\) at time \(t\) is given by the solution of the free Schrödinger equation:

\[
(6.33) \quad i\hbar \frac{\partial \psi}{\partial t} + \frac{\hbar^2}{2m} \Delta \psi = 0, \quad x \in \mathbb{R}^n, t \in \mathbb{R}
\]
with initial condition \( \psi(x, 0) = \psi_0(x) \) (here we can assume \( \psi_0(x) \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n) \)).

Let us first consider the momentum of a free particle. The wavefunction in momentum representation is the **Fourier transform** of the state in position representation

\[
(6.34) \quad \hat{\psi}(p, t) = \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} \psi(x, t) e^{-ipx} dx.
\]

It is not hard to show, using the Schrödinger equation, that

\[
(6.35) \quad \hat{\psi}(p, t) = e^{-i\hbar|p|^2/2m}\hat{\psi}_0(p).
\]

Thus, \(|\psi(p, t)|^2 = |\hat{\psi}_0(p)|^2\). Hence the probability density of the momentum \(p\) is not changing at all in time. This is what we expect, since the momentum of a free particle is a constant of motion.

The free Schrödinger equation (6.33) has an explicit solution as inverse Fourier transform of (6.34)

\[
(6.36) \quad \psi(x, t) = \frac{1}{(2\pi)^n/2} \int_{\mathbb{R}^n} \hat{\psi}_0(p)e^{i\frac{\hbar}{m}|p|^2+i}px\hat{\psi}_0(p) dp.
\]

The above formula can be interpreted as a superposition of plane waves with amplitude determined by the initial condition (in momentum space) \(\hat{\psi}_0(p)\). Using the convolution theorem for the Fourier transform, we can write

\[
(6.37) \quad \psi(x, t) = \left(\frac{2m}{4\pi i\hbar}\right)^{n/2} \int_{\mathbb{R}^n} e^{im|x-y|^2}\hat{\psi}_0(y)dy = \int_{\mathbb{R}^n} G(x, t; y, 0)\psi_0(y)dy
\]

We see that the solutions is given by the **convolution** of the free particle propagator

\[
G(x, t; x', t') = \left(\frac{m}{2\pi i\hbar(t-t')}\right)^{n/2} e^{im|x-x'|^2/(2\hbar(t-t'))}
\]

and the initial condition \(\psi_0\). It is easy to show using this explicit representation of the solution that for any \(\psi_0 \in L^2(\mathbb{R}^n)\), \(\|\psi_0\|_\infty \to 0\) as \(t \to \infty\). So, the probability density \(|\psi(x, t)|^2\) of the position flattens as \(t \to \infty\). In other words, there is no limiting probability distribution for the position of the particle; it spreads out across the whole space.

**6.4. Scattering processes.** Problems for which the energy eigenvalues are continuously distributed usually arise in connection with the collision of a particle with a force field. The method of approach is different from that employed for systems with discrete levels. There the boundary conditions were used to determine the discrete energy levels of the particle. In a **scattering** problem, the energy is specified in advance, and the behaviour of the wavefunction at great distances (the boundary values) is found in terms of it. This asymptotic behaviour can then
be related to the amount of scattering of the particle by the force field. We discuss now a couple of exact solutions for simplified models that are routinely used for approximate calculations on more complicated systems.

Consider the one-dimensional scattering of a particle with a potential \( V(x) \). Suppose that \( V(x) \) is nondecreasing and that

\[
\lim_{x \to +\infty} V(x) = V_0 > 0, \\
\lim_{x \to -\infty} V(x) = 0.
\]

In classical mechanics, a particle coming from the left with energy \( E < V_0 \) can not reach \( +\infty \): the potential is a barrier. In quantum mechanics there are new phenomena:

(i) If \( E < V_0 \) the particle can penetrate the barrier. In general, there is nonzero probability that the particle is in the classical forbidden region \( \{ x : E < V(x) \} \);

(ii) If \( E > V_0 \) there is the possibility that the particle is reflected by the barrier.

These properties will be shown by calculating the probabilities of transmission and reflection.

The wavefunction of a particle moving to the right, have the asymptotic form

\[
\psi(x) \xrightarrow{x \to +\infty} \alpha e^{ik_2x}, \quad \text{with} \quad k_2 = \sqrt{\frac{2m(E - V_0)}{\hbar}}
\]

(a particle with energy \( E \), moving in a constant potential \( V_0 \), with positive momentum \( p \).) For large negative \( x \), the particle is approximately free. The wave function has the asymptotic form

\[
\psi(x) \xrightarrow{x \to -\infty} e^{ik_1x} + \beta e^{-ik_1x}, \quad \text{with} \quad k_1 = \sqrt{\frac{2mE}{\hbar}}
\]

(a superposition of an incident wave with intensity 1 and a reflected wave) The probability current \( \vec{J} \propto (\psi \partial_x \psi^* - \psi^* \partial_x \psi) \) (see Exercise 5.1) is

- incident wave: \( |\vec{J}| \propto k_1 \)
- reflected wave: \( |\vec{J}| \propto |\beta|^2 k_1 \)
- transmitted wave: \( |\vec{J}| \propto |\alpha|^2 k_2 \).

The transmission and reflection coefficients are defined as

\[
T = \frac{|\vec{J}_t|}{\vec{J}_i} = \frac{k_2}{k_1} |\alpha|^2, \quad R = \frac{|\vec{J}_r|}{\vec{J}_i} = |\beta|^2.
\]

Since the probability is conserved, we must have \( R + T = 1 \).
EXAMPLE 6.4. Consider a particle moving along the x-axis with energy $E \geq V_0 > 0$ in a potential

$$V(x) = \begin{cases} V_0 & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases}$$

We compute the reflection coefficient. The (exact, not asymptotic!) solution of the Schrödinger equation is

$$\psi(x) = \begin{cases} \alpha e^{ik_2 x} & \text{if } x > 0, \\ e^{ik_1 x} + \beta e^{-ik_1 x} & \text{if } x < 0. \end{cases}$$

If we impose the continuity of $\psi$ and $\psi'$ at $x = 0$, we determine

$$\beta = \frac{k_1 - k_2}{k_1 + k_2}.$$ 

Therefore

$$R = |\beta|^2 = \left| \frac{k_1 - k_2}{k_1 + k_2} \right|^2.$$

Since $k_1 \neq k_2$, in general there is reflection, even if the energy $E > V_0$ is large enough for a classical particle to overcome the barrier. If $E = V_0$ ($k_2 = 0$), there is total reflection. There is total reflection even if $E < V_0$ ($k_2$ purely imaginary). Finally, if $E \gg V_0$, $R \sim V_0/4E^2 \to 0$, there is total transmission.

EXAMPLE 6.5. Consider a potential barrier: the potential energy is positive in the interval $0 < x < a$ and is zero outside this region

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

We examine the case $E > V_0$ first. In this case

$$\psi(x) = \begin{cases} e^{ik_1 x} + Ae^{-ik_1 x} & \text{if } x < 0, \\ Be^{ik_2 x} + B'e^{-ik_2 x} & \text{if } 0 < x < a, \\ Ce^{ik_1 x} & \text{if } x > 0. \end{cases}$$

Imposing the condition that $\psi$ and $\psi'$ are continuous at $x = 0$ and $x = a$ we obtain a linear systems for the coefficients $A, B, B', C$:

$$1 + A = B + B'$$

$$k_1(1 - A) = k_2(B - B')$$

$$Ce^{ik_1 a} = B e^{ik_2 a} + B'e^{-ik_2 a}$$

$$Ck_1 e^{ik_1 a} = k_2(B e^{ik_2 a} - B'e^{-ik_2 a}).$$
If we set $\kappa = k_1/k_2$, then

\[ B = \frac{2\kappa(1 + \kappa)}{(1 + \kappa)^2 - (1 - \kappa)^2 e^{2ik_2a}} \]

\[ B' = \frac{2\kappa(1 - \kappa)e^{2ik_2a}}{(1 + \kappa)^2 - (1 - \kappa)^2 e^{2ik_2a}} \]

\[ C = \frac{4\kappa e^{i(k_2-k_1)a}}{(1 + \kappa)^2 - (1 - \kappa)^2 e^{2ik_2a}}. \]

It follows that the transmission coefficient is

\[ T = |C|^2 = \frac{4k_1^2k_2^2}{4k_1^2k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(k_2a)} = \left(1 + \frac{V_0^2 \sin^2(k_2a)}{4E(E - V_0)}\right)^{-1}. \]

Note a new phenomenon: the barrier becomes transparent ($T = 1$) if $k_1a = n\pi$.

If the case $E < V_0$ we have $k_2 = i\sqrt{2m|E - V_0|/\hbar^2} = ix_2$. Using the previous formulae we find the transmission coefficient

\[ T = \frac{4k_1^2x_2^2}{4k_1^2x_2^2 + (k_1^2 + x_2^2)^2 \sinh^2(x_2a)} = \left(1 + \frac{V_0^2 \sinh^2(x_2a)}{4E(V_0 - E)}\right)^{-1}. \]

In classical mechanics the particle would not penetrate the barrier. In quantum mechanics, the solution of the Schrödinger equation shows that, contrary to classical expectations, there is a finite probability that an incident particles will be transmitted $T > 0$. This effect goes under the name of tunnel effect.

When $x_2a \to \infty$ it is easy to find the asymptotic formula

\[ T \sim \frac{16E(V_0 - E)}{V_0^2} e^{-\sqrt{2m(V_0-E)}a}. \]

**Exercise 6.6.** Find an expression for the transmission coefficient in the limit $V_0 \to +\infty$, $a \to 0$, with $V_0a = K$ constant. (This corresponds to a scattering process by a $\delta$-potential.)

**6.5. The one-dimensional crystal.** A crystal consisting of a vast number of positive ions and electrons is a paradigmatic example of many-body quantum mechanical system. It is possible to derive some of its salient features from a simplified model of a quantum particles moving in a periodic potential

\[ H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x), \quad \text{with} \quad V(x + d) = V(x), \quad x \in [0, Nd]. \]

In words, we model the crystal lattice of ions as a crystal ring with $N$ sites with a periodic potential with period $d$ (the lattice spacing).

The fact that the potential $V(x)$ is periodic of period $d$ does not mean that the eigenfunctions have to be periodic with the lattice period. The Schrödinger
equation remains unchanged if we replace $V(x)$ with $V(x+d)$, but this leaves a constant factor in its eigenfunctions undetermined, so that
\[ \psi(x+d) = \mu_1 \psi(x) \]
where $\mu$ is a constant. If we move over one more lattice space, the same argument holds true. Hence
\[ \psi(x+2d) = \mu_1 \mu_2 \psi(x). \]
If we continue $N$ times, we get back to where we were, so that
\[ \psi(x+Nd) = \mu_1 \mu_2 \cdots \mu_N \psi(x) = \psi(x) \Rightarrow \mu_1 \mu_2 \cdots \mu_N = 1. \]
Because of the symmetry of the crystal ring, it cannot make any difference at which lattice point we started, and we conclude that all the $\mu_k$'s must be equal. Hence
\[ \psi(x+md) = e^{2\pi in/N} \psi(x). \]
We have, therefore,
\[ \psi(x+md) = e^{2\pi in/N} \psi(x). \]
This statement can also be expressed in the following form
\[ (6.42) \quad \psi(x) = e^{2\pi i n x / N} u(x) = e^{ikx} u(x) \]
where $k = 2\pi n /Nd$ and $u(x)$ is periodic of period $d$. Eigenfunctions of this form are called Bloch's waves and it can be proved that they form a basis (this is known as Bloch's theorem or Floquet's theorem). Although $\psi(x)$ is not $d$-periodic, it is easy to see that the probability density $|\psi(x)|^2$ is periodic with the lattice period.

