Quantum Mechanics

Ronald Kleiss (R.Kleiss@science.ru.nl)
IMAPP, FNWI
version of June 25, 2013

$$\hbar = 1.0545716\ \ 10^{-34}\ \mathrm{kg\ m^2\ sec^{-1}}$$

Contents

0 Words **5**
- 0.1 Quantum mechanics past, leading to present . . . 5
- 0.2 A road map . . . 6
- 0.3 Exercise 0 . . . 8

1 States **9**
- 1.1 The state of a system : ket notation . . . 9
- 1.2 The superposition principle . . . 10
- 1.3 Matrix notation . . . 12
- 1.4 Bases . . . 12
- 1.5 The probability interpretation . . . 13
- 1.6 The measurement problem . . . 14
- 1.7 Bra's and inner products : the Hilbert space . . . 15
- 1.8 Playing with light . . . 18
- 1.9 Exercises 1 to 6 . . . 20

2 Observables **23**
- 2.1 A measurement with a specific result . . . 23
- 2.2 A measurement with a general result . . . 24
- 2.3 Using probability theory . . . 25
- 2.4 Hermitean operators and eigenstuff . . . 26
- 2.5 Observables as matrices . . . 28
- 2.6 This is not a measurement ! . . . 29
- 2.7 Degeneracy of eigenvalues . . . 30
- 2.8 What is a classical observable ? . . . 30
- 2.9 Exercises 7 to 16 . . . 32

3 Systems of observables **36**
- 3.1 More than one property . . . 36
- 3.2 The commutator of two operators . . . 36
- 3.3 Commensurable observables and common bases . . . 37
- 3.4 Complete commensurable sets of observables . . . 38
- 3.5 Incommensurability and information loss . . . 39
- 3.6 Commensurability and state labelling . . . 40
- 3.7 The Heisenberg inequality . . . 41
- 3.8 **Extra !** The Mermin-Peres magic square . . . 43
- 3.9 Exercises 17 to 21 . . . 44

4 Does reality have properties ? **47**
- 4.1 The question at issue . . . 47
- 4.2 The EPR paradox . . . 47
  - 4.2.1 EPR's thought experiment . . . 47
  - 4.2.2 Entangled states . . . 48
- 4.3 Bell's inequalities . . . 49

4.3.1 The experiment : Obama *vs* Osama . . . 49
4.3.2 Hidden variables : Bell's inequality . . . 49
4.3.3 QM's predictions . . . 51
4.3.4 Faster than light ? . . . 52
4.3.5 **Extra !** A loophole ? The no-clone theorem . . . 53

5 The Schrödinger equation 55
5.1 The Schrödinger equation . . . 55
5.2 Preservation of the norm . . . 56
5.3 The time-independent Schrödinger equation . . . 56
5.4 Time evolution : the Heisenberg equation . . . 58
5.5 The free particle and the fundamental commutator . . . 58
5.6 The Ehrenfest theorem . . . 60
5.7 **Extra !** Maximum speed : the Margolus-Levitin theorem . . . 61
5.8 Exercises 22 to 29 . . . 63

6 The wave function 66
6.1 The position representation . . . 66
6.2 Position and momentum for wave functions . . . 67
6.3 Inner products for wave functions . . . 68
6.4 The Schrödinger equation for wave functions . . . 69
6.5 Doing the wave . . . 70
6.6 **Extra !** Probability currents . . . 71
6.7 Particle in a box : energy quantisation . . . 73
6.8 **Extra !** Antimatter . . . 75
6.9 Exercises 30 to 45 . . . 79

7 The Harmonic oscillator 84
7.1 A general phenomenon . . . 84
7.2 Raising and lowering operators . . . 85
7.3 Matrix notation . . . 86
7.4 A Hamiltonian and its spectrum . . . 87
7.5 The Hamiltonian in operator form . . . 87
7.6 Wave functions for the harmonic oscillator . . . 88
7.7 Expectation values of position and momentum . . . 90
7.8 **Extra !** Coherent states . . . 91
7.9 Exercises 46 to 52 . . . 92

8 Other one-dimensional systems 96
8.1 A kinky function . . . 96
8.2 Pinning down a particle : the single pin . . . 97
8.3 The double-pin model . . . 98
8.4 A single barrier . . . 100
8.5 The double barrier : a wave filter . . . 103
8.6 Tunneling . . . 104
8.7 **Extra !** Periodic potentials . . . 106

8.7.1 General properties . . . 106
8.7.2 Infinitely many Dirac pins . . . 108
8.7.3 Band structure . . . 109
8.8 Exercises 53 to 60 . . . 111

9 Two-particle systems 114
9.1 Two particles in one dimension . . . 114
9.2 Translation invariance and momentum conservation . . . 114
9.3 Global and internal observables . . . 115
9.4 Separable Hamiltonians . . . 116
9.5 Three dimensions . . . 117
9.6 Exercises 62 to 69 . . . 118

10 Discrete probability 122

11 The Dirac delta function 123

12 The Gaussian integral 125

13 The Gamma function 126

14 Solutions to the Exercises 128
14.1 Excercises 1 to 6 . . . 128
14.2 Excercises 7 to 16 . . . 131
14.3 Excercises 17 to 21 . . . 134
14.4 Excercises 22 to 29 . . . 136
14.5 Excercises 30 to 45 . . . 139
14.6 Excercises 46 to 52 . . . 145
14.7 Excercises 53 to 61 . . . 147

0 Words

— All science is either physics or stamp collecting.
Ernest Rutherford, 1st Baron Rutherford of Nelson
(Nobel Prize for *Chemistry*, 1908)

0.1 Quantum mechanics past, leading to present

Quantum mechanics[1] is a way of thinking, and talking, about the world that leads to so radical a departure with what is considered *classical*, or 'everyday' physics, that some of its consequences appear outrageous[2]. Nevertheless, as far as we can see it is a more *correct* way[3], and the classical world pictures of Newton and Maxwell are seen to be only approximations — albeit often very excellent ones.

The historical roots of the development of quantum mechanics, around the turn of the 20$^{\text{th}}$ century, are diverse and have been fed by dissatisfaction with the performance of classical physics in the description of a number of phenomena. The most important of these are the behaviour of black-body radiation, that led Max Planck to introduce his famous *Planck's constant* ; the unexplained but undeniable stability of atoms that motivated Niels Bohr to invent *quantization rules* ; the photo-electric effect that induced Albert Einstein to postulate that light *waves* are actually composed of *particles*, the *photons* ; and on top of that the complementary idea of Louis de Broglie[4] that, by symmetry, *particles* might as well be thought of as *waves*. Of these, black-body radiation is too difficult, and photons are both too simple and too relativistic, to be discussed in a very elementary text as this ; the quantum-mechanical hydrogen atom will be discussed in some detail in following courses ; the particle-wave *duality* will be treated in these lectures. The *primordial* (but unaware) founding father of quantum mechanics is Josiah Willard Gibbs[5] who, in studying the *Gibbs paradox* about the entropy of gas mixtures, gave the world the notion that *a finite list of properties can be enough to specify a system completely*[6], by introducing the notion of *indistinguishable* particles[7].

A *chronological* description of quantum mechanics, following its historical development, makes for entertaining and instructive reading, but from the point of view of equipping physics students with *what they need to know* it makes as much sense as teaching people to read and write by first explaining to them cuneiform and hieroglyphs, Phœnician and Greeks alphabets, majuscules and

[1] We shall restrict ourselves to *nonrelativistic* quantum mechanics in these notes.

[2] Such as the consequences of the work of John Bell, which allow us to propose that reality may not have any properties of itself.

[3] In the sense that, where classical and quantum predictions differ, quantum *always* turns out to be right — so far.

[4] In full : Louis-Victor-Pierre-Raymond, 7me duc de Broglie.

[5] Less universally famous than he would deserve to be, not least for his statement that "A mathematician may say anything he pleases, but a physicist must be at least partially sane".

[6] Look up any textbook on statistical physics on this.

[7] Which we nowadays know as fermions and bosons.

minuscules[8], before starting on *reading* modern letters and texts. The *history* of science, like the history of writing, is eminently interesting, but it does not belong *here*. I aim to present the status of our understanding of quantum mechanics *as it is now*, and produce a student or reader sufficiently educated to ponder the Bohr-Einstein debate or the Heisenberg-Schrödinger diatribe with the understanding of hindsight[9].

On the way by which quantum mechanics ought to be first introduced to students there are two main schools of thought. Many teachers hold the particle-wave duality dear, and talk about quantum mechanics starting from the concept of the wave function : a minority prefers to start with the superposition principle which leads naturally to the concepts of quantum states and the Hilbert space. I choose the latter road : I find it more fun.

0.2 A road map

In our approach to quantum physics the most fundamental idea is that physical systems can be described as *states* that can be *combined* in various ways (sec. 1). In my view it is *this* that makes quantum mechanics utterly and fundamenally different from classical mechanics. As an ilustration, we discuss some observations made with the help of slices of polaroid.

Having discussed states we will have to accommodate the notion of an *observable*, that is any physical property that can in principle be measured (sec. 2). We shall see that the *mathematical* form of an observable is — almost ! — forced upon us from the properties of the states.

Of course we also look at systems with more than just one measurable property (sec. 3) and we shall find that, again in dramatic contrast to the classical situation, there are properties that can never be both precisely defined at the same time[10].

For many of you the first three sections will look suspiciously like what you have learned during courses on linear algebra. These suspicions are founded : it *is* linear algebra, probably with a slightly modified notation ; and quantum mechanics is one of the reasons why linear algebra is so important. The treatment given in these notes is self-contained, however, and everything about linear algebra that is needed in this course will be introduced in the text.

Perhaps the most radical and disturbing aspects of quantum mechanics is the light it sheds on the nature of reality itself (sec. 4). A few simple thought experiments lead us to the dizzying conclusion that **reality *has* no properties, but *acquires* properties if we *measure* those properties**.

Up to this point our treatment is *static*, in the sense that we only consider systems at a single given moment, as it were in a snapshot picture[11]. Any good physics must of course also describe the *dynamics*, that is a system's time evolution. In classical physics Newton's laws do the trick ; in quantum

[8] And the historical emergence of the sign @ in medieval merchant's administration.
[9] Which is always more useful, if less exciting, than eyewitness' confusion.
[10] The famous *uncertainty relations*.
[11] With the exception of exercise 6 .

mechanics their rôle is taken over by the *Schrödinger equation* (sec. 5). This is also the place where we shall confront what is possibly the most disturbing and counter-intuitive result of all : the relative *incommensurability* of the position and velocity[12] of any given particle. Nevertheless, we shall also see that in many cases classical and quantum time-evolution laws give very similar results[13].

Quite often we assign special importance to quantum states in which a particle has a well-defined position, and all other states are then described in terms of these, using *wave functions* (sec. 6). We shall then see how our quantum mechanics looks in terms of these wave functions. This is where we make contact with the other popular *entrée* into quantum mechanics[14], the one that bases itself on the wave function and views states as the more abstract quantities. The quantum system of a *free* particle, as well as that of a particle shut up inside a box of finite size, provide fine examples of the uses of the wave-function approach.

After all this we are ready to study a very important example of a physical system, namely that of the *harmonic oscillator* (sec. 7). We shall restrict ourselves to the one-dimensional case for simplicity : but we shall find a number of interesting properties even so. To study this system , we shall make much use of the notion of *creation* and *annihilation* operators[15].

In section 8 we shall look at a number of one-dimensional systems of interest. The most physically relevant[16] of these is that in which a particle interacts with a *barrier* of finite height : this system is at the heart of the *quantum tunneling* phenomenon which is widely applied for sub-atomic-precision microscopy. This discussion is in hardcore wave-function language throughout.

So far, all discussions are about quantum systems consisting of only a single particle, and these are certainly very important ; on the other hand, we can only try to begin to understand things like the proton-electron combination that makes up the hydrogen atom if we can handle more than just a single particle. Section 9 is therefore dedicated to two-particle systems.

Some more advanced material, that does not properly belong in the course but which I have included because I like it, is found in the sections beyond 9. It deals with the two-dimensional harmonic oscillator (sec. 10), a first step in more-dimensional quantum mechanics which can be seen as a first introduction into the physics of *angular monentum* and of *spin*; and with the three-dimensional harmonic oscillator (sec. 11), which has an even better claim to realism. Here, the complete angular-momentum structure appears. The world of angular momentum and spin is further investigated in a more formal way in section 12.
Finally, mathematical remarks about probability theory, the Dirac delta func-

[12] Actuallly, the *momentum*.

[13] This goes under the name of the Ehrenfest theorem.

[14] And the way in which, presumably, you will see it presented in your next QM course.

[15] These operators play an important rôle in the *relativistic quantum field theory* of particle physics, as well as in condensed matter physics. For whatever reason, many teachers tend to consider them too abtruse to display to junior students. Prove these many teachers wrong !

[16] Probably.

tion, Gaussian integrals and the like are given in the appendices.

There are a number of sections marked with the blurb **Extra** *!* . These are not required reading but they may be fun nonetheless. At least, I had fun writing them.

On exercises : I have included exercises intended to help you get to grips both with the physics and with the mathematics of 'quantum'. In the margins of the text I have indicated the exercises apposite to what is mentioned in the text, indicated by **E** א where 'א' is the number of the exercise.

E 0

Two things are to be noted about a course such as this : the first is that *it is not trivial*. Even coming to grips with the apparently counter-intuitive results that we shall find[17] is far from easy if you have not yet become used to them by regular exposure[18]. The second is that it is *steep* : the mathematics grows heavier, and the exercises more complicated, as you progress. This is a useful thing : suffering is good for the soul[19].

0.3 Exercise 0

Exercise 0 : **Look it up**
Choose an introductory textbook on quantum mechanics[20], and read what it has to say about blackbody radiation, the photoelectric effect and Bohr's hydrogen atom. Then, decide for yourself if that was useful.

[17] Such a the measurement problem, or the existence of incommensurable properties.

[18] After some time, you will ; many physical and mathematical insights simply take time to be accepted. The mathematician Hippasus of Metapontum demonstrated the existence of irrational numbers, only to be thrown into the crater of Mount Ætna by enraged followers of Pythagoras. In 1545, Gerolamo Cardano published a book describing the use of imaginary numbers, but it was some 250 years before people started to take them seriously. One of the reasons Ludwig Boltzmann killed himself was that he felt that he could not convince people that atoms really exist.

[19] Charles V 'the Wise', king of France from 1356 (1364) to 1380.

[20] Such as, for instance:
B.H. Bransden and C.J. Joachain, *Quantum Mechanics* (Prentice-Hall);
E. Merzbacher, *Quantum Mechanics* (Wiley);
N. Zettili, *Quantum Mechanics* (Wiley)

1 States

— Can nature possibly be so absurd as it seemed
to us in these atomic experiments ?
W. K. Heisenberg, *Physics and Philosophy : The Revolution in Modern Science* (1958)

1.1 The state of a system : ket notation

A *system* is the object of physics, quantum or not. It is considered to be isolated from the rest of the universe[21]. A system can be small and simple, such as a single electron : or large and complicated, such as a star, a lump of iron, or a cat[22]. A system can only be described by *what is known about it*. In quantum mechanics we denote this by a *ket*, an object of the form $|\ \rangle$:

$$|\text{“a description of what is known about the system”}\rangle$$

or

$$|\text{“a list of known properties of the system”}\rangle$$

Here, it is important to realize what we mean by this : *what is known about some property of the system* simply says that if ever we decide to observe or measure that property at some future moment, we already know the outcome in advance. What we know about the system can be obtained in two ways : either we can *observe* the system by performing a *measurement* upon it ; or we can *prepare* the system by using some available technology to bring it into a state in which we can predict the measurement outcome in advance to some extent[23]. If we do not know anything much about the system it is usual to denote its state with an arbitrary symbol such as[24]

$$|\psi\rangle$$

As a simple example we can consider a particle with no other property than that it can occupy one of two possible positions, which we shall label 1 and 2. Some notational conventions are in order : the *notion* of an observable quantity

[21] This is of course an idealization. Note that it is precisely *this* assumption that fails at the moment of observation !

[22] In increasing order of complexity.

[23] As we shall see, *measuring* a property of a system is, in a sense, also a way of *preparing* it. This is of course the same in classical physics : we can decide to measure properties of many identical copies of the system and simply jettison those with property values we don't like. For instance, imagine that you have a set of 100,000 particles moving about with all kinds of different velocities. Inspect them all, and destroy (or disregard) those particles of which the velocity differs too much from a pre-set value. You are then left with a smaller set of 100 particles, say, all with more or less the same velocity. Now you have 'prepared' a system of particles with a given velocity !

[24] It is customary to use the symbol ψ to denote quantum states, but probably for no reason other than that Greek letters make you look intellectual. In fact, really brave people denote the general state simply by $|\rangle$.

such as 'position' will be denoted by an upper case letter, for instance X ; the *value* of the observable 'position' is denoted by a lower case letter, x in this case. In our example, we may indicate the two possible outcomes of a position measurement by, *e.g.* '$x = 1$' or '$x = 2$' and the corresponding states would then be written as

$$|x = 1\rangle \quad , \quad |x = 2\rangle \quad , \text{ respectively,}$$

or simply as

$$|1\rangle \quad , \quad |2\rangle \quad , \text{ respectively.}$$

The *meaning* of what is written inside the ket must be clear from the context[25]. For instance, if the particle can occupy any position in three dimensions, the state could be indicated by $|x^1, x^2, x^3\rangle$ or $|\vec{x}\rangle$: whereas in another setting the notation $|x^1, x^2, x^3\rangle$ might refer to the distinct (one-dimensional) positions of three separate particles. In our example, the statement

$$|\psi\rangle = |1\rangle$$

means that, *if* we perform the measurement, we will *certainly* find the particle at position 1, and certainly *not* at position 2. If a system is in a state for which the outcome of the measurement of an observable A will *certainly* be a number a, we say that that observable is *well-defined* for that state. Of course our particle can have more properties, for instance it may also have a colour. The notion of colour may be denoted by C, and its value by c. If the particle can be either black or white, we would then have '$c = b$' or '$c = w$' as possible results of our looking at its colour. We would then have four natural possibilities :

$$|1, b\rangle \quad , \quad |1, w\rangle \quad , \quad |2, b\rangle \quad , \quad |2, w\rangle \quad .$$

It is easy to generalize this idea. If we only know that position measurement will certainly give $x = 2$ but have no information as to the particle's colour, the state is simply written as $|2\rangle$: on the other hand, if we know that the particle will certainly turn out to be white but know nothing about its position, we would write its state as $|\psi\rangle = |w\rangle$.

1.2 The superposition principle

What we have said so far holds for any kind of physics, also for classical physics[26]. But *now* comes the point of divergence between classical and quantum physics : the *superposition principle*. This is one of the fundamental postulates of quantum physics :

[25] It has this in common with any other mathematics, and indeed with any other language.

[26] The ket notation is not popular in classical physics because we cannot *do* much with it : it is unhelpful in improving our *classical* understanding.

If $|\psi_1\rangle$ and $|\psi_2\rangle$ are two possible states of a system, then

$$|\psi_3\rangle = z_1\,|\psi_1\rangle + z_2\,|\psi_2\rangle \quad ,$$

where $z_{1,2}$ are complex numbers not *both* equal to zero,
is *also* a possible state of the system.

In our two-position example, possible states of the system are for instance

$$3\,|1\rangle \quad , \quad \frac{1}{\sqrt{2}}\Big(|1\rangle - |2\rangle\Big) \quad , \quad \pi\,|1\rangle + e^{137.036}\,|2\rangle \quad .$$

Zero times *any* state of course gives the same zero, which is not considered to be a state at all : $0\times|\psi\rangle = 0$ is not a state[27], and that is why z_1 and z_2 cannot both be zero. Multiplying a state by a (nonzero) complex number does *not* change the physical information in the ket, and therefore $|1\rangle$ and $3i\,|1\rangle$ describe the identically same system, that of a particle that upon measurement will reveal its position value to be $x = 1$. Adding zero to a state does not change it.

At this point it is necessary to comment on the fact that in *classical* physics the superposition principle is absent[28]. The point is that for the usual macroscopic objects that classical physics deals with it is very hard to *make* a nontrivial superposition, rather than that such a superposition cannot exist. In other words, it is a matter of *technology* rather than of *philosophy*. In general, the smaller the system is, the easier it is isolated and manipulated into a nontrivial superposition. As a consequence, there is a chronological order to be found. The famous 'double-slit' experiments are, in fact, nothing but a technological device to put objects in a superposition of different 'to-be-observed' positions on the detection screen. The simplest such objects are photons, and they have been double-slitted the early 19$^{\text{th}}$ century[29]. Electrons are harder to double-slit[30], and this was first done in the 1970s[31]. Nowadays, we can double-slit quite complicated objects such as Buckminsterfullerene molecules[32]. That is

[27] Since $0\times|\psi_1\rangle = 0\times|\psi_2\rangle$, where $|\psi_1\rangle$ and $|\psi_2\rangle$ may have completely different physical content. The statement $|\psi\rangle = 0$ must not be confused $|\psi\rangle = |0\rangle$ which simply means that there is an observable A which, if we measure it for the system in the state $|0\rangle$, will certainly come out with value $a = 0$.

[28] Notice the malicious way in which I put the problem with classical physics, rather than with quantum physics ! Nevertheless, the point is valid : if it be conceded that quantum physics is more fundamental than classical physics, it becomes be the classical child's burden to explain its degeneracy from its quantum parent.

[29] By Thomas Young, who also pioneered the decipherment of Egyptian hieroglyphics, as well as teach-yourself-from-a-booklet tightrope walking.

[30] Mainly because of their electric charge, which makes them very sensitive to disturbing influences such as light.

[31] Many pictures of double-slit experiments with electrons that you see in older textbooks are, in fact, computer simulations. Kind of cheating, really...

[32] C_{60}, the 'Buckyball' molecules.

why one generally states that quantum effects show up at the 'atomic scale', or rather for *microscopic systems*. We must realize, however, that classical physics cannot account for such well-known facts as the stability of *macroscopic systems* of matter : a big lump of iron is solid and hard *because* the electrons in it are in very nontrivial superpositions *all the time*.

1.3 Matrix notation

It is often convenient to employ a matrix notation for kets. In our two-position example we can for instance introduce the notation

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad .$$

The three examples of superpositions mentioned above are then written as

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} \quad , \quad \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \quad , \quad \begin{pmatrix} \pi \\ \exp(137.036) \end{pmatrix} \quad .$$

In the position-plus-colour case we could decide to write

$$|1,b\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |1,w\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |2,b\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |2,w\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad ,$$

and then a possible superposition would be

$$|\aleph\rangle = |1,b\rangle + 0.45\,|2,b\rangle - 2.3i\,|2,w\rangle = \begin{pmatrix} 1 \\ 0 \\ 0.45 \\ \text{-}2.3i \end{pmatrix} \quad . \tag{1}$$

In this case, a state such as $|b\rangle$ would have to be written as

$$|b\rangle = \begin{pmatrix} \alpha \\ 0 \\ \beta \\ 0 \end{pmatrix} \quad ,$$

where we have to live with the fact that α and β are simply unknown complex numbers (but not both zero).

1.4 Bases

The superposition principle states that we can combine states into other states. We can therefore imagine a minimal independent set of states, by taking superpositions of which we can form all other possible states. Such a 'minimal' set

of states is called a *basis*. In our two-position model, the two states $|1\rangle$ and $|2\rangle$ form a basis. The three states

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad |\Omega\rangle = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

do *not* form a basis since $|\Omega\rangle = 2\,|1\rangle + 3\,|2\rangle$ can itself be formed from $|1\rangle$ and $|2\rangle$ and is therefore not independent of them. On the other hand, $|2\rangle$ and $|\Omega\rangle$ *do* form a basis since we might as well decide to leave out $|1\rangle$ from the basis : after all, since $|1\rangle = (|\Omega\rangle - 3\,|2\rangle)/2$ it is not independent from these two either. In our two-position model, the basis consisting of $|1\rangle$ and $|2\rangle$ appears to be the most natural choice, but a choice like

$$|\Psi\rangle = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad , \quad |\Xi\rangle = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

is also perfectly acceptable and might be more convenient under some circumstances. We then simply have

$$|1\rangle = (|\Psi\rangle + |\Xi\rangle)/\sqrt{2} \quad , \quad |2\rangle = (|\Psi\rangle - |\Xi\rangle)/\sqrt{2} \ .$$

Deciding on what kind of basis to take is often an important step in studying a quantum system[33]. **E 3**

1.5 The probability interpretation

We have to face the following problem. Let both z_1 and z_2 be nonzero. We must find an *interpretation* for the system in the state

$$|\psi\rangle = z_1\,|1\rangle + z_2\,|2\rangle \quad . \tag{2}$$

A position measurement will *not* give $x = 1$ with certainty, since then the state would be $|1\rangle$; nor will it give $x = 2$ with certainty. Instead, a position measurement will give $x = 1$ *or* $x = 2$ with *un*certainty, that is, both outcomes have a certain *probability* of occurring ! Let us denote the probability[34] to find $x = 1$ by $P(1)$ and that of finding $x = 2$ by $P(2)$. The interpretation of the above state $|\psi\rangle$ is then that

$$P(1) = \frac{|z_1|^2}{|z_1|^2 + |z_2|^2} \quad , \quad P(2) = \frac{|z_2|^2}{|z_1|^2 + |z_2|^2} \ . \tag{3}$$

These probabilities have the correct properties of being nonnegative and adding up to unity : after all, the probability of finding the particle in any position whatsoever must be one hundred percent[35]. Also the limit $z_1 \to 0$, for instance,

[33] This will become especially clear in our discussion of the position representation, see section 6 .

[34] For some notions about probability theory, see 10.

[35] A system in which the particle can disappear completely is possible, but it is *not* the one we are discussing here.

gives the correct result $P(1) = 0, P(2) = 1$.

The above interpretation is of course easily extended to more complicated cases : in the state $|\aleph\rangle$ of Eq.(1), the probability to find a black particle at position 2 would be

$$P(2, b) = \frac{|0.45|^2}{|1|^2 + |0|^2 + |0.45|^2 + |-2.3i|^2} \approx 0.0312 \ . \tag{4}$$

The situation where a particle can occupy a continuum of positions will be discussed in section 6. **E 1**

An important thing to note about the superposition coefficients is that they can depend on many things : importantly, they can be *time-dependent*. We then have a superposition whose composition may change as a function of time :

$$|\psi\rangle = z_1(t)\,|1\rangle + z_2(t)\,|2\rangle$$

is an example. The various probabilities may then also be time-dependent.

Given the right technology, we may be able to prepare a system into a given superposition $|\psi\rangle$ again and again ; but for two systems in exactly the same state the outcome of a measurement may still be different ! After all, that is what we *mean* by 'probability'. Quantum physics contains an *inherent* randomness that is absent in classical physics[36] : in fact, the only real randomness that exists in nature[37].

1.6 The measurement problem

The notion of superpositions leads to an interesting problem. Take again our two-position particle and suppose that it is in the state $|\psi\rangle$ given by Eq.(2), with $z_{1,2}$ both nonzero. Let us now perform the measurement, and assume that the result is $x = 1$. If we were now to *immediately* repeat the measurement (maybe simply as a check on the measurement just made), what would the outcome be ? Surely $x = 1$ again since we now *know* where the particle is. But that means that we do not really have to do the repeat measurement, we already know its outcome beforehand[38]. But *that* means that we *know* the state of the system : it must be in state $|1\rangle$. The mere fact that we obtained information about the system has changed its state ! We say that because of the measurement with result $x = 1$ the state has *collapsed* from $|\psi\rangle$ to $|1\rangle$.

How the state comes to collapse[39] is *the* fundamental problem in the study of the foundations of quantum physics, and generations of smart people have

[36] Doch würfelt, lieber Albert, unser guter Herr.