If we want to go on with the solution of the Schrödinger equation, we have to specify $V(x)$. A mathematically convenient choice is the Kronig-Penney (1931) potential (a succession of square wells potential of width $a$ and depth $v$). Considering a single period of the potential, the periodic solution $u(x)$ is
\[ (6.43) \quad u(x) = \begin{cases} 
A e^{i(\alpha-k)x} + B e^{-i(\alpha-k)x}, & \alpha^2 = \frac{2mE}{\hbar^2} \\
C e^{i(\beta-k)x} + D e^{-i(\beta-k)x}, & \beta^2 = \frac{2m(v-E)}{\hbar^2}
\end{cases} \]
with the periodic condition $u(0) = u(d)$, $u'(0) = u'(d)$, and the continuity conditions of $u(x)$ and $u'(x)$ at $x = a$. These conditions provide four linear homogeneous equations to determine the four constants $A, B, C$ and $D$. To have a nontrivial solution the determinant of the matrix coefficient of linear system must vanish. For $E < v$ we get
\[ \frac{\beta^2 - \alpha^2}{2\alpha\beta} \sinh(\beta(d-a)) \sin(\alpha a) + \cosh(\beta(d-a)) \cos(\alpha a) = \cos(ka). \]
The problem simplifies in the limit $a \to d, v \to \infty$ with $v(d-a) = K$ fixed (the potential becomes a periodic array of repulsive $\delta$-functions). In this limit, it turns
out that the condition for the existence of nontrivial solutions of the Schrödinger equation is

\[(6.44) \quad P \sin(\alpha d) + \cos(\alpha d) = \cos(kd), \quad \text{with} \quad \alpha = \sqrt{\frac{2mE}{\hbar^2}}, \quad \text{and} \quad P = \frac{Km}{\alpha \hbar^2}.\]

The left-hand side (a function of \(E\)) can assume values larger than one, the right-hand side can not (and it does not depend on \(E\)). Therefore, the relation can be satisfied only for those values of \(E = \frac{\hbar^2(\alpha d)^2}{2md^2}\) for which the left-hand side remains between \(-1\) and \(+1\). In other words, in a crystal only certain energy bands are allowed. (A particle trapped in a potential well can have only discrete energy values. In a periodic potential these discrete energies are replaced by energy bands.)

It is customary to present the results as a plot of the allowed energies \(E\) against \(kd\). Note that, the larger \(K\), i.e., the higher the potential separating the regions of zero potential, the narrower the energy bands. For \(K \to \infty\) the crystal reduces to a series of independent quantum boxes, and the energy bands become discrete energy levels \(E = \frac{\hbar^2 k^2}{2md^2}\). It is instructive to compare the results with that of a free electron. We can obtain this result by letting \(v = 0\) (i.e. \(K = 0\), so that \(E = \frac{\hbar^2 k^2}{2m}\)).

**Exercise 6.7.** Derive the relation (6.44) by considering the one-dimensional Schrödinger equation with potential

\[V(s) = K \sum_{n \in \mathbb{Z}} \delta(x - nd).\]

### 6.6. The hydrogen atom.

We consider the hydrogen atom as a system of an electron in three dimension moving in the attractive Coulomb potential generated by a fixed nucleus. Let us, for generality, ascribe to the nucleus charge \(+Z > 0\) and unit charge \(-1\) to the electron. The classical potential energy of the system is \(-\frac{Z}{|x|}\), in which \(|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}\) is the distance between the electron and the nucleus. For convenience we can assume that the nucleus is fixed at the origin of the coordinate system. The classical Hamiltonian is

\[(6.45) \quad H(p, x) = \frac{p^2}{2m} - \frac{Z}{|x|}, \quad p = (p_1, p_2, p_3) \in \mathbb{R}^3, \quad x = (x_1, x_2, x_3) \in \mathbb{R}^3.\]

The corresponding Schrödinger operator, in suitable units, is

\[(6.46) \quad H = -\Delta - \frac{Z}{|x|},\]

and the time-independent Schrödinger equation is

\[(6.47) \quad -\Delta \psi(x) - \frac{Z}{|x|} \psi(x) = E \psi(x).\]
The spherical symmetry of the problem suggests to use spherical coordinates

\[
\begin{align*}
    x_1 &= r \sin \theta \cos \varphi \\
    x_2 &= r \sin \theta \sin \varphi \quad \text{with } r > 0, -\pi \leq \theta \leq \pi, \text{ and } 0 \leq \varphi < 2\pi. \\
    x_3 &= r \cos \varphi
\end{align*}
\] (6.48)

\[
\begin{align*}
    x_1 &= r \sin \theta \cos \varphi \\
    x_2 &= r \sin \theta \sin \varphi \\
    x_3 &= r \cos \varphi
\end{align*}
\]

**EXERCISE 6.8.** Show that the Laplacian in spherical coordinates is

\[
\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}.
\] (6.49)

The Schrödinger equation in spherical coordinates for a generic symmetric potential \( V(x) = V(|x|) \) is

\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \varphi^2} + (E - V(r)) \psi = 0.
\]

The above partial differential equation can be separated into ordinary equations using the ansatz

\[
\psi(r, \theta, \varphi) = R(r) \Theta(\theta) \Phi(\varphi).
\]

On introducing this into the equation and dividing by \( R \Theta \Phi \) we get

\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial R}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \Theta}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \varphi^2} + \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) \left( E + \frac{Z}{r} \right) + \frac{1}{\Theta \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \Theta}{\partial \theta} \right) - \frac{m^2}{r^2 \sin \theta} = 0.
\]

On multiplying through by \( \sin^2 \theta \), the remaining part of the third term, \( \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \varphi^2} \) which could only be a function of the independent variable \( \varphi \), is seen to be equal to terms independent of \( \varphi \). Hence this term must be equal to a constant, which we call \(- m^2\):

\[
\frac{\partial^2 \Phi}{\partial \varphi^2} + m^2 \Phi = 0.
\]

The equation in \( r \) and \( \theta \) then can be written as

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial R}{\partial r} \right) + \frac{1}{\Theta \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \Theta}{\partial \theta} \right) - \frac{m^2}{r} \frac{1}{\sin \theta} = 0.
\]

If we set the \( \theta \) terms equal to the constant \(- l(l + 1)\), (and consequently the \( r \) terms equal to \( l(l + 1) \)) we eventually obtain

\[
\begin{align*}
    - \frac{\partial^2 \Phi}{\partial \varphi^2} &= m^2 \Phi \\
    - \frac{1}{\sin \theta} \left( \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{m^2}{\sin \theta} \right) \Theta &= l(l + 1) \Theta \\
    \left( - \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial R}{\partial r} \right) + \frac{l(l + 1)}{r^2} + V(r) \right) R &= ER(r).
\end{align*}
\]
These equations are now to be solved in order to determine the allowed values of the energy. The sequence of solution is the following: We first find that the $\Phi$-equation possesses acceptable solutions only for certain values of the parameter $m$. Introducing these in the $\Theta$-equation, we find that it then possesses acceptable solutions only for certain values of $l(l + 1)$. (The form $m^2$ and $l(l + 1)$ chosen for the constants is for convenience later on.) Finally, we introduce these values of $l$ in the $R$-equation and find that it possesses acceptable solutions only for certain values of $E$. These are the energy levels for the stationary states of the system.

**REMARK.** Note that the above considerations are valid for a generic radially symmetric potential $V(x) = V(|x|)$.

6.6.1. The $\Phi$-equation. The solutions of the $\Phi$-equation are

\begin{equation}
\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}} e^{im\varphi}.
\end{equation}

In order for the function to be periodic, $m$ must be an integer, i.e., $m \in \mathbb{Z}$. The constant $m$ is called **magnetic number**.

The factor $1/\sqrt{2\pi}$ is chosen in order to satisfy the normalisation condition. Actually, they form an orthonormal system

\begin{equation}
\int_0^{2\pi} \Phi_m^*(\varphi)\Phi_n(\varphi) = \delta_{nm}.
\end{equation}

It may pointed out that except for the lowest eigenvalue $m^2 = 0$, all eigenvalues are twice degenerate (the two functions $\Phi_m$ and $\Phi_{-m}$ satisfy the same differential equation if $m > 0$).

6.6.2. The $\Theta$-equation. Now we turn to the equation for $\Theta(\theta)$. In order to solve the equation, it is convenient for us to introduce the new independent variable $t = \cos \theta \in [-1, 1]$, and consider the function $P(t) = \Theta(\theta(t))$. The equation transforms to the somewhat simpler equation

\[-\frac{d}{dt}(1 - t^2) \frac{dP(t)}{dt} - \frac{m^2}{1 - t^2} P(t) = l(l + 1)P(t).\]

For $l = 0, 1, 2, \ldots$, the above equation is solved by the associated Legendre polynomials

\begin{equation}
P_l^m(t) = (-1)^m (1 - t^2)^{m/2} \frac{d^m}{dt^m} \frac{1}{2^m m!} \frac{d^l}{dt^l} (x^2 - 1)^l, \quad l \in \mathbb{Z}_+, |m| \leq l.
\end{equation}

$l$ is called **azimuthal quantum number**.

6.6.3. The $R$-equation. The equation in $r$ is

\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial R}{\partial r} \right) + \left( -\frac{l(l + 1)}{r^2} + E + V(r) \right) R = 0.
\]
For generic $V(r)$ the problem is intractable. The special case $V(r) = -Z/r$ is solvable. In this case the $R$-equation is

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial R}{\partial r} \right) + \left( -\frac{l(l+1)}{r^2} + E - \frac{Z}{r} \right) R = 0.$$ 

Let us consider the case $E < 0$. Introducing the symbols

$$\alpha^2 = -E, \quad \lambda = \frac{Z}{2E},$$

and the new independent variable $\rho = 2\alpha r > 0$, the wave equation for $S(\rho) = R(\rho(r))$ becomes

$$\frac{1}{\rho^2} \frac{\partial}{\partial \rho} \left( \rho^2 \frac{\partial S}{\partial \rho} \right) + \left( -\frac{1}{4} - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} \right) S = 0.$$ 

As in the treatment of the harmonic oscillator, we first discuss the asymptotic equation. For $\rho$ large, the equation approaches the form

$$\frac{1}{\rho^2} \frac{\partial}{\partial \rho} \left( \rho^2 \frac{\partial S}{\partial \rho} \right) = \frac{1}{4} S$$

the acceptable solution of which is $S = e^{-\rho/2}$.

We now assume that the solution of the complete equation in the whole region $\rho > 0$ has the form

$$S(\rho) = e^{-\rho/2} \rho^s L(\rho),$$

in which $L(\rho)$ is a power series in $\rho$ with a non-vanishing constant term

$$L(\rho) = \sum_k a_k \rho^k, \quad a_0 \neq 0.$$ 

After some calculations we find the conditions on $s$,

$$s = l$$

and a differential equation for $L(\rho)$:

$$\rho L'' + (2(l + 1) - \rho) L' + (\lambda - l - 1) L = 0.$$ 

We now introduce the series representation of $L$ and obtain the recursion

$$(2(k + 1)(l + 1) + k(k + 1)) a_{k+1} + (\lambda - l - 1 - k) a_k.$$ 

It can be shown by an argument similar to that used for the harmonic oscillator that for any values of $\lambda$ and $l$ the series whose coefficients are determined by this formula leads to a function $S$ unacceptable as a wavefunction, unless it breaks off. The condition that it breaks off after the term in $\rho^{n'}$ is

$$\lambda - l - 1 - n' = 0$$

or

$$\lambda = n, \quad \text{where } n = n' + l + 1.$$
\( n' \) is called **radial quantum number**, and can assume values \( n' = 0, 1, 2, \ldots, n \) is called **total quantum number**.

**6.6.4. Energy levels and eigenfunctions.** Summarising, if we introduce the quantum numbers \( n, l, \) and \( m \) as subscript (using \( n \) in preference of \( n' \)), the wavefunctions we have found as acceptable solutions may be written as

\[
\psi_{nlm}(r, \theta, \varphi) = R_{nl}(r)\Theta_{lm}(\theta)\Phi_m(\varphi),
\]

with

\[
\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}} e^{im\varphi},
\]

\[
\Theta_{lm}(\theta) = \sqrt{\frac{(2l + 1)!}{2(l + |m|)!}} P^{|m|}_l(\cos \theta),
\]

\[
R_{nl}(r) = -\sqrt{\frac{(n - l - 1)!}{2n^4}} e^{-\frac{Ze}{2n}} \left( \frac{Ze}{n} \right)^l L_{n+l}^{2l+1} \left( \frac{Ze}{n} \right).
\]

The wavefunctions corresponding to distinct sets of values for \( n, l, \) and \( m \) are independent and the energy value corresponding to \( \psi_{nlm} \) is

\[
E_n = -\left( \frac{Z}{2n} \right)^2.
\]

In particular, the lowest eigenvalue \( E_1 = -Z^2/4 \) is simple and the corresponding eigenfunction \( \psi_{nlm}(r, \theta, \varphi) = \sqrt{\frac{Z^3}{2}} e^{-Ze/2} \) is positive.

The allowed values of these quantum numbers we have determined to be

\[
m = 0, \pm 1, \pm 2, \ldots
\]
\[
l = |m|, |m| + 1, |m| + 2, \ldots
\]
\[
n = l + 1, l + 2, l + 3, \ldots
\]

This we may rewrite as

- **total quantum number** \( n = 1, 2, 3, \ldots \)
- **azimuthal quantum number** \( l = 0, 1, 2, \ldots, n - 1 \)
- **magnetic quantum number** \( m = -l, \ldots, -1, 0, +1, \ldots, +l \)

There are consequently \( 2l + 1 \) independent wave functions with given values of \( n \) and \( l \), and \( n^2 \) independent wave functions with a given value of \( n \), that is, with the same energy value.
7. Matrix Mechanics

Heisenberg formulated and successfully attacked the problem of calculating values of the frequencies and intensities of the spectral lines which a system could emit or absorb; that is, of the energy levels and the electric-moment integrals which we have been discussing. He did not use wave functions and wave equations, but instead developed a formal mathematical method for calculating values of these quantities.

Heisenberg invented the new type of algebra as he needed it; it was immediately pointed out by Born and Jordan, however, that in his new quantum mechanics Heisenberg was making use of quantities called matrices, and that his newly invented operations were those of matrix algebra which had been already discussed by mathematicians.

7.1. Matrices, their relation to wavefunctions, and the rules of matrix algebra. Let us consider a set of orthogonal wavefunctions $\Psi_0, \Psi_1, \ldots, \Psi_n, \ldots$ and a dynamical quantity $f(q_i, p_i)$. The corresponding operator is $f_{op} = f \left( q_i, \frac{\hbar}{i} \frac{\partial}{\partial q_i} \right)$.

In the foregoing sections we have often made use of integrals such as

\[ f_{mn} = \int \Psi^* f_{op} \Psi; \]  \hspace{1cm} \text{(7.1)}

for example, we have given $f_{nn}$ the physical interpretation of the average value of $f$ when the system is in the $n^{th}$ stationary state. Let us now arrange the numbers $f_{mn}$ (the values of the integrals) in a square array $F = (f_{mn})$ ordered according to $m$ and $n$.

We can construct similar arrays $G = (g_{mn}), H = (h_{mn})$ etc. for other dynamical quantities. It is found that the symbols $F, G, H$ etc. representing such arrays can be manipulated by an algebra closely related to ordinary algebra, differing from it mainly in the process of multiplication. The rules of this algebra can be easily derived from the properties of wave functions, which we already know.

Now let us derive some rules of the new algebra. For example, the sum of two such arrays is an array each of whose elements is the sum of the corresponding elements of the two arrays

\[ \{f + g\}_{mn} = \{f + g\}_{mn} = f_{mn} + g_{mn}. \]  \hspace{1cm} \text{(7.2)}

The addition of arrays is commutative: $F + G = G + F$.

On the other hand, multiplication is not commutative: the product $FG$ is not necessarily equal to the product $GF$. Let us evaluate the $mn^{th}$ element of the array $FG$. It is

\[ \{fg\}_{mn} = \int \Psi^*_m f_{op} g_{op} \Psi_n. \]
Now we can express the quantity $g_{op}\Psi_n$ in terms of the functions $\Psi_k$ with constant coefficients, obtaining

$$g_{op}\Psi_n = \sum_k g_{kn}\Psi_k.$$  

That the coefficients are the quantities $g_{kn}$ is seen on multiplying by $\Psi^*_k$ and integrating. Introducing this in the integral for $\{fg\}_{mn}$ we obtain

$$\{fg\}_{mn} = \sum_k \int \Psi^*_m f_{op}\Psi_k g_{kn};$$

since $\int \Psi^*_m f_{op}\Psi_k = f_{mk}$, this becomes

$$\{fg\}_{mn} = \sum_k f_{mk}g_{kn}. \tag{7.3}$$

We see that the arrays $F$, $G$, etc. can be manipulated according to the rules of matrices.

The non-commutative nature of the multiplication of matrices is of great importance in matrix mechanics.

7.2. Diagonal matrices and their physical interpretation. A diagonal matrix is a matrix whose elements $f_{mn}$ are all zero except those with $m = n$. The unit matrix, 1, is a special kind of diagonal matrix, all the diagonal elements being equal to unity. A constant matrix is equal to a constant times the unit matrix. Application of the rule for matrix multiplication shows that the square (or any power) of a diagonal matrix is also a diagonal matrix, its diagonal elements being the squares (or other powers) of the corresponding elements of the original matrix.

In discussing the physical interpretation of the wave equation, we saw that the fundamental postulate regarding physical interpretation requires a dynamical quantity $f$ to have a definite value for a system in the state represented by the wave function $\Psi_n$ only when $f_{nn}$ is qual to $(f_{nn})^r,$ for all values of $r$. We can now express this in terms of matrices: if the dynamical quantity $f$ is represented by a diagonal matrix $F$, then this dynamical quantity has the definite value $f_{nn}$ for the state corresponding to the wavefunction $\Psi_n$ of the set $\Psi_0, \Psi_1, \ldots$.

**Example 7.1.** The solutions $\Psi_n = \psi_n(x)e^{-\frac{i}{\hbar}E_n t}$ of the Schrödinger equation correspond to a diagonal energy matrix $H = \text{diag}(E_0, E_1, E_2, \ldots)$, so that, the system in a physical condition (i.e. a state) represented by one of these wavefunctions has a fixed value of the total energy.

In the case of a system with one degree of freedom no other quantity (except functions of $H$, such as $H^2$) is represented by a diagonal matrix; with more degrees of freedom there are other diagonal matrices.

**Example 7.2.** The surface-harmonic wavefunctions $\Theta_{lm}(\theta)\Phi_m(\varphi)$ for the hydrogen atom and any other spherical symmetric system make the matrices for
the square of the total angular momentum and the \( z \) component of the angular momentum diagonal. These quantities thus have definite values for \( \Theta_{lm}(\theta) \Phi_m(\varphi) \).

7.3. Matrix elements and their physical interpretation. The principle of interference, which applies to all quantum systems (photons, electrons, atoms, . . . ), suggests that transition probabilities \( P_{nm} \) from a state \( \psi_n \) to a state \( \psi_m \) are the squares of transition amplitudes, and that the latter combine in a linear fashion. Likewise, for the observable \( f \) we interpret the matrix element \( f_{nm} \) as transition amplitudes. We will not discuss their use in this module.

7.4. The Heisenberg Uncertainty Relations. Given a state \( |\psi\rangle \), one can give the probabilities for the possible outcomes of a measurement of an observable \( A \). The probability distribution has a mean or expectation value

\[
\langle A \rangle = \langle \psi | A | \psi \rangle
\]

and an uncertainty or deviation about this mean

\[
\Delta A = \sqrt{\langle \psi | (A - \langle A \rangle)^2 | \psi \rangle}.
\]

There are states for which \( \Delta A = 0 \), and these are the eigenstates of \( A \).

If we consider two Hermitian operators \( A \) and \( B \), they will have some uncertainties \( \Delta A \) and \( \Delta B \) in a given state. The Heisenberg uncertainty relations provide a lower bound on the product of uncertainties \( \Delta A \Delta B \). Generally the lower bound depends not only on the operators but also on the state. Of particular interest are those cases in which the lower bounds is independent of the state.

Let \( A \) and \( B \) be two Hermitian operators, with commutator

\[
[A, B] = iC.
\]

It is easy to verify that \( C \) is also Hermitian. Fix a state \( |\psi\rangle \). Note that the pair \( A' = A - \langle A \rangle \) and \( B' = B - \langle B \rangle \) has the same commutator as \( A \) and \( B \) (verify this). Then, by Schwartz inequality:

\[
(\Delta A)^2(\Delta B)^2 = \langle \psi | A'^2 | \psi \rangle \langle \psi | B'^2 | \psi \rangle = \langle A' \psi | A' \psi \rangle \langle B' \psi | B' \psi \rangle \geq |\langle A' \psi | B' \psi \rangle|^2.
\]

We can write the above inequality in terms of commutators and anticommutators

\[
(\Delta A)^2(\Delta B)^2 \geq |\langle \psi | A'B' | \psi \rangle|^2 = |\langle \psi | \frac{1}{2} \{A', B'\} + \frac{1}{2} [A', B'] | \psi \rangle|^2.
\]

Since \( [A', B'] = iC \) has pure imaginary expectation value, and \( \{A', B'\} \) has real expectation value, we get

\[
(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4} |\langle \psi | \{A', B'\} | \psi \rangle|^2 + \frac{1}{4} |\langle \psi | C | \psi \rangle|^2.
\]

(7.4)
This is the general uncertainty relation between any two Hermitian operators and is evidently state dependent. Consider now canonically conjugate operators, for which \( C = \hbar \). In this case,

\[
(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4} \left| \langle \psi | \{ A', B' \} | \psi \rangle \right|^2 + \frac{1}{4} \hbar^2 \geq \frac{1}{4} \hbar^2,
\]

or

\[
(7.5) \quad \Delta A \Delta B \geq \hbar/2.
\]

8. Mathematical Foundations of Quantum Mechanics

In classical probability theory one assumes that all the events concerning a statistical experiment form a Boolean \( \sigma \)-algebra and defines a probability measure as a completely additive non-negative function which assigns the value unity for the identity element of the \( \sigma \)-algebra. Typically, the \( \sigma \)-algebra is the Borel \( \sigma \)-algebra \( \mathcal{B}_X \) of a nice topological space \( X \). Under very general conditions it turns out that all probability measures on \( \mathcal{B}_X \) constitute a convex set whose extreme points are degenerate probability measures.

We observe that \( \mathcal{B}_X \) admits a null element, namely \( \emptyset \), a unit element, namely \( X \), a partial order \( \subset \) and operations union (\( \cup \)), intersection (\( \cap \)) and complementation (\( ^c \)). Quantum theory can be thought of as a non-commutative version of probability theory where the \( \sigma \)-algebra \( \mathcal{B}_X \) of events is replaced by a lattice of projections on a Hilbert space \( \mathcal{H} \). Most of the computations of quantum mechanics is done in such a lattice.

Let \( \mathcal{H} \) be a real or complex separable Hilbert Space and let \( P(\mathcal{H}) \) denote the set of all orthogonal projection operators on \( \mathcal{H} \), where \( 0 \) denotes the null operator and \( I \) denotes the identity operator. If \( P_1, P_2 \in P(\mathcal{H}) \) we say that \( P_1 \leq P_2 \) if the range of \( P_1 \) is contained in the range of \( P_2 \). Then \( \leq \) makes \( P(\mathcal{H}) \) a partially ordered set. For any linear operator \( A \) on \( \mathcal{H} \) let \( R(A) \) denote its range. For two orthogonal projections \( P_1, P_2 \), let \( P_1 \vee P_2 \) be the orthogonal projection on the closed linear span of the subspaces \( R(P_1) \cup R(P_2) \). Let \( P_1 \wedge P_2 \) be the orthogonal projection on \( R(P_1) \cap R(P_2) \). For any \( P \in \mathcal{L}(\mathcal{H}) \), \( I - P \) is the orthogonal projection on the orthogonal complement \( R(P)^\perp \) of the range of \( P \).

We may compare \( 0, I, \leq, \vee, \wedge \) and the map \( P \mapsto I - P \) in \( \mathcal{L}(\mathcal{H}) \) with \( \emptyset, X, \subset, \cup, \cap \) and complementation \( A \mapsto X \setminus A \) of standard probability theory on \( \mathcal{B}_X \). The chief distinction lies in the fact that \( \cup \) distributes with \( \cap \) but \( \vee \) need not distribute with \( \wedge \).

8.1. The postulates of quantum theory. A consistent framework of quantum theory can be introduced through a sequence of postulates.