[37] As far as we know. One might consider building the 'perfect ultimate random number generator' using quantum effects : in practice it is not so easy although not *a priori* impossible.

[38] Assuming, of course, that we made the first position measurement correctly.

[39] The term 'collapse of the wave function' is also used, see section 6.

tried to either explain, evade, or reformulate it : it goes under the name of the *measurement problem*. Clearly, it must have to do with the interaction between system and outside world, since at the moment of measurement the idealization that the system is isolated *must* fail, otherwise no information could be obtained. The precise mechanism is still far from clear[40].

1.7 Bra's and inner products : the Hilbert space

The set of kets with all their superpositions must be endowed with even more structure, since it ought to be possible to isolate and compute the coefficients $z_{1,2}$ in some way. This is done by the introduction of *bra*'s, and of an *inner product*. We shall state this as follows. For every *ket* $|\psi\rangle$ there is a *bra* $\langle\psi|$ which is the *conjugate* of $|\psi\rangle$. We write the relation between them using a dagger symbol :

$$\langle\psi| = |\psi\rangle^\dagger \quad , \quad |\psi\rangle = \langle\psi|^\dagger \quad . \tag{5}$$

For a superposition's ket :

$$|\psi_3\rangle = z_1 \, |\psi_1\rangle + z_2 \, |\psi_2\rangle \tag{6}$$

the corresponding bra is :

$$\langle\psi_3| = {z_1}^* \, \langle\psi_1| + {z_2}^* \, \langle\psi_2| \tag{7}$$

where the asterisks denote complex conjugation. The rule is simple : the conjugate of a superposition is the (complex conjugated) superposition of the conjugates. Below we give the recipe for taking the conjugate of any quantum object, where we have anticipated a little on the next section.

To take the conjugate :

- Replace each ket by its bra ;
- Replace each bra by its ket ;
- Reverse the order of every string of bras and kets ;
- Replace i by $-i$ (except in the labels inside kets and bras since these label are only descriptions, not formulæ).

[40] Ideas abound. Maybe quantum gravity plays a rôle ; perhaps the universe splits, at every measurement, into alternative realities in every one of which one of the possible outcomes is obtained ; conceivably the split-up between system and observer is more subtle than these notes might seem to imply since the system-observer combination might be just a single quantum system from the point of view of *another* observer... Add your guess, get it right, and win the Galactic Institute's Prize for Extreme Cleverness (see the *Hitch Hiker's Guide to the Galaxy*'s invention of the Infinite Improbability Drive).

A bra $\langle\chi|$ and a ket $|\psi\rangle$ can form an inner product $\langle\chi|\,\psi\rangle$, with the rule you might expect :

$$|\psi_3\rangle = z_1\,|\psi_1\rangle + z_2\,|\psi_2\rangle \quad , \quad |\chi_3\rangle = y_1\,|\chi_1\rangle + y_2\,|\chi_2\rangle \tag{8}$$

give the inner product[41]

$$\begin{aligned}\langle\chi_3|\,\psi_3\rangle \;\; = \;\; & y_1^*\,z_1\;\langle\chi_1|\,\psi_1\rangle + y_2^*\,z_1\;\langle\chi_2|\,\psi_1\rangle \\ & + y_1^*\,z_2\;\langle\chi_1|\,\psi_2\rangle + y_2^*\,z_2\;\langle\chi_2|\,\psi_2\rangle\end{aligned} \tag{9}$$

The *value* of an inner product is a complex number with the properties

$$\begin{aligned}\langle\psi|\,\chi\rangle &= \langle\chi|\,\psi\rangle^* \;\; , \\ \langle\psi|\,\psi\rangle &\geq 0 \;\; , \\ \langle\psi|\,\psi\rangle &= 0 \quad \Leftrightarrow \quad |\psi\rangle = 0 \;\; .\end{aligned} \tag{10}$$

A *normalized* state is a state $|\psi\rangle$ with $\langle\psi|\,\psi\rangle = 1$. If $|\psi_{1,2,3}\rangle$ are all normalized and $|\psi_3\rangle = z_1\,|\psi_1\rangle + z_2\,|\psi_2\rangle$ while $\langle\psi_1|\,\psi_2\rangle = 0$, then we must have $|z_1|^2 + |z_2|^2 = 1$, and the probability rules become particularly simple: $P(1) = |z_1|^2$, $P(2) = |z_2|^2$. **E 2**

Now for some nomenclature : if two states $|\psi\rangle$ and $|\chi\rangle$ have a vanishing inner product, $\langle\psi|\,\chi\rangle$, then they are said to be *orthogonal*. A basis of states in which the inner product between any two basis states is zero is called an *orthogonal basis*. If two *normalized* states are orthogonal we call them *orthonormal*. If the basis states of an orthogonal basis are all normalized, that basis is called an *orthonormal basis*.

The inner product of two states can be seen as a measure of the *compatibility* of the physical information between them. In our two-position example, the physical information in states $|1\rangle$ and $|2\rangle$ is completely incompatible, and therefore we have **E 5**

$$\langle 1|\,1\rangle = 1 \quad , \quad \langle 2|\,2\rangle = 1 \quad , \quad \langle 1|\,2\rangle = 0 \;\; . \tag{11}$$

In the position-and-colour case, states $|1\rangle$ and $|b\rangle$ may be compatible or not, and the inner product can be zero (if $|b\rangle$ is actually $|2, b\rangle$) or can have an absolute value of up to one (if both $|1\rangle$ and $|b\rangle$ are actually equal to $|1, b\rangle$). If the state happens to be the $|\aleph\rangle$ of Eq.(1), we have

$$\langle\aleph|\,2, w\rangle = 2.3i \;\; . \tag{12}$$

The inner product $\langle\chi|\,\psi\rangle$ between two states $|\psi\rangle$ and $|\chi\rangle$ is also called the *overlap* between these two states. Its absolute value squared, $|\,\langle\chi|\,\psi\rangle\,|^2$, is a probability : the probability that, once the system is in state $|\psi\rangle$, upon observation it will

[41] The *bracket*, hence 'ket' and 'bra'. We apologize for the sense of humour of P.A.M. Dirac, who invented these names.

show precisely the properties of state $|\chi\rangle$ (and, of course, *vice versa*).

The matrix notation for bra's is that of a row vector. In the two-position model we have

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \Leftrightarrow \quad \langle\psi| = (\alpha^* \;\; \beta^*) \quad . \tag{13}$$

For example,

$$\begin{aligned} \langle 1| = |1\rangle^\dagger &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}^\dagger = (1 \quad 0) \quad , \\ \langle 2| = |2\rangle^\dagger &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}^\dagger = (0 \quad 1) \quad . \end{aligned} \tag{14}$$

For the more complicated state of Eq.(1), we have

$$\langle \aleph| = |\aleph\rangle^\dagger = (1 \quad 0 \quad 0.45 \quad 2.3i) \quad . \tag{15}$$

The rule for the inner product is the usual one from linear algebra : **E 4**

$$\begin{aligned} |\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad , \quad |\chi\rangle = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \quad \Rightarrow \\ \langle\chi|\,\psi\rangle = (\gamma^* \;\; \delta^*) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \gamma^*\alpha + \delta^*\beta \quad . \end{aligned} \tag{16}$$

The generalization to more complicated cases is of course straightforward.

As we have mentioned above, the inner product gives us a way of determining the probabilities : assuming $|1\rangle$ and $|2\rangle$ to be normalized, then if our two-position particle is in the normalized state $|\psi\rangle = z_1\,|1\rangle + z_2\,|2\rangle$ we have

$$P(1) = |\,\langle\psi|\,1\rangle\,|^2 \quad , \quad P(2) = |\,\langle\psi|\,2\rangle\,|^2 \quad . \tag{17}$$

Mathematically speaking, the set of all possible states of a system is seen to form a (complex) linear vector space with an inner-product structure. This is called the *Hilbert space* of the system[42]. Finding the Hilbert space of a system is, to a large extent, what quantum physics is about. **E 6**

We have now finished the first introduction with the concept of quantum states. Our treatment has been quite abstract, so abstract in fact that you might be tempted to see all the above as just so much juggling with kets, bras and other symbols. It is, however, very important that you realize that *quantum states are real*. A ket is a very concrete thing referring to a real, physical system. You yourself are described by a ket, addmittedly a very big one. Superpositions of dead cats and live ones are *not* a crazy notion : quantum states and superpositions are *how the world works*.

[42]The importance of Hilbert space is one of the reasons why linear algebra is such an important topic in the curriculum's first year.

1.8 Playing with light

An interesting experiment, which is particularly nice since it needs no complicated apparatus, is the following. As you know[43] light consists of indivisible particles called *photons*. Photons have a colour[44], with more energetic ones looking bluer[45] but they also have another property called *polarization*[46]. Photons with the same colour can come in different polarizations, and this is what we shall investigate. Humans are (almost) completely insensitive to the polarization of light, although some animal species[47] can see it.

Now there exist materials called polarizing materials, or *polaroid*, which are for instance used in the fabrication of sunglasses. If photons impinge on polaroid, they either pass through or they don't. How do we know that ? Well, looking through the polaroid we notice that it seems somewhat dark, so that some light, at least, is stopped. You might think that maybe some photons have only *part* of their energy transmitted[48], but in that case the *colour* of the transmitted light would change towards the red, and that is not what we see.

Without having to know anything about polarization or the specific structure of the material, we can set up the following quantum-mechanical investigation. Let us take a few slices of polaroid[49]. Superimpose two slices, called A and B. Look through them at a light source and rotate B with respect to A. For one orientation of B, which we shall call '0 degrees' you will see that all[50] light passing through A also passes through B, while a rotation over 90 degrees makes the set of slices completely opaque (*i.e.* no photons that passed A also passed B). Rotating further to 180 degrees makes the two slices transparent again. From this we can immediately conclude that the polaroid has some internal orientation that influences whether photons are transmitted or absorbed. This is a measurement[51] ! Photons that are transmitted by A are therefore in a (collapsed) quantum state which we shall denote by $|0\rangle$. Since A and B are made of the same material, a rotation of B by 90 degrees would therefore bring any transmitted photon in the state that we may call $|90\rangle$. As no photons are, in fact transmitted, we find that for these two normalized quantum states

$$\langle 0|\,0\rangle = \langle 90|\,90\rangle = 1 \quad , \quad \langle 0|\,90\rangle = 0 \ , \tag{18}$$

[43] Or ought to know : it is what got Albert Einstein his 1921 Nobel prize.

[44] Obviously !

[45] *Red* photons have 2eV of energy each, and *purple* ones about 3 eV.

[46] Never mind what that property actually *is* : the important thing here is that we can distinguish photons with different polarization, as we shall describe. In courses on electromagmetism, the polarization will be related to the orientation of the electromagnetic field describing the photons ; again, at this point this is immaterial.

[47] For instance bees, and notably the so-called mantis shrimp, a nasty submarine predator, known for killing prey with sonic shock waves and breaking through aquarium glass.

[48] In which case you might imagine such a photon 'split up', with part of it stopped and part of it transmitted.

[49] Which you can buy in specialist shops, or 'borrow' from your practical-course tutor.

[50] Or *almost* all : we neglect the effects of absorption that are to be expected for non-ideal materials.

[51] Namely, to be or not to be transmitted.

since (apparently) the states $|0\rangle$ and $|90\rangle$ have no overlap, for the square of the overlap gives the probability of photon transmission through B, that is, the intensity of the light. We also see that the overlap between $|0\rangle$ and $|180\rangle$ is maximal, as (therefore) is that between $|90\rangle$ and $|270\rangle$. We have now established that photons come in (superpositions of) exactly *two* basis states, for which we may take $|0\rangle$ and $|90\rangle$: light comes in two polarizations !

Suppose that we keep B at 90 degrees, so no light is passed through. Now, we slide a third slice C in between A and B, oriented at 45 degrees. What will be the result ? Slice C will pass photons in the state $|45\rangle$, and absorb photons in the state $|135\rangle$. Now $|45\rangle$ must be a superposition of $|0\rangle$ and $|90\rangle$: by symmetry it must have the form

$$|45\rangle = \frac{1}{\sqrt{2}} \Big(|0\rangle + |90\rangle \Big) \quad , \tag{19}$$

and then

$$|135\rangle = \frac{1}{\sqrt{2}} \Big(|0\rangle - |90\rangle \Big) \quad , \tag{20}$$

since these two states are orthogonal[52]. Therefore,

$$\langle 0|\, 45\rangle = 1/\sqrt{2} \ , \tag{21}$$

so exactly half the photons transmitted by A make it through C. On the other hand, from Eq.(19) and (20) we find

$$|0\rangle = \frac{1}{\sqrt{2}} \Big(|45\rangle + |135\rangle \Big) \quad , \quad |90\rangle = \frac{1}{\sqrt{2}} \Big(|45\rangle - |135\rangle \Big) \quad , \tag{22}$$

so that

$$\langle 45|\, 90\rangle = 1/\sqrt{2} \tag{23}$$

and half of the photons that survived C will emerge out of B. The effect of inserting C at 45 degrees is therefore that 25% of the photons now actually make it ! Classically, this is weird : if the combination AB is an insurmountable barrier for the light, you would never expect that adding a *third* barrier would make it *easier* for the photons ! Nevertheless, this is exactly what happens if you actually do the experiment.

You may be unsurprised but disappointed that only a quarter of the photons that pass through A make it through the whole setup ACB. We can do better than that, though. A slice oriented at a general angle θ (in radians) will pass photons in the state[53]

$$|\theta\rangle = \cos\theta \ |0\rangle + \sin\theta \ |90\rangle \tag{24}$$

[52] We can always choose the complex phases of $|0\rangle$ and $|90\rangle$ so as to make Eq.(19) correct. The point is that the two coefficients must have the *same* absolute value. Eq.(20) is then automatic.

[53] The appearance of the cosine and sine in the equation given below looks reasonable, at least to students who have been sufficiently brainwashed by physics courses. However, by using a slightly more sophisticated set-up that can quantitatively measure the amount of light transmitted, we can actually determine the form of (the absolute value squared of) the coefficients experimentally.

The probability of a photon passing A and then passing a slice oriented by θ is therefore $(\cos\theta)^2$. Now, let us replace slice C by $n-1$ slices C_1, C_2, ... C_{n-1}, where C_1 is oriented over an angle $\alpha = \pi/2n$, slice C_2 over an angle 2α, and so on. The total transmission probability for any photon transmitted by A is then, for large n, given by[54]

$$(\cos\alpha)^{2n} \approx \left(1 - \frac{\pi^2}{8n^2}\right)^{2n} \approx 1 - \frac{\pi^2}{4n} \quad , \tag{25}$$

so that for really large n essentially no photons are lost : we have managed to twist the polarization of all photons over 90 degrees[55] !

1.9 Exercises 1 to 6

Exercise 1 : **Three-position model**
Develop the matrix formalism for a particle that can have not two but *three* different positions 1,2, and 3. Find the most general form of a state in which the probabilities to find the particles at these positions are all equal.

Exercise 2 : **Inner products**
Assume that two quantum states, $|\psi\rangle$ and $|\chi\rangle$, are normalized. Show that we must have $0 \leq |\langle\chi|\psi\rangle| \leq 1$, and that $|\langle\chi|\psi\rangle| = 1$ implies that $|\psi\rangle$ and $|\chi\rangle$ are equal up to a complex phase. (Hint : consider $|\phi\rangle = |\psi\rangle - e^{i\phi}|\chi\rangle$.) In linear algebra, this result is known as the *Cauchy-Schwartz inequality*.

Exercise 3 : **Some states**
If you want, you can do this excercise in matrix notation. Consider a system with 4 basis states, which we shall label $|\psi_1\rangle, \ldots, |\psi_4\rangle$. Now we form the superpositions

$$\begin{aligned}
|\chi_1\rangle &= |\psi_1\rangle + |\psi_2\rangle + |\psi_3\rangle + |\psi_4\rangle \quad , \\
|\chi_2\rangle &= |\psi_1\rangle + |\psi_2\rangle - |\psi_3\rangle - |\psi_4\rangle \quad , \\
|\chi_3\rangle &= |\psi_1\rangle - |\psi_2\rangle + |\psi_3\rangle - |\psi_4\rangle \quad , \\
|\chi_4\rangle &= |\psi_1\rangle - |\psi_2\rangle - |\psi_3\rangle + |\psi_4\rangle \quad .
\end{aligned}$$

1. Show that the states $|\chi_1\rangle, \ldots, |\chi_4\rangle$ are mutually *orthogonal*.
2. Normalize them so that they are mutually *orthonormal*, that is, they have zero product with each other and unit product with themselves.
3. Write the states $|\psi_1\rangle, \ldots, |\psi_4\rangle$ as linear combinations of the *orthnormal* versions of the states $|\chi_1\rangle, \ldots, |\chi_4\rangle$.

[54] We use the fact that for small z, $\cos z \approx 1 - z^2/2$.
[55] Of course, in practice, as mentioned above each slice will absorb *some* photons even if they have the right polarization, so this is a theoretical rather than a practical exercise.

Exercise 4 : **Some inner products**
Consider the following quantum states, in matrix notation :

$$|\psi_1\rangle = \begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \\ 2 \end{pmatrix} \quad , \quad |\psi_2\rangle = \begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \\ -2 \end{pmatrix} \quad , \quad |\psi_3\rangle = \begin{pmatrix} 1 \\ i \\ \sqrt{2} \\ 2i \end{pmatrix} \quad , \quad |\psi_4\rangle = \begin{pmatrix} 1 \\ -1 \\ -2 \\ \sqrt{2} \end{pmatrix} \quad .$$

Compute the inner products of all these states with each other, and with themselves.

Exercise 5 : **Orthonormality can persist**
As remarked above, two states $|1\rangle$ and $|2\rangle$ are called *orthonormal* if

$$\langle 1|\,1\rangle = \langle 2|\,2\rangle = 1 \quad , \quad \langle 1|\,2\rangle = 0 \ ;$$

that is, they are normalized physical states without any overlap. Let the states evolve as a function of time t as

$$|1\rangle \ \rightarrow \ e^{i\alpha\, t}\,|1\rangle \quad , \quad |2\rangle \ \rightarrow \ e^{i\beta\, t}\,|2\rangle \quad ,$$

where α and β are some real numbers.

1. Show that the *physical* content of the states $|1\rangle$ and $|2\rangle$ does not change with time.

2. Show that the states remain orthonormal at all times. Give a simple physical argument why you should have expected this anyway.

3. Extend this reasoning to an arbitrarily large set of mutually orthonormal states (this is *easy*).

Exercise 6 : **A simple model for neutrino oscillations**
Neutrinos (denoted by the symbol ν) are subatomic particles that can occur in several species[56]. They can be produced (and observed) by the so-called *weak interaction*, and then come in two types (states) which we denote by $|\nu_e\rangle$ and $|\nu_\mu\rangle$. Several years ago it was established that these states are, in fact, superpositions of two simpler states, $|\nu_1\rangle$ and $|\nu_2\rangle$:

$$\begin{aligned} |\nu_e\rangle &= \cos\theta\,|\nu_1\rangle + \sin\theta\,|\nu_2\rangle \quad , \\ |\nu_\mu\rangle &= -\sin\theta\,|\nu_1\rangle + \cos\theta\,|\nu_2\rangle \quad , \end{aligned}$$

where θ is the so-called *mixing angle*.

1. Assuming that the states $|\nu_1\rangle$ and $|\nu_2\rangle$ are orthonormal, show that the states $|\nu_e\rangle$ and $|\nu_\mu\rangle$ are also orthonormal.

[56] Actually, three, but for simplicity we shall take just two.

2. Assume that at time $t = 0$ the state $|\nu_e\rangle$ has been prepared (by selecting the correct weak-interaction process, which makes only $|\nu_e\rangle$ and never $|\nu_\mu\rangle$). As time goes by, the two states $|\nu_{1,2}\rangle$ have their own time evolution :

$$\begin{aligned} |\nu_1\rangle &\rightarrow e^{i\alpha\, t}\,|\nu_1\rangle \quad , \\ |\nu_2\rangle &\rightarrow e^{i\beta\, t}\,|\nu_2\rangle \quad . \end{aligned}$$

Show that the states $|\nu_{e,\mu}\rangle$ evolve therefore as

$$\begin{aligned} |\nu_e\rangle &\rightarrow \quad e^{i\alpha\, t}\cos\theta\,|\nu_1\rangle + e^{i\beta\, t}\sin\theta\,|\nu_2\rangle \equiv |\nu_e;t\rangle \quad , \\ |\nu_\mu\rangle &\rightarrow -e^{i\alpha\, t}\sin\theta\,|\nu_1\rangle + e^{i\beta\, t}\cos\theta\,|\nu_2\rangle \equiv |\nu_\mu;t\rangle \quad , \end{aligned}$$

and show that these are orthonormal at all times.

3. At some time t the neutrino is observed by letting it partake in some weak-interaction process which selects only neutrinos of the pure type $|\nu_\mu\rangle$. The overlap between $|\nu_e;t\rangle$ and the pure state $|\nu_\mu\rangle$ is given by

$$\langle\nu_\mu|\,\nu_e;t\rangle \quad .$$

Compute this overlap.

4. Show that the *probability* that, starting out with a neutrino of type $|\nu_e\rangle$, we shall find it to have emerged as one of type $|\nu_\mu\rangle$ at time t is given by

$$P(\nu_e \rightarrow \nu_\mu) = \left(\,\sin(2\theta)\;\sin\left(\frac{1}{2}(\alpha-\beta)t\right)\right)^2 \quad .$$

Also compute the probability that $|\nu_e\rangle$ persists as itself (this is *easy*).

5. Explain why it is resonable that the probabilities depend on α and β only through the combination $|\alpha - \beta|$, and that it vanishes whenever θ is a multiple of $\pi/2$.

These 'neutrino oscillations' have been experimentally observed.

2 Observables

— To observe is to be selective.
Caspar Lavater, *Essays on Physiognomy* (1855)

2.1 A measurement with a specific result

So far we have discussed the *states* of a system ; but of course we also must implement the *observables* into our mathematical description. The most important consideration is that a measurement (observation) made upon a system in *any* state leaves it again in *some* state, possibly the same but often another one ; an observable is therefore best described as an *operator*, that is, an *operation* that transforms states into states. An observable is *not* istelf a state, nor is it a *number* : but of course its *value* upon measurement is a number[57]. In the next few sections, we shall argue our way towards a mathematical implementation of the idea of an observable in quantum mechanics[58].

For the sake of concreteness, we extend out two-position model to that of a particle that can be found to occupy a number of discrete coordinate positions x_j, $j = 1, 2, \ldots, N$, along a line. The normalized state where the coordinate of the particle is certainly x_j will be denoted by $|j\rangle$ (we might as well have used the notation $|x_j\rangle$ but that would lead to more confusing typography in what follows). We then have an orthonormal set of position states :

$$\langle j|\, j\rangle = 1 \quad , \quad \langle j|\, k\rangle = 0 \ \text{ if } k \neq j \ , \tag{26}$$

for which we introduce the *Kronecker symbol* $\delta_{j,k}$. It is defined as

$$\delta_{m,n} = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n \end{cases} \quad , \tag{27}$$

so that we can write $\langle j|\, k\rangle = \delta_{j,k}$. Next, we consider a generic normalized superposition of our states :

$$|\psi\rangle = \sum_{j=1}^{N} c_j \ |j\rangle \quad \text{with} \quad \sum_{j=1}^{N} |c_j|^2 = 1 \ . \tag{28}$$

That this state is, indeed, properly normalized follows from the following short calculation, that we present in great detail to show how the Kronecker symbol works :

$$\langle \psi|\, \psi\rangle \quad = \quad \left(\sum_j c_j \, |j\rangle \right)^{\dagger} \left(\sum_k c_k \, |k\rangle \right)$$

[57] It is important to always keep a clear distinction between the *observable* itself and its *value* in a given measurement.

[58] Note that I carefully refrain from the word 'derive'.

$$
\begin{aligned}
&= \left(\sum_j c_j^* \, \langle j| \right) \left(\sum_k c_k \, |k\rangle \right) \\
&= \sum_{j,k} c_j^* c_k \, \langle j|\, k\rangle \;=\; \sum_{j,k} c_j^* c_k \; \delta_{j,k} \\
&= \sum_j c_j^* \; \left(\sum_k c_k \; \delta_{j,k} \right) \\
&= \sum_j c_j^* \; \Big(c_j \Big) \;=\; \sum_j |c_j|^2 \;=\; 1 \;. \qquad (29)
\end{aligned}
$$

As we have stated before, the probability to find the particle at position x_n is then $|c_n|^2$. Now suppose that we carry out the measurement, and that its result is that $x = x_3$. We then know[59] that the particle must be left in the state $|3\rangle$ after the measurement : we therefore propose that *the measurement with given result* $x = x_3$ be represented by an object like

$$|3\rangle \, \langle 3|$$

which must *not* be confused with the inner product $\langle 3|\, 3\rangle$! We can multiply such an object with a state such as our $|\psi\rangle$, as follows :

$$
\begin{aligned}
\Big(|3\rangle \, \langle 3| \Big) \; |\psi\rangle \;&=\; |3\rangle \; \Big(\langle 3|\, \psi\rangle \Big) \;=\; |3\rangle \; \sum_{j=1}^{N} \; c_j \; \langle 3|\, j\rangle \\
&=\; |3\rangle \; \sum_{j=1}^{N} \; c_j \; \delta_{3,j} \;=\; c_3 \; |3\rangle \;\;, \qquad (30)
\end{aligned}
$$

since in the sum over j only one term survives. We see that this multiplication transforms the original state $|\psi\rangle$ into $|3\rangle$, as wanted ; and we also find the correct coefficient c_3 into the bargain.

2.2 A measurement with a general result

Of course, a reasoning as presented above must also hold if the measurement result is $1, 2, 4, 5, \ldots, N$, and therefore it is reasonable to propose that the general observable X of position ought to be represented, in our mathematics, by some combination of these individual measurements-with-result, which we shall denote by writing the observable's symbol crowned with a 'hat' :

$$X \;\;\Leftrightarrow\;\; \hat{X} \;\equiv\; \sum_{j=1}^{N} \; \alpha_j \; \; |j\rangle \, \langle j| \;\;, \qquad (31)$$

where we still have to work out what the coefficients α_j are. We shall do this in the next section ; at this point it is important to see that we could measure

[59] From our discussion of the measurement problem.

not just the position coordinate X itself, but for instance also its *square*, X^2. State $|5\rangle$, for instance, would then give the measurement result $x^2 = (x_5)^2$. It is reasonable to expect that the corresponding operator ought to be the square of $\hat{X}$:

$$\begin{aligned} \widehat{X^2} &= (\hat{X})^2 = \left(\sum_{j=1}^{N} \alpha_j \ |j\rangle\langle j|\right)\left(\sum_{k=1}^{N} \alpha_k \ |k\rangle\langle k|\right) \\ &= \sum_{j=1}^{N}\sum_{k=1}^{N} \alpha_j\alpha_k \ |j\rangle \ \langle j|\,k\rangle \ \langle k| = \sum_{j=1}^{N} {\alpha_j}^2 \ |j\rangle\langle j| \ , \end{aligned} \tag{32}$$

where we have again used the Kronecker symbol result $\langle j|\,k\rangle = \delta_{j,k}$. This is of course easily extended : the operator $\widehat{X^p}$ corresponding to the observable X^p is given by

$$\widehat{X^p} = (\hat{X})^p = \sum_{j=1}^{N} {\alpha_j}^p \ |j\rangle\langle j| \ . \tag{33}$$

It is interesting to notice that in this result we do not have to restrict ourselves to $p = 1, 2, 3, \ldots$ but that p could for instance be a fraction, or even $p = 0$. In that case, the observable is just the number 1, since $x^0 = 1$ for all x. We therefore have automatically

$$1 = \hat{1} = \sum_{j=1}^{N} |j\rangle\langle j| \ ,$$

which is called the *completeness relation*. **E 7**

2.3 Using probability theory

Here, we shall use several probabilistic notions : if you don't know much about probability theory, *this* is the moment to have a look at appendix 10.