P1 The state of a physical system is described by a normalised vector \( \psi \) in a separable complex Hilbert space \( \mathcal{H} \).
P2 To each (real-valued) observable corresponds a self-adjoint operator $A$ on $\mathcal{H}$, or on a suitable subspace of $\mathcal{H}$ (in fact, the observables of a physical system are elements of a non-commutative $C^*$-algebra).

P3 If the observable $A$ is measured, then the possible outcomes of the measurement are the eigenvalues of $A$.

P4 Suppose we measure $A$ on the state $\psi$. Then, the probability that the observed value is $\lambda \in \mathbb{R}$ is

$$\sum_{i: \lambda_i = \lambda} |\langle x_i, \psi \rangle|^2,$$

where $\{x_i\}$ are the normalised eigenvectors of $A$ and $\{\lambda_i\}$ the corresponding eigenvalues.

P5 The state of a (closed) system evolves in time as

$$\psi(t) = U(t) \psi_0$$

for some strongly continuous one-parameter unitary group that only depends on the system (and not on the state).

8.2. Hilbert spaces. Suppose $\mathcal{H}$ is a complex vector space. A map $\langle \cdot | \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is called a sesquilinear form if it is conjugate linear in the first argument and linear in the second. A positive definite sesquilinear form is called a scalar product. Associated with every scalar product is a norm

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle}.$$ 

The triangle inequality follows from the Cauchy-Schwarz inequality:

$$|\langle \psi | \varphi \rangle| \leq \|\psi\| \|\varphi\|$$

with equality if and only if $\psi$ and $\varphi$ are parallel. If $\mathcal{H}$ is complete with respect to the above norm, it is called a Hilbert space. It is no restriction to assume that $\mathcal{H}$ is complete since one can easily replace it by its completion.

Example 8.1. $\mathcal{H} = \mathbb{C}^n$ with the usual scalar product

$$\langle \psi | \varphi \rangle = \sum_{j=1}^n \psi_j^* \varphi_j$$

is a (finite dimensional) Hilbert space. Similarly, the set of all square summable sequences $\ell^2(\mathbb{N})$ is a Hilbert space with scalar product

$$\langle \psi | \varphi \rangle = \sum_{j \in \mathbb{N}} \psi_j^* \varphi_j.$$

Example 8.2. The space $\mathcal{H} = L^2(M, d\mu)$ is a Hilbert space with scalar product given by

$$\langle \psi | \varphi \rangle = \int_M \psi^*(x) \varphi(x) d\mu(x).$$
(Note that the Example 8.1 is a special case of the this one; take $M = \mathbb{R}$ and $\mu$ a sum of Dirac measures.)

**Remark.** Even though the elements of $L^2(M, d\mu)$ are, strictly speaking, equivalence classes of functions, we will still call them functions for notational convenience. However, note that for $f \in L^2(M, d\mu)$ the value $f(x)$ is not well-defined (unless there is a continuous representative and different continuous functions are in different equivalence classes, e.g., in the case of Lebesgue measure).

A vector $\psi \in \mathcal{H}$ is called normalized or a unit vector if $\|\psi\| = 1$. Two vectors $\psi, \varphi \in \mathcal{H}$ are called **orthogonal** ($\psi \perp \varphi$) if $\langle \psi | \varphi \rangle = 0$.

Suppose $\psi$ is a unit vector. Then the projection of $\psi$ in the direction of $\varphi$ is given by and $\psi_{\parallel}$ defined via

$$\psi_{\parallel} = \langle \varphi | \psi \rangle \varphi,$$

and $\psi_{\perp}$ defined via

$$\psi_{\perp} = \psi - \langle \varphi | \psi \rangle \varphi,$$

is perpendicular to $\varphi$.

These results can also be generalised to more than one vector. A set of vectors $\{\varphi_j\}$ is called an **orthonormal set** if $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$.

**Exercise 8.1.** Show that every orthonormal set is linearly independent.

**Theorem 8.1** (Pythagorean theorem). Suppose $\{\varphi_j\}_{j=0}^n$ is an orthonormal set. Then every $\psi \in \mathcal{H}$ can be written as

$$\psi = \psi_{\parallel} + \psi_{\perp}, \quad \psi_{\parallel} = \sum_{j=0}^n \langle \varphi_j | \psi \rangle \varphi_j,$$

where $\psi_{\parallel}$ and $\psi_{\perp}$ are orthogonal. Moreover, $\langle \varphi_j | \psi_{\perp} \rangle = 0$ for all $1 \leq j \leq n$. In particular

$$\|\psi\|^2 = \sum_{j=0}^n |\langle \varphi_j | \psi \rangle|^2 + \|\psi_{\perp}\|^2.$$

Moreover, every $\hat{\psi}$ in the span of $\{\varphi_j\}_{j=0}^n$ satisfies

$$\|\psi - \hat{\psi}\| \geq \|\psi_{\perp}\|$$

with equality holding if and only if $\hat{\psi} = \psi_{\parallel}$. In other words, $\psi_{\parallel}$ is uniquely determined as the vector in the span of $\{\varphi_j\}_{j=0}^n$ closest to $\psi$.

From (8.9) we obtain **Bessel’s inequality**

$$\sum_{j=0}^n |\langle \varphi_j | \psi \rangle|^2 \leq \|\psi\|^2.$$
with equality holding if and only if \( \psi \) lies in the span of \( \{ \varphi_j \}_{j=0}^n \).

A scalar product can be recovered from its norm by virtue of the **polarization identity**

\[
(8.12) \quad \langle \varphi | \psi \rangle = \frac{1}{4} \left( \| \varphi + \psi \|_2^2 - \| \varphi - \psi \|_2^2 + i \| \varphi - i \psi \|_2^2 - i \| \varphi + i \psi \|_2^2 \right).
\]

A bijective linear operator \( U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2) \) is called **unitary** if \( U \) preserves scalar products:

\[
(8.13) \quad \langle U\varphi | U\psi \rangle_2 = \langle \varphi | \psi \rangle_1, \quad \text{for all } \varphi, \psi \in \mathcal{H}_1.
\]

By the polarization identity, this is the case if and only if \( U \) preserves norms:

\[
\| U\psi \|_2 = \| \psi \|_1 \text{ for all } \psi \in \mathcal{H}_1.
\]

The two Hilbert spaces \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) are called **unitarily equivalent** in this case.

**Exercise 8.2.** The shift operator

\[
S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N}), \quad (a_1, a_2, a_3, \ldots) \mapsto (0, a_1, a_2, \ldots)
\]
satisfies \( \| Sa \| = \| a \| \). Is it unitary?

By continuity of the scalar product we see that Theorem 8.1 can be generalised to arbitrary orthonormal sets \( \{ \varphi_j \}_{j \in J} \). Note that from Bessel’s inequality, it follows that the map \( \psi \mapsto \| \psi \| \) is continuous.

An orthonormal set which is not a proper subset of any other orthonormal set is called an **orthonormal basis** due to the following result.

**Theorem 8.2.** For an orthonormal set \( \{ \varphi_j \}_{j \in J} \), the following conditions are equivalent:

(i) \( \{ \varphi_j \}_{j \in J} \) is a maximal orthonormal set;

(ii) For every \( \psi \in \mathcal{H} \) we have

\[
(8.14) \quad \sum_{j \in J} \langle \varphi_j | \psi \rangle \varphi_j;
\]

(iii) For every \( \psi \in \mathcal{H} \) we have **Parseval’s relation**

\[
(8.15) \quad \| \psi \|_2^2 = \sum_{j \in J} | \langle \varphi_j | \psi \rangle |^2;
\]

(iv) \( \langle \varphi_j | \psi \rangle = 0 \) for all \( j \in J \) implies \( \psi = 0 \).

**Example 8.3.** The set of functions

\[
\varphi_m(x) = \frac{1}{\sqrt{2\pi}} e^{imx}, \quad m \in \mathbb{Z},
\]

forms an orthonormal basis for \( \mathcal{H} = L^2(0, 2\pi) \). The corresponding orthogonal expansion is just the ordinary Fourier series.
A Hilbert space is separable if and only if there is a countable orthonormal basis. In fact, if $\mathcal{H}$ is separable, then there exists a countable total set $\{\psi_j\}_{j=0}^N$. Here $N \in \mathbb{N}$ if $\mathcal{H}$ is finite dimensional and $N = \infty$ otherwise. After throwing away some vectors, we can assume that $\psi_{n+1}$ cannot be expressed as a linear combination of the vectors $\psi_0, \ldots, \psi_n$. Now we can construct an orthonormal basis as follows: we begin by normalising $\psi_0$,

$$\varphi_0 = \frac{\psi_0}{\|\psi_0\|}.\quad (8.17)$$

Next we take $\psi_1$ and remove the component parallel to $\varphi_0$ and normalise again:

$$\varphi_1 = \frac{\psi_1 - \langle \varphi_0 | \psi_1 \rangle \varphi_0}{\|\psi_1 - \langle \varphi_0 | \psi_1 \rangle \varphi_0\|}.\quad (8.18)$$

Proceeding like this, we define recursively

$$\varphi_n = \frac{\psi_n - \sum_{j=0}^{n-1} \langle \varphi_j | \psi_1 \rangle \varphi_j}{\psi_n - \sum_{j=0}^{n-1} \langle \varphi_j | \psi_1 \rangle \varphi_j}.\quad (8.19)$$

This procedure is known as **Gram-Schmidt orthogonalization**. The result of the procedure is the orthonormal basis $\{\varphi\}_{j=0}^N$.

**Exercise 8.3.** In $L^2(-1, 1)$, we can orthogonalise the monomials $f_n(x) = x^n$. The resulting polynomials are up to a normalisation equal to the Legendre polynomials $\varphi_n(x) = \sqrt{n + 1}/2P_n(x),\quad (8.20)$

$$P(x) = 1, \quad P(x) = x, \quad P(x) = \frac{1}{2} (3x^2 - 1), \ldots$$

**Theorem 8.3.** If $\mathcal{H}$ is separable, then every orthonormal basis is countable.

In quantum theory it is assumed that the Hilbert space associate to a quantum system is separable.

In particular, it can be shown that $L^2(M, d\mu)$ is separable. Moreover, it turns out that, up to unitary equivalence, there is only one separable infinite dimensional Hilbert space: Let $\mathcal{H}$ be an infinite dimensional Hilbert space and let $\{\varphi_j\}_{j \in \mathbb{N}}$ be any orthogonal basis. Then the map $U : \mathcal{H} \to \ell^2(\mathbb{N})$, $\psi \mapsto (\langle \varphi_j | \psi \rangle)_{j \in \mathbb{N}}$ is unitary. In particular,

**Theorem 8.4.** Any separable infinite dimensional Hilbert space is unitarily equivalent to $\ell^2(\mathbb{N})$.

**Exercise 8.4.** Let $\{\varphi_j\}$ be some orthonormal basis. Show that a bounded linear operator $A$ is uniquely determined by its matrix elements $A_{jk} = \langle \varphi_j | A \varphi_k \rangle$ with respect to this basis.
8.3. The projection theorem and the Riesz lemma. Let $M \subset \mathcal{H}$ be a subset. Then $M^\perp = \{ \psi \in \mathcal{H} : \langle \varphi | \psi \rangle = 0, \forall \varphi \in M \}$ is called the orthogonal complement of $M$. By continuity of the scalar product it follows that $M^\perp$ is a closed linear subspace and by linearity that $(\text{span}(M))^\perp = M^\perp$. For example, we have $\mathcal{H}^\perp = \{0\}$ since every vector in $\mathcal{H}^\perp$ must be in particular orthogonal to all vectors in some orthonormal basis.

**Remark.** Note that if $M \subset \mathcal{H}$ is closed, it is a Hilbert space and has an orthonormal basis $\{ \varphi_j \}_{j \in J}$.

**Theorem 8.5 (Projection theorem).** Let $M$ be a closed linear subspace of a Hilbert space $\mathcal{H}$. Then every $\psi \in \mathcal{H}$ can be uniquely written as $\psi = \psi_\parallel + \psi_\perp$ with $\psi_\parallel \in M$ and $\psi_\perp \in M^\perp$. One writes

\begin{equation}
M \oplus M^\perp = \mathcal{H}
\end{equation}

in this situation.