We again use the general normalized state $|\psi\rangle$ of Eq.(28). If our particle is in this state, the probability to find position coordinate $x = x_j$ is given by $|c_j|^2$, where $c_j = \langle j|\,\psi\rangle$: by now we know this quite well[60]. Therefore, the probability density $P(x_j)$ for finding the result x_j is

$$P(x_j) = |c_j|^2 = |\,\langle j|\,\psi\rangle\,|^2 = \langle\psi|j\rangle \ \langle j|\,\psi\rangle \ . \tag{34}$$

We can write this in different ways : **E 8**

$$\langle\psi|\,j\rangle \ \langle j|\,\psi\rangle = \langle\psi|\,\Big(\,|j\rangle \ \langle j|\,\psi\rangle\Big) = \langle\psi|\,\Big(|j\rangle\,\langle j|\,\Big)\ |\psi\rangle \ , \tag{35}$$

and we recognize this as the operator for the measurement-with-result-'$x = x_j$' acting on the state ψ, where the resulting state is then multiplied by the bra

[60] I hope, and trust.

$\langle\psi|$ of the original state $|\psi\rangle$. The *expectation value* of the value x of the position observable X is

$$\begin{aligned}\langle x\rangle &= \sum_{j=1}^{N} P(x_j)\, x_j \;=\; \sum_{j=1}^{N} \langle\psi|\,\Big(|j\rangle\,\langle j|\Big)\,|\psi\rangle\; x_j \\ &= \langle\psi|\,\left(\sum_{j=1}^{N} x_j\;|j\rangle\,\langle j|\right)\,|\psi\rangle \quad . \end{aligned} \tag{36}$$

This strongly suggests that the observable X ought to be obtained by choosing for the coefficients α_j in Eq.(31) simply the *measurement results* x_j. Therefore, we decide that our position observable X has the corresponding operator

$$\hat{X} = \sum_{j=1}^{N} x_j\;|j\rangle\,\langle j| \quad . \tag{37}$$

Then, the expectation value[61] is given by the simple rule

$$\langle x\rangle = \langle\psi|\,\hat{X}\,|\psi\rangle \quad , \tag{38}$$

and we see that this can consistently be extended to *powers* of the observable: **E 9**

$$\langle x^p\rangle = \langle\psi|\,\widehat{X^p}\,|\psi\rangle = \langle\psi|\,\hat{X}^p\,|\psi\rangle \quad . \tag{39}$$ **E 10**

In particular we can find the *variance* of the observable for the system in the state $|\psi\rangle$: it is

$$\sigma(X)^2 = \big\langle (X-\langle X\rangle)^2\big\rangle = \langle\psi|\,\Big(\hat{X}-\langle X\rangle\Big)^2\,|\psi\rangle \quad , \tag{40}$$

which we can also write as **E 11**

$$\sigma(X)^2 = \big\langle X^2\big\rangle - \langle X\rangle^2 = \langle\psi|\,\hat{X}^2\,|\psi\rangle - \langle\psi|\,\hat{X}\,|\psi\rangle^2 \quad . \tag{41}$$

2.4 Hermitean operators and eigenstuff

Observables' operators are built up from numbers, bras and kets as we have seen. Since bras and kets are each other's conjugates, we can also talk about the conjugates of such operators. To conjugate any combination of bras, kets and complex numbers, we replace each bra by its ket ; each ket by its bra ; reverse their ordering ; and replace each complex number by its complex conjugate[62]. As an example, of the object

$$W \equiv (2+i)\;|a_1\rangle\,\langle z| \;+\; e^{4i}\;|5\rangle\,\langle n_2| \;-\; 3\,\langle b|\,v\rangle$$

[61] In the notation for the expectation value, we use $\langle X\rangle$ for the *observable* X somewhat interchangeably with $\langle x\rangle$ for the *value* x. This does not, usually, lead to confusion.

[62] Note that this rule says that

$$\langle\psi|\,\chi\rangle^\dagger = \langle\chi|\,\psi\rangle = \langle\psi|\,\chi\rangle^* \quad ,$$

as it should.

(whatever all its ingredients may mean) the conjugate version (again denoted by a dagger) is

$$W^\dagger = (2-i)\ |z\rangle\langle a_1|\ \ +\ \ e^{-4i}\ |n_2\rangle\langle 5|\ \ -\ \ 3\,\langle v|b\rangle$$

Let us now perform this conjugation for our position observable X, with operator

$$\hat{X} = \sum_{j=1}^{N} x_j\ \ |j\rangle\langle j| \tag{42}$$

as before. Since the position values are real numbers, ${x_j}^* = x_j$ and we see that conjugation changes *nothing* here :

$$\hat{X}^\dagger = \hat{X}\ \ . \tag{43}$$

Observables' operators are *self-conjugate*, or *Hermitean*[63].

Another useful notion is obtained when we let the operator $\hat{X}$ act on a state for which the position value x will be found to be some x_n with certainty :

$$\begin{aligned} \hat{X}\ |n\rangle &= \left(\sum_{j=1}^{N} x_j\ \ |j\rangle\langle j|\right)\ |n\rangle \\ &= \sum_{j=1}^{N} x_j\ \ |j\rangle\ \langle j|n\rangle\ \ =\ \ x_n\ |n\rangle\ \ , \end{aligned} \tag{44}$$

again via the Kronecker symbol. We see that the net effect of $\hat{X}$ acting on $|n\rangle$ is to simply multipy it by x_n. We say that the state $|n\rangle$ is an *eigenstate* of the operator $\hat{X}$, and we call x_n its corresponding *eigenvalue*. **E 12**

We can now formulate the second fundamental postulate of quantum mechanics :

To every observable A
corresponds a Hermitean operator $\hat{A}$:
The operator's eigenvalues are the possible outcomes
of a measurement of the value a of A ;
their corresponding eigenstates are
the states with certain outcome.

[63] In the mathematical literature the term *self-adjoint* is more commonly used. The subtle differences between these notions are not important for us here.

Since we can always perform an observation[64] of A, the operator $\hat{A}$ acts on all possible states, *i.e.* on the whole Hilbert space of the system. The collection of all possible outcomes for a is called the *spectrum* of the operator $\hat{A}$: in the case of our position observable X it is the set $\{x_1, x_2, x_3, \ldots, x_N\}$. Often the operator is the thing that is given, and then it is our task to find its spectrum and eigenstates. Once this is done, the *quantum structure* of the system is essentially determined, and we can then examine the physics for any state in which the system may find itself[65].

Before finishing this train of reasoning which we started in section 2.1, it is important that we realize that we have not *proven* or *derived* the operator form of observables : rather, we have looked at *plausible arguments* that lead us to suspect that things can probably not be different from this. Still, the notion of observables as self-adjoint linear operators is as much a postulate as the superposition principle. It is the *successes* of this approach that must convince us that we are on the right track : after all, we are supposed to be doing *physics*, not *mathematics*. **E 13**

2.5 Observables as matrices

We have seen that kets can often be represented by column vectors, and bras by row vectors : what is the corresponding representation of an observable's operator? An operator contains both bras and kets, that is, both row and column vectors : therefore operators are *matrices* if we adopt matrix notation. It is useful to return to our two-position particle again, for which we introduced the matrix notation to start with. Given two arbitrary states

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{and} \quad |\chi\rangle = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \quad , \tag{45}$$

we can form the combination[66]

$$|\psi\rangle\langle\chi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} \gamma^* & \delta^* \end{pmatrix} = \begin{pmatrix} \alpha\gamma^* & \alpha\delta^* \\ \beta\gamma^* & \beta\delta^* \end{pmatrix} \quad . \tag{46}$$

Note that its conjugate,

$$|\chi\rangle\langle\psi| = \Big(|\psi\rangle\langle\chi|\Big)^\dagger = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} = \begin{pmatrix} \gamma\alpha^* & \gamma\beta^* \\ \delta\alpha^* & \delta\beta^* \end{pmatrix} \quad , \tag{47}$$

[64] What should we think of a version of quantum mechanics that would *forbid* us to measure an observable at any time we like ? Nonetheless, some such world views have been proposed.

[65] Although not precisely rigorous, the notion of 'spectrum' is also sometimes applied not to the operator $\hat{A}$ but to the corresponding observable A. This does not, usually, lead to confusion.

[66] This combination is called the *tensor product*, not to be confused (as we have said before) with the *inner* (or vector) product. The rule is simple : a bra followed by a ket is the inner product, a simple number ; a ket followed by a bra is the tensor product, a matrix.

is nothing but the *Hermitean conjugate*, obtained by interchanging rows and columns, and complex-conjugating.

For our two-position model we can thus find several cases:

$$|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad |2\rangle\langle 2| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad , \tag{48}$$

and

$$|1\rangle\langle 2| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad |2\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad ; \tag{49}$$

here we see explicitly that $|2\rangle\langle 1|$ is the Hermitean conjugate of $|1\rangle\langle 2|$.

In matrix notation, the *eigenstates* of an operator are nothing but the *eigenvectors* : and the *eigenvalues* are exactly what we mean by that term in matrix theory. As an example, let us define an observable Z that takes the value 1 if the particle is seen to be in position 1 and -1 if it seen to be in position 2. Its corresponding operator in our matrix notation is then

$$\hat{Z} = (+1)\,|1\rangle\langle 1| + (-1)\,|2\rangle\langle 2| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad . \tag{50}$$

The eigenvalues of this matrix are, indeed, 1 and -1, so that the spectrum of the observable is $\{1, -1\}$: and the respective eigenvectors are

$$\text{for } z = 1 : \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \text{for } z = -1 : \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \tag{51}$$

as expected[67].

2.6 This is not a measurement !

We have seen that to an observable B there must correspond a Hermitean operator $\hat{B}$, that acts on the Hilbert space of the possible quantum states of the system. That is, given some specific quantum state $|\psi\rangle$ of the system, we can define (and possibly even compute) the resulting state

$$|\chi\rangle = \hat{B}\,|\psi\rangle \quad . \tag{52}$$

Now, do ***not*** make the mistake of assuming that $|\chi\rangle$ is 'the result of measuring B on the system in state $|\psi\rangle$' ! That this cannot possibly be the interpretation of Eq.(52) is immediately obvious if we realize that, in general, various results for the value of B are possible. If we find $B = b_1$, say, the system will be left in the state $|B = b_1\rangle$; whereas if we would have found $B = b_2 \neq b_1$, the system would have ended up in state $|B = b_2\rangle \;\neq\; |B = b_1\rangle$. The state resulting from a measurement is therefore not fixed, while of course the result $|\chi\rangle$ of applying $\hat{B}$ is always the same[68], no matter how often we perform Eq.(52). In a very real sense, the operator $\hat{B}$ is rather *the collection of all possible measurement results.*

[67] For some reason that I don't quite fathom, nature *likes* linear algebra.

[68] This is what mathematics is all *about.*

2.7 Degeneracy of eigenvalues

To every eigenvalue of an observable's operator there must correspond an eigenstate. But such eigenstates are not necessarily unique. There is the trivial fact that when $|\psi\rangle$ is an eigenstate of $\hat{A}$ with eigenvalue λ, then so is $z\,|\psi\rangle$ for any nonzero z ; but it is also quite possible that there is a quite *different* eigenstate $|\chi\rangle$ of $\hat{A}$ that happens to have the *same* eigenvalue λ. In such a case the two states $|\psi\rangle$ and $|\chi\rangle$ contain, apart from the fact that the value a of A is λ, additional physical information about the system. If this additional information is incompatible between the states, we have $\langle\chi|\,\psi\rangle = 0$: if it is partly compatible, $|\psi\rangle$ and $|\chi\rangle$ are independent eigenstates. This is the phenomenon of *degeneracy* : an eigenvalue is said to be *n times (or n-fold) degenerate* if there exist precisely n independent normalized eigenstates with the same eigenvalue. As an illustration, consider a particle that can move around in one dimension. If we know the particle's kinetic energy we know the *magnitude* of its velocity as well – but *not* the direction in which it is moving. The energy eigenvalue is therefore two-fold degenerate, the additional information being whether it moves to the right or to the left. If the particle moves in two or three dimensions, the energy eigenvalues are even infinitely degenerate ! **E 14**

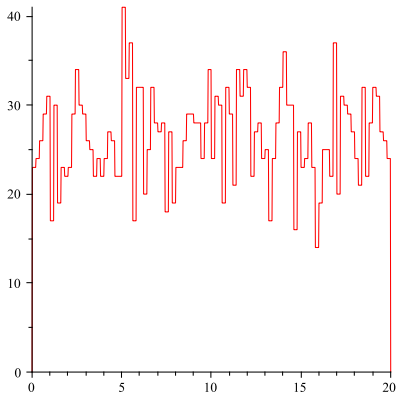
Another example is provided by our two-position-two-colour model. There are *two* independent states that have well-defined position $x = 1$: they are distinguished by the *other* observable value, namely $c = b$ or $c = w$. The eigenvalue $x = 1$ is therefore two-fold degenerate[69]. **E 15** **E 16**

2.8 What is a classical observable ?

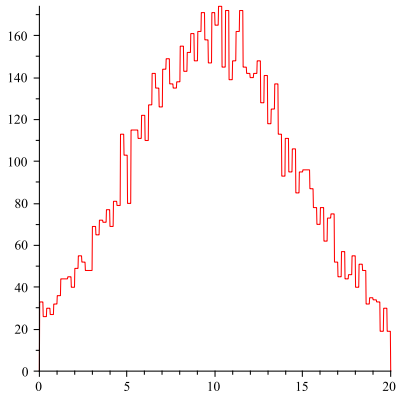
We have now constructed the mathematical form of an observable in quantum mechanics. It is clear that this looks *nothing at all* like what you would call an observable in classical physics. How, then, are we to understand/interpret what we mean by a *classical* observable, which is supposed to have one unique (although possibly unknown) value ? To gain some insight, recall that the experimental information about the quantum observable for a system in a given quantum state $|\psi\rangle$ can be characterized by drawing a histogram of the measurement results for a large number of measurements all performed on identical copies of the system in state $|\psi\rangle$[70]. Let us look at various possibilities for such a histogram.

[69] In fact, in that model *all* the eigenvalues are two-fold degenerate.

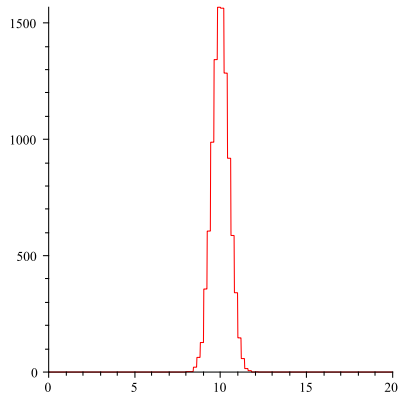
[70] Do not make the mistake of thinking of a large number of measurements, performed one after the other on a single system ! The first measurement will generally change the state of the system as we have discussed ; and subsequent measurements will then simply repeat the same outcome.



The distribution of measurement values is in this case very wide, and we cannot possibly maintain that there is anything like a unique outcome of the measurement. The observable has, for a system in this quantum state, *no* classical interpretation.



The distribution of measurement values is in this case more or less centered around a central value, but still has a appreciable spread. One would hesitate to assign a classical character to the observable unless in a very rough, qualitative manner of speaking.



The distribution of measurement values is in this case very concentrated around a central value. Almost every measurement returns a value close to the central value. In this case we can consider the central value as *the* value of the classical observable.

From these pictures we conclude that an observable A can be interpreted in the classical way if the *width* (standard deviation) of the possible measurement results for A is very small compared to, say, the precision of the measuring apparatus, and then in that case the classical observable $A_{\rm cl}$ has the value $\langle A \rangle$. A point to be kept in mind is that, depending on the quantum state of the system, it is quite possible that one observable looks quite classical while another observable does not look classical at all. Indeed, as we shall see in the following, this is sometimes even unavoidable[71]! For many (mostly macroscopic) systems the standard deviations are usually very small indeed, and then we can approximate the physics of the system by classical mechanics[72].

2.9 Exercises 7 to 16

Exercise 7 : The completeness relation
Prove that if $|\psi\rangle$ is any superposition of the states $|j\rangle$ of section 2.1, then the operator

$$\hat{1} = \sum_j |j\rangle \langle j|$$

is indeed the unit operator in the sense that

$$\hat{1} \ |\psi\rangle = |\psi\rangle$$

provided that the sum runs over *all* j.

Exercise 8 : Using completeness
Use the completeness relation to prove that for the probabilities $P(x_j)$ defined

[71] This remark is a sneak preview of the Heisenberg uncertainty relations, to be discussed below.
[72] We shall return to this in our discussion of the Ehrenfest theorem.

in Eq.(34) the sum is correct :

$$\sum_j P(x_j) = 1$$

Exercise 9 : **Working with operators**
Consider again the operator $\hat{X}$ defined in Eq.(31). Show that, if we choose as our observable not simply X itself but rather $X^p + X^q$ (where p and q are nonnegative numbers), the corresponding operator must be

$$\widehat{(X^p + X^q)} = \hat{X}^p + \hat{X}^q = \sum_{j=1}^{N} \left({a_j}^p + {a_j}^q \right) \; |j\rangle \langle j| \quad .$$

Use this to show that the operator corresponding to the observable $\exp(zX)$, where z is some number, is given by

$$\widehat{\exp(zX)} = \exp(z\hat{X}) = \sum_{j=1}^{N} \exp(z\,\alpha_j) \; |j\rangle \langle j| \quad .$$

Note : the notion of a *function* of an operator, such as $\exp(\hat{X})$, is given by the Taylor expansion. That is, if

$$f(z) = c_0 + c_1 z + c_2 z^2 + c_3 z^3 + \cdots$$

then

$$f(\hat{A}) = c_0 + c_1 \hat{A} + c_2 \hat{A}^2 + c_3 \hat{A}^3 + \cdots$$

Exercise 10 : **Some algebra**
Prove that Eq.(39) is indeed correct.

Exercise 11 : **Two formulas for the variance**
Prove that Eq.(40) van indeed be rewritten as Eq.(41).

Exercise 12 : **Hermitean operators**
Consider a self-adjoint operator $\hat{Q}$. Let $|n\rangle$ be an eigenstate of $\hat{Q}$ with eigenvalue λ_n, and $|k\rangle$ an eigenstate with eigenvalue λ_k.

1. By considering $\langle n| \hat{Q} |n\rangle$, prove that $\lambda_n = {\lambda_n}^*$: all eigenvalues of $\hat{Q}$ must be *real* numbers.

2. By considering $\langle k| \hat{Q} |n\rangle$, prove that $\lambda_k \neq \lambda_n$ implies that $\langle k| n\rangle = 0$: eigenstates with different eigenvalues must be orthogonal[73].

[73] In this proof, note that $\lambda_n = \lambda_k$ does not *forbid* orthogonality of the eigenstates : this is of course the case when the eigenvalue is degenerate.

Exercise 13 : **Classical and quantum dice**
We first consider a *classical* die that we do not throw ourselves : assuming someone has done it for us, we simply *look* at which of the numbers 1,2,3,4,5 or 6 it shows. For an unloaded[74] die the probability of its showing k is $P(k) = 1/6$ for $k = 1, \ldots, 6$.

1. Compute the expectation value $\langle k \rangle$.
2. Compute the standard deviation $\sigma(k)$ (see appendix 10).

We now turn the die into a *quantum* one : the observable K is still the number shown when we look at the die, and a normalized state in which the die will certainly show value k is denoted, of course, by $|k\rangle$.

1. Construct the operator $\hat{K}$ corresponding to K.
2. Let the state of our quantum die be

$$|\psi\rangle = \frac{1}{\sqrt{6}} \sum_{k=1}^{6} |k\rangle \quad .$$

   Show that this state is normalized. Show, by computing the probabilities, that this state cannot be distinguished from the classical die.

3. Show that the state

$$|\psi\rangle = \frac{1}{\sqrt{6}} \sum_{k=1}^{6} e^{(k^2 - 2/k)i} \; |k\rangle$$

   can *also* not be distinguished from a classical die (this is *easy*).

4. Let the state of our quantum die now be

$$|\psi\rangle = \frac{1}{\sqrt{21}} \sum_{k=1}^{6} \sqrt{k} \; |k\rangle \quad .$$

   Show that this is a normalized state. Compute $\langle k \rangle$ and $\sigma(k)$ and show that this cannot be obtained by an unloaded classical die.

5. Find a quantum state in which the expectation value $\langle k \rangle$ is equal to that of the classical die, but its standard deviation $\sigma(k)$ is larger than that of the classical die.

Exercise 14 : **Explain this !**
Explain why in two or more dimensions the eigenvalues of the kinetic-energy observable are almost all infinitely degenerate — and find the single exception.

[74] Note for non-gamblers: 'loading' dice is fiddling with them so that the probabilities for the outcomes are no longer equal.

Exercise 15 : **Eigenvalues and degeneracies**
For a quantum system with two basis states, we can define three observables M_z, M_x and M_y, the operators of which are

$$\hat{M}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad , \quad \hat{M}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \hat{M}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Prove that these are indeed Hermitean, and compute their eigenvalues.

Exercise 16 : **A more complicated observable**
Next, consider a quantum system with three basis states. For this sysem we define an observable Ω whose operator $\hat{\Omega}$ has the following matrix form :

$$\hat{\Omega} = \begin{pmatrix} 1 & 0 & (1+i)/2 \\ 0 & 1 & 1/\sqrt{2} \\ (1-i)/2 & 1/\sqrt{2} & 1 \end{pmatrix}$$

Show that this operator is, indeed, self-adjoint, and compute its eigenvalues with their degeneracies (if any).

3 Systems of observables

— At first, I was deeply alarmed.
I had the feeling that, through the surface of atomic
phenomena, I was looking at a strangely beautiful interior,
and felt almost giddy at the thought that I now had to probe
this wealth of mathematical structures that nature
had so generously spread out before me.
Werner Karl Heisenberg, *Physics and Beyond : Encounters and Conversations* (1971)

3.1 More than one property

A general system can, of course, have more than just one single measureable property of interest ; we must therefore consider *systems of observables*. For simplicity we shall consider two observables, A and B : these could, for instance, be the position and colour of our two-position-two-colour toy model, or the x and y positions of a particle in two dimensions, or whatever.

In classical physics, systems with more than one given property are of course very common. Indeed, for a particle in classical mechanics we can only compute the trajectory if *both* its position and its velocity are precisely known at some given instant[75]. Nothing forbids a *classical* system to have any number of well-defined properties at one and the same time.

In *quantum* physics the situation is fundamentally different. If the observable A is well-defined for some state, the system must be in an eigenstate of $\hat{A}$: but it is not *a prori* obvious that such an eigenstate is *also* an eigenstate of $\hat{B}$! For both A and B to be well-defined the system must be in a *common* eigenstate of both $\hat{A}$ and $\hat{B}$, and it depends on the mathematical form of these operators whether or not such common eigenstates exist. We shall find that for many very important observables such common eigenstates do, in fact, *not* exist. We must now examine this question somewhat more quantitatively.

3.2 The commutator of two operators

Let us denote the eigenvalues of $\hat{A}$ by $a_{1,2,3,\ldots}$ and those of $\hat{B}$ by $b_{1,2,3,\ldots}$. In the most 'user-friendly' case[76], both the value of A and that of B may be well-defined. If we know that measurement of A will certainly give $a = a_n$ while at the same time measurement of B will certainly give $b = b_k$, the system must be in a state which we can obviously denote by $|a_n, b_k\rangle$. We then also have

$$\hat{A}\ |a_n, b_k\rangle = a_n\ |a_n, b_k\rangle \quad , \quad \hat{B}\ |a_n, b_k\rangle = b_k\ |a_n, b_k\rangle \quad . \tag{53}$$

[75] Mathematically speaking, Newton's laws of motion have the form of *second-order* differential equations for the position as a function of time : such equations need *two* initial conditions to be solved. As we shall see, in quantum mechanics the equations of motion are differential equations of only *first order* in time.

[76] Assuming the user to be classically-minded.

But this implies that

$$\hat{A}\,\hat{B}\;|a_n, b_k\rangle = \hat{A}\,b_k\;|a_n, b_k\rangle = b_k\,\hat{A}\;|a_n, b_k\rangle = b_k\,a_n\;|a_n, b_k\rangle \quad , \tag{54}$$

since b_k is just a number ; while on the other hand

$$\hat{B}\,\hat{A}\;|a_n, b_k\rangle = \hat{B}\,a_n\;|a_n, b_k\rangle = a_n\,\hat{B}\;|a_n, b_k\rangle = a_n\,b_k\;|a_n, b_k\rangle \quad . \tag{55}$$

We see that the *order* in which we apply $\hat{A}$ and $\hat{B}$ does not matter :

$$\begin{aligned} \hat{A}\,\hat{B}\;|a_n, b_k\rangle - \hat{B}\,\hat{A}\;|a_n, b_k\rangle &= a_n\,b_k\;|a_n, b_k\rangle - b_k\,a_n\;|a_n, b_k\rangle \\ &= (a_n b_k - b_k a_n)\;|a_n, b_k\rangle = 0 \quad . \end{aligned} \tag{56}$$

We have here introduced an important notion, the *commutator* of two operators. This is denoted by square brackets as follows, for arbitrary operators $\hat{A}$ and $\hat{B}$:

$$[\hat{A}, \hat{B}] \equiv \hat{A}\,\hat{B} - \hat{B}\,\hat{A} \quad . \tag{57}$$

The commutator of two operators is itself also an operator. Some properties of commutators are

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \quad , \tag{58}$$

$$[\hat{A}, z_1\hat{B}_1 + z_2\hat{B}_2] = z_1\,[\hat{A}, \hat{B}_1] + z_2\,[\hat{A}, \hat{B}_2] \quad , \tag{59}$$

$$[\hat{A}, \hat{B}]^\dagger = -[\hat{A}^\dagger, \hat{B}^\dagger] \quad , \tag{60}$$

$$\text{if } \hat{A}^\dagger = \hat{A}\ ,\ \hat{B}^\dagger = \hat{B}\ ,\ [\hat{A}, \hat{B}] = i\hat{C} \text{ then } \hat{C}^\dagger = \hat{C} \quad , \tag{61}$$

$$[\hat{A}, [\hat{B}, \hat{C}]\,] + [\hat{B}, [\hat{C}, \hat{A}]\,] + [\hat{C}, [\hat{A}, \hat{B}]\,] = 0 \quad , \tag{62}$$

$$[\hat{A}, \hat{B}\,\hat{C}] = [\hat{A}, \hat{B}]\,\hat{C} + \hat{B}\,[\hat{A}, \hat{C}] \quad . \tag{63}$$

For two given operators $\hat{P}$ and $\hat{Q}$ it *may* be the case that the order in which they are applied on *any* state never matters. In that case **E 17**

$$[\hat{P}, \hat{Q}] = 0 \tag{64}$$

and the operators are said to *commute* : they are then *commuting* operators. If the commutator does not vanish identically, the operators are said to be *noncommuting*[77]. **E 18**

3.3 Commensurable observables and common bases

We present here without proof the following theorem from linear algebra : *if two Hermitean operators commute, then it is possible to find a basis of common eigenstates that is orthonormal.* For a system in any such a common eigenstate,

[77] The mention of '*any* state' in the above is important. If two operators $\hat{P}$ and $\hat{Q}$ are noncommuting, there *may* of course be still one or more particular states $|\zeta\rangle$ for which the ordering *happens* not to matter, such that $\hat{P}\hat{Q}\,|\zeta\rangle = \hat{Q}\hat{P}\,|\zeta\rangle$: such states are simply eigenstates of the commutator, with eigenvalue equal to zero. The commutator is only said to vanish if *all* states are eigenstates with zero eigenvalue.

the two corresponding observables are simultaneously well-defined ; this was the situation discussed in the above. Of course, the system may also be in a superposition of common eigenstates such that one or both observables are not well-defined : but the important point is that they *can* be simultaneously well-defined. In that case the observables are called *commensurable*. If the commutator does *not* vanish, then a complete basis of common eigenstates can *not* be found[78], and the observables are called *incommensurable*[79].