Theorem 8.5 implies that to every $\psi \in \mathcal{H}$ we can assign a unique vector $\psi_\parallel \in M$ closest to $\psi$. The rest, $\psi - \psi_\parallel$, lies in $M^\perp$. The operator $P_M$ defined as $P_M \psi = \psi_\parallel$ is called the orthogonal projection corresponding to $M$. Note that

\begin{equation}
P_M^2 = P_M, \quad \text{and} \quad \langle P_M \psi | \varphi \rangle = \langle P_M \varphi | \psi \rangle.
\end{equation}

Clearly, $P_M \psi = \psi - P_M \psi = (I - P_M) \psi$.

Finally we turn to linear functionals, that is, to linear operators $\ell : \mathcal{H} \to \mathbb{C}$. By the Cauchy-Schwarz inequality we know that $\ell : \psi \mapsto \langle \varphi | \psi \rangle$ is a bounded linear functional (with norm $||\varphi||$). It turns out that, in a Hilbert space, every bounded linear functional can be written in this way.

**Theorem 8.6 (Riesz lemma).** Suppose $\ell$ is a bounded linear functional on a Hilbert space $\mathcal{H}$. Then there is a unique vector $\varphi \in \mathcal{H}$ such that $\ell(\psi) = \langle \varphi | \psi \rangle$ for all $\psi \in \mathcal{H}$. In other words, a Hilbert space is equivalent to its own dual space $\mathcal{H}^\ast \simeq \mathcal{H}$ via the map $\varphi \mapsto \ell_\varphi = \langle \varphi | \cdot \rangle$ which is a conjugate linear isometric bijection between $\mathcal{H}$ and $\mathcal{H}^\ast$.

**Proof.** If $\ell = 0$ we can choose $\varphi = 0$. Otherwise $\text{Ker}(\ell) = \{ \psi \in \mathcal{H} : \ell(\psi) = 0 \}$ is a proper subspace of $\mathcal{H}$ and we can find a unit vector $\tilde{\varphi} \in \text{Ker}(\ell)^\perp$. For every $\psi \in \mathcal{H}$, $\ell(\tilde{\varphi}) \psi - \ell(\psi) \tilde{\varphi} \in \text{Ker}(\ell)$ and hence

\[0 = \langle \tilde{\varphi} | \ell(\tilde{\varphi}) \psi - \ell(\psi) \tilde{\varphi} \rangle = \ell(\tilde{\varphi}) \langle \tilde{\varphi} | \psi \rangle - \ell(\psi) \langle \tilde{\varphi} | \tilde{\varphi} \rangle.
\]

Therefore we can choose $\varphi = \ell(\tilde{\varphi})^{-1} \tilde{\varphi}$. To see uniqueness, let $\varphi_1, \varphi_2$ be two such vectors. Then $\langle \varphi_1 - \varphi_2 | \psi \rangle = \langle \varphi_1 | \psi \rangle - \langle \varphi_2 | \psi \rangle = \ell(\psi) - \ell(\psi) = 0$ for all $\psi \in \mathcal{H}$. Therefore $\varphi_1 - \varphi_2 \in \mathcal{H}^\perp = \{0\}$. \qed
In view of the isomorphism $\mathcal{H}^* \simeq \mathcal{H}$, one often uses Dirac’s notation and denotes vectors $\psi \in \mathcal{H}$ by the ket symbol $|\psi\rangle$ and linear functionals $\ell_\varphi = \langle \varphi | \cdot \rangle \in \mathcal{H}^*$ using the bra symbol $\langle \varphi |$. The contraction $\ell_\varphi(\psi)$ of a covector $\varphi$ and a vector $\psi$ is denoted by the bracket (scalar product) $\langle \varphi | \psi \rangle$. One writes the above isomorphism as

$$\mathcal{H} \ni |\psi\rangle \mapsto \langle \psi | \in \mathcal{H}^*.$$

The ket-bra $|\psi\rangle\langle \varphi |$ denotes the rank-one linear operator $|\psi\rangle \mapsto \ell_\varphi(\psi) = \langle \varphi | \psi \rangle |\psi\rangle$.

**Exercise 8.5.** Suppose $U: \mathcal{H} \to \mathcal{H}$ is unitary and $M \subset \mathcal{H}$. Show that $UM^\perp = (UM)^\perp$.

### 8.4. The $C^*$ algebra of bounded linear operators.

Observables like momentum and spin component are to be represented by operators acting on states (vectors). Almost all operators that appear in quantum theory are linear. A linear operator $A: \mathcal{H} \to \mathcal{H}$ is called **bounded** if the operator norm

$$\|A\| = \sup_{\|\psi\|=1} \|A\psi\|$$

is finite. We write $A \in \mathcal{L}(\mathcal{H})$.

**Example 8.4.** An orthogonal projection $P_M \neq 0$ has norm one.

**Lemma 8.7.**

$$\|A\| = \sup_{\|\psi\|=\|\varphi\|=1} |\langle \varphi | A\psi \rangle|.$$  

We start by introducing a conjugation for operators on a Hilbert space $\mathcal{H}$. Let $A \in \mathcal{L}(\mathcal{H})$. Then the **adjoint operator** $A^*$ is defined via

$$\langle \varphi | A^* \psi \rangle = \langle A\varphi | \psi \rangle.$$

**Example 8.5.** If $\mathcal{H} = \mathbb{C}^n$, then a linear operator is represented by a matrix $A = (a_{ij})_{i,j=1}^n$. Note that

$$\|A\| = \sup_{\|\psi\|=1} \sqrt{\sum_{ijk} a_{ik}^* a_{ij} \psi^*_k \psi_j}.$$  

Then

$$\langle A\varphi | \psi \rangle = \sum_i \left( \sum_j a_{ij} \varphi_j \right)^* \psi_i = \sum_{ij} a_{ij}^* \varphi_j^* \psi_i = \sum_i \varphi_i^* \psi_j \sum_j a_{ij}^* \psi_j$$

Hence, $A^* = (a_{ji}^*)_{i,j=1}^n$.

**Lemma 8.8.** Let $A, B \in \mathcal{L}(\mathcal{H})$ and $\alpha \in \mathbb{C}$. Then

1. $(A + B)^* = A^* + B^*$, $(\alpha A)^* = \alpha^* A^*$;
2. $A^{**} = A$;
(iii) \((AB)^* = B^*A^*\);
(iv) \(\|A^*\| = \|A\|\) and \(\|A\|^2 = \|A^*A\| = \|AA^*\|\).

**Remark.** As a consequence of \(\|A^*\| = \|A\|\), observe that taking the adjoint is continuous.

The algebra \(\mathcal{L}(\mathcal{H})\) of bounded operators with the involution \(*\) forms a C*-algebra.

**Example 8.6.** The continuous functions \(C(I)\) on a bounded interval \(I \subset \mathbb{R}\) together with complex conjugation form a commutative C*-algebra.

An operator \(A \in \mathcal{L}(\mathcal{H})\) is called **normal** if \(AA^* = A^*A\), **self-adjoint** if \(A = A^*\), **unitary** if \(A^*A = AA^* = I\), an (orthogonal) **projection** if \(A = BB^*\) for some \(B \in \mathcal{L}(\mathcal{H})\). Clearly both self-adjoint and unitary elements are normal.

**Exercise 8.6.** Consider the Pauli matrices \(\sigma_1, \sigma_2, \sigma_3 \in \mathcal{L}(C^2)\):
\[
\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

They are Hermitian \(\sigma_i^* = \sigma_i\) and also unitary \(\sigma_i^*\sigma_i = \sigma_i\sigma_i^* = \sigma_i^2 = 1\).

**Exercise 8.7.** Compute the adjoint of the shift operator \(S: \ell^2(\mathbb{N}) \rightarrow \ell^2(\mathbb{N})\), \((a_1, a_2, a_3, \ldots) \mapsto (0, a_1, a_2, \ldots)\).

### 8.5. Resolvents and spectra.

Let \(A\) be a (densely defined) closed operator. The **resolvent set** of \(A\) is the subset of \(\mathbb{C}\) defined by
\[
\rho(A) = \{z \in \mathbb{C}: (A - z)^{-1} \in \mathcal{L}(\mathcal{H})\}.
\]
The **spectrum** of \(A\) is the complement of the resolvent set of \(A\):
\[
\sigma(A) = \mathbb{C} \setminus \rho(A).
\]
In particular, \(\lambda \in \sigma(A)\), if and only if \((A - \lambda)\psi = 0\) for some nonzero vector \(\psi\). In this case \(\psi\) is called an **eigenvector** of \(A\) corresponding to the **eigenvalue** \(\lambda\).

We can characterise the spectra of self-adjoint operators.

**Proposition 8.9.** Let \(A\) be self-adjoint. Then all eigenvalues are real and eigenvectors corresponding to distinct eigenvalues are orthogonal.

**Proof.** If \(A\psi_j = \lambda_j\psi_j, j = 1, 2,\) we have
\[
\lambda_1 \|\psi_1\|^2 = \langle \psi_1 | \lambda_1 \psi_1 \rangle = \langle \psi_1 | A \psi_1 \rangle = \langle A \psi_1 | \lambda_1 \psi_1 \rangle = \langle \lambda \psi_1 | \psi_1 \rangle = \lambda_1 \|\psi_1\|^2
\]
and
\[
(\lambda_1 - \lambda_2) \langle \psi_1 | \psi_2 \rangle = \langle A \psi_1 | \psi_2 \rangle - \langle A \psi_1 | \psi_2 \rangle = 0.
\]
\(\square\)
The result does not imply that two linearly independent eigenfunctions to the same eigenvalue $\lambda$ are orthogonal. However, it is no restriction to assume that they are since we can use Gram-Schmidt to find an orthonormal basis for $\text{Ker}(A - \lambda)$. If $\mathcal{H}$ is finite dimensional, we can always find an orthonormal basis of eigenvectors. (In the infinite dimensional case this is no longer true in general.)

**Theorem 8.10** (Spectral theorem for self-adjoint bounded operators). Let $A$ be bounded and self-adjoint. Then,

$$A = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$$

where $\lambda_k \in \mathbb{R}$ and $|\psi_k\rangle \in \mathcal{H}$ are the eigenvalues and corresponding normalised eigenvectors of $A$. Moreover, $\{|\psi_k\rangle\}$ is an orthonormal basis of $\mathcal{H}$. The above spectral decomposition of $A$ provides a unique norm continuous, $\ast$-homomorphism from the set of all bounded measurable functions $B(\mathbb{R})$ to $L(\mathcal{H})$. For every function $f \in B(\mathbb{R})$, we have

$$f(A) = \sum_k f(\lambda_k) |\psi_k\rangle\langle\psi_k|.$$ 

**Remark.** We can also characterise the spectra of unitary operators. If $U$ is a unitary operator, then all eigenvalues have modulus one and eigenvectors corresponding to distinct eigenvalues are orthogonal.

### 8.6. Time evolution.

Now let us turn to the time evolution of such a quantum mechanical system. Given an initial state $\psi(0)$ of the system, there should be a unique $\psi(t)$ representing the state of the system at time $t \in \mathbb{R}$. We will write

$$\psi(t) = U(t)\psi(0). \quad (8.29)$$

Moreover, it follows from physical experiments that superposition of states holds; that is, $U(t)(\alpha_1\psi_1(0) + \alpha_2\psi_2(0)) = \alpha_1\psi_1(t) + \alpha_2\psi_2(t)$. In other words, $U(t)$ must be a linear operator. Moreover, since $\psi(t)$ is a normalised state $\|\psi(t)\| = 1$, we have

$$\|U(t)\psi\| = \|\psi\|. \quad (8.30)$$

Therefore $U(t)$ is a unitary operator. Next, since we have assumed uniqueness of solutions to the initial value problem, we must have

$$U(0) = 1, \quad U(t + s) = U(t)U(s). \quad (8.31)$$

A family of unitary operators $U(t)$ having this property is called a **one-parameter unitary group**. In addition, it is natural to assume that this group is strongly continuous; that is,

$$\lim_{t \to 0} U(t)\psi = \psi, \quad \text{for all } \psi \in \mathcal{H}. \quad (8.32)$$
Each such group has an \textbf{infinitesimal generator}, defined by

\begin{equation}
H\psi = \lim_{t \to 0} \frac{i}{t} (U(t)\psi - \psi), \quad \text{for all } \psi \in D(H)
\end{equation}

where

\begin{equation}
D(H) = \left\{ \psi \in \mathcal{H}: \lim_{t \to 0} \frac{i}{t} (U(t)\psi - \psi) \text{ exists} \right\}.
\end{equation}

This operator is called the Hamiltonian and corresponds to the energy of the system. If $\psi(0) \in D(H)$, then $\psi(t)$ is a solution of the Schrödinger equation (in suitable units)

\begin{equation}
i\frac{d}{dt}\psi(t) = H\psi(t).
\end{equation}

\textbf{Example 8.7.} Suppose that $\mathcal{H} = \mathbb{C}$ and consider the scalar equation

\begin{equation}
i\frac{d}{dt}\psi(t) = H\psi(t), \quad \psi(0) = \psi_0 \in \mathbb{C}
\end{equation}

where $H \in \mathbb{R}$. It is immediate to integrate the equation and find $\psi(t) = e^{-iHt}\psi_0$.