We can easily extend the notion of commensurability to three or more observables. For n observables $\hat{A}_1, \ldots, \hat{A}_n$ to be commensurable, it is necessary that each pair of the corresponding operators commute :

$$[\hat{A}_i, \hat{A}_j] = 0 \quad , \quad i, j \in \{1, 2, 3, \ldots n\} \ . \tag{65}$$

3.4 Complete commensurable sets of observables

As we have seen, eigenvalues of a single obervable's operator may be degenerate if there exist eigenstates with the same value for that observable, which differ in some other physical property. If we consider that other physical property as well, we split the various eigenstates by distinguishing them further, using the value of the second observable. This is called *lifting* the degeneracy of the eigenvalue. We are thus naturally led to the concept of *complete commensurable sets* of observables : such a set is the smallest set of mutually commensurable observables (mutually commuting operators) such that there is precisely *one* eigenstate for each possible combination of values in the spectra of the observables. In other words : if $\hat{A}_1$, $\hat{A}_2$ and $\hat{A}_3$ form a complete commensurable set, there is precisely *one* eigenstate[80] that has precisely the eigenvalues a_1, a_2, a_3 of their three operators. We then have the maximal number of observables that can be simultaneously well-defined[81], and the degeneracy is completely lifted. Once the values of all the observables in the complete commensurable set are specified, we know *everything* it possible to know about the system in that state. An example : an electron is (in three dimensions) completely specified if we are provided its three-dimensional position and the value of its *spin*[82] in a given direction. For any quantum system it is always useful to determine its complete commensurable sets of observables[83].

[78] The proof of this is quite trivial : if there *were* a basis of common eigenstates the commutator would *have* to vanish since we can write all other states as superpositions of these basis states.

[79] The word 'commensurable' comes from Latin *cum* (with, at the same time as) and *mensura* (measure, measuring). Note the distinction : *observables* are *(in)commensurable*, while *operators* are *(non)commuting*.

[80] As always, up to a nonzero complex overall factor.

[81] Of course, if A is part of such a set, we might always choose to use A^2 and the like instead of A, but obviously this does not give any new information. Also we do not consider the identity operator $\hat{1}$ to be part of a complete commensurable set since its corresponding observable is just the number one.

[82] To be treated in your next, more advanced courses.

[83] Very often, one of the observables is the *total energy* of the system. As we shall see in our discussion of the Schrödinger equation, all observables that are commensurable with the

It is one of the peculiar features of quantum mechanics that a system may have more than one such a complete commensurable set. Those sets are then of course not commensurable : if $A_{1,2,\ldots,n}$ and $B_{1,2,\ldots,n}$ are two complete commensurable sets, then some $\hat{A}_i$ must be noncommuting with at least some of the $\hat{B}$'s, and *vice versa*[84]. We see that we have here ended up with a curious situation, one that is classically impossible ! By specifying the well-defined values of all the A's we have a *complete* description of the system ; likewise, by specifying the well-defined values of all the B's we also have a *complete* description. Nevertheless these two *complete* descriptions are *inæquivalent* : the well-defined values of the A's do *not* allow us to compute any well-defined values for the B's. As an example, we might decide to describe an electron not in terms of its position, but in terms of its (three-dimensional) velocity : these two complete commensurable sets give inæquivalent descriptions of the electron since (as we shall see) position and velocity do *not* commute.

3.5 Incommensurability and information loss

We must come to grips with the following seeming contradiction : suppose A and B are incommensurable and suppose (for simplicity) that the commutator $[\hat{A}, \hat{B}]$ never gives zero when applied to any state. Then *no* eigenstate of $\hat{A}$ is an eigenstate of $\hat{B}$, and *vice versa*. On the other hand, who can forbid us to measure A and then, very quickly afterwards, B ? To see what this situation entails, let us suppose that A has the spectrum $\{a_1, a_2, a_3, \ldots\}$ and B the spectrum $\{b_1, b_2, b_3, \ldots\}$. Let us now measure A and suppose that the result is $a = a_4$. The system is then in the eigenstate $|a_4\rangle$ of $\hat{A}$, which is *not* an eigenstate of $\hat{B}$: therefore the value of B is not well-defined. We now measure B, and find $b = b_7$, say. The system is *now* in an eigenstate of $\hat{B}$ and therefore *no longer* in an eigenstate of $\hat{A}$: the value of A is no longer well-defined and a new measurement of A may easily give the outcome $a = a_1$, say, rather than the previous result $a = a_4$. If we would repeat the series of measurements, alternating a measurement of A with one of B, we would therefore end up with a series of results that might look like this :

total energy are *conserved*, that is, their probability distributions do not change with time : certainly something very useful to know !

[84]Suppose that $\hat{B}_1$, say, is not one of, or a combination of, the $\hat{A}$'s but commutes with all of them : then the set $A_{1,2,\ldots,n}$ would not be complete since we could enlarge it by writing $B_1 = A_{n+1}$, and so on.

$$
\begin{array}{lll}
\text{the state is} & |\psi\rangle & \\
& \Downarrow & \text{measure } A\text{, with result } a = a_4 \\
\text{the state is} & |a_4\rangle & \\
& \Downarrow & \text{measure } B\text{, with result } b = b_7 \\
\text{the state is} & |b_7\rangle & \\
& \Downarrow & \text{measure } A\text{, with result } a = a_1 \\
\text{the state is} & |a_1\rangle & \\
& \Downarrow & \text{measure } B\text{, with result } b = b_{13} \\
\text{the state is} & |b_{13}\rangle & \\
& \Downarrow & \text{measure } A\text{, with result } a = a_5 \\
\text{the state is} & |a_5\rangle & \\
& \vdots &
\end{array}
$$

It is now clear what happens : by measuring B we destroy the certainty we had about the value of A, and by then re-measuring A the information we had about the value of B becomes useless, and so forth. The statement that two observables are incommensurable does not imply that we cannot *measure* them both, but rather that the results that we obtain cannot be *combined* in any useful manner. **E 19**

3.6 Commensurability and state labelling

It is useful to pause here for a moment and contrast the cases of commensurable and incommensurable observables. Let us consider a simple quantum system that has only two observables of interest, again called A and B. There are, as we have seen, two possbilities.

(I) A **and** B **are commensurable**. In this case their operators commute : $[\hat{A}, \hat{B}] = 0$. An eigenstate of $\hat{A}$, $|a_n\rangle$, is then in general degenerate since there are several independent eigenstates of $\hat{B}$ with this eigenvalue for $\hat{A}$. A *complete* labelling of the state is only obtained once we specify both the eigenvalues a_n and b_k. The system state is therefore $|a_n, b_k\rangle$, and such states are orthonormal :

$$\langle a_{n'}, b_{k'} | \, a_n, b_k \rangle = \delta_{n,n'} \, \delta_{k,k'} \quad .$$

We say that the system has *two degrees of freedom*.

(II) A **and** B **are incommensurable**. The operators do *not* commute : $[\hat{A}, \hat{B}] \neq 0$. This implies that the system has *one* degree of freedom: a single piece of information suffices to specify the state completely. For this information we can choose either the value of A or the value of B. If

we choose A, a complete description of the system is given in terms of the states $|a_n\rangle$, that are orthonormal :

$$\langle a_{n'} | a_n \rangle = \delta_{n',n} \quad ;$$

if we choose B, the states b_k *also* describes the system completely, and they, as well. form an orthonormal set :

$$\langle b_{k'} | b_k \rangle = \delta_{k',k} \quad .$$

But these two descriptions are *not* æquivalent[85] : the eigenstates of the one operator are *superpositions* of the eigenstates of the other one, so that

$$|a_n\rangle = \sum_k \beta_k \, |b_k\rangle \quad , \quad |b_k\rangle = \sum_n \alpha_n \, |a_n\rangle$$

with the α's and β complex numbers for which

$$\sum_n |\alpha_n|^2 = \sum_k |\beta_k|^2 = 1 \quad .$$

3.7 The Heisenberg inequality

Consider two observables A and B. If our system is in a normalized state $|\psi\rangle$ this implies a probability distribution $P(a)$ for the values of A, and likewise a probability distribution $P(b)$ for the values of B. As discussed before, we can compute the various expectation values and variances for the system in this state :

$$\begin{aligned} \langle A \rangle = \langle \psi | \, \hat{A} \, | \psi \rangle \quad , \quad \sigma(A)^2 = \langle \psi | \, \hat{A}^2 \, | \psi \rangle - \langle A \rangle^2 \quad , \\ \langle B \rangle = \langle \psi | \, \hat{B} \, | \psi \rangle \quad , \quad \sigma(B)^2 = \langle \psi | \, \hat{B}^2 \, | \psi \rangle - \langle B \rangle^2 \quad . \end{aligned} \tag{66}$$

Let us now assume A and B to be incommensurable. If $|\psi\rangle$ is an eigenstate of $\hat{A}$, so that A is well-defined, then $\sigma(A) = 0$; it is then impossible that $\sigma(B) = 0$. There must therefore be some relation between $\sigma(A)$ and $\sigma(B)$ which keeps them from being both zero. The *Heisenberg uncertainty relations*, which we shall now derive, provide a quantitative form of this idea.

We shall start with assuming (for ease of calculation, not for any deeper reason) that $\langle A \rangle$ and $\langle B \rangle$ are both zero[86]. This means, of course, that

$$\sigma(A)^2 = \langle \psi | \, \hat{A}^2 \, | \psi \rangle \quad , \quad \sigma(B)^2 = \langle \psi | \, \hat{B}^2 \, | \psi \rangle \quad . \tag{67}$$

For later use, we shall define an operator $\hat{C}$ by

$$[\hat{A}, \hat{B}] = i\hat{C} \quad . \tag{68}$$

[85] The idea of inæquivalent descriptions that are both *complete* has been known to boggle the mind of philosophers and theologians alike ; but not, hopefully, the physicist's mind.

[86] For any given state we can always accomplish this by redefining A and B. For instance if $\langle A \rangle \neq 0$ we could shift our attention to $A' \equiv A - \langle A \rangle$, which *does* have $\langle A' \rangle = 0$: and $\sigma(A) = \sigma(A')$.

We see that $\hat{C}$ is self-adjoint, and thus (in principle) corresponds to some observable C.

The derivation of the Heisenberg relations hinges on the clever choice of *another* quantum state, one which we build using $|\psi\rangle$, $\hat{A}$, $\hat{B}$ and an arbitrary *real* number z. This state, which we shall denote by $|\zeta(z)\rangle$, is defined as

$$|\zeta(z)\rangle \equiv \left(\hat{A} + iz\,\hat{B}\right)\,|\psi\rangle \quad . \tag{69}$$

No matter that this is an unlikely-looking expression : it is a *bona fide* quantum state, and therefore its inner product with itself must be a real, nonnegative number. Explicitly, this product reads

$$\begin{aligned}
\langle\zeta(z)|\,\zeta(z)\rangle &= \langle\psi|\left(\hat{A} + iz\,\hat{B}\right)^{\dagger}\left(\hat{A} + iz\,\hat{B}\right)|\psi\rangle \\
&= \langle\psi|\left(\hat{A} - iz\,\hat{B}\right)\left(\hat{A} + iz\,\hat{B}\right)|\psi\rangle \\
&= \langle\psi|\left(\hat{A}^2 + z^2\hat{B}^2 + iz\left(\hat{A}\hat{B} - \hat{B}\hat{A}\right)\right)|\psi\rangle \\
&= \sigma(A)^2 + z^2\sigma(B)^2 - z\,\langle C\rangle \quad .
\end{aligned} \tag{70}$$

The last line no longer contains states or operators, but only *numbers* ; it is simply a quadratic function of z which *cannot* be negative. It must therefore have a minimum value that can be found by differentiating to z :

$$\frac{d}{dz}\,\langle\zeta(z)|\,\zeta(z)\rangle = 2z\sigma(B)^2 - \langle C\rangle \quad , \tag{71}$$

and we see that the minimum is reached for $z = z_0 = \langle C\rangle\,/(2\sigma(B)^2)$. The value of this minimum is of course *also* at least zero, so that we can state that

$$\langle\zeta(z_0)|\,\zeta(z_0)\rangle = \sigma(A)^2 - \frac{\langle C\rangle^2}{4\sigma(B)^2} \geq 0 \quad . \tag{72}$$

We now have found the above-mentioned interplay between the two variances, in the form of the Heisenberg uncertainty relation[87] :

$$\sigma(A)\,\sigma(B) \;\geq\; \frac{1}{2}\left|\,\langle C\rangle\right| \quad . \tag{73}$$

As $\sigma(B)$ becomes smaller, so that the outcome of a measurement of B becomes better and better defined, the spread in measurement values for A must become larger and larger. In many important cases, $\hat{C}$ is in fact proportional to the identity so that $\langle C\rangle$ is a fixed number ; then the uncertainty relations take on an absolute form. **E 20**

[87] The word 'uncertainty' here does *not* mean that the outcome of one or more measurements are uncertain : rather, it refers to the uncertainty as another word for the spread in measurement outcomes.

3.8 Extra ! The Mermin-Peres magic square

A very instructive example of the fundamental difference between quantum observable-operators and classical observables is given by the following devious construction. Consider a quantum system with 4 orthonormal basis states[88]. In matrix notation, such states can of course be represented by column vectors with 4 components, and observables by Hermitian 4×4 matrices. Let us choose the following 9 matrices :

$$\hat{M}_{1,1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \hat{M}_{1,2} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\hat{M}_{1,3} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \hat{M}_{2,1} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\hat{M}_{2,2} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \hat{M}_{2,3} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\hat{M}_{3,1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \hat{M}_{3,2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\hat{M}_{3,3} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{74}$$

These matrices are Hermitian[89], so correspond to observables. Let us group them in the following way, in the form of three rows and three columns, the so-called *Mermin-Peres magic square* :

$\hat{M}_{1,1}$	$\hat{M}_{1,2}$	$\hat{M}_{1,3}$
$\hat{M}_{2,1}$	$\hat{M}_{2,2}$	$\hat{M}_{2,3}$
$\hat{M}_{3,1}$	$\hat{M}_{3,2}$	$\hat{M}_{3,3}$

[88] Such a system is usually discussed as consiting of two electrons, the spin of each of which can take two values. In fact, this is totally unimportant : the important fact is that there are 4 physically distinct states.

[89] In fact, real and symmetric.

It is easy to check that every one of these matrices has unit square :

$$\hat{M}_{k,j}^2 = \hat{1} \quad k,j = 1,2,3 \tag{75}$$

so that their eigenvalues are either +1 or -1. The nine observables corresponding to these matrices[90] therefore can take the values +1 or -1 upon observation. Furthermore, you can also verify that the operators in each row, and in each column, of the square, are mutually *commuting* : **E 21**

$$\begin{aligned} &[\hat{M}_{k,1}, \hat{M}_{k,2}] = [\hat{M}_{k,1}, \hat{M}_{k,3}] = [\hat{M}_{k,2}, \hat{M}_{k,3}] = 0 \quad , \\ &[\hat{M}_{1,k}, \hat{M}_{2,k}] = [\hat{M}_{1,k}, \hat{M}_{3,k}] = [\hat{M}_{2,k}, \hat{M}_{3,k}] = 0 \quad , \quad k = 1,2,3 \end{aligned} \tag{76}$$

so that each row, and each column, corresponds to three *commensurable* observables. In addition, you can see that the product of the matrices in each row[91] is equal to *one*, and the product in each column is equal to *minus one* :

$$\begin{aligned} &\hat{M}_{k,1}\hat{M}_{k,2}\hat{M}_{k,3} = \hat{1} \quad , \\ &\hat{M}_{1,k}\hat{M}_{2,k}\hat{M}_{3,k} = -\hat{1} \quad , \quad k = 1,2,3 \end{aligned} \tag{77}$$

So far, this is simply a (maybe vaguely interesting) system of observable-operators : for our system in a given state, we can compute expectation values, uncertainties, and what have you. The crucial step comes now. Imagine that we try to describe our system by classical observables rather than quantum ones. Obviously, such observables must take on the value ± 1 only, otherwise they could not possibly mimic the quantum system. The observables along each row and column in our square are commensurable, just what you want for classical observables. However it is *impossible* to fill a 3×3 square with only 1's or -1's in such a way that the product along each row is 1 while simultaneously the product along each column is -1 ! We have here a system that has *no classical counterpart* no matter what state it is in[92]. The quantum observable-operators simply have a much richer mathematical structure than the classical observables.

3.9 Exercises 17 to 21

Exercise 17 : Properties of commutators

[90] Whatever they are!

[91] Because of the mutual commutation, the order in which we multiply the matrices is irrelevant.

[92] This is of course due to the devious choice of the matrices, but also to the fact that we have 3×3 of them. For instance, the squares

1	1
-1	-1

and

1	1	-1	-1
1	1	1	1
1	1	1	1
-1	-1	1	1

do have products 1 along rows and -1 along columns.

1. Prove the identities given in Eqs.(58)-(63).

2. We can extend the last of these identities as follows. Let $\hat{A}$ and $\hat{B}_{1,2,3,\ldots,N}$ be operators : then

$$\begin{aligned}[\hat{A}, \hat{B}_1\hat{B}_2\hat{B}_3\hat{B}_4\cdots\hat{B}_N] = \\ = \quad [\hat{A}, \hat{B}_1]\,\hat{B}_2\hat{B}_3\hat{B}_4\cdots\hat{B}_N + \hat{B}_1\,[\hat{A}, \hat{B}_2]\,\hat{B}_3\hat{B}_4\cdots\hat{B}_N \\ +\hat{B}_1\hat{B}_2\,[\hat{A}, \hat{B}_3]\,\hat{B}_4\cdots\hat{B}_N + \cdots + \hat{B}_1\hat{B}_2\cdots\hat{B}_{N-1}, [\hat{A}, \hat{B}_N]\end{aligned}$$

   where the operator $\hat{A}$ is seen to 'move through' the product of the $\hat{B}$'s. Prove this.

3. Assume that two operators, $\hat{X}$ and $\hat{P}$, have a commutator that is constant *i.e.* it is proportional to the identity operator :

$$[\hat{X}, \hat{P}] = k$$

   Use the previous result to prove that

$$[\hat{X}, \hat{P}^n] = n\,k\,\hat{P}^{n-1} \quad , \quad [\hat{X}^n, \hat{P}] = n\,k\,\hat{X}^{n-1}$$

Exercise 18 : **An important set of operators**
Consider a quantum system with two basis states (the two-position model is an example). For this system we consider the three observables $M_{x,y,z}$ introduced in exercise 15, with their operators as given in that exercise.

1. Prove that

$$\hat{M}_x^2 = \hat{M}_y^2 = \hat{M}_z^2 = \hat{1}$$

2. Prove the following commutation relations:

$$[\hat{M}_x, \hat{M}_y] = 2i\hat{M}_z \quad , \quad [\hat{M}_y, \hat{M}_z] = 2i\hat{M}_x \quad , \quad [\hat{M}_z, \hat{M}_x] = 2i\hat{M}_y$$

These operators are very important in the description of the electron spin.

Exercise 19 : **Alternating measurements**
Consider a system with two basis states. We can study the Hermitean operators $\hat{M}_{x,y,z}$ of exercise 18. We shall contemplate measuring M_x, M_y and M_z on this system.

1. Compute the eigenvectors of $\hat{M}_x$, $\hat{M}_y$, and $\hat{M}_z$ (the eigenvalues have already been computed in exercise 15).

2. Assume that a measurement of M_z has been succesfully made. Compute the probability distribution $P(m_x)$ for the values of M_x, and the probability distribution $P(m_y)$ for the values of M_y, after this measurement.

3. Do the analogous calculation assuming a measurement of M_x has been made ; and again, assuming a measurement of M_y has been made.

Exercise 20 : **An explicit Heisenberg inequality**
We assume the quantum system to be in a state $|\chi\rangle$ with matrix notation

$$|\chi\rangle = \begin{pmatrix} a\, e^{i\phi_a} \\ b\, e^{i\phi_b} \end{pmatrix}$$

where a, b, and $\phi_{a,b}$ are real numbers. We again consider the observables $M_{x,y,z}$.

1. Show that the state $|\chi\rangle$ is normalized if $a^2 + b^2 = 1$.
2. Compute the expectation values $\langle M_z\rangle$, $\langle M_x\rangle$ and $\langle M_y\rangle$.
3. Compute the variances $\sigma(M_x)^2$ and $\sigma(M_y)^2$.
4. Show that the Heisenberg inequality for M_x and M_z can be written as

$$\sigma(M_z)^2\sigma(M_x)^2 - \langle M_y\rangle^2 \geq 0$$

Express this inequality in terms of a, b, and $\phi_{a,b}$, and prove it by explicit inspection ; also, show that the inequality becomes an equality for $a = 0$, $b = 0$ or $\sin(2(\phi_a - \phi_b)) = 0$.

Exercise 21 : **Multiple multiplications**
Check the statements about the Mermin-Peres matrices by explicit matrix multiplication and commutation.

4 Does reality have properties ?

— For me, it is so reasonable to assume that the photons in those experiments carry with them programs, which have been correlated in advance, telling them how to behave. This is so rational that I think that when Einstein saw that, and the others refused to see it, he was the rational man. The other people, although history has justified them, were burying their heads in the sand... So for me, it is a pity that Einstein's idea doesn't work. The reasonable thing just doesn't work.
John S. Bell

4.1 The question at issue

We have seen that at the moment of measurement[93] the state of a system collapses into another state, which predicts a definite outcome for that measurement. For instance, consider the polaroid experiment discussed in chapter 1. A photon can be in any polarization state $|\psi\rangle$, but once it has been observed to have passed a polarization filter oriented vertically (defined as an angle of 0 degrees) it is definitely in a state $|0\rangle$[94]. After that, it will *certainly* not pass through a polarimeter oriented at $\pi/2$. The philosphical question[95] is then : did the photon know, somehow, *in advance*, how to react to the polarimeter ? If yes, we could say that the photon had a, somehow, *predetermined* idea of its polarization ; if no, its polarization must have come about, in some way, at the moment of measurement. To put it bluntly : does reality have properties of it own, without us looking at it ?

4.2 The EPR paradox

4.2.1 EPR's thought experiment

In 1935, Einstein, Podolsky, and Rosen proposed a thought experiment. Without going into their parlance, we can reformulate it into a slightly better[96] setting. We consider an atom in some excited state, which decays by emitting two photons. In terms of polarization, the two-photon state is[97]

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(\,|1=v, 2=h\rangle - |1=h, 2=v\rangle\Big) \quad ,$$

[93] Or somewhere around that time — this is exactly one of the unresolved questions.
[94] The entry in the ket will, here, denote the orientation of the polarimeter.
[95] *i.e.* a question asked by idling philosphers rather than by bread-earning physisicts.
[96] Better adapted to the language of these courses.
[97] Do not worry about technology : such atomic systems, and their two-photon decays into this particular two-photon state, are eminently feasible.

where h denotes horizontal, and v vertical polarization[98] : the labels 1 and 2 refer to the photons moving in two opposite directions. Now suppose that the two photons move apart (at the speed of light !) for some time (a second, say), and then photon 1 is passed through a polarimeter oriented v. Supose it passes. Then photon 2 is certainly in a state $|h\rangle$: we have performed a measurement on photon 2, without touching it in any way ! And that photon is 600,000 kilometers away ! And nothing, *no information at all*, can move faster than the speed of light to inform photon 2 that it has, actually, been measured ! So quantum mechanics is not the whole story : the photons must have had some 'secret instructions' before they set out. These 'secret instructions' are called *hidden variables*. Einstein and his cronies did not concoct this set-up to show that quantum mechanics is *wrong* : rather, they aimed at proving that it is *incomplete*, and something like hidden variables (an unknown, and possibly unknoweable, additional physical ingredient in the theory) is necessary to allow for correlations that might otherwise seem to propagate at speeds faster than that of light. At the time, this was considered a very strong argument indeed.

4.2.2 Entangled states

The basic idea behind EPR is that the two photons are in such a state that information about the one implies information about the other. In constrast, if the state were simply $|1 = h, 2 = v\rangle$ this would not be the case, and measuring one photon would not allow us to improve a prediction for the outcome of the measurement on the other photon, so that we might simply write

$$|1 = h, 2 = v\rangle = |1 = h\rangle\,|2 = v\rangle \quad .$$

States such as the EPR one are called *entangled* : they cannot be written in such a 'factorizing' form, as we can easily see. Suppose that we consider states $|v'\rangle$ and $|h'\rangle$ which are appropriate to a situation were we rotate the polarimeters over some angle α. As we have seen in Eq.(24), we have

$$\begin{aligned} |v'\rangle &= \cos\alpha\,|v\rangle + \sin\alpha\,|h\rangle \quad , \\ |h'\rangle &= -\sin\alpha\,|v\rangle + \cos\alpha\,|h\rangle \quad , \end{aligned} \tag{78}$$

and, conversely,

$$\begin{aligned} |v\rangle &= \cos\alpha\,|v'\rangle - \sin\alpha\,|h'\rangle \quad , \\ |h\rangle &= \sin\alpha\,|v'\rangle + \cos\alpha\,|h'\rangle \quad . \end{aligned} \tag{79}$$

It is then a trivial matter to check that

$$\Big(|1 = v, 2 = h\rangle - |1 = h, 2 = v\rangle\Big) = \Big(|1 = v', 2 = h'\rangle - |1 = h', 2 = v'\rangle\Big) \quad : \tag{80}$$

no angular redefinition of the states' basis can get rid of the entanglement.

[98] In the sense of certainly passing through a suitably oriented polarimeter.

4.3 Bell's inequalities

In 1964, John Bell of CERN put paid to the EPR 'paradox' by taking the hidden-variables idea seriously. In one of the many ways this can be done, let us formulate things as follows.

4.3.1 The experiment : Obama *vs* Osama

Let us assume that there is a source of photons pairs, regularly emitting two photons in the state 4.2.1. Now we take two observers, Obama[99] and Osama[100] who each are at one end of the system. Obama holds one polarimeter up to his photons, and records whether they pass or not, accordingly noting down a '1' for passing and a '0', for not doing so : on the other side, Osama does the same with *his* polarimeter. We then obtain Obama's list of 1's and 0's (O_1) and Osama's list (O_2), and (after the lists have been brought together[101]) we can compute the products of their results. This will look something like this :

O_1	1	0	1	0	0	1	0	1	1	...
O_2	1	1	1	1	0	0	1	1	0	...
O_1O_2	1	0	1	0	0	0	0	1	0	...

After a great number of observations, we can then compute the average value of O_1O_2. For an angular setting θ_1 of Obama's polarimeter, and a setting θ_2 of Osama's one, we call this average $C(\theta_1, \theta_2)$ (equal, of course, to $C(\theta_2, \theta_1)$).

4.3.2 Hidden variables : Bell's inequality

Now suppose EPR's idea of hidden variables is actually correct. This means that Obama's photons carry, in addition to the quantum state, some hidden variable(s) w that tell it how to react to a polarimeter oriented at an angle θ, namely whether to pass (1) or to be absorbed (0). That, is, there must be some (as yet unknown) function of w and the polarimeter setting θ that results in either 0 or 1 :

$$A(\theta_1, w) = 0 \text{ or } 1 \ . \tag{81}$$

Similarly, Osama's photons carry a similar set of instructions :

$$B(\theta_2, w) = 0 \text{ or } 1 \ . \tag{82}$$

Obviously, we do not *know* these secret instructions : indeed, we do not have slightest idea about w, whether it is just a simple number, or a set of numbers, or whatever. The whole point is that w must be the same for Obama's and Osama's photons, since they are 'assigned' their instructions at the moment of

[99] Obviously, named after Ricardo Mangue Obama Nfubea (1961), prime minister of Equatorial Guinea from 2006 to 2008.