Recall that the exponential $e^A$ of a matrix $A$ is defined by the series

\[ e^A = \sum_{k \geq 0} \frac{A^k}{k!}. \]

\textbf{Example 8.8.} Suppose that $\mathcal{H} = \mathbb{C}^n$ and consider the vector equation

\begin{equation}
i\frac{d}{dt}\psi(t) = H\psi(t), \quad \psi(0) = \psi_0 \in \mathbb{C}^n
\end{equation}

where $H \in \mathbb{C}^{n \times n}$ is Hermitian $H^* = H$. Using the definition of exponential of a matrix, it is easy to verify that the solution of the problem is $\psi(t) = e^{-iHt}\psi_0$.

There is a one-to-one correspondence between one-parameter strongly continuous unitary groups of operators on $\mathcal{H}$ and self-adjoint operators on $\mathcal{H}$.

\textbf{Theorem 8.11 (Stone's theorem).} Let $H$ be self-adjoint and let $U(t) = \exp(-iAt)$.

(i) $U(t)$ is a strongly continuous one-parameter unitary group.
(ii) The limit $\lim_{t \to 0} \frac{1}{i}(U(t)\psi - \psi)$ exists if and only if $\psi$ is in the domain of $H$, $\psi \in D(H)$, in which case $\lim_{t \to 0} \frac{1}{i}(U(t)\psi - \psi) = H\psi$.
(iii) $U(t)D(H) = D(H)$ and $HU(t) = U(t)H$. 
8.7. Orthogonal sums and tensor products. Given two Hilbert spaces \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \), we define their orthogonal sum \( \mathcal{H}_1 \oplus \mathcal{H}_2 \) to be the set of all pairs \((\psi_1, \psi_2) \in \mathcal{H}_1 \times \mathcal{H}_2\) together with the scalar product

\[
(\langle (\varphi_1, \varphi_2)|((\psi_1, \psi_2)) = \langle \varphi_1|\psi_1 \rangle + \langle \varphi_2|\psi_2 \rangle.
\]

It is left as an exercise to verify that \( \mathcal{H}_1 \oplus \mathcal{H}_2 \) is again a Hilbert space. Moreover, \( \mathcal{H}_1 \) can be identified with \( \{ (\psi_1, 0) : \psi_1 \in \mathcal{H}_1 \} \), and we can regard \( \mathcal{H}_1 \) as a subspace of \( \mathcal{H}_1 \oplus \mathcal{H}_2 \), and similarly for \( \mathcal{H}_2 \). It is customary to write \( \psi_1 + \psi_2 \) instead of \( (\psi_1, \psi_2) \).

More generally, for a countable collection \( (\mathcal{H}_j)_{j=1}^\infty \) of Hilbert spaces, the set

\[
\bigoplus_{j=1}^\infty \mathcal{H}_j = \left\{ \sum_{j=1}^\infty \psi_j : \psi_j \in \mathcal{H}_j, \sum_{j=1}^\infty \|\psi_j\|^2 < \infty \right\},
\]

with scalar product

\[
\left\langle \sum_{j=1}^\infty \varphi_j \bigg| \sum_{j=1}^\infty \psi_j \right\rangle = \sum_{j=1}^\infty \langle \varphi_j|\psi_j \rangle
\]

is a Hilbert space.

**Example 8.9.** \( \bigoplus_{j=1}^\infty \mathbb{C} = \ell^2(\mathbb{N}) \).

Similarly, if \( \mathcal{H} \) and \( \mathcal{K} \) are two Hilbert spaces, we define their tensor product as follows: we start with the set of all finite linear combinations of elements of \( \mathcal{H} \times \mathcal{K} \):

\[
F(\mathcal{H}, \mathcal{K}) = \left\{ \sum_{j=1}^n \alpha_j (\psi_j, \varphi_j) : (\psi, \varphi) \in \mathcal{H} \times \mathcal{K}, \alpha_j \in \mathbb{C} \right\}.
\]

Since we want \((\psi_1 + \psi_2) \otimes \varphi = \psi_1 \otimes \varphi + \psi_2 \otimes \varphi, \psi \otimes (\varphi_1 + \varphi_2) = \psi \otimes \varphi_1 + \psi \otimes \varphi_2,\) and \((\alpha \psi) \otimes \varphi = \psi \otimes (\alpha \varphi),\) we consider the quotient \( F(\mathcal{H}, \mathcal{K})/N(\mathcal{H}, \mathcal{K}),\) where

\[
N(\mathcal{H}, \mathcal{K}) = \text{span} \left\{ \sum_{j,k=1}^n \alpha_j \beta_k (\psi_j, \varphi_k) - \left( \sum_{j=1}^n \alpha_j \psi_j, \sum_{k=1}^n \beta_k \varphi_k \right) \right\}
\]

and write \( \psi \otimes \varphi \) for the equivalence class of \((\psi, \varphi)\).

Next, define

\[
\langle (\psi_1 \otimes \varphi_1)|((\psi_2 \otimes \varphi_2) = \langle \psi_1|\psi_2 \rangle \langle \varphi_1|\varphi_2 \rangle
\]

which extends to a sesquilinear form on \( F(\mathcal{H}, \mathcal{K})/N(\mathcal{H}, \mathcal{K}) \). It is easy to show that this sesquilinear form is a scalar product. The completion of \( F(\mathcal{H}, \mathcal{K})/N(\mathcal{H}, \mathcal{K}) \) with respect to the induced norm is called tensor product \( \mathcal{H} \otimes \mathcal{K} \) of \( \mathcal{H} \) and \( \mathcal{K} \).

**Lemma 8.12.** If \( \{ \psi_j \} \) and \( \{ \varphi_k \} \) are orthonormal bases for \( \mathcal{H}, \mathcal{K}, \) respectively, then \( \{ \psi_j \otimes \varphi_k \} \) is an orthonormal basis for \( \mathcal{H} \otimes \mathcal{K} \).
EXAMPLE 8.10. $\mathcal{H} \otimes \mathbb{C}^n = \mathcal{H}^n$. In particular, $\mathbb{C}^n \otimes \mathbb{C}^m = \mathbb{C}^{nm}$.

EXAMPLE 8.11. $L^2(M, d\mu) \otimes L^2(N, d\nu) = L^2(M \times N, d\mu \times d\nu)$.

9. Angular momentum and spin

In classical mechanics, the components of the angular momentum of a point particle are related to the coordinates $x$, $y$ and $z$ and momenta $p_x$, $p_y$ and $p_z$ by

\[
\begin{align*}
L_x &= yp_z - zp_y, \\
L_y &= zp_x - xp_z, \\
L_z &= xp_y - yp_x.
\end{align*}
\]

To find quantum mechanical operators for the orbital angular momentum, use is made of the assumption that the correspondence principle must be satisfied. Thus any relation which appears in classical mechanics must be valid after quantisation of the canonical variables. As an example, the resulting expression for the $z$-component of orbital angular momentum is

\[
L_z = x (-i\hbar \partial_y) - y (-i\hbar \partial_x) = -i\hbar (x \partial_y - y \partial_x).
\]

EXERCISE 9.1. Verify the commutation relations

\[
\begin{align*}
[L_z, x] &= i\hbar y, & [L_z, y] &= -i\hbar x, & [L_z, z] &= 0, \\
[L_z, p_x] &= i\hbar p_y, & [L_z, p_y] &= -i\hbar p_x, & [L_z, p_z] &= 0.
\end{align*}
\]

The commutation relations among the various components of the angular momentum can be obtained:

\[
[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y. \quad \text{etc.}
\]

Note that the operators for the three components of angular momentum do not commute with one another (they are not simultaneously measurable). The commutation relations can be written in a compact notation using the Levi-Civita antisymmetric tensor

\[
[L_k, L_h] = i\hbar \epsilon_{jkl} L_l, \quad j, k, l = 1, 2, 3.
\]

Another physical quantity of considerable interest is the square of the magnitude of the angular momentum. The corresponding operator is defined through

\[
L^2 = L_x^2 + L_y^2 + L_z^2.
\]

It is easy to check that $L^2$ commutes with all three of the components:

\[
[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0.
\]

Since the $z$-component and the square of the angular momentum commute with each other, it is possible to choose diagonalise simultaneously both operators. We
then have
\[ L^2 \psi = a \psi \]
\[ L_z \psi = b \psi. \]

Of course \( \langle L^2 \rangle \geq \langle L_z^2 \rangle \), so that \( a \geq b^2 \).

It is useful at this point to define two operators which play a role similar to that of the ladder operators of the simple harmonic oscillator. They are
\[ L_{\pm} = L_x \pm iL_y. \]

It can be shown that
\[ [L_z, L_{\pm}] = \pm \hbar L_{\pm}. \]

This equation says that \( L_+ \) and \( L_- \) play the role of ladder operators with the regard to the eigenvalues of \( L_z \). In fact,
\[ L_+(L_+ \psi) = (b + \hbar)(L_+ \psi). \]

Since \( L^2 \) commutes with all three components of the angular momentum, we get
\[ L^2(L_+ \psi) = a(L_+ \psi). \]

Thus the operator \( L_+ \) operating on an eigenfunction of \( L_z \) and \( L^2 \) generates a new simultaneous eigenfunction of these two operators for which the eigenvalue of \( L^2 \) is left unchanged but for which the eigenvalue of \( L_z \) is increased by \( \hbar \). The eigenvalue \( b \) has an upper bound; otherwise the inequality \( a \geq b^2 \) would be violated. If we assume that \( b \) is the largest eigenvalue satisfying the inequality, then it must be \( L_+ \psi = 0 \). If we multiply both sides by \( L_- \) we get
\[ L_-L_+ \psi = (L^2 - L_z^2 - \hbar L_z) \psi = 0, \]
so that \( a = b(b + \hbar) \). In a similar manner it is easy to show that \( L_+^n \psi \) is an eigenfunction of \( L_z \):
\[ L_z(L_+^n \psi) = (b - nh)L_+^n \psi. \]

Take \( n \) to be the largest integer for which the inequality \( a \geq (b - nh)^2 \) is satisfied. In this case \( L_-L_+^n \psi = 0 \). If we multiply by \( L_+ \) we get,
\[ L_+L_-L_+^n \psi = (L^2 - L_z^2 + \hbar L_z)L_+^n \psi. \]

From this, \( a = (b - nh)^2 - (b - nh)h \). Combining with the previous relation between \( a \) and \( b \) we finally get
\[ (9.1) \quad b = \frac{1}{2}(n - 1)h = lh. \]

From this, \( l \) is nonnegative and either integer or half-integer (depending on whether \( n \) is even or odd). We will show later that for the orbital angular moments, \( l \) takes only integral values
\[ l = 0, 1, 2, 3, \ldots \]
The corresponding value of $a$ is

$$a = l(l + 1)\hbar^2.$$  

We can summarise as

$$L^2 Y_{lm} = l(l + 1)\hbar^2 Y_{lm},$$
$$L_z Y_{lm} = m\hbar Y_{lm}.$$  

$l$ can take on nonnegative integral values, and $m$ can take positive or negative integral values such that $l \geq |m|$. Note that these properties result directly from the commutation relations of the angular momentum, and hence follow from only the algebraic properties of the operators.