[100] Named, of course, after Majd ad-Din Osama ibn Murshid ibn 'Ali ibn Munqidh al-Kinani (1095-1188), poet, author, warrior and diplomat.

[101] This is really important !

their creation[102]. Since not all observations give the same answer, we must also conclude that the value(s) of w differ from event to event, and therefore the w's must be governed by some distribution[103], $\rho(w)$, for which, obviously[104],

$$\rho(w) \geq 0 \quad , \quad \int \rho(w)\, dw = 1 \ . \tag{83}$$

We have one fact to trade on : namely, that if Obama's and Osama's settings are equal, then if Obama's photon passes, Osama's will not, and *vice versa* : therefore,

$$B(\theta, w) = 1 - A(\theta, w) \ . \tag{84}$$

Now, we are in a position to say something about $C(\theta_1, \theta_2)$, namely that it must be given by

$$\begin{aligned} C(\theta_1, \theta_2) &= \int \rho(w)\ A(\theta_1, w)\, B(\theta_2, w)\, dw \\ &= \int \rho(w)\ A(\theta_1, w)\ \Big(1 - A(\theta_2, w)\Big)\ dw \\ &= \int \rho(w)\ A(\theta_1, w)\ \ \mathtt{XOR}\Big(A(\theta_1, w), A(\theta_2, w)\Big)\ dw \ . \end{aligned} \tag{85}$$

Here we have introduced the function `XOR`:

$$\mathtt{XOR}(a, b) = \begin{cases} 1 & \text{if } a \neq b \\ 0 & \text{if } a = b \end{cases} \tag{86}$$

Bell's ingenious idea is to compare this combination of settings with *another* such combination :

$$\begin{aligned} C(\theta_1, \theta_2) - C(\theta_1, \theta_3) &= \int \rho(w)\ A(\theta_1, w)\ \Big(- A(\theta_2, w) + A(\theta_3, w)\Big)\ dw \\ &= \int \rho(w)\ A(\theta_1, w)\ \Big(2A(\theta_3, w) - 1\Big)\ \ \mathtt{XOR}\Big(A(\theta_2, w), A(\theta_3, w)\Big)\ dw \end{aligned} \tag{87}$$

This last line can be verified by simply checking the possible values of the A's. Now, note that $\rho(w)$ and all A's are never negative. Then, if we replace $2A(\theta_3, w) - 1$ (which is 1 or -1) by $A(\theta_3, w)$ (which is 1 or 0), the integral can

[102] You might argue that the two photons carry different hidden variables, w_1 and w_2, say. But that is simply evaded by combining w_1 and w_2 into a larger hidden variable w. It is as if both photons are given, at their birth, an identical booklet with instructions, including the following one : "if you are photon 1, follow the instructions on page 4-10, if you are photon 2, follow the instructions on page 11-17".

[103] Whether this is some probability distribution (although this, of course, would pain Einstein), or obeys some regular law, or is, indeed, determined from far off by an alien intelligence, does not matter in the slightest ; the important fact is that the w's are determined *somehow*.

[104] For the first property, note that $\rho(w)$ is the *fraction of the number of times* that this w comes up and thus cannot be negative ; the second property follows from the fact that every event must carry *some* value of w.

never decrease :

$$C(\theta_1,\theta_2) - C(\theta_1,\theta_3) \leq \int \rho(w)\; A(\theta_1,w)\, A(\theta_3,w)\, \texttt{XOR}\Big(A(\theta_2,w), A(\theta_3,w)\Big) \quad , \tag{88}$$

Similarly, if we replace $A(\theta_1, w)$ by 1 the integral can never decrease, and we find

$$C(\theta_1,\theta_2) - C(\theta_1,\theta_3) \leq C(\theta_2,\theta_3) \quad , \tag{89}$$

or

$$C(\theta_1,\theta_2) \leq C(\theta_1,\theta_3) + C(\theta_3,\theta_2) \quad : \tag{90}$$

the triangle inequality ! This is (a form of) Bell's inequality. Notice its generality : we do not have to know anything about the w's or the A's !

4.3.3 QM's predictions

Using what we have learned about quantum mechanics, we can of course obtain our own prediction of $C(\theta_1, \theta_2)$. Let us call $|v\rangle$ the state of a photon sure to pass Obama's polarimeter, oriented at angle θ_1, and v' the state of one sure to pass Osama's polarimeter, oriented at and angle θ_2, and let us denote $\theta_1 - \theta_2$ by α. Now, since Obama's measurement will assign 1 to a photon in the state $|v\rangle$ and 0 to one in the state $|h\rangle$, its operator $\hat{O}_1$ must read

$$\hat{O}_1 = |1 = v\rangle \langle 1 = v| \quad . \tag{91}$$

Similarly, Osama's observable-operator must read

$$\hat{O}_2 = |2 = v'\rangle \langle 2 = v'| \quad . \tag{92}$$

Their combined observable $O_1 O_2$ therefore has the operator $\hat{O}_1 \hat{O}_2$, with the understanding that $\hat{O}_1$ acts on the state of photon 1, and $\hat{O}_2$ on that of photon 2[105]. The expectation value $C(\theta_1, \theta_2)$ is, for the two-photon system in a state $|\psi\rangle$, of course given by

$$C(\theta_1,\theta_2) = \langle\psi|\, \hat{O}_1 \hat{O}_2\, |\psi\rangle \quad , \tag{93}$$

and we shall now compute this for

$$|\psi\rangle = \frac{1}{\sqrt{2}} \Big(|1 = v, 2 = h\rangle - |1 = h, 2 = v\rangle \Big) \quad . \tag{94}$$

In the first place, we immediately see that

$$\hat{O}_1\, |\psi\rangle = \frac{1}{\sqrt{2}}\, |1 = v, 2 = h\rangle \quad , \tag{95}$$

[105] It follows that $\hat{O}_1$ and $\hat{O}_2$ commute. This is a good thing, since from special relativity we know that, in the given experimental setup, it is *fundamentally impossible* to decide whether Obama or Osama is the first to make the measurement ; it depends on the state of motion of the referee observing them !

so that

$$\begin{aligned}\hat{O}_2\hat{O}_1\,|\psi\rangle &= \frac{1}{\sqrt{2}}\,|2=v'\rangle\,\langle 2=v'|\,|1=v,2=h\rangle \\ &= \frac{1}{\sqrt{2}}\,\langle 2=v'|\,2=h\rangle\;\;|2=v',1=v\rangle \\ &= \frac{\sin\alpha}{\sqrt{2}}\;\;|2=v',1=v\rangle\quad . \end{aligned} \tag{96}$$

Finally, multiplying this with $\langle\psi|$ we find

$$\langle\psi|\,\hat{O}_2\hat{O}_1\,|\psi\rangle = \frac{\sin\alpha}{2}\,\langle 1=v|\,1=v\rangle\;\;\langle 2=h|\,2=v'\rangle = \frac{1}{2}(\sin\alpha)^2\quad . \tag{97}$$

This, then, is quantum mechanics' prediction[106] of $C(\theta_1,\theta_2)$. Is it compatible with Bell's inequality ? It is easy to find a situation in which it is certainly *not* : for instance, choose

$$\theta_1 = 0\quad ,\quad \theta_3 = \frac{\pi}{6}\quad ,\quad \theta_2 = \frac{\pi}{3}\quad :$$

then

$$C(\theta_1,\theta_3) = C(\theta_3,\theta_2) = \frac{1}{2}\sin\left(\frac{\pi}{6}\right)^2 = \frac{1}{8}\quad , \tag{98}$$

while

$$C(\theta_1,\theta_2) = \frac{1}{2}\sin\left(\frac{\pi}{3}\right)^2 = \frac{3}{8}\quad , \tag{99}$$

and the inequality is *violated*. The conclusion is inescapable : if we take quantum mechanics seriously, then we cannot say that (at least this part of) reality has properties 'of its own', that are independent of whether a measurement is made on it[107].

4.3.4 Faster than light ?

You might be tempted to think that we have gotten ourselves into a nice mess at this point. EPR advocated *realism*, that is, systems having properties 'of their own', encoded by hidden variables. They used the idea of *locality* to arrive at this conclusion : no influences are supposed to be able to move faster than light[108]. Quantum mechanics appears to tell us exactly the contrary : hidden variables, and hence realism, are down the drain, and locality as well ! Things are not so

[106] This is, of course, compatible with what you might expect. For instance, if $\alpha = 0$ or π then Obama's 1 will always result in Osama's 0 and *vice versa*; for $\alpha = \pi/2$ they will both find either 1 or 0.

[107] Such a discussion simply cries out for an experiment ! Indeed, experiments (notably those of Alain Aspect and collaborators) show the expected violations of Bell's inequality. Nowadays (at least in my view) arguing in favor of hidden variables is a rearguard action at best.

[108] Like most of us, Einstein was uncomfortable with the idea of an immediately-acting influence-at-a-distance, for which he coined the wonderfully poetic phrase *Spukhafte Fernwirkung*.

bleak, however. In the first place, the violations of locality (namely the fact that Osama's photons are influenced as soon as Obama has made a measurement, and the other way around) can only be seen once we have actually computed the average of $O_1 O_2$, and this can only be done once the lists of O_1 and O_2 have been brought together, by radio signals, by fast cars, or by snail : at any rate this cannot happen faster than with the speed of light. In the second place, locality does not state that there are no phenomena moving at faster than light speed, but rather that *information* and energy (which is, in fact, just a form of information) cannot propagate faster than light. The question is therefore : can Obama use his measurements to send information to Osama ?

Let us assume, again, that the two-photon state $|\psi\rangle$ of Eq.(94) presents itself to Obama, but that he declines to make any measurement. The state at Osama's side is therefore, by Eq.(80),

$$|\xi\rangle = \frac{1}{\sqrt{2}} \Big(|1 = v', 2 = h'\rangle - |1 = h', 2 = v'\rangle \Big) \quad , \tag{100}$$

and Osama's average is, using Eq.(92), given by

$$\langle O_2 \rangle = \langle \xi | \, \hat{O}_2 \, | \xi \rangle = \frac{1}{2} \quad , \tag{101}$$

as you might have guessed. Now, suppose that Obama *does* make the measurement. Since also $\langle O_1 \rangle = 1/2$, the state after Obama's measurement will be $|\psi_1\rangle = |1 = v, 2 = h\rangle$ or $|\psi_2\rangle = |1 = h, 2 = v\rangle$, each in 50% of the cases. Now Osama's average has to be recomputed : it is given by

$$\begin{aligned} \langle O_2 \rangle &= \frac{1}{2} \langle \psi_1 | \, \hat{O}_2 \, | \psi_1 \rangle + \frac{1}{2} \langle \psi_2 | \, \hat{O}_2 \, | \psi_2 \rangle \\ &= \frac{1}{2} \langle 2 = h | \, 2 = v' \rangle \, \langle 2 = v' | \, 2 = h \rangle + \frac{1}{2} \langle 2 = v | \, 2 = v' \rangle \, \langle 2 = v' | \, 2 = v \rangle \\ &= \frac{1}{2} \Big((\sin \alpha)^2 + (\cos \alpha)^2 \Big) = \frac{1}{2} \quad ! \end{aligned} \tag{102}$$

Clearly, Osama will, by just looking at his list of 1's and 0's, not be able to tell whether Obama made a measurement or not. The EPR setup cannot be used for superluminal signalling[109].

4.3.5 Extra ! A loophole ? The no-clone theorem

The essential reason why Osama cannot see whether Obama did a measurement or not lies in the fact that a photon's polarization can be measured only *once* : after the measurement, its state is no longer what it was, in general. Suppose, for the sake of argument, that Osama *were* able to repeat his measurement without destroying the state ; then he would be able to distinguish between

[109] Again, we repeat that this is a good thing, since if it *were* possible, signals would, from some observers' point of view, actually arrive *before* they were sent !

the state $|\psi\rangle$ of Eq.(100), and the states $|\psi_1\rangle$ and $|\psi_2\rangle$[110]. In that way, *one* bit of information would have been transferred from Obama to Osama, namely whether Obama did the measurement or did not.

Could there not exist a 'quantum copier', which makes copies of states ? In that case Osama could quickly make many copies of the state of the photon he received, and pass them all through his polarimeter, and so find out what the original photon's state actually was. Through that loophole, faster-than-light signaling might be possible, after all !

Such a 'quantum copier' would be an operator[111] $\hat{X}$, that takes as argument an 'original' state, $|a\rangle$, to be copied, and an arbitrary state $|s\rangle$, and yield as output twice the state $|a\rangle$:

$$|a\rangle\,|s\rangle \stackrel{\hat{X}}{\to} |a\rangle\,|a\rangle \quad . \tag{103}$$

The net effect is that state $|s\rangle$ has been transformed into state $|a\rangle$. Obviously, to be a good copier, it should allow for many possible originals :

$$|a_1\rangle\,|s\rangle \stackrel{\hat{X}}{\to} |a_1\rangle\,|a_1\rangle \quad , \quad |a_2\rangle\,|s\rangle \stackrel{\hat{X}}{\to} |a_2\rangle\,|a_2\rangle \quad . \tag{104}$$

Now, suppose that the original submitted to the copier is the superposition

$$z_1\,|a_1\rangle + z_2\,|a_2\rangle \quad .$$

Assuming linearity[112], we then have the effect

$$\Big(z_1\,|a_1\rangle + z_2\,|a_2\rangle\Big)\,|s\rangle \stackrel{\hat{X}}{\to} z_1\,|a_1\rangle\,|s\rangle + z_2\,|a_2\rangle\,|s\rangle \quad . \tag{105}$$

But this is not good ! What we were aiming for was, of course,

$$\Big(z_1\,|a_1\rangle + z_2\,|a_2\rangle\Big)\,|s\rangle \stackrel{\hat{X}}{\to} \Big(z_1\,|a_1\rangle + z_2\,|a_2\rangle\Big)\,\Big(z_1\,|a_1\rangle + z_2\,|a_2\rangle\Big) \quad . \tag{106}$$

We conclude, that even if there would exist a quantum cloner that makes perfect copies of $|\psi_1\rangle$ and of $|\psi_2\rangle$, it would be hopeless at making copies of their superposition $|\xi\rangle$, and poor Osama is doomed to ignorance after all[113] : the no-clone theorem makes superluminal communication impossible[114].

[110] For instance, let $\alpha = 0$ so that $v = v'$ and $h = h'$. Then, if the state were $|\psi_1\rangle$, Osama would find 0 *every* time ; for $|\psi_2\rangle$, he would find 1 *every* time ; but for $|\psi\rangle$ he would find 0 as often as 1.

[111] You are allowed 1 guess as to why I use this letter.

[112] Obviously, the quantum copier cannot involve a measurement, since that destroys states. In the absence of state collapse, all time evolution in quantum mechanics is indeed linear (as we shall see in the next chapter), in the sense that the time evolution of a superposition of states is the superposition of the time-evolving states.

[113] To avoid confusion, the following. Once you *know* the state of a system, there is of course nothing to keep you from making many copies of that system in the same state, if your technology, patience, and budget allows. The real problem is that you cannot, by making a *single* measurement, figure out what *was* the quantum state of the system *before* the measurement, but only what *is* the quantum state *after* the measurement.

[114] Farewell, dream of many a science-fiction author...

5 The Schrödinger equation

— I hope I can report soon in a little more detailed and understandable way about the matter. At present, I must learn a little more mathematics [...]
Erwin R.J.A. Schrödinger, cited in *"Schrödinger"* by Walter J. Moore

5.1 The Schrödinger equation

Physics is no good if we can only describe what we see at a given moment : we ought to also be able to make predictions ! Therefore we need the quantum æquivalent of Newton's laws, an æquivalent that describes the *time evolution* of states. This means that we have to consider *time-dependent* states. Although it is not very usual, we occasionally denote the time dependence of a state explicitly by not writing $|\psi\rangle$ but $|\psi\rangle(t)$: hopefully this use will be clear from the context. The time-evolution equation of nonrelativistic[115] quantum mechanics is the *Schrödinger equation* :

$$i\hbar \frac{d}{dt} |\psi\rangle(t) = \hat{H} \; |\psi\rangle(t) \quad , \tag{107}$$

where $\hat{H}$ is the *Hamiltonian* operator of the system ; its observable is nothing else than the *total energy* of the system[116]. The fact that $\hbar$, which makes its appearance here for the first time, has the units of energy×time allows the dimensions on both sides to work out correctly[117]. Since it involves an explicit time dependence, the Schrödinger equation is also sometimes called the *time-dependent Schrödinger equation*. **E 22**

It is very important to stress, once more, that the Schrödinger equation is not a *provable* equation, but that by proposing it Schrödinger sticks his neck out, ready to be disproved by experiment — except that he hasn't been so far[118].

Finally, it must be remarked that the Schrödinger equation can*not* describe the collapse of states upon observation : what happens during measurement is a

[115] In fact, Schrödinger originally set out to write a *relativistic* equation but for various reasons (which you will learn about in later courses) this was not a complete succes : he therefore opted for the nonrelativistic version which bears his name, and got him his Nobel prize.

[116] This is a, historically grown, exception to our notation convention since the observable is usually denoted by E, and its value is denoted by the same capital letter as well.

[117] In the physicist's everyday parlance, $\hbar$ is called 'Planck's constant' but this is not really correct. Planck introduced a constant h but it later became clear that the combination $h/(2\pi)$ occurs very frequently, and the notation $\hbar \equiv h/(2\pi)$, which seems due to Dirac, is much the more common tool. It should properly be called the *reduced Planck's constant*, or *Dirac's constant* : but it is considered very pedantic to actually use these names. At any rate, since h is a very dull letter and $\hbar$ a more sexy symbol which is, moreover, unique to quantum physics, it is $\hbar$ that has become the corporate logo of 'quantum', in the same way that the object $\sqrt{1 - v^2/c^2}$ is an infallible sign that special relativity is being perpetrated. In short : argue about this, and be a nerd !

[118] That is, as far as we limit ourselves to *nonrelativistic* physics.

fundamentally different phenomenon[119]. The Schrödinger equation tells us hat happens *in between* measurements.

5.2 Preservation of the norm

The Schrödinger equation for kets can of course also be turned into one for bras, by conjugation :

$$-i\hbar \frac{d}{dt} \langle\psi| = \langle\psi| \hat{H} \quad . \tag{108}$$

This allows us to work out how the inner product of a state with itself (the 'square' of the state) evolves with time :

$$\begin{aligned} \frac{d}{dt} \langle\psi|\psi\rangle &= \left(\frac{d}{dt}\langle\psi|\right)|\psi\rangle + \langle\psi|\left(\frac{d}{dt}|\psi\rangle\right) \\ &= \left(\frac{i}{\hbar}\langle\psi|\hat{H}\right)|\psi\rangle + \langle\psi|\left(-\frac{i}{\hbar}\hat{H}|\psi\rangle\right) \\ &= \frac{i}{\hbar}\langle\psi|\left(\hat{H}-\hat{H}\right)|\psi\rangle = 0 \quad . \end{aligned} \tag{109}$$

The square of all states is constant in time : if a state is normalized to unity, it remains so ! Note that this is owed to the occurrence of 'i' in the Schrödinger equation, which is there for precisely *this* reason.

5.3 The time-independent Schrödinger equation

The Schrödinger equation is an attractive equation : simple, linear, and — importantly — of *first order*. This means that we only have to know the state $|\psi\rangle(t_0)$ at some time t_0, and *not* also its 'velocity' $d|\psi\rangle(t_0)/dt_0$ in order to predict its future. On the other hand, all this does not mean that it is always *easy* to solve the Schrödinger equation.

We shall now describe an approach that is almost universally used. In this approach, we first look for solutions to what is called the time-*in*dependent Schrödinger equation :

$$\hat{H}|\psi\rangle = E|\psi\rangle \quad . \tag{110}$$

It is, in fact, is nothing else than the eigenvalue equation for $\hat{H}$. The eigenvalue E is then, of course, the value of the system's total energy[120]. Solving the time-independent Schrödinger equation may already be difficult enough, but at least it doesn't involve time as well.

[119] As we have mentioned before, the assumption of isolatedness of the system fails at precisely that instant.

[120] A word of caution is in order here. From the usage of the words, it might seem that there are two Schrödinger equations, a time-dependent and a time-independent one, implying two types of physics. This is *not* true : there is only *one* Schrödinger equation, namely that of Eq.(108). The time-independent Schrödinger equation is just a tool to help with its solution. Unfortunately, many elementary texts treat the time-independent Schrödinger equation as the real thing, but beware ! It only tells you the energy values of the system, and not how the system behaves. This really confused me when I was an undergraduate student.

Suppose that the spectrum of E consists of the values $\{E_1, E_2, E_3, \ldots\}$, and that we have found these values, together with the corresponding eigenstates $|E_1\rangle, |E_2\rangle, |E_3\rangle, \ldots$. That is, we have the complete solutions of

$$\hat{H}\ |E_n\rangle = E_n\ |E_n\rangle \quad , \quad n = 1, 2, 3, \ldots \tag{111}$$

Of course, many (or most) of the energy eigenvalues E_n may be degenerate, so that there are more than just a single $|E_n\rangle$ for given n : we shall simply assume that we have found them all, and allow for a p-fold degenerate value of E by including that value of E precisely p times in the spectrum[121].

For a single energy eigenstate, the Schrödinger equation can be solved in an almost trivial manner : since

$$i\hbar \frac{d}{dt}\ |E_n\rangle = \hat{H}\ |E_n\rangle = E_n\ |E_n\rangle \quad , \tag{112}$$

where the last lemma does not contain an *operator* $\hat{H}$ but only the *real number* E_n, the correct time dependence is simply

$$|E_n\rangle\,(t) = e^{-iE_n t/\hbar}\ |E_n\rangle\,(0) \quad . \tag{113}$$

We see that the time dependence of an energy eigenstate is very simple : the state just rocks around the clock in complex space with a fixed speed that depends on its energy. Such an overall complex phase cannot influence the physics of a single energy eigenstate ; but since different energy eigenstates evolve with different speeds things may be quite another matter for a *superposition* of energy eigenstates.

We can now find the solution to the Schrödinger equation. Suppose that at time $t = 0$ we start with a given quantum state $|\psi\rangle\,(0)$. Since the Hamiltonian must have an orthonormal and complete basis of energy eigenstates, we can write $|\psi\rangle\,(0)$ as a superposition of these basis states :

$$|\psi\rangle\,(0) = \sum_n\ c_n\ |E_n\rangle\,(0) \quad . \tag{114}$$

The trick to find $|\psi\rangle\,(t)$ is now to simply replace the $|E_n\rangle\,(0)$ by *their* time-dependent forms $|E_n\rangle\,(t)$. More explicitly, we have the expression

$$|\psi\rangle\,(t) = \sum_n\ e^{-iE_n t/\hbar}\, c_n\ |E_n\rangle\,(0) \quad . \tag{115}$$

We see that the solutions to the time-independent Schrödinger equation provide also solutions to the full Schrödinger equation : once the time-independent Schrödinger equation has been solved, the rest of the job is fairly *straightforward* — altough seldomly *easy* ! **E 23**

[121] This may sound more daunting than it is : the following example shows you how it is done. Suppose there are three energy eigenvalues, E_a, E_b and E_c, and that E_b is two-fold degenerate. That is, there is one eigenstate $|E_a\rangle$, and one eigenstate $|E_c\rangle$, but two eigenstates for E_b, which we denote by $\left|E_{b,1}\right\rangle$ and $\left|E_{b,2}\right\rangle$. We then simply say that the spectrum is $\{E_1, E_2, E_3, E_4\}$, with $E_1 = E_a$, $E_2 = E_b$, $E_3 = E_b$, $E_4 = E_c$; and $|E_1\rangle = |E_a\rangle$, $|E_2\rangle = \left|E_{b,1}\right\rangle$, $|E_3\rangle = \left|E_{b,2}\right\rangle$, $|E_4\rangle = |E_c\rangle$.

5.4 Time evolution : the Heisenberg equation

The fully-fledged, time-evolving expression for a complicated quantum state is of course a beautiful thing to have — but it is not very often what we want. After all, what we can *measure* about a state are the various *probabilities* ; so therefore it might be even more directly useful to see how these probabilities evolve with time under the Schrödinger equation. More in particular, let us study the time evolution of *expectation values*. We can do this in a way similar to that by which we came to. Eq.(109). The expectation value of an observable A for a system in the state $|\psi\rangle$ is given, as we have seen, by

$$\langle A \rangle = \langle \psi | \, \hat{A} \, | \psi \rangle \quad . \tag{116}$$

The time evolution of this object is governed by the Schrödinger equation as follows[122] :

$$\begin{aligned} \frac{d}{dt} \langle A \rangle &= \left(\frac{d}{dt} \langle \psi | \right) \hat{A} \, | \psi \rangle + \langle \psi | \, \hat{A} \left(\frac{d}{dt} | \psi \rangle \right) \\ &= \frac{i}{\hbar} \left(\langle \psi | \, \hat{H} \right) \hat{A} \, | \psi \rangle - \frac{i}{\hbar} \, \langle \psi | \, \hat{A} \left(\hat{H} \, | \psi \rangle \right) \\ &= \frac{i}{\hbar} \, \langle \psi | \left(\hat{H} \, \hat{A} - \hat{A} \, \hat{H} \right) | \psi \rangle \quad , \end{aligned} \tag{117}$$

so that we arrive at the following expression[123], the *Heisenberg equation* :

$$\frac{d}{dt} \langle A \rangle = \frac{1}{\hbar} \left\langle i[\hat{H}, \hat{A}] \right\rangle \quad . \tag{118}$$

We can immediately draw a few conclusions. In the first place, if the system is an energy eigenstate, then all expectation values are time-independent ; in fact this already follows from Eq.(113). Also, every observable B commensurable with the total energy has an expectation value (for *any* state of the system) that is constant : $d \langle B \rangle / dt = 0$; moreover since, when $\hat{B}$ commutes with $\hat{H}$ then also $\hat{B}^2, \hat{B}^3, \ldots$ commute with $\hat{H}$, the whole probability distribution $P(b)$ is time-independent. We see that a complete commensurable set that includes the total energy gives us a complete set of *constants of motion*[124]. **E 24**

5.5 The free particle and the fundamental commutator

It is now time to apply what we have learned to some more-or-less realistic system. The very simplest physical system is that of a single particle, moving

[122] We assume that the observable A does *not* depend explicitly on time. Of course, the 'time of day' is also something we can observe, but it evolves according to the *clock*, not according to *Schrödinger*.

[123] Note that $i[\hat{H}, \hat{A}]$ itself is Hermitean and that therefore it corresponds, in principle, to some observable.

[124] Constants of motion are, classically, those properties of a system that do not change with time. Examples are (in many systems) the total energy, the total momentum, the total *angular* momentum.

in one dimension without any forces acting upon it : the *one-dimensional free particle*. In classical physics, the total energy of such a particle is simply the *kinetic energy* :

$$E_{\rm cl} = \frac{1}{2m}\, {p_{\rm cl}}^2 \quad , \tag{119}$$

where m is the particle's mass[125], and $p_{\rm cl}$ is the momentum. What shall we take for the quantum-mechanical æquivalent ? The simplest idea is to replace the *classical* momentum by the quantum-mechanical momentum *operator*, whatever *that* may turn out to be[126] :

$$\hat{H} = \frac{1}{2m}\, \hat{p}^2 \quad . \tag{120}$$

We now come to a crucial step. If there is any justice in the world, then a quantum description ought to at least be able to be *compared* to a classical description : but what are the æquivalents of the classical notions of position and momentum ? Surely not simply the *operators*, since in classical physics position and the like are just *numbers*. We therefore assert, remembering the discussion in section 2.8, that *the classical observables ought to correspond to the expectation values of the quantum observables*, whenever such a comparison makes sense[127]. That is, the expectation values of position and momentum, $\langle x \rangle$ and $\langle p \rangle$, ought to behave in the manner that the *classical* numbers $x_{\rm cl}$ and $p_{\rm cl}$ are wont to do. For instance, momentum and position are, for a classical particle, related by Newton's very own definition :

$$\frac{d}{dt}\, x_{\rm cl} = \frac{1}{m}\, p_{\rm cl} \quad ; \tag{121}$$

we therefore require things to be such that

$$\frac{d}{dt}\, \langle x \rangle = \frac{1}{m}\, \langle p \rangle \quad . \tag{122}$$

Now, the Heisenberg equation, Eq.(118), tells us that

$$\frac{d}{dt}\, \langle x \rangle = \frac{i}{\hbar} \left\langle [\hat{H}, \hat{x}] \right\rangle = \frac{i}{2m\hbar}\, \langle [\hat{p}^2, \hat{x}] \rangle \quad , \tag{123}$$

[125] In nonrelativistic quantum mechanics, the particle mass is assumed to be a constant, and hence not in itself a very interesting observable ; we shall take it to be simply a *number*.

[126] This question, what to take for the quantum version of the classical energy, is one of the most fundamental problems in constructing a quantum theory. The usual approach is to do what we do here : simply take the classical notions of position and momentum, and replace them by their respective quantum operators. This procedure is not free from ambiguities. However, we are helped here by the fact that we are *physicists* and not *mathematicians* : we do not have to *derive* the Hamiltonian operator from something else ; we simply *postulate* it, and compare the experimental consequences of that choice with what observation of the world tells us. If we turn out to have chosen unfortunately, we can opt for another Hamiltonian. To some extent, this is what being a physicist (at least a theoretical physicist) is *about*.

[127] This is based on a *feeling* that for classical-looking states, position and momentum and such ought to be pretty narrowly defined, in which case the expectation value is what you're most likely to get from a measurement anyway.

and by combining these two last equations we find that we must have

$$\left\langle [\hat{p}^2, \hat{x}] \right\rangle = -2i\hbar \, \langle \hat{p} \rangle \quad . \tag{124}$$

Moreover, since the definition of momentum as given by Newton is so very general (holding also for particles upon which forces act, and so on) that we really must have

$$[\hat{p}^2, \hat{x}] = -2i\hbar \, \hat{p} \ ; \tag{125}$$

and the only simple way to obtain this is to decide that $\hat{x}$ and $\hat{p}$ satisfy the *fundamental commutator* :

$$[\hat{x}, \hat{p}] = i\hbar \quad . \tag{126}$$

This is an impressively curious result ! Two of the most fundamental observables in all of physics are *incommensurable*. Nevertheless, it must be true since Heisenberg's equation tells us that, if $\hat{x}$ and $\hat{p}$ were to commute, *no motion would be possible* in quantum mechanics ! This last statement holds even if we could devise a clever alternative to the commutator (126).

If we move to the realm of three dimensions, the above discussion is easily generalized. Assuming three position coordinates x, y and z, and three components of the momentum vector p_x, p_y and p_z, the fundamental commutators are

$$[\hat{x}, \hat{p}_x] = [\hat{y}, \hat{p}_y] = [\hat{z}, \hat{p}_z] = i\hbar \quad , \tag{127}$$

while all other commutators between position and/or momentum components are zero[128] . **E 25**

5.6 The Ehrenfest theorem

It is of course not only the position of which the expectation value has a time dependence : we may also be interested in the momentum's behaviour. To make things a bit less trivial, we shall study a particle that moves, still in one dimension, under the influence of a potential. Classically, we then have

$$E_{\rm cl} = \frac{1}{2m} {p_{\rm cl}}^2 + V(x_{\rm cl}) \ ; \tag{128}$$

and the logical quantum version will therefore be

$$\hat{H} = \frac{1}{2m} \hat{p}^2 + V(\hat{x}) \quad . \tag{129}$$

The Heisenberg equation now gives for the momentum :

$$\begin{aligned} \frac{d}{dt} \langle p \rangle &= \frac{i}{\hbar} \left\langle [\hat{H}, \hat{p}] \right\rangle = \frac{i}{\hbar} \, \langle [V(\hat{x}), \hat{p}] \rangle \\ &= \frac{i}{\hbar} \, i\hbar \, \langle V'(\hat{x}) \rangle = - \langle V'(\hat{x}) \rangle \quad , \end{aligned} \tag{130}$$

[128] This is an *assumption*. Perfectly consistent versions of quantum physics exist in which the position operators, say, do *not* commute ; in such theories, there are constants $\theta^{jk} = -\theta^{kj}$ such that $[\hat{x}^j, \hat{x}^k] = \theta^{jk}$. The θ's would be *new* constants of nature or such. But no evidence of this 'non-commuting geometry' in quantum mechanics has yet been found.

where we have used the fundamental commutator (126) and the result of exercise 17[129]. This result is called the *Ehrenfest theorem*. It is the quantum version of Newton's law : **E 26**

$$\frac{d}{dt}\, p_{\rm cl} = -V'(x_{\rm cl}) \ . \tag{131}$$

The conclusion therefore is : as long as the probability distributions of position and momentum are very sharply peaked around some values, these values obey the laws of classical mechanics[130]. **E 27**

5.7 Extra ! Maximum speed : the Margolus-Levitin theorem

Under the Schrödinger equation, states evolve into other states, and so the physical content of the state changes[131]. How much time will it take for an initial state to be changed 'beyond recognition' ? In other words, how fast can a given initial state evolve into another, physically incompatible state ? Or, more precisely, what is the minimum time T for which the initial state $|\psi_0\rangle \equiv |\psi\rangle\,(0)$ and the evolved state $|\psi\rangle\,(T)$ can be orthogonal ? The Margolus-Levitin theorem supplies an answer. For simplicity, we assume that the energy spectrum is discrete, and we label the energy eigenstates by $|E_n\rangle$, where E_n is the energy eigenvalue. The initial state $|\psi_0\rangle$ can then be written as

$$|\psi_0\rangle = \sum_n z_n\ |E_n\rangle \quad , \tag{132}$$

and the Schrödinger equation then tells us that at some later time t we have

$$\begin{aligned} |\psi\rangle\,(t) &= \sum_n z_n\ \exp\left(-i\frac{E_n t}{\hbar}\right)\ |E_n\rangle \\ &= \exp\left(-i\frac{E_m t}{\hbar}\right) \sum_n z_n\ \exp\left(-i\frac{\tilde{E}_n t}{\hbar}\right)\ |E_n\rangle \quad , \end{aligned} \tag{133}$$

where

$$\tilde{E}_n = E_n - E_m \tag{134}$$

and E_m is the *minimum* of all energy eigenvalues, so that $\tilde{E}_n \geq 0$ for all n. We now construct the following object :

$$\begin{aligned} A(t) &\equiv \langle\psi_0|\,\psi\rangle\,(t)\ \exp\left(+i\frac{E_m t}{\hbar}\right) \\ &= \sum_n |z_n|^2\ \exp\left(-i\frac{\tilde{E}_n t}{\hbar}\right) \ . \end{aligned} \tag{135}$$

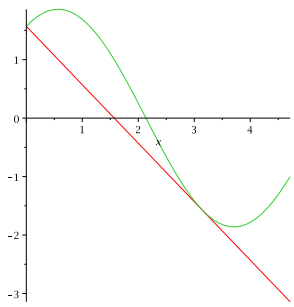
[129] A word of caution : it is easy to confuse $\langle V'(\hat{x})\rangle$ with $V'(\langle\hat{x}\rangle)$. The first is correct, and the second is *wrong*. On the other hand, the two expressions can be very close in value, especially if X is quite well-defined.

[130] It might seem that we have somehow *derived* Newton's second law, which in classical physics has to be *postulated* ; but in fact we have of course only replaced it by another postulate of motion, namely the Schrödinger equation.

[131] Unless, of course, the state is an energy eigenstate.

It follows that

$$\begin{aligned} \mathrm{Re}A(t) &= \sum_n |z_n|^2 \cos\left(\frac{\tilde{E}_n t}{\hbar}\right) \ , \\ \mathrm{Im}A(t) &= -\sum_n |z_n|^2 \sin\left(\frac{\tilde{E}_n t}{\hbar}\right) \ . \end{aligned} \tag{136}$$



Now, the clever trick is to use the following inequality, valid for *positive* x :

$$\sin(x) + \frac{\pi}{2}\cos(x) \geq \frac{\pi}{2} - x \ . \tag{137}$$

E 28

The picture to the left displays the two sides of the inequality for x between 0 and $3\pi/2$. At $x = \pi$ the inequality is saturated.

Therefore,

$$\begin{aligned} &-\mathrm{Im}A(t) + \frac{\pi}{2}\mathrm{Re}A(t) = \\ &= \sum_n |z_n|^2 \left\{ \sin\left(\frac{\tilde{E}_n t}{\hbar}\right) + \frac{\pi}{2}\cos\left(\frac{\tilde{E}_n t}{\hbar}\right) \right\} \\ &\geq \sum_n |z_n|^2 \left(\frac{\pi}{2} - \frac{\tilde{E}_n t}{\hbar}\right) \\ &= \sum_n |z_n|^2 \left(\frac{\pi}{2} - \frac{t}{\hbar}\left(E_n - E_m\right)\right) = \frac{\pi}{2} - \frac{t}{\hbar}\left(\langle E\rangle - E_m\right) \ , \end{aligned} \tag{138}$$

where $\langle E\rangle$ is of course the energy's expectation value. Now, let us assume that there is some time T such that at $t = T$ the overlap vanishes ; therefore both the real and the iaginary part of $A(T)$ must be zero at time T. But that means that

$$\frac{\pi}{2} - \frac{T}{\hbar}\left(\langle E\rangle - E_m\right) \leq 0 \ , \tag{139}$$

or

$$T \geq \frac{\pi\hbar}{2\left(\langle E\rangle - E_m\right)} \ . \tag{140}$$

This is the shortest time in which a given system can 'really change'. Note that this is *not* an uncertainty relation of the Heisenberg type ! That the lower bound in Eq.(140) is essentially the best possible is shown in excercise 29. Also, note that the occurrence of E_m is necessary in Eq.(140), since otherwise we could change the physics by just adding an irrelevant constant to all energies (see also excercise 22).

E 29

5.8 Exercises 22 to 29

Exercise 22 : **A constant energy addition term**
In classical physics, the total energy of a system is only defined up to an additive constant : adding a constant C to the total energy of a system does not change its physics. And so it is in the quantum world, too. To see this, consider the alternative Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi_c\rangle = \left(\hat{H} + C\right) |\psi_c\rangle$$

with some constant number C. We assume that $\hat{H}$ does not depend explicitly on any time derivatives[132]. Show that *every* solution to this equation can be written as

$$|\psi_c\rangle = e^{-iCt/\hbar} |\psi\rangle$$

where $|\psi\rangle$ is the solution to the Schrödinger equation with $C = 0$. Such a universal overall complex factor is not physically noticeable, as we have seen.

Exercise 23 : **Checking the solution**
Verify that Eq.(115) is indeed the solution to the Schrödinger equation by inserting Eq.(115) into Eq.(108), and working it out.

Exercise 24 : **More time dependence**

1. Prove that all expectation values are time-independent if a system is in an energy eigenstate.
2. Apply the Heisenberg equation (118) to prove that

$$\frac{d^2}{dt^2} \langle A \rangle = -\frac{1}{\hbar^2} \left\langle [\hat{H}, [\hat{H}, \hat{A}]] \right\rangle$$

   and extend this to higher derivatives.

3. Prove that

$$\frac{d}{dt}\sigma(A)^2 = \frac{i}{\hbar} \left(\left\langle [\hat{H}, \hat{A}]\hat{A} \right\rangle - \left\langle [\hat{H}, \hat{A}] \right\rangle \langle A \rangle + \left\langle \hat{A}[\hat{H}, \hat{A}] \right\rangle - \langle A \rangle \left\langle [\hat{H}, \hat{A}] \right\rangle \right)$$

Exercise 25 : **Rotating the commutators**
From Eq.(127) it might appear that there are specially favored directions in space. This exercise proves that that is not the case[133]. Imagine that we make

[132] This is usually the case.
[133] Fortunately ! It would be surprising, to say the least, if there were such a thing as *the* x-axis of the universe.

a rotation in the $x-y$ plane over an angle θ. We then have new position and momentum operators given by

$$\begin{aligned}
\hat{x}' &= \hat{x}\,\cos\theta + \hat{y}\,\sin\theta \ , \\
\hat{y}' &= -\hat{x}\,\sin\theta + \hat{y}\,\cos\theta \ , \\
\hat{p}'_x &= \hat{p}_x\,\cos\theta + \hat{p}_y\,\sin\theta \ , \\
\hat{p}'_y &= -\hat{p}_x\,\sin\theta + \hat{p}_y\,\cos\theta \ .
\end{aligned}$$

Prove that

$$[\hat{x}', \hat{p}'_x] = [\hat{y}', \hat{p}'_y] = i\hbar \quad , \quad [\hat{x}', \hat{y}'] = [\hat{x}', \hat{p}'_y] = [\hat{y}', \hat{p}'_x] = [\hat{p}'_x, \hat{p}'_y] = 0 \ .$$

In other words, the fundamental commutators have exactly the same form in both coordinate systems.

Exercise 26 : **Motion in three dimensions**
Derive the three-dimensional analogue of Eq.(130).

Exercise 27 : **The viral theorem**
We shall derive an instance of the so-called *virial theorem*. To this end we consider a one-dimensional system of a particle of mass m in a potential given by $V(\hat{x}) = k\,\hat{x}^n$, with $k > 0$, and $n > 0$ even.

1. Show that the particle is always in a bound state.

2. Let $\hat{T}$ be the operator for the observable of kinetic energy. Write down $\hat{T}$ and the Hamiltonian $\hat{H}$.

3. Consider the following quantity : $\hat{W} = \hat{x}\hat{p} + \hat{p}\hat{x}$. Show that this corresponds, in principle, to some observable W, and can therefore have a (possibly time-dependent) expectation value. Give the Heisenberg equation for $d\langle W\rangle/dt$.

4. We introduce the notion of the *time average* $\{A\}$ of a time-dependent observable $A(t)$ as follows :

$$\{A\} = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau A(t)\,dt \ .$$

   Assuming $\langle W\rangle$ never to be infinite, prove that

$$\{d\langle W\rangle/dt\} = 0 \ .$$

5. Compute $[\hat{H}, \hat{W}]$.

6. Now prove the virial theorem :

$$2\{\langle T\rangle\} = n\{\langle V\rangle\} \ .$$

Exercise 28 : **Touch, don't cut**
Verify that the inequality of Eq.(137) is exactly an equality for $x = \pi$, in other words, show that the two curves touch but do not intersect at that value.

Exercise 29 : **The Margolus-Levitin theorem in a simple case**
Consider a system with a basis of energy eigenstates. Consider the normalized states $|1\rangle$ and $|2\rangle$ that have energy eigenvalues E_1 and E_2, respectively, with $E_2 \neq E_1$.

1. Show that the states

$$|+\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle) \quad \text{and} \quad |-\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)$$

   are orthonormal.

2. Show that $\langle E\rangle = (E_1 + E_2)/2$ for both $|+\rangle$ and $|-\rangle$.

3. Show that if the system is in state $|+\rangle$ at some time, it will be in state $|-\rangle$ (up to an irrelevant overall complex phase factor) after a time interval of precisely $\pi\hbar/(2(\langle E\rangle - E_m)$.

This system shows that the Margolus-Levitin bound cannot be improved.

6 The wave function

— The box [...] is a method of solitary confinement used [...] as a method of punishment.
Wikipedia

6.1 The position representation

So far, our discussion has been quite abstract. States and observables have been introduced, but it may not be so easy to *visualize* them. The *position representation* is somtimes helpful in this respect : it consists in taking the *position* of a particle as an important (maybe *the* important) property. The basis states of choice are, then, the eigenstates of position. Let us denote the state in which the particle is certainly going to be found[134] at coordinate x by $|x\rangle$; then every state of the system[135] can be written as follows :

$$|\psi\rangle = \sum_x c_x \ |x\rangle \quad ,$$

for given coefficients c_x. Taking into account that the position coordinates usually form a *continuum* rather than a discrete set of values, we had better replace the sum by an integral, so that we have

$$|\psi\rangle = \int dx \ \ c(x) \ |x\rangle \quad ,$$

where now the coeficient $c(x)$ has become a *function* of x. It is most common to denote the coefficient by $\psi(x)$ rather than by $c(x)$, and so we arrive at the final form

$$|\psi\rangle = \int dx \ \ \psi(x) \ |x\rangle \quad . \tag{141}$$

Note the following important distinction : $|\psi\rangle$ is a *quantum state*, an element of the Hilbert space, whereas $\psi(x)$ is a *function*, a complex number assigned to every value of x. From the one, we can compute the other :

$$\psi(x) = \langle x|\,\psi\rangle \tag{142}$$

by the (by now, hopefully) standard rules[136]. The function $\psi(x)$ is called the *wave function* of the system. It is clear that $\psi(x)$ contains exactly the same

[134] For now, again in one dimension for simplicity.

[135] Note that this approach disregards all other properties of the particle, such as for instance spin : in more sophisticated treatments, the representation becomes accordingly more complictated.

[136] Actually these rules have to be modified a little bit. In the *discrete* cases we have studied so far the orthonormal basis states have inner products that are given by the *Kronecker delta* of Eq.(27). The states $|x\rangle$, on the other hand, form a *continuous* set, and must therefore be orthonormal with some other rule: in fact we have

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

an expression that involves the *Dirac delta function* (or *Delta distribution*). More on this important function can be found in appendix 11. For us, at this point, these are fairly unimportant mathematical details.

information about the system as does $|\psi\rangle$ itself, so that we could construct quantum mechanics completely in terms of the wave function alone[137].

From what we have learned about states, we can immediately form the following statement : *the probability (density) to find the particle at position x must be given by $|\psi(x)|^2$*. This is extremely reminiscent of what is known from electrodynamics : a light wave is described by a (complex) wave function, and the *intensity* of the light (that is, the probability to find photons) is given by its modulus squared[138].

Let us state the important fact again : the wave function is a *description* of the quantum state for those who want to use position eigenstates. Other representations are eminently possible ; we might decide to eploy *momentum* eigenstates as well, or use any other complete commensurable set of observables.

6.2 Position and momentum for wave functions

If we decide to use wave functions, we of course have to work out what the observable-operators do to those wave functions. Many observables are built up from position and momentum ; these we shall consider in detail. Let us remember that position eigenstates are just that, *position* eigenstates : they satisfy

$$\hat{x}\ |x\rangle = x\ |x\rangle \quad , \tag{143}$$

where the $\hat{x}$ occuring on the left-hand side is an *operator*, while the x on the right-hand side is a *number*, namely the *value* of the position coordinate which is well-defined for that eigenstate. For a general state, we therefore have

$$\hat{x}\ |\psi\rangle = \hat{x}\ \int\ dx\ \psi(x)\ |x\rangle = \int\ dx\ \psi(x)\ \hat{x}\ |x\rangle = \int\ dx\ \psi(x)\ x\ |x\rangle \quad , \tag{144}$$

so that we can specify what the position operator does to the wave function ; we could write it as

$$\hat{x}\ \psi(x) \equiv x\ \psi(x) \ . \tag{145}$$

The position operator $\hat{x}$ simply multiplies the wave function by x. The action of the *momentum* operator is then determined by the requirement that $[\hat{x}, \hat{p}] = i\hbar$. The rule is that

$$\hat{p}\ \psi(x) \equiv -i\hbar\ \frac{\partial}{\partial x}\ \psi(x) \quad , \tag{146}$$

since it gives us

$$\begin{aligned} [\hat{x}, \hat{p}]\ \psi(x) &= \hat{x}\,\hat{p}\,\psi(x)\ -\ \hat{p}\,\hat{x}\,\psi(x) \\ &= -i\hbar\ x\ \frac{\partial}{\partial x}\psi(x) + i\hbar\frac{\partial}{\partial x}\ \Big(x\,\psi(x)\Big) \\ &= i\hbar\,\psi(x) \quad , \end{aligned} \tag{147}$$

[137] In fact, this is what Schrödinger did. The notion of 'state' is not his, but rather Heisenberg's and Dirac's.

[138] Not surprisingly, it is this idea that gave rise to the probability interpretation of the wave function in the first place.

as desired. We see that $\hat{x}$ *supplies* one power of x to the wave function, while $\hat{p}$ *detracts* one power of x.

The actions of the two fundamental operators
on wave functions are

$$\hat{x}\ \psi(x) = x\ \psi(x)$$

and

$$\hat{p}\ \psi(x) = -i\hbar\ \frac{\partial}{\partial x}\ \psi(x)$$

6.3 Inner products for wave functions

As stated above, the aim of this chapter is to express quantum mechanics in wave-function language. We must therefore consider what the inner product of two states looks like in this language. Let us consider two states, $|\psi\rangle$ and $|\phi\rangle$, with their corresponding wave functions :

$$|\psi\rangle\ \leftrightarrow\ \psi(x)\quad ,\quad |\phi\rangle\ \leftrightarrow\ \phi(x)\ . \tag{148}$$

The secret of working out the wave-function form of these states' inner product is in the realization that the set of all position states must be a *complete* basis, that is,

$$\int dx\ |x\rangle\ \langle x|$$

acts like the identity operator on any state. That this must be so can be seen from

$$\left(\int dx\ |x\rangle\ \langle x|\right)|\psi\rangle = \int dx\ |x\rangle\ \langle x|\,\psi\rangle = \int dx\ \psi(x)\ |x\rangle = |\psi\rangle\quad , \tag{149}$$

using the definition of the wave function that we already encountered in Eq.(141). We can now work out how the inner product must be formed :

$$\begin{aligned}\langle\phi|\,\psi\rangle &= \langle\phi|\left(\int dx\ |x\rangle\ \langle x|\right)|\psi\rangle \\ &= \int dx\ \langle\phi|\,x\rangle\ \langle x|\,\psi\rangle \\ &= \int dx\ \phi(x)^*\ \psi(x)\ .\end{aligned} \tag{150}$$

The inner product, translated into the language of wave functions, is simply the integral over the product of the wave functions : the only thing to be remembered is that of the wave function of the *bra* we must take the complex **E 30**

conjugate, while for the *ket* we take the wave function itself.

It is now a simple matter to work out products of states with operators between them. For instance,

$$\langle\phi|\,\hat{x}\,|\psi\rangle = \int dx\ \phi(x)^*\ x\ \psi(x)\ \ , \tag{151}$$

and

$$\langle\phi|\,\hat{p}\,|\psi\rangle = \langle\phi|\ \Big(\hat{p}\ |\psi\rangle\Big) = \int dx\ \phi(x)^*\ (-i\hbar)\ \psi'(x)\ \ . \tag{152}$$

Expectation values are particularly simple : **E 31**

$$\langle x\rangle = \int dx\ x\ |\psi(x)|^2 \tag{153}$$

for a system in state $|\psi\rangle$; which is of course precisely what you would expect. For the momentum we have

$$\langle p\rangle = -i\hbar\ \int dx\ \psi(x)^*\ \psi'(x) \tag{154}$$

and for the *kinetic energy* : **E 32**

E 33

$$\langle E_{\rm kin}\rangle = -\frac{\hbar^2}{2m}\int dx\ \psi(x)^*\ \psi''(x)\ \ . \tag{155}$$

6.4 The Schrödinger equation for wave functions

We are now in a position to work out the form of the Schrödinger equation for the wave function. Inserting the time dependence into the wave function, we see that for a one-dimensional system of a particle in a potential $V(x)$ we must have[139]

$$i\hbar\ \frac{d}{dt}\,\psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\,\psi(x,t) + V(x)\,\psi(x,t)\ \ . \tag{156}$$

The somewhat-simpler time-independent Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\,\psi(x) + V(x)\,\psi(x) = E\,\psi(x)\ \ , \tag{157}$$

and it is this equation that occupies a lot of the theorists' time and ingenuity. On differential equations like this, second-order, linear and homogeneous, there exists a vast body of knowledge. We must mention three important issues here.

[139] Note that this differential equation is of *second* order in space, but of only *first* order is time. It is therefore *not* adapted to special relativity's demand that space and time are treated on an equal footing. Schrödinger originally chose an equation that was also of *second* order in *time*, and was dissatisfied with it, although later on Klein and Gordon used it with good results ; Dirac opted instead for an equation that was of *first* order in *space*, constructed the *Dirac equation* that describes electrons — and discovered the existence of *antimatter* into the bargain.

In the first place, solutions to the (time-independent or not) Schrödinger equation must be *continuous* : if there were a discontinuity in $\psi(x)$ at some x, the *derivative* of $\psi(x)$ would be infinite there, and consequently the momentum would also be infinite ! We therefore will only be satisfied with *continuous* solutions. On the other hand, *kinks* (that is, discontinuities in the derivative itself) are acceptable.

Secondly, the position representation relates $|\psi(x)|^2$ to the probability of finding the particle at x : since the *total* probability to find the particle *somewhere* must be unity, we require that

$$\int dx\ |\psi(x)|^2 = 1\ . \tag{158}$$

If the integral comes out not to unity but to some other number k, we can always rescale, that is, multiply the wave function (or the state!) by $1/\sqrt{k}$ to make things work out ; but it often happens that the solution to Eq.(157) has an integral that is infinite. Such solutions are not acceptable, and the requirement that the wave function be *quadratically integrable* often serves to restrict the possible values of the energy eigenvalue E : this is, in wave-function langauge, the origin of *energy quantisation*.

E 34

E 35

Thirdly, and most importantly, a *second*-order homogeneous differential equation like the time-independent Schrödinger equation has *two* linearly independent basic solutions ; *all* other solutions are linear combinations of these two[140]. Which solutions you take to be the 'basic' ones is a matter of choice, since every solution can be formed from any two linearly independent ones. There is one *caveat*, in particular in the case of systems with bound states : that a given wave function $\psi(x)$ is a solution of Eq.(157) does *not* mean that it is physically acceptable, since it must also be normalizable according to Eq.(158). It frequently happens[141] that out of two basic solutions only *one* is normalizable, and the other one isn't. In that case, we discard the unnormalizable solution[142].

6.5 Doing the wave

Let us now again consider the free particle, this time in the language of wave functions. Moreover, let us assume that the particle is in a *momentum* eigenstate, *i.e.* its momentum is well-defined[143]. Then,

$$\hat{p}\ \psi(x,t) = -i\hbar\ \frac{\partial}{\partial x}\ \psi(x,t) = p\,\psi(x,t)\ \ , \tag{159}$$

[140] This is, of course, the superposition principle ! If two states of a system are possible, then *any* linear combination of them is also, in principle, possible.

[141] And we shall meet this !

[142] Although in many cases - as you will see in our treatment of the harmonic oscillator - we will arrange things such that we never even ever come *close* to the unmentionable unnormalizable.

[143] The fact that we are working in the position *representation* does *not* mean that all states must be position *eigenstates* ! The position representation just decides to write *all* states as *superpositons* of position eigenstates.

where p is the value of the momentum. We can easily solve this first-order differential equation :

$$\psi(x,t) \;\propto\; \exp\left(\frac{i}{\hbar}px\right) \;\; . \tag{160}$$

On the other hand, if the free-particle system is in an eigenstate of momentum, it is of course also in an eigenstate of the energy operator $\hat{H} = \hat{p}^2/2m$, so that the Schrödinger equation gives

$$i\hbar\,\frac{d}{dt}\,\psi(x,t) = E\,\psi(x,t) \quad , \quad E = p^2/2m \;\; , \tag{161}$$

with the solution

$$\psi(x,t) \;\propto\; \exp\left(-\frac{i}{\hbar}Et\right) \;\; . \tag{162}$$

We see that the complete solution in this case reads

$$\psi(x,t) \;\propto\; \exp\left(-\frac{i}{\hbar}(Et - px)\right) \;\; . \tag{163}$$

This is where the notion 'wave function' comes from ! Free particles with well-defined momentum are described by *plane waves*[144]. On the other hand, we find that for such systems $|\psi(x,t)|^2$ is a constant, so that the particle's *position* is completely indeterminate[145]. The best 'picture' is therefore not that of a single particle with a given momentum, but rather a continuous 'stream' of particles moving along like a train, with constant density[146]. To build a state that looks more like an individual particle, that is, a more-or-less localized 'lump' moving about, we must employ a more complicated superposition of plane waves.