Note that only one component of the angular momentum may be precisely specified at a time. Although a simultaneous knowledge of the two other components is impossible, it is possible to say something about their expectation values. For example, for a particle in the angular-momentum state $Y_{lm}$ (eigenfunction of $L^2$ and $L_z$),

$$\langle L_x \rangle = \langle L_y \rangle = 0.$$  

Also,

$$\langle L_x^2 \rangle = \langle L_y^2 \rangle = \frac{1}{2}(L^2 - L_z^2) = \frac{1}{2}(l(l + 1) - m^2)\hbar^2.$$  

Note that when the angular momentum “parallel” to the $z$-axis ($m = l$), the $x$- and $y$-components are still not zero.

It is helpful to visualise the results of this section of the aid of a geometrical model. Consider the length of the angular momentum vector $(L_x, L_y, L_z)$ to be $\sqrt{l(l + 1)\hbar}$. The $2l + 1$ allowed projections of this on the $z$-axis are given by $m\hbar$, with $m = 0, \pm 1, \pm 2, \ldots, \pm l$. Note that the projection on the $z$-axis never exceeds the length of the vector. The angular momentum thus may be visualised as lying on the surface of a cone having the $z$-axis for its axis and an altitude of $m\hbar$. All positions in the surface are equally likely.

9.1. Orbital angular momentum. Consider now the orbital angular momentum wavefunctions which are simultaneously eigenfunctions of $L^2$ and $L_z$. It is helpful to introduce spherical coordinates in the usual fashion $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, $z = r \cos \theta$. The operator $L_z$ in spherical coordinates takes the form

$$L_z = -i\hbar \frac{\partial}{\partial \phi}.$$  

We see that the simultaneous eigenfunctions of $L_z$ and $L^2$ are

(9.2)  

$$Y_{lm} = \exp(im\phi)\Theta(\theta),$$
where $m_l$ must take integer values if the resulting function is to be single-valued. In a similar manner, the operator $L^2$ takes the form

$$L^2 = -\hbar^2 \left[ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \theta^2} \right].$$

It can be seen that the operator for $L^2$ is essentially the angular part of the Laplacian operator:

$$-\frac{\hbar^2}{2m} \Delta = -\frac{\hbar^2}{2m} \left( \frac{1}{r} \frac{\partial}{\partial r} r \right)^2 + \frac{L^2}{2mr^2}.$$

The wavefunctions $Y_{lm}(\theta, \phi)$ are called spherical harmonics. They form an orthonormal set of functions on the sphere in the sense that

$$(9.3) \quad \langle Y_{lm} | Y'_{l'm'} \rangle = \int Y_{lm}(\theta, \phi)^* Y'_{l'm'}(\theta, \phi) d\phi \sin \theta d\theta = \delta_{ll'} \delta_{mm'}.$$

Consequently, one can expand any wavefunction $\psi(r, \theta, \phi)$ as

$$\psi(r, \theta, \phi) = \sum_{l,m} a_{lm}(r) Y_{lm}(\theta, \phi).$$

The matrix elements for the $z$-component of the angular momentum in this representation are ($Y_{lm} = |l, m\rangle$)

$$(9.4) \quad (L_z)_{lm,l'm'} = \langle l, m | L_z | l', m' \rangle = mh \delta_{ll'} \delta_{mm'}.$$

In a similar manner, the matrix elements of the square of the angular momentum are

$$(9.5) \quad (L^2)_{lm,l'm'} = \langle l, m | L^2 | l', m' \rangle = l(l+1)\hbar^2 \delta_{ll'} \delta_{mm'}.$$

The (infinite) matrices representing $L_z$ and $L^2$ have thus been evaluated in the representation in which they are diagonal

$$L_z = \hbar \begin{pmatrix} 0 & 1 & 2 & \cdots & -1 & -2 & -3 & \cdots \end{pmatrix}.$$
The next problem is that of calculating the matrix elements of the operators \( L_x \) and \( L_y \). To do this, use is made of the ladder operators \( L_\pm \) that satisfy

\[
L_+ Y_{lm} = c Y_{l,m+1}, \\
L_- Y_{lm} = c' Y_{l,m-1}.
\]

\( c \) and \( c' \) are the only nonzero matrix elements of \( J_+ \) and \( J_- \):

\[
c = \langle Y_{l,m+1} | L_+ Y_{lm} \rangle, \\
c' = \langle Y_{l,m-1} | L_- Y_{lm} \rangle.
\]

It turns out that

\[
L_- Y_{lm} = \hbar \sqrt{l(l+1) - m(m-1)} Y_{l,m-1}, \\
L_+ Y_{lm} = \hbar \sqrt{l(l+1) - m(m+1)} Y_{l,m+1}.
\]

In matrix form:

\[
L^- = \hbar \begin{pmatrix}
0 & \sqrt{2} & 0 & \sqrt{2} & & & \\
0 & 0 & \sqrt{4} & 0 & \sqrt{2} & & \\
0 & \sqrt{4} & 0 & \sqrt{6} & 0 & \sqrt{2} & \\
0 & \sqrt{6} & 0 & \sqrt{6} & 0 & \sqrt{6} & \\
0 & \sqrt{6} & 0 & \sqrt{10} & 0 & \sqrt{10} & \\
0 & \sqrt{10} & 0 & \sqrt{12} & 0 & \sqrt{12} & \\
0 & \sqrt{12} & 0 & \sqrt{12} & 0 & \sqrt{12} & \\
\cdots & & & & & & \cdots
\end{pmatrix}.
\]

Note that \( L_- \) and \( L_+ \) are not self-adjoint.
EXERCISE 9.2. Prove the above formulae. (Hint: Use the fact that \( L_+ \) and \( L_- \) are Hermitian adjoints \( L_+ = L_+^* \), and the identity \( L^2 = L_+ L_- + L_-^2 - \hbar L_z \) to compute first \( \langle Y_{l,m}|L_+ L_- Y_{l,m}\rangle = [l(l + 1) - m(m - 1)] \hbar^2 \).

From the definition \( L_x = \frac{L_+ + L_-}{2} \) and \( L_y = \frac{L_+ - L_-}{2i} \) we get the nonzero matrix elements of \( L_x \) and \( L_y \):

\[
\langle Y_{l,m-1}|L_x Y_{l,m}\rangle = \sqrt{(l + m)(l - m + 1)} \frac{\hbar}{2},
\]

\[
\langle Y_{l,m+1}|L_x Y_{l,m}\rangle = \sqrt{(l + m + 1)(l - m)} \frac{\hbar}{2},
\]

\[
\langle Y_{l,m-1}|L_y Y_{l,m}\rangle = \sqrt{(l + m)(l - m + 1)} \frac{i\hbar}{2},
\]

\[
\langle Y_{l,m+1}|L_y Y_{l,m}\rangle = \sqrt{(l + m)(l - m + 1)} \frac{\hbar}{2i}.
\]

9.2. Spin. Thus far, only orbital angular momentum has been dealt with explicitly. However, it was seen that the formalism based on the commutation relations permitted either half-integral or integral values for \( l \): the restriction to integral \( l \)-values resulted from the explicit form of the operators \( L_z = xp_y - yp_x \), etc., and the requirement of a single-values wavefunction.

Experiments confirmed that particles have an internal degree of freedom that obey the commutation relations of an angular moment but they are not associated to the ‘motion’ of the particles. This degree of freedom is called spin angular momentum of the particle. For spin angular momentum, it is found empirically that the quantum number may take on either integral or half-integral values. The relations obtained for the spin can be derived form the commutation relations. We therefore have the result that the simultaneous eigenfunctions of the square of the spin, denoted by \( S^2 \), and the \( z \)-component of the spin, denoted by \( S_z \), are given by

\[
S^2 \psi_{s m_s} = s(s + 1) \hbar^2 \psi_{s m_s},
\]

\[
S_z \psi_{s m_s} = m_s \hbar \psi_{s m_s},
\]

where \( s \) may take on either integral or half-integral values depending on the nature of the particle and \( m_s = -s, -s + 1, \ldots, +s - 1, +s \). It is also found that for a given particle, \( s \) is fixed (a number characterizing the particle), while \( L^2 \) can assume any value \( l(l + 1)\hbar^2 \).

Note that the spin is a purely quantum observable. This can be seen heuristically: as \( \hbar \to 0 \), \( S^2 \to 0 \). In the classical limit, \( \hbar \to 0 \), but \( l \) can tend to infinity in such a way that \( l(l + 1)\hbar \) remains finite; this is not possible for the spin since \( s \) is fixed.

9.3. Total angular momentum. Consider the problem of the addition of two kinds of angular momenta, the orbital angular momentum and the spin of a particle.
The relations which we will obtain, however, are valid for any two commuting angular momenta (e.g., total spin of a system of two particles).

The total angular momentum $J$ can be written as the sum of the orbital and the spin angular momenta:

$$J = L + S,$$

where $J$ has components

$$J_x = L_x + S_x, \quad J_y = L_y + S_y, \quad J_z = L_z + S_z.$$

Since the angular and spin momenta commute, $[L_j, S_k] = 0$, the total angular momentum $J$ satisfies the same commutation relations of $L$ and $S$. Therefore $J^2$ and $J_z$ commute

$$J^2 \psi_{jm} = j(j + 1)\hbar^2 \psi_{jm},$$

$$J_z \psi_{jm} = m_j \hbar^2 \psi_{jm},$$

Consider first the eigenvalues of

$$J_z = L_z + S_z.$$

They are given by the sum $m_j = m_l + m_s$ of the eigenvalues of $L_z$ and $S_z$ (because $[L_z, S_z] = 0$). $L^2$ and $S^2$ have eigenvalues $l(l + 1)\hbar^2$ and $s(s + 1)\hbar^2$, respectively, and the square of the total angular momentum is

$$J^2 = L^2 + S^2 + 2(L_xS_x + L_yS_y + L_zS_z).$$

We want to determine the eigenvalues of $J^2$. First, note that $J^2$, $L^2$, $S^2$, $J_z$ all commute with one another. It is also clear that $L^2$, $L_z$, $S^2$, $S_z$ form another mutually commuting set of observables. Therefore we have (at least) two alternative sets of four operators which are mutually commuting. If we fix the values of $l$ and $s$, the corresponding subspace has dimension

$$(2l + 1) \times (2s + 1).$$

(This is the number of possible ‘orientations’ of $m_l$ and $m_s$.)

### 9.4. Sum of angular momenta

Consider a system characterised by two angular moments $J_1$ and $J_2$ with $[J_1, J_2] = 0$. We want to determine the eigenvalues of

$$J = J_1 + J_2 = J_1 \otimes 1 + 1 \otimes J_2.$$
From the previous considerations we have that

\[ J_{1z} \text{ has eigenvalues } \hbar m_1 \]
\[ J_{2z} \text{ has eigenvalues } \hbar m_2 \]
\[ J_z \text{ has eigenvalues } \hbar m = \hbar (m_1 + m_2) \]
\[ J_1^2 \text{ has eigenvalues } \hbar^2 j_1 (j_1 + 1) \]
\[ J_2^2 \text{ has eigenvalues } \hbar^2 j_2 (j_2 + 1) \]
\[ J^2 \text{ has eigenvalues } \hbar^2 j(j + 1). \]

A complete orthonormal set of the Hilbert space \( \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \) of the system is

\[ \{ |j_1, j_2; m_1, m_2\rangle \} = \{ |j_1, m_1\rangle \otimes |j_1, m_2\rangle \} \]

Consider the subspace of \( \mathcal{H} \) generated by the system

\[ S_{j_1j_2} = \{ |j_1, j_2; m_1, m_2\rangle : |m_1| \leq j_1, |m_2| \leq j_2 \}. \]

for fixed values of \( j_1 \) and \( j_2 \). The dimension of this subspace is \( (2j_1 + 1)(2j_2 + 1) \).

Consider another basis in this subspace, namely

\[ T_{j_1j_2} = \{ |j_1, j_2; j, m\rangle \}. \]

Note that \( j \) and \( m \) are defined by the conditions that \( \hbar^2 j(j + 1) \) is an eigenvalue of \( J_2^2 \), and \( \hbar m \) is an eigenvalue of \( J_z \). The eigenvalue \( m \) assume the \( 2j + 1 \) values, from \(-j\) to \(+j\). It remains to find the range of values of \( j \).