E 36

E 37

6.6 Extra ! Probability currents

The picture of a wave function as representing, somehow, a stream of particles can be further refined, by saying that the probability density, $|\psi(x)|^2$, informs us about the *density* of particles, *i.e.* indeed the probability of finding particles (in a unit volume, for instance). As a quantum system evolves with time, the probability density at a given point may of course change, and this is what we shall investigate. Let us denote the probability density at point x and time t by

$$\rho(x,t) \equiv |\psi(x,t)|^2 \;\; , \tag{164}$$

[144] Note one curious fact : although unintended, the 'special-relativity' combination $Et - xp$ makes its appearance here. There is some conspiracy about : relativity and quantum cannot very well live without one another, although their marriage is not completely consummated yet.

[145] As it should, according to the Heisenberg inequality.

[146] This is actually a quite good metaphor. Imagine that you are waiting at a railroad crossing, while a large freight train is passing. All you can see is wagon after wagon running by ; the *speed* of the train is easily seen, but as long as you can't see the front end or the rear end, the *location* of the train is undetermined.

where we have indicated explicitly the time and space dependence. Let us now see how this behaves when we set the clock running. Using the Schrödinger equation and denoting the potential energy $V(x)$ by V, we can write

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\rho &= \left(i\hbar\frac{\partial}{\partial t}\psi\right)\psi^* - \psi\left(i\hbar\frac{\partial}{\partial t}\psi\right)^* \\ &= \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi + V\psi\right)\psi^* - \psi\left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi + V\psi\right)^* \\ &= -\frac{\hbar^2}{2m}\left\{\left(\frac{\partial^2}{\partial x^2}\psi\right)\psi^* - \psi\left(\frac{\partial^2}{\partial x^2}\psi\right)^*\right\} \\ &= -\frac{\hbar^2}{2m}\frac{\partial}{\partial x}\left\{\left(\frac{\partial}{\partial x}\psi\right)\psi^* - \psi\left(\frac{\partial}{\partial x}\psi\right)^*\right\} \quad , \end{aligned} \tag{165}$$

where we have used the fact that V must be real. If we now define

$$J(x,t) \equiv -i\frac{\hbar}{2m}\left\{\left(\frac{\partial}{\partial x}\psi\right)\psi^* - \psi\left(\frac{\partial}{\partial x}\psi\right)^*\right\} \quad , \tag{166}$$

then we see that

$$\frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}J(x,t) = 0 \quad . \tag{167}$$

This is a so-called *continuity equation* : J is called the *probability current* and represents the 'streaming' of probability from one place to another with time. The continuity equation tells us that a decrease of probability at a given point can only come about by that probability streaming away to somewhere else, and of course an increase in probability must come from a stream towards the point of interest. In other words, no probability can diasppear into thin air, or pop up from the vacuum. The *total* probability is conserved[147]. **E 38**

For a free particle of well-determined momentum the wave function was seen to read

$$\psi(x,t) \propto \exp\left(-\frac{i}{\hbar}(Et - px)\right) \quad , \tag{168}$$

and the resulting probability current comes out as

$$J = \frac{p}{m} \quad : \tag{169}$$

here, it is nothing but the velocity of the particle ! Even in this case, where ρ is everywhere constant, we can nevertheless 'see the particles move'. **E 39**

Finally, note that all this depends sensitively on the fact that V is real. If we relax this (that is, for Hamiltonians that have a nonzero imaginary part) probability *can* disappear into nowhere, or appear from nowhere. We usually try to avoid this, but it may be used in the description of *unstable* particles, where indeed the total particle density may decrease by 'radioactive' decay of the particles.

[147] Remember the discussion in section 5.2, where the same conclusion was drawn from, indeed, the same Schrödinger equation.

6.7 Particle in a box : energy quantisation

We shall now discuss what is probably the simplest one-dimensional system after the free particle : the *particle in a box* of size L. This means that there is some interval $0 < x < L$ where the particle may be found, and it can *not* be found at $x \leq 0$ or $x \geq L$. Inside the box, the particle is free (no forces are acting on it). We can envisage this as a potential energy that is *zero* inside the box, and jumps to *infinity* outside it. This is admittedly a very drastic and crude model for a system in which a particle is *bound* to a certain space region, but many of its consequences can be taken over (at least qualitatively) to more realistic systems.

We shall now solve the time-independent Schrödinger equation for this system. The restriction of the particle to the inside of the box implies two things. In the first place $\psi(x) = 0$ for $x \leq 0$ and $x \geq L$, since we have decided that it is *impossible* to find the particles outside the box. In the second place, since discontinuities in the wave function cannot be accepted, the wave function *inside* the box must go to zero at the two endpoints $x = 0$ and $x = L$. Now we are ready to start. The Hamiltonian inside the box, since the particle is free there, is simply

$$\hat{H} = \frac{1}{2m}\,\hat{p}^2 = -\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2} \quad , \tag{170}$$

so that the time-independent Schrödinger equation reads

$$\frac{\hbar^2}{2m}\,\psi''(x) = -E\,\psi(x) \quad . \tag{171}$$

This is the equation of sines and cosines, so that the general solution is

$$\psi(x) = A\,\sin(c\,x) + B\,\cos(c\,x) \quad , \tag{172}$$

for some numbers A, B and c. Since $\cos(0) = 1$ and $\sin(0) = 0$, we immediately see that the cosine term is unacceptable because it leads to a discontinuity at $x = 0$; hence $B = 0$. On the other side of the box, the wave function must also vanish at $x = L$, so we require

$$\sin(c\,L) = 0 \quad . \tag{173}$$

This means that $c\,L$ must be a multiple of π, hence we can write **E 40**

$$\psi_n(x) = A\,\sin\left(\frac{n\,\pi}{L}\,x\right) \quad , \quad n = 1, 2, 3, 4 \ldots \tag{174}$$

The normalization factor A can be worked out simply, and is in fact independent of n for this system. The energy eigenfunctions are [148] **E 41**

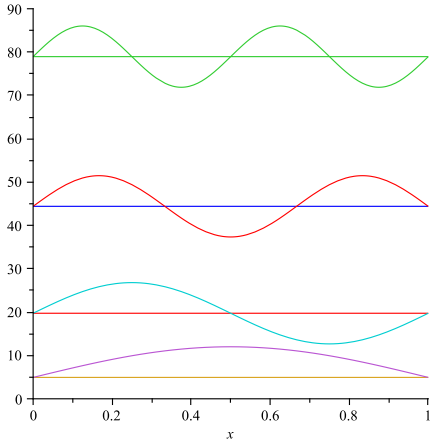
$$\psi_n(x) = \sqrt{\frac{2}{L}}\,\sin\left(\frac{n\,\pi}{L}\,x\right) \quad , \quad n = 1, 2, 3, 4 \ldots \tag{175}$$

[148] The fact that only sine functions occur is owed to our choice of the interval $(0, L)$ for the box. Many textbooks take the interval $(-L, L)$ for some reason. The formula for the energy then of course picks up extra factor 1/4, but there is a clumsy drawback : with that alternative choice of interval, the eigenfunctions are *cosines* for odd n, and *sines* for even n.

Each energy eigenfunction has its own eigenvalue[149], which we obtain by inserting the form (175) into the time-independent Schrödinger equation (170) :

$$E_n = \frac{\hbar^2 \pi^2}{2mL^2} n^2 \quad , \quad n = 1, 2, 3, 4 \ldots \tag{176}$$

Here we have, for the first time, a more-or-less realistic system with energy quantization. In this case, the reason for the occurrence of discrete possible energies is the fact that the wave function must vanish at the endpoints of the box ; for 'milder' potentials similar results come from the requirement that the wave function should not blow up at infinity, but rather go to zero in an appropriate way. **E 42**



Artist's impression of the wave functions $\psi_n(x)$ for $n = 1, 2, 3, 4$. The horizontal lines around which the wave functions oscillate denote the energy levels. For dramatic purposes I have exaggerated the wave functions' amplitudes by a factor 5. The plots were made with $m = L = \hbar = 1$.

The energy eigenvalues increase rapidly as n increases, which is due to the steepness of the 'potential' decribing the box ; for 'milder' potentials the increase is less pronounced[150]. **E 44** **E 45**

A last remark is in order here, when we consider our system 'from the outside', as it were. Suppose the particle is in a state ψ_n, so that its energy is given by

$$E = \frac{\hbar^2 \pi^2 n^2}{2mL^2} \quad . \tag{177}$$

Now consider what happens if we move the wall at $x = L$ over a distance dL, ever so slightly. Assuming that the quantum number n is not influenced by such an infinitesimal displacement[151], the energy of the system will change according

[149] There is no degeneracy in this model. This is only to be expected : in this one-dimensional system the only measurable quantity is the particle's position, so that there can be no *additional* property by which we might distinguish states with the same energy.

[150] In fact, for potentials that go to a constant at infinity, such as the Coulomb potential of the hydrogen atom, the energy levels become more and more *closely* spaced for larger n.

[151] This is called the *adiabatic* approximation.

to

$$\frac{dE}{dL} = -\frac{2E}{L} \quad . \tag{178}$$

That is, the particle exerts a *force* on the wall. This force has magnitude $2E/L$ and is directed outwards[152]. Things become really interesting once we start to compare this with a *classical* system. Let us assume that we have a classical particle bouncing back and forth inside the box, with a velocity $v_{\rm cl}$. If we observe the system during a time interval Δt, the particle will bounce off the wall at $x = L$ a number of times N, given by

$$N = \frac{v_{\rm cl}}{2L}\,\Delta t \quad . \tag{179}$$

Each time the particle hits the wall, its momentum is changed by an amount $2mv_{\rm cl}$. That is, in the time interval Δt the wall transfers onto the particle an amount of momentum $-\Delta p_{\rm cl}$, given by

$$\Delta p_{\rm cl} = 2mv_{\rm cl}N = \frac{m{v_{\rm cl}}^2}{L}\Delta t = \frac{2E_{\rm cl}}{L}\Delta t \quad . \tag{180}$$

The (time-)average force exerted by the particle on the wall is therefore

$$F_{\rm cl} = \lim_{\Delta t\to\infty}\frac{\Delta p_{\rm cl}}{\Delta t} = \frac{2E_{\rm cl}}{L} \quad : \tag{181}$$

exactly the quantum result ! If we can only observe the box from the outside, it is not possible to distinguish a quantum particle from a classical one[153].

6.8 Extra ! Antimatter

Although in these notes we ought to restrict ourselves to *non*relativistic quantum mechanics, it is very interesting to study one consequence of a move toward a *relativistic* theory. Let us consider again the wave function for a free particle of mass m :

$$\psi(\vec{x},t) \propto \exp\left(-\frac{i}{\hbar}(Et - \vec{p}\cdot\vec{x})\right) \quad , \tag{182}$$

where we have switched to three space dimensions for good measure. Now, in nonrelativistic mechanics, the (kinetic) energy is given by

$$E = \frac{1}{2m}|\vec{p}|^2 \quad , \tag{183}$$

and is of course never negative[154]. In special relativity, however[155], the defining relation is[156]

$$E^2 = c^2|\vec{p}|^2 + m^2c^4 \quad , \tag{184}$$

[152] If the system is not in a pure state, we ought rather to speak of the *expectation value* of the energy. But, since all energy eigenvalues are proportional to L^{-2}, the result for the force remains the same.

[153] This is not automatic : for instance, it does not hold for the harmonic oscillator that we discuss in the next chapter. For that system, we could in principle tell the difference.

[154] Negative kinetic energy ! The very *idea* !...

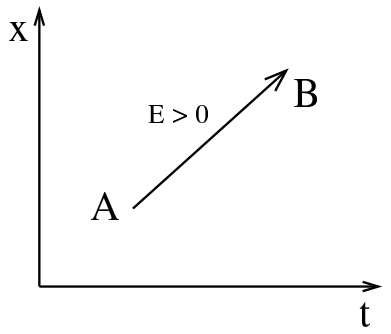
[155] You have learned this already !

[156] Here, c is the speed of light, which we use here for the first (and only) time.

and there appears to be no compelling *a priori* reason, other than provincial yokelish prejudice, to forbid *negative* values of E. Never mind that we never see particles with negative kinetic energy : the *mathematics* allow them, and taking this very seriously is what led Dirac to his famous prediction of antimatter, and Nobel honours[157]. From the simple fact that

$$Et = (-E)(-t) \tag{185}$$

we must conclude that the wave function of Eq.(182) may describe both particles with your everyday, normal, *positive* kinetic energy that move in your everyday, normal way *forward* in time[158], and particles with *negative* kinetic energy moving *backwards* in time ! Rather than blaming this notion on the physicists' exuberant imagination[159], let us try to imagine what such a particle would *look* like. To this end, let us use so-called *space-time diagrams*[160] where we denote a particle moving between two spacetime points, one at position x_A and time t_A, and one at position x_B and time t_B.



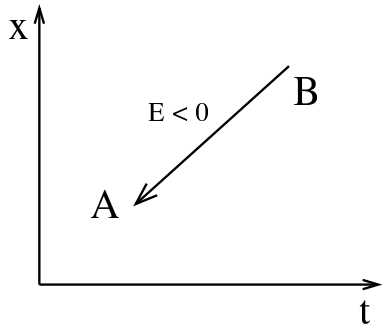
A particle with positive energy $|E|$ moves from A to B, forward in time. At position x_A, the energy *decreases*, by an amount $|E|$, and, sometime later, the energy at position x_B *increases* by the same amount $|E|$.

[157] Forget what you may have read about the Dirac equation, and maybe the Dirac sea : the whole idea is much simpler, although the presentation given here is rather in the spirit of Ernst Carl Gerlach Stückelberg and Richard Phillips Feynman than of Paul Adrien Maurice Dirac himself.

[158] Like you and me.

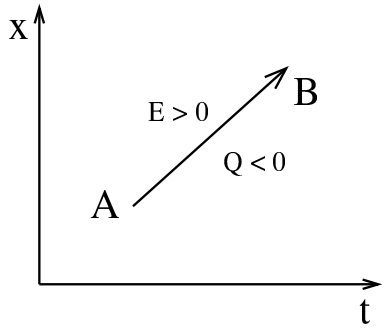
[159] Or abuse of illegal substances.

[160] Which have been further developed and are now known as Feynman diagrams — see (much) later courses.



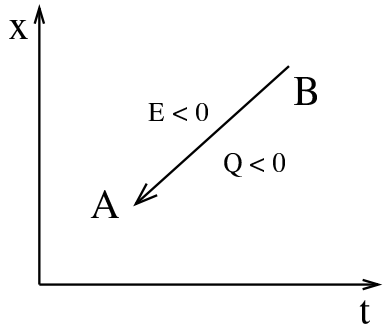
A particle with negative energy $-|E|$ moves from B to A, backwards in time. At position x_A, the energy *decreases*, by an amount $|E|$, and, sometime later, the energy at position x_B *increases* by the same amount $|E|$.

We see that, from the point of energy bookkeeping, *there is absolutely no distinction* between the two situations ! Does this mean that the two cannot be distinguished at all, and perhaps half of the particles in our bodies are, in fact, moving backwards in time[161] ? Fortunately not, since particles may have many more properties than just an an energy – for instance, they may carry an electric charge. Let us include the charge in our pictures, using the fact that an electron (say) carries a *negative* charge Q.



A particle with positive energy $|E|$ and negative charge Q moves from A to B, forward in time. At position x_A, the charge *increases*, by an amount $|Q|$, and, sometime later, the charge at position x_B *decreases* by the same amount $|Q|$. We can interpret this as, indeed, a negatively charged particle, of charge $-|Q|$, moving from A to B.

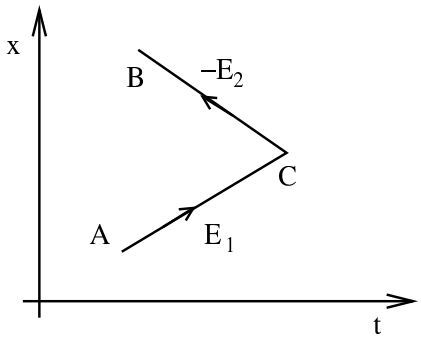
[161] Consider the blinding headaches you will develop if you try to describe chemical reactions between molecules moving forward and molecules moving backwards in time...Try to apply this to a description of your digestive process, and come to appreciate our human language's total incapacity to deal with these ideas.



A particle with negative charge $-|Q|$ moves from B to A, backwards in time. At position x_A, the charge *decreases*, by an amount $|Q|$, and, sometime later, the energy at position x_B *increases* by the same amount $|Q|$. For all purposes, this looks like a *positively* charged particle, of charge $+|Q|$, moving from A to B.

The conclusion is that particles with negative energy, moving backwards in time, may very well exist – but we shall, adhering to the most comfortable way in which we hope to see the world, *interpret* them as particles with positve energy, moving forwards in time, but with opposite charge[162] : this is called **antimatter**. The antiparticle of the electron exists, and is called the positron. It is very well known, and studied, and used[163]. So every particle known to science has its antiparticle[164]. Keep the following in mind : antiparticles do *not* move backwards in time ; they are our *interpretation* of negative-energy, backwards-moving particles, but any positron you ever meet moves forward in time, with positive energy, only with opposite charge.

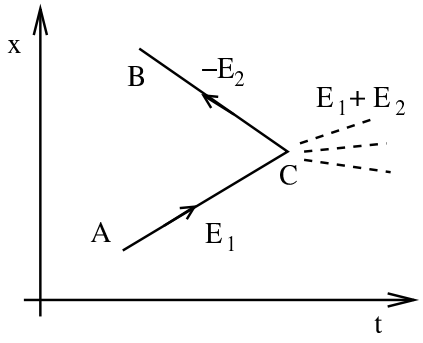
Let us now consider the following, somewhat more complicated situation.

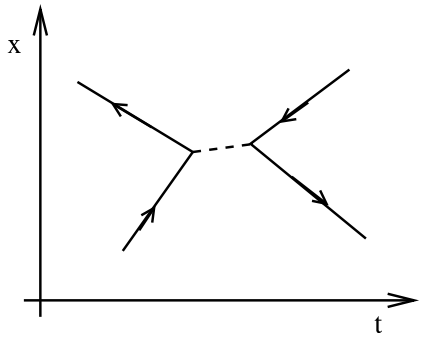


A particle starts out at A with positive energy E_1. Upon reaching C it reverses its time direction, and then moves backwards, with negative energy $-E_2$, until it reaches B. This is *one* way of telling the story : the alternative, easier version is that a particle starts out at A with positive energy E_1, and an *anti*particle starts out at B, with positive energy E_2. The particle-antiparticle pair meet at C. The two stories are equally valid, the second one is simply easier to tell. We, the observers, move along the time axis along with the rest of the universe. After we have passed time t_C, both particles have *disappeared* as you can see from the diagram : we say that the particle and

[162] And other properties, such as spin, colour, or simply the *direction* in which they move.
[163] In medical applications, such as the PET (Positron Emission Tomography) scan.
[164] Sometimes the antiparticle quantum state is the same as the particle quantum state ; a photon is an example. Obviously, such particles that are their own antiparticles must be electrically neutral.

antiparticle have *annihilated*. However, we have to take into account that total energy must be conserved ; and, importantly, that *energy is not an object but a property that is carried by particles.* The *real* story must therefore be that at the moment of annihilation, *other* particles must be created, that carry away the total energy. Often these will be photons, as is most commonly seen in the annihilation of electrons and positrons ; but protons and antiprotons will usually annihilate into so-called *pions*. Other possibilities exist : an electron can annihilate with an anti*neutrino* into a so-called W^- boson. Which processes are possible and which not depends on the various existing interactions[165]. The reverse process of annihiliation, *pair creation*, is of course also possible, so that the even more interesting scenario comes to mind: annihilation of a particle-antiparticle pair, followed by creation of a *different*, possible never-before-seen particle-antiparticle pair. This is a common way of producing new and interesting particles at high-energy colliders. The spacetime diagram depicting such a process is an elementary form of a so-called *Feynman diagram*.





6.9 Exercises 30 to 45

Exercise 30 : **Conjugates of products**
Write the identity

$$\langle\psi|\,\phi\rangle = \langle\phi|\,\psi\rangle^*$$

in the language of wave functions.

Exercise 31 : **Momentum is Hermitean**
Consider the following alternative form of Eq.(152) :

$$\langle\phi|\,\hat{p}\,|\psi\rangle = \Big(\hat{p}\;|\phi\rangle\Big)^\dagger\;|\psi\rangle$$

and show, by partial integration, that this gives the same expression. This proves that the momentum operator is indeed Hermitean, also for wave functions (NB : you may assume that the wave function vanishes at the integral's endpoints).

[165] For instance, an electron cannot annihilate with an antiproton.

Exercise 32 : **A nice and simple theorem**
Use Eq.(152) to prove the following : if the wave function has a *constant* complex phase (for instance, if it is real), then $\langle p \rangle = 0$.

Exercise 33 : **On kinetic energy**
Prove by partial integration that the expectation value of the kinetic energy can be written as

$$\langle E_{\rm kin} \rangle = \frac{\hbar^2}{2m} \int dx \, |\psi'(x)|^2$$

and that this is always *positive*.

Exercise 34 : **Wave functions in three dimensions**
Generalize the form of Eqs. (156), (157) and (158) to systems in three dimensions.

Exercise 35 : **The Wronskian**
The so-called *Wronskian* of two functions $f_1(x)$ and $f_2(x)$ is given by

$$W(f_1, f_2) = f_1(x) f_2'(x) - f_1'(x) f_2(x)$$

Let $\psi_1(x)$ and $\psi_2(x)$ be two solutions of the time-independent Schrödinger equation with the same energy eigenvalue. Show that $W(\psi_1, \psi_2)$ is constant, that is, independent of x.

Exercise 36 : **Plane waves in three dimensions**
Generalize the form of Eq.(163) to the case of three dimensions.

Exercise 37 : **A moving lump**
This is a challenging exercise ! You can use the following general formula[166] :

$$\int_{-\infty}^{\infty} dx \, \exp\left(-a(x-z)^2\right) = \sqrt{\frac{\pi}{a}}$$

Here, a can be a complex number with *positive* real part, and z can be an *arbitrary* complex number.

We consider the following wave function of a free particle of mass m in one dimension : it is a superposition of plane waves with the correct relation $E = p^2/2m$ between energy and momentum, and the various values of p are weighted with a normal distribution centered around p_0, with width σ. This means that

$$\psi(x,t) \; \propto \int_{-\infty}^{\infty} dp \, \exp\left(-\frac{(p-p_0)^2}{4\sigma^2}\right) \exp\left(-\frac{i}{\hbar}\left(\frac{p^2 t}{2m} - px\right)\right)$$

Note that here p is not an *operator*, but just a *number* over which we integrate.

[166] If you haven't learned this important equality, have a look at appendix 12.

1. Perform the p integral and show that

$$\psi(x,t) \;\propto\; \exp\left(\frac{-m\sigma^2x^2 + i\hbar m p_0 x - i\hbar {p_0}^2 t/2}{\hbar(m\hbar + 2i\sigma^2 t)}\right)$$

2. Show that

$$|\psi(x,t)|^2 = N \;\exp\left(-\frac{2\sigma^2(mx - p_0 t)^2}{m^2\hbar^2 + 4\sigma^4 t^2}\right)$$

   and compute N by requiring the correct normalization of the wave function.

3. Show that the peak of this lump moves with constant velocity p_0/m.

4. Show that

$$\langle x\rangle = \frac{p_0}{m}t \quad , \quad \sigma(x)^2 = \frac{\hbar^2}{4\sigma^2} + \frac{\sigma^2 t^2}{m^2}$$

5. Show that the lump gradually spreads out as it moves ; and show why this is only to be expected.

Exercise 38 : **Going 3D**
Derive the continuity equation in the three-dimensional case.

Exercise 39 : **No flow**
Show that $J = 0$ for real-valued wave functions, such as those for bound states.

Exercise 40 : **Only positive n?**
Show that for the n-values of Eq.(174) only the positive integers are to be chosen, *i.e.* that there are good reasons not to take $n = 0$ or $n = -1, -2, -3, \ldots$.

Exercise 41 : **Normalization factor for the particle in the box**
Show that the normalization factor $A = \sqrt{2/L}$ in Eq.(175) is correct.

Exercise 42 : **A paradox in the box?**
Show that the spread $\sigma(x)^2$ for the particle in the box can never be infinite. Use the Heisenberg inequality to show that the system can therefore *never* be in an eigenstate[167] of $\hat{p}$. Discuss how it can be possible that the system is, nevertheless, in an eigenstate of $\hat{p}^2$.

[167] This only goes to show that the box is not a totally realistic system, since an Hermitean operator such as $\hat{p}$ *ought* to have eigenstates.

Exercise 43 : **Motion in the box**
Show that the most general time-dependent state of a particle in the box is given by the wave function

$$\psi(x,t) = \sum_{n\geq 1} c_n \, \sin\left(\frac{n\pi}{L}\, x\right) \, \exp\left(-i\frac{\hbar\pi^2 n^2}{2mL^2}t\right) \quad , \quad \sum_{n\geq 1} |c_n|^2 = \frac{2}{L} \ .$$

Next, assume that $c_1 = c_2 = 1/\sqrt{L}$ and all other c_n are zero, and compute $\langle x \rangle$ as a function of time.

Exercise 44 : **Real-life box models ?**
We shall employ some realistic numbers in order to see how the particle-in-a-box model fits into physics experience.

1. Consider a particle with a mass of 1 gram, moving about in a (one-dimensional) box of size 1 cm, with a velocity of 1 cm per second. Compute its kinetic energy, and determine the value of n it must have to attain the same energy in the quantum model.

2. We may try to model an atom such as the hydrogen atom as a (one-dimensional) box in which an electron moves about. The 'typical energy scale' governing the energy difference between states is of course $E_1 = \hbar^2\pi^2/(2mL^2)$. The electron has a mass of about 10^{-30} kg, and the atom has a size of about $2\ 10^{-10}$ m. Compute E_1, and compare this with what you know about typical atomic energy levels[168].

Exercise 45 : **The quantum bead**
This exercise contains some serious calculation about another system. We consider a quantum system consisting of a 'bead' of mass m that can move on a cricle (a 'wire') of radius R around some origin. The *classical* position of the bead is then given by its angular coordinate $\phi_{\rm cl}$. No force is acting on the bead as long as it is on the circular wire. The Cartesian coordinates of the bead can be written as

$$x_{\rm cl} = R\,\sin\phi_{\rm cl} \quad , \quad y_{\rm cl} = R\,\cos\phi_{\rm cl}$$

1. Show that the classical *angular momentum* can be written as

$$(p_\phi)_{\rm cl} = mR^2\,\dot{\phi}_{\rm cl}$$

where the dot denotes the time derivative[169]. Also show that the kinetic energy can be written as

$$E_{\rm cl} = \frac{1}{2mR^2}{(p_\phi)_{\rm cl}}^2$$

[168] One electronvolt is $1.6\ 10^{-19}$ Joule.
[169] This is not a *quantum* exercise item, but simply one from *classical* mechanics.