We first compute the largest value of \( j_{\text{max}} \) of \( j \). Suppose that \( j = j_{\text{max}} \) and \( m = j \); then \( m = m_{\text{max}} = (m_1 + m_2)_{\text{max}} = (m_1)_{\text{max}} + (m_2)_{\text{max}} = j_1 + j_2 \). Hence,

\[ j_{\text{max}} = j_1 + j_2 \]

and

\[ |j_1, j_2; m_1 = j_1, m_2 = j_2\rangle = |j_1, j_2; j = j_1 + j_2, m = j\rangle. \]

We show now that the second largest value of \( j \) is \( j = j_1 + j_2 - 1 \). Consider the eigenvalue \( \hbar m = \hbar (m_1 + m_2 - 1) \). There are two corresponding eigenstates in the set \( S_{j_1j_2} \):

\[ |j_1, j_2; m_1 = j_1 - 1, m_2 = j\rangle \quad \text{and} \quad |j_1, j_2; m_1 = j, m_2 = j - 1\rangle. \]

In other words, the subspace generated by the eigenstates of \( J_z \) with eigenvalue \( \hbar m = \hbar (j_1 + j_2 - 1) \) has dimension 2. Therefore, there are two eigenstates of \( J_z \) in the set \( T_{j_1j_2} \). One is

\[ |j_1, j_2; j = j_1 + j_2, m = j - 1\rangle; \]

the other must be an eigenstate of \( J_2^2 \) corresponding to another eigenvalue and the only possibility is

\[ |j_1, j_2; j = j_1 + j_2 - 1, m = j\rangle. \]
Thus, $j_1 + j_2 - 1$ is an eigenvalue of $J^2$. This reasoning can be reiterated: $j$ can assume values $j_1 + j_2$, $j_1 + j_2 - 1$, $j_1 + j_2 - 2$, etc. until a minimum value $j_{\text{min}}$. To determine $j_{\text{min}}$ we reason as follows. The subspace generated by $S_{j_1,j_2}$ has dimension $(2j_1 + 1)(2j_2 + 1)$. The number of vectors in the second basis $T_{j_1,j_2}$ must be the same. Therefore

$$(2j_1 + 1)(2j_2 + 1) = \sum_{j = j_{\text{min}}}^{j_1 + j_2} (2j + 1).$$

This is an equation in the unknown $j_{\text{min}}$ whose solution is

$$j_{\text{min}} = |j_1 - j_2|.$$ 

We conclude that for the sum of two angular momenta, the eigenvalues of $J^2 = (J_1 + J_2)^2$ has eigenvalues $\hbar^2 j(j + 1)$ where $j = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2|$.

9.4.1 Clebsch-Gordan coefficients. We found that

$$S_{j_1,j_2} = \{|j_1,j_2;m_1,m_2\} : |m_1| \leq j_1, |m_2| \leq j_2\}.$$ 

and

$$T_{j_1,j_2} = \{|j_1,j_2;j,m\} : j = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2|, |m| \leq j\}.$$ 

are two orthonormal sets of the same Hilbert space describing two angular momenta (the angular momenta of two particles, or the spin and orbital angular momentum of the same particle, etc.). Therefore, any vector in $T_{j_1,j_2}$ can be written as linear combination of vectors is $S_{j_1,j_2}$:

$$|j_1,j_2;j,m\rangle = \sum_{m_1,m_2} |j_1,j_2;m_1,m_2\rangle \langle j_1,j_2;m_1,m_2|j_1,j_2;j,m\rangle$$

(9.6)

The coefficients $\langle j_1,j_2;m_1,m_2|j_1,j_2;j,m\rangle$ of this expansion are called Clebsch-Gordan (CG) coefficients. We can suppress the label $j_1,j_2$ from the ket for convenience.

Suppose without loss of generality that $j_1 \geq j_2$. From the properties of addition of angular momenta we have

$$\langle j_1,j_2;m_1,m_2|j,m\rangle \neq 0 \iff j_1 - j_2 \leq j \leq j_1 + j_2$$

$$\langle j_1,j_2;m_1,m_2|j,m\rangle \neq 0 \iff m = m_1 + m_2.$$ 

(The first condition is called triangle condition, for geometrically it means that we must be able to form a triangle with sides $j_1$, $j_2$ and $j$.) Moreover we can always choose the CG coefficients so that

$$\langle j_1,j_2;m_1,m_2|j,m\rangle \in \mathbb{R}$$

$$\langle j_1,j_2;j_1,j-j_1|j,j\rangle > 0.$$ 

Another useful property is

$$\langle j_1,j_2;m_1,m_2|j,m\rangle = (-1)^{j_1+j_2-j} \langle j_1,j_2;-m_1,-m_2|j,-m\rangle.$$ 

This relation halves the work in the computations.

The calculations of the CG coefficients proceed as follows. First, from the identity
\[ |j_1, j_2; m_1 = j_1, m_2 = j_2⟩ = |j_1, j_2; j = j_1 + j_2, m = j⟩ \]
we have
\[ ⟨j_1, j_2; j_1, j_2|j_1 + j_2, j_1 + j_2⟩ = 1. \]
Then we apply the operator \( J_{−} = J_{1,−} ⊗ 1 + 1 ⊗ J_{2,−} \) to both sides and we recall that
\[ J_{−} |j, m⟩ = ℏ \sqrt{j(j+1)} - m(m-1) |j, m−1⟩, \]
\[ J_{1,−} |j_1, m_1⟩ = ℏ \sqrt{j_1(j_1+1) - m_1(m_1-1)} |j_1, m_1−1⟩, \]
\[ J_{2,−} |j_2, m_2⟩ = ℏ \sqrt{j_2(j_2+1) - m_2(m_2-1)} |j_2, m_2−1⟩. \]
The other coefficients can be found by symmetry.

If we arrange the CG coefficients into a matrix, we find it is unitary. This follows from the fact that it relates one orthonormal set to another. Since the CG are all real, this matrix is orthogonal.

Note that the relation (9.6) can be inverted (using the orthogonality of the CG coefficients matrix). We can write
\[ |j_1, j_2; m_1, m_2⟩ = \sum_j ξ_j |j_1, j_2; j, m = m_1 + m_2⟩, \]
where \( ξ_j = ⟨j_1, j_2; j, m = m_1 + m_2|j_1, j_2; m_1, m_2⟩ \). In this way, the tensor product \( |j_1, m_1⟩ ⊗ |j_2, m_2⟩ \) can be written as linear combination of the vectors \( |j, m = m_1 + m_2⟩ \), where \( j = j_1 + j_2, \ldots, |j_1 - j_2|. \)

When \( j_1 = j_2 = 1/2, \) we write symbolically
\[ \frac{1}{2} ⊗ \frac{1}{2} = 1 ⊕ 0; \]
when \( j_1 = 1, j_2 = 1/2: \)
\[ 1 ⊗ \frac{1}{2} = \frac{3}{2} ⊕ \frac{1}{2}; \]
if \( j_1 = 1, j_2 = 1: \)
\[ 1 ⊗ 1 = 2 ⊕ 1 ⊕ 0, \]
and so on.

**Example 9.1.** Let \( \vec{J} = \vec{J}_1 + \vec{J}_2 \) be the sum of two (commuting) angular momenta. Suppose that \( j_1 = 1/2 \) and \( j_2 = 1/2. \) If we measure \( J^2, \) the possible
outcomes are \( \hbar^2 j(j+1) \) with \( j = 0, 1 \). The nonzero CG coefficients are

\[
\begin{align*}
\langle \frac{1}{2}, \frac{1}{2}; +\frac{1}{2}, +\frac{1}{2} | 1, 1 \rangle &= 1 \\
\langle \frac{1}{2}, \frac{1}{2}; +\frac{1}{2}, -\frac{1}{2} | 1, 0 \rangle &= +\frac{1}{\sqrt{2}} \\
\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, +\frac{1}{2} | 1, 0 \rangle &= +\frac{1}{\sqrt{2}} \\
\langle \frac{1}{2}, \frac{1}{2}; +\frac{1}{2}, -\frac{1}{2} | 0, 0 \rangle &= +\frac{1}{\sqrt{2}} \\
\langle \frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, +\frac{1}{2} | 0, 0 \rangle &= -\frac{1}{\sqrt{2}}.
\end{align*}
\]

The CG coefficients can be represented in a square matrix:

\[
\begin{pmatrix}
|1, 1\rangle & |1, 0\rangle & |1, -1\rangle & |0, 0\rangle \\
\end{pmatrix} = 
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\
0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\
\end{pmatrix} 
\begin{pmatrix}
|+\frac{1}{2}, +\frac{1}{2}\rangle \\
|+\frac{1}{2}, -\frac{1}{2}\rangle \\
|1, +\frac{1}{2}\rangle \\
|1, -\frac{1}{2}\rangle \\
\end{pmatrix}.
\]

10. Identical particles

Although the definition of identical particles is the same classically and quantum mechanically, the implications are different in the two cases. When a system contains a number of particles of the same kind, e.g. a number of electrons, the particles are indistinguishable one from another. No observable change is made when two of them are interchanged. This circumstance gives rise to some curious phenomena in quantum mechanics having no analogue in the classical theory, which arise from the fact that in quantum mechanics a transition may occur resulting in merely the interchange of two identical particles, which transition then could not be detected by any observational means. A satisfactory theory ought, of course, to count two indistinguishable states as the same state and to deny that any transition does occur when two identical particles are swapped.

Consider, for simplicity, two quantum particles. Their state is represented by the wavefunction \( \psi(x_1, x_2) \). Consider the operator \( P \) that swap the two particles \( P\psi(x_1, x_2) = \psi(x_2, x_1) \). If the particles are indistinguishable, then \( \psi(x_2, x_1) = e^{i\alpha} \psi(x_1, x_2) \). If we swap again the particles we get

\[
P^2\psi(x_1, x_2) = e^{2i\alpha} \psi(x_1, x_2) = \psi(x_1, x_2),
\]

so that \( e^{i\alpha} = \pm 1 \). This can be extended to an arbitrary number \( N \) of indistinguishable particles. The state is \( \psi(x_1, \ldots, x_N) \). If we permute the \( i \)-th and the \( j \)-th
 particles

\[ P_{ij} \psi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N) = \psi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N) = e^{2i\alpha} \psi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N). \]

Again, the only possibilities (if we assume that the wavefunction is single-valued) are \( e^{2i\alpha} = \pm 1 \). In other words, the state of \( N \) indistinguishable particles is totally symmetric (\( e^{2i\alpha} = +1 \)) or totally antisymmetric (\( e^{2i\alpha} = -1 \)) under permutations. Indistinguishable particles with symmetric states are called \textbf{bosons}; particles with antisymmetric states are called \textbf{fermions}.

10.1. \textbf{Bosonic and Fermionic Hilbert Spaces}. Two identical bosons will always have symmetric state vectors and two identical fermions will always have antisymmetric state vector. Let us call the Hilbert space of symmetric vectors \( V_S \) and the Hilbert space of antisymmetric vectors \( V_A \). We now examine the relations between these two spaces and the tensor product \( V_1 \otimes V_2 \). The space \( V_1 \otimes V_2 \) consists (in finite dimension) of all linear combinations of vectors of the form \( |v_1v_2\rangle = |v_1\rangle \otimes |v_2\rangle \). Suppose that the vectors \( \{|e_i\}\) is a basis of \( V_1 \) (and of \( V_2 \)). To each pair of vectors \( |e_i e_j\rangle \) and \( |e_j e_i\rangle \) \((i \neq j)\) there is one bosonic vector \( \frac{1}{\sqrt{2}} (|e_i e_j\rangle + |e_j e_i\rangle) \), and one fermionic vector \( \frac{1}{\sqrt{2}} (|e_i e_j\rangle - |e_j e_i\rangle) \). If \( e_i = e_j \), the vector \( |e_i e_i\rangle \) is already symmetric, i.e. bosonic. There is no corresponding fermionic vector (the Pauli principle). We express this relation as

\[ V_1 \otimes V_2 = V_S \oplus V_A. \]

The case of \( N = 2 \) particles lacks one feature that is found at larger \( N \). For every \( N! \) product vectors, we get only two acceptable (bosonic or fermionic) vectors. Hence, the tensor product \( V^\otimes N \) for \( N \geq 3 \) is larger (in dimensionality) than \( V_S \otimes V_A \).

References

Here are the main sources I used during the preparation of these lecture notes.