2. Show that

$$\phi_{\rm cl} = \arctan(x_{\rm cl}/y_{\rm cl})$$

and that

$$\frac{\partial}{\partial \phi_{\rm cl}} = \frac{\partial x_{\rm cl}}{\partial \phi_{\rm cl}} \frac{\partial}{\partial x_{\rm cl}} + \frac{\partial y_{\rm cl}}{\partial \phi_{\rm cl}} \frac{\partial}{\partial y_{\rm cl}} = y_{\rm cl} \frac{\partial}{\partial x_{\rm cl}} - x_{\rm cl} \frac{\partial}{\partial y_{\rm cl}}$$

3. Let us now, as usual, replace the classical observables by quantum operators : this means $x_{\rm cl} \to \hat{x}$ and $y_{\rm cl} \to \hat{y}$, but also $\phi_{\rm cl} \to \hat{\phi}$. In the wave-function formalism, we had found that the momentum operators are

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x} \quad , \quad \hat{p}_y = -i\hbar \frac{\partial}{\partial y}$$

Use the results of the previous step to show that we can similarly introduce the *angular momentum operator* $\hat{p}_\phi$ as

$$\hat{p}_\phi = -i\hbar \frac{\partial}{\partial \phi}$$

by showing[170] that it has the correct commutation relation :

$$[\hat{\phi}, \hat{p}_\phi] = \left[\arctan\left(\frac{x}{y}\right), -i\hbar \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) \right] = i\hbar$$

4. We are now ready to describe the quantum system of the bead in wave-function language. The Hamiltonian is

$$\hat{H} = \frac{1}{2mR^2} (\hat{p}_\phi)^2 \quad , \quad \hat{p}_\phi = -i\hbar \frac{\partial}{\partial \phi}$$

and the wave function is of course a function of ϕ : we write it as $\psi(\phi)$. Write the time-independent Schrödinger equation for this system.

5. Keep in mind that the wave function must be periodic in ϕ :

$$\psi(\phi + 2\pi) = \psi(\phi)$$

Show that the normalized energy eigenstates and the corresponding energy eigenvalues can be written as

$$\psi_n(\phi) = \frac{1}{\sqrt{2\pi R}} e^{in\phi} \quad , \quad E_n = \frac{\hbar^2 n^2}{2mR^2} \quad n = 0, \pm 1, \pm 2, \pm 3, \ldots$$

[170] You can use the fact that

$$\frac{d}{dz} \arctan(z) = \frac{1}{1+z^2} \ .$$

Also, do not forget that we are dealing with *operators* here : you have to let the commutator work on an *arbitrary function* of x and y.

7 The Harmonic oscillator

—A five-year-old kid from Minnesota has patented a way of swinging on a child's swing. The US Patent Office issued patent 6,368,227 on 9 April to Steven Olson of St Paul, Minnesota for a "method of swinging on a swing".
New Scientist, April 2002

7.1 A general phenomenon

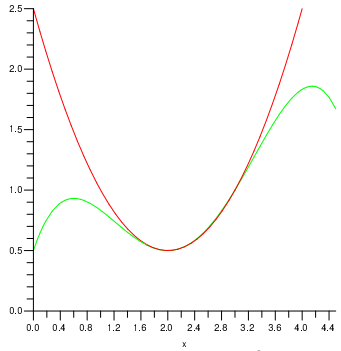
A very important system in quantum mechanics, as well as in classical mechanics, is that of a particle in a quadratic potential. The classical expression for the energy of such a system (in one dimension) is given by

$$E_{\rm cl} = \frac{1}{2m}\, {p_{\rm cl}}^2 + \frac{1}{2} m\omega^2\, {x_{\rm cl}}^2 \quad . \tag{186}$$

The *reason* why this is so general a system is that it describes very many potentials *in the neighbourhood of a minimum*. Let a general potential be given by $V(x)$, and let $x = x_0$ be the position of a minimum of this potential. Then, Taylor expansion around $x = x_0$ gives us

$$\begin{aligned} V(x) &= V(x_0) + (x - x_0)\, V'(x_0) + \frac{1}{2}(x - x_0)^2\, V''(x_0) + \dots \\ &= V(x_0) + \frac{1}{2}(x - x_0)^2\, V''(x_0) + \dots \quad , \end{aligned} \tag{187}$$

since the first derivative vanishes at the minimum. As long as we consider only small deviations from the equilibrium value x_0, the potential is well approximated by a quadratic form.



A generic one-dimensional potential, together with its quadratic approximation as given by the Taylor expansion. For small deviations around the position of the (local) minimum, $x = 2$ in this case, the approximation is quite good.

We define $V''(x_0) = m\omega^2$, where ω has the dimension of a *frequency*. Classically, all motions in such a potential are expressed in terms of sines and cosines of ωt, that is, *harmonic* motions : hence the name *harmonic oscillator* for this system. **E 46**

We shall now discuss the quantum mechanical harmonic oscillator, which must have the following Hamiltonian, in analogy to Eq.(186) :

$$\hat{H} = \frac{1}{2m}\,\hat{p}^2 + \frac{1}{2}m\omega^2\,\hat{x}^2 \quad . \tag{188}$$

We shall do so in a slightly roundabout manner ; but on the way we will learn some interesting things.

7.2 Raising and lowering operators

Let us consider a one-dimensional quantum system with an *infinite* number of non-degenerate energy eigenstates[171], that we label by an index n starting from $n = 0$. The normalized states will be denoted, accordingly, by $|n\rangle$. For such a system we can define the *annihilation* (or *lowering*) operator $\hat{a}$ as

$$\hat{a} = \sum_{n\geq 0} \sqrt{n+1}\ |n\rangle\,\langle n+1| \quad . \tag{189}$$

We see that this operator, when applied on a state $|q\rangle$, results in a state proportional to $|q-1\rangle$; when applied to $|0\rangle$ it gives zero. The operator $\hat{a}$ is not Hermitean — in fact it does not even *commute* with its conjugate, the *creation* (or *raising*) operator :

$$\hat{a}^\dagger = \sum_{n\geq 0} \sqrt{n+1}\ |n+1\rangle\,\langle n| \quad . \tag{190}$$

The products of $\hat{a}$ and $\hat{a}^\dagger$ are

$$\begin{aligned}
\hat{a}\,\hat{a}^\dagger &= \sum_{n,n'\geq 0} \sqrt{(n+1)(n'+1)}\ |n\rangle\,\langle n+1|\,n'+1\rangle\,\langle n'| \\
&= \sum_{n\geq 0} (n+1)\ |n\rangle\,\langle n| \quad ; \\
\hat{a}^\dagger\hat{a} &= \sum_{n,n'\geq 0} \sqrt{(n+1)(n'+1)}\ |n+1\rangle\,\langle n|\,n'\rangle\,\langle n'+1| \\
&= \sum_{n\geq 1} n\ |n\rangle\,\langle n| \ = \sum_{n\geq 0} n\ |n\rangle\,\langle n| \quad .
\end{aligned} \tag{191}$$

Hence the *commutator* of $\hat{a}$ and $\hat{a}^\dagger$ is[172]

$$[\hat{a}, \hat{a}^\dagger] = \hat{a}\,\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = \sum_{n\geq 0} |n\rangle\,\langle n| = \hat{1} \quad , \tag{192}$$

[171] This can be seen as inspired by the particle-in-a-box system. If a potential goes all the way up to infinity, so that a particle can never escape, there are indeed infinitely many *bound states* ; but many other systems, including the hydrogen atom, also have an infinite number of bound states.

[172] In a sense it is this commutation relation that *defines* $\hat{a}$ and $\hat{a}^\dagger$ to be creation and annihilation operators.

and the *anticommutator* (denoted by curly brackets) is

$$\{\hat{a}, \hat{a}^\dagger\} = \hat{a}\,\hat{a}^\dagger + \hat{a}^\dagger\hat{a} = \sum_{n\geq 0} (2n+1)\ |n\rangle\,\langle n| \quad . \tag{193}$$

Starting from the lowest-energy normalized state $|0\rangle$, we can build all higher-energy states by repeated application of the creation operator. For instance:

$$\hat{a}^\dagger\ |0\rangle = |1\rangle \quad , \quad \hat{a}^\dagger\ |1\rangle = \sqrt{2}\ |2\rangle \quad , \quad \hat{a}^\dagger\ |2\rangle = \sqrt{3}\ |3\rangle \quad , \tag{194}$$

and so on : the general rule is

$$|n\rangle = \frac{1}{\sqrt{n!}}\ \left(\hat{a}^\dagger\right)^n\ |0\rangle \quad . \tag{195}$$

7.3 Matrix notation

If we want we can also express the findings of the previous section in matrix notation, as long as we are willing to live with *infinitely* big vectors. The various states may then be written as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad , \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} \quad , \tag{196}$$

and so on. The lowering operator $\hat{a}$ will then of course an infintely big *matrix* :

$$\hat{a} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & 0 & 0 & \sqrt{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad , \tag{197}$$

and the raising operator is its *transpose*[173] :

$$\hat{a}^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad . \tag{198}$$

These matrices have their nozero elements just off the diagonal. You can easily check their (anti)commutation relations by explicit matrix multiplication. Since these matrices are infinitely large, they are however somewhat clumsy to use for explicit calculation, and the ket/bra notation is probably preferable.

E 47

[173] Since these matrices are purely real, their Hermitean conjugation amounts to simply interchanging rows and columns.

7.4 A Hamiltonian and its spectrum

Let us *suppose* that the Hamiltonian of the system *happens* to be proportional to the anticommutator of Eq.(193) :

$$\hat{H} = K \{\hat{a}, \hat{a}^\dagger\} \ ; \tag{199}$$

then we know immediately the spectrum of this system : for we have

$$\hat{H} \ |n\rangle = K \ (2n+1) \ |n\rangle \quad . \tag{200}$$

The energy spectrum[174] is then given by the values $K(2n+1)$, $n = 0, 1, 2, 3, \ldots$.

7.5 The Hamiltonian in operator form

We shall now show that the Hamiltonian of the harmonic oscillator can, indeed, be written using the anticommutator. It stands to reason that the (non-Hermitean!) operator $\hat{a}$ must be formed from position and momentum operators, since those are the fundamental building blocks of our quantum world[175]. We therefore write

$$\hat{a} = A\hat{x} + iB\hat{p} \quad , \quad \hat{a}^\dagger = A\hat{x} - iB\hat{p} \ , \tag{201}$$

where A and B are some real constants[176]. The commutation relation of Eq.(192) tells us something about these constants :

$$[\hat{a}, \hat{a}^\dagger] = -2iAB \ [\hat{x}, \hat{p}] = 2\hbar \, AB \ , \tag{202}$$

so that we must satisfy

$$AB = 1/(2\hbar) \tag{203}$$

if the commutator is indeed equal to the identity operator. Another relation follows from the assumption we made on the Hamiltonian :

$$\hat{H} = K\{\hat{a}, \hat{a}^\dagger\} = 2KA^2 \, \hat{x}^2 + 2KB^2 \, \hat{p}^2 \quad . \tag{204}$$

If we now compare this to Eq.(188), we see that we can indeed get the correct form for the Hamiltonian if

$$2KA^2 = \frac{1}{2} m\omega^2 \quad , \quad 2KB^2 = \frac{1}{2m} \quad . \tag{205}$$

Multiplying these two results together, we get

$$(2KA^2)(2KB^2) = 4K^2(AB)^2 = \frac{K^2}{\hbar^2} = \frac{\omega^2}{4} \quad , \tag{206}$$

[174] We must observe that there *must* be *some* state with lowest energy. The potential energy is never negative because of its quadratic form : and we have proven in exercise 33 that the kinetic energy can never be negative either. We are therefore allowed to number states starting at some lowest value, for which we take, of course, $n = 0$.

[175] So far, at least.

[176] This disregards a possible overall complex phase in the definition of $\hat{a}$: as we shall see we don't need such a phase anyway.

in other words, $K = \hbar\omega/2$. From this it easily follows[177] that

$$A^2 = \frac{m\omega}{2\hbar} \quad , \quad B^2 = \frac{1}{2m\hbar\omega} \ , \tag{207}$$

and that *we have now solved the system completely* ! The Hamiltonian is given by Eq.(188) ; the annihiliation and creation operators are given by **E 48**

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} + \frac{i}{\sqrt{2m\hbar\omega}}\,\hat{p} \quad , \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} - \frac{i}{\sqrt{2m\hbar\omega}}\,\hat{p} \ ; \tag{208}$$

the energy eigenvalues are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad , \quad n = 0, 1, 2, \ldots \tag{209}$$

and the corresponding energy eigenstates are given by Eq.(195), which can be used as soon as we know the state $|0\rangle$. **E 49**

7.6 Wave functions for the harmonic oscillator

It is not a very diffcult matter to find the wave functions for the harmonic oscillator. In fact, we only have to find the wave function for the ground state. All the other ones will then follow automatically from blind application of Eq.(195). The ground-state wave function is found from the requirement that the ground state $|0\rangle$ is indeed the state with lowest energy, so that trying to lower the energy further is lethal :

$$\hat{a}\ |0\rangle = 0 \ . \tag{210}$$

For the annihiliation operator in wave-function language we simply insert what the position and momentum operators do to wave functions :

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} + \frac{i}{\sqrt{2m\hbar\omega}}\,\hat{p} = \sqrt{\frac{m\omega}{2\hbar}}\,x + \sqrt{\frac{\hbar}{2m\omega}}\,\frac{\partial}{\partial x} \ ; \tag{211}$$

The ground-state wave function $\psi_0(x)$ must therefore satisfy the following form of Eq.(210) :

$$\left(\sqrt{\frac{m\omega}{2\hbar}}\,x + \sqrt{\frac{\hbar}{2m\omega}}\,\frac{\partial}{\partial x}\right)\psi_0(x) = 0 \ . \tag{212}$$

This simple linear homogeneous first-order equation is immediately solved : **E 50**

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) \ . \tag{213}$$

As mentioned before, the higher-energy states are then given by

$$\psi_n(x) = \frac{1}{\sqrt{n!}}\left(\sqrt{\frac{m\omega}{2\hbar}}\,x - \sqrt{\frac{\hbar}{2m\omega}}\,\frac{\partial}{\partial x}\right)^n \psi_0(x) \ . \tag{214}$$

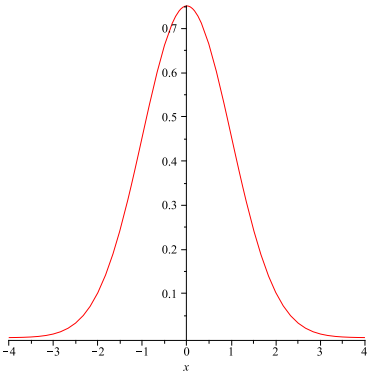
[177] Note that Eq.(202) tells us that A and B must have the same sign.

Note that we have now *constructed* all the normalized solutions to the time-independent Schrödinger equation

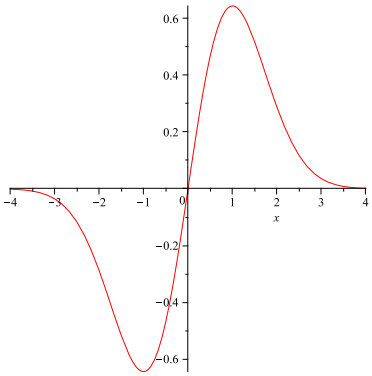
$$-\frac{\hbar^2}{2m}\psi_n''(x) + \frac{m\omega^2}{2}\psi_n(x) = E_n\psi(x) \quad , \tag{215}$$

and that we have also determined that the *only* possible energy eigenvalues are those given by Eq.(209) — quite an achievement ! Below, we give the first few wave functions $\psi_0, \psi_1, \psi_2, \psi_3$.

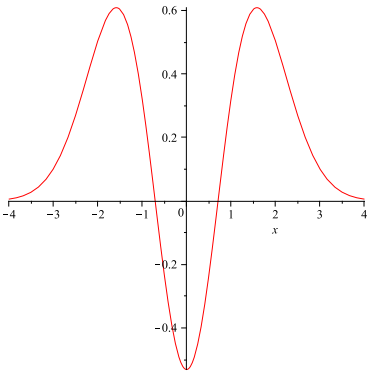
$\psi_0(x)$



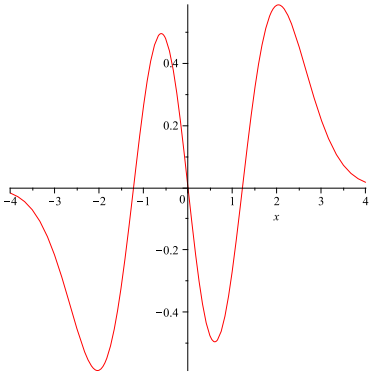
$\psi_1(x)$



$\psi_2(x)$

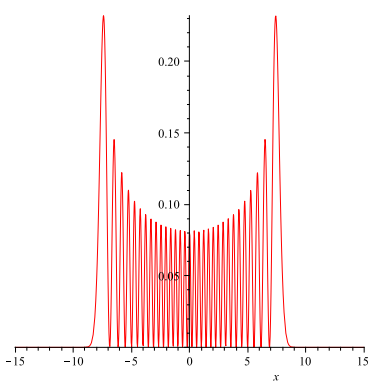


$\psi_3(x)$



in these plots, I used $m = \omega = \hbar = 1$. As n increases, the number of nodes keeps pace, as in fact it did in the case of the particle-in-the-box of the previous section.

Another issue of interest becomes apparent if we take n quite a bit higher.



Here we plot $|\psi_{30}(x)|^2$. The probability density for the position oscillates wildly, but shows an average behaviour that tells us that the particle is more likely to be found at the 'endpoints' than around the origin. In fact, *this is also the classical behaviour* ! If you swing on a swing, you pass the lowest point quickly, and spend more time at the extreme points of your trajectory, simply because you are moving more slowly there.

E 51

7.7 Expectation values of position and momentum

We may be interested not so much in the wave functions themselves but rather in the various expectation values. Of course, the result for $\langle x \rangle$ can be obtained by performing the integral

$$\int_{-\infty}^{\infty} dx \; x \; |\psi_n(x)|^2$$

and such, but again the creation and annihilation operators can be used to advantage ! This we do by expressing $\hat{x}$ and $\hat{p}$ in terms of these operators :

$$\hat{x} = \frac{1}{2A}\left(\hat{a} + \hat{a}^\dagger\right) \quad , \quad \hat{p} = \frac{1}{2iB}\left(\hat{a} - \hat{a}^\dagger\right) \quad , \tag{216}$$

where A and B are given by Eq.(207). For the system in a state $|\psi\rangle$ we then find, for $\langle x \rangle$:

$$\langle x \rangle = \langle \psi | \, \hat{x} \, | \psi \rangle = \frac{1}{2A}\left(\langle \psi | \, \hat{a} \, | \psi \rangle + \langle \psi | \, \hat{a}^\dagger \, | \psi \rangle \right) = \frac{1}{A} \, \mathrm{Re} \, \langle \psi | \, \hat{a} \, | \psi \rangle \quad , \tag{217}$$

and for $\langle p \rangle$, in an analogous manner :

$$\langle p \rangle = \frac{1}{B} \, \mathrm{Im} \, \langle \psi | \, \hat{a} \, | \psi \rangle \quad . \tag{218}$$

We also have

$$\hat{x}^2 = \frac{1}{4A^2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + \{\hat{a}, \hat{a}^\dagger\}\right) \quad , \quad \hat{p}^2 = \frac{1}{4B^2}\left(-\hat{a}^2 - \hat{a}^{\dagger 2} + \{\hat{a}, \hat{a}^\dagger\}\right) \quad . \tag{219}$$

Let us now assume that $|\psi\rangle$ is the energy eigenstate $|n\rangle$. Since $\hat{a} \, |n\rangle$ is proportional to $|n-1\rangle$, we find immediately that

$$\langle x \rangle = \langle p \rangle = 0 \quad , \quad \langle x^2 \rangle = \frac{2n+1}{4A^2} \quad , \quad \langle p^2 \rangle = \frac{2n+1}{4B^2} \quad , \tag{220}$$

and we can check Heisenberg's inequality :

$$\sigma(x)\,\sigma(p) = \frac{2n+1}{4AB} = (2n+1)\frac{\hbar}{2} \quad . \tag{221}$$

With increasing energy the uncertainty product in the eigenstate becomes larger and larger. It is possible to construct states with a smaller uncertainty product, as we shall see in the next section.

7.8 Extra ! Coherent states

In this section, let us step back from the harmonic oscillator, and simply consider a system with an infinity of states labelled as $|0\rangle, |1\rangle, |2\rangle, \ldots$. As before we have annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$, respectively, as given by Eqs.(189) and (190). These are not Hermitean and do not correspond to observables ; but the two combinations

$$\hat{A} = \frac{1}{2}\left(\hat{a} + \hat{a}^\dagger\right) \quad , \quad \hat{B} = \frac{1}{2i}\left(\hat{a} - \hat{a}^\dagger\right) \tag{222}$$

are Hermitean and therefore in principle measurable[178]. Their commutator is given by

$$[\hat{A}, \hat{B}] = i/2 \quad . \tag{223}$$

Now, let us quickly redo the argument of sec.3.7. We construct the state

$$|\zeta(z)\rangle = \left(\hat{A} - \langle A\rangle + iz(\hat{B} - \langle B\rangle)\right)|\psi\rangle \quad , \tag{224}$$

where $\langle A\rangle$ and $\langle B\rangle$ are the expectation values for the normalized state $|\psi\rangle$:

$$\langle A\rangle = \langle\psi|\,\hat{A}\,|\psi\rangle \quad , \quad \langle B\rangle = \langle\psi|\,\hat{B}\,|\psi\rangle \quad . \tag{225}$$

By squaring the state $|\zeta(z)\rangle$, we have

$$0 \leq \langle\zeta(z)|\,\zeta(z)\rangle = \sigma(A)^2 + z^2\sigma(B)^2 - \frac{z}{2} \quad . \tag{226}$$

The minimum is reached for $z = z_0 = 1/(4\sigma(B)^2)$, so that

$$\sigma(A)\sigma(B) \geq \frac{1}{4} \quad . \tag{227}$$

We can now ask ourselves the following question : what is the state that looks *most classical* ? For this state, the uncertainty product in Eq.(227) must surely take the minimal possible value 1/4, and this is reached for $\langle\zeta(z_0)|\,\zeta(z_0)\rangle = 0$, or $|\zeta(z_0)\rangle = 0$. For no other good reason than symmetry we will assume that $\sigma(A) = \sigma(B)$ so they both equal 1/2, and hence $z_0 = 1$. Inserting this in Eq.(224), we see that the condition $|\zeta(1)\rangle = 0$ actually means that

$$\hat{a}\,|\psi\rangle = \Big(\,\langle A\rangle + i\,\langle B\rangle\Big)\,|\psi\rangle \quad . \tag{228}$$

[178] In fact, for the harmonic oscillator $\hat{A} \propto \hat{x}$ and $\hat{B} \propto \hat{p}$.

We conclude that for a quantum system with annihilation and creation operators the state that is 'most classical' is an *eigenstate* of the *annihilation* operator. Its eigenvalue is determined by the expectation values of the observables A and B (position and momentum, in the case of the harmonic oscillator). Such states are called *coherent states*. Excercise 52 shows how to compute it explicitly for the harmonic oscillator. **E 52**

7.9 Exercises 46 to 52

Exercise 46 : **Dimensional arguments**
Show that the quantity ω defined in section 7.1 has the dimension of a *frequency*, that is inverse seconds. Show that this is (up to a constant factor) the *only* quantity with the dimension of a frequency that can be formed from the objects in the classical harmonic oscillator.

Exercise 47 : **Explicit work with infinite matrices**
Using the explicit matrix notation for $\hat{a}$ and $\hat{a}^\dagger$ as given in section 7.3, show that

$$[\hat{a},\hat{a}^\dagger] = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad , \quad \{\hat{a},\hat{a}^\dagger\} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 3 & 0 & 0 & \cdots \\ 0 & 0 & 5 & 0 & \cdots \\ 0 & 0 & 0 & 7 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Exercise 48 : **A dimensional analysis**
Use dimensional arguments to show the following :

1. The annihilation operator $\hat{a}$ and the creation operator $\hat{a}^\dagger$ must be dimensionless ;
2. The constant K of Eq.(199) must be proportional to $\hbar\omega$ in the sense that this is the only object in our system[179] with the dimension of an energy ;
3. The constants A and B of Eq.(201) do indeed have the dimension of an inverse momentum and an inverse length, respectively, and that they are the only possible such objects, again up to a constant factor.

We see that, in fact, the whole problem in solving the harmonic oscillator is in determining where some factors of 2 have to go...

Exercise 49 : **A more complicated spectrum**
We can use the formalism of creation and annihiliation operators for more complicated (but less useful) systems as well. Let us take $\hbar = m = \omega = 1$ for

[179] Not containing the unknowns X and P.

simplicity. Suppose that we want to find a one-dimensional model for which the (nondegenerate) spectrum has the form

$$E_n = E_0(n+1)^2 \quad , \quad n = 0, 1, 2, 3, \ldots$$

1. The Hamiltonian must have the form

$$\hat{H} = \sum_{n \geq 0} E_0(n+1)^2 \ |n\rangle \langle n|$$

Figure out how to write the Hamiltonian in terms of $\hat{a}$ and $\hat{a}^\dagger$, using the relations

$$[\hat{a}, \hat{a}^\dagger] = \sum_{n \geq 0} |n\rangle \langle n| \quad , \quad \{\hat{a}, \hat{a}^\dagger\} = \sum_{n \geq 0} (2n+1) \ |n\rangle \langle n|$$

2. Write the creation and annihilation operators in terms of $\hat{x}$ and $\hat{p}$, analogous to the way they where introduced in Eq.(201), and write down what $\hat{H}$ now looks like in terms of the operators $\hat{x}$ and $\hat{p}$. Is this an energy operator you expect to be likely to encounter in physical problems ?

Exercise 50 : **Normalizing the ground state**
Show that the ground-state wave function of the harmonic oscillator in Eq.(213) is correctly normalized.

Exercise 51 : **Timing your swing**
Consider a classical particle that undergoes harmonic motion with frequency ω and amplitude a:

$$x_{\text{cl}}(t) = a \sin(\omega t)$$

Compute dx_{cl}/dt and, from that dt/dx_{cl}, and express the result as a function of x_{cl} rather than as a function of t. Sketch the result.

Exercise 52 : **Coherent states**
This exercise should please you if you like serious calculation : we are going to compute the coherent states of the harmonic oscillator. If it helps you may take $m = \hbar = \omega = 1$.

1. Prove that (somewhat surprisingly) the creation operator $\hat{a}^\dagger$ has *no* eigenstates[180].

2. Compute the eigenstates of $\hat{a}$ by the following steps.

[180] It is not Hermitean and therefore does not correspond to an observable : there is therefore no *a priori* reason to expect that it refers to any measurable property that could be well-defined. And behold ! it doesn't.

(a) Suppose that $\hat{a}$ has an eigenstate of the form

$$|s\rangle = \sum_{n\geq 0} c_n \ |n\rangle$$

where s is the eigenvalue : it is in principle a *complex number*. We have to determine the coefficients c_n.

(b) Write out the eigenvalue equation

$$\hat{a} \ |s\rangle = s \ |s\rangle$$

and show that this implies

$$c_{n+1} \ \sqrt{n+1} = s \ c_n \quad , \quad n \geq 0$$

(c) Show that the coefficients must be given by the equation

$$c_n = c_0 \ \frac{s^n}{\sqrt{n!}}$$

(d) Use the normalization $\langle s|\, s\rangle = 1$ to prove that

$$c_0 = \exp\left(-\frac{1}{2}|s|^2\right)$$

3. Use the techniques of section 7.7 to compute[181] $\langle x\rangle$, $\langle p\rangle$, $\sigma(x)$ and $\sigma(p)$, all these as a function of s. Show that these states satisfy the Heisenberg inequality in the best possible way, namely that $\sigma(x)\,\sigma(p) = \hbar/2$.

4. Show that we can write

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\right) \quad , \quad \hat{H}^2 = (\hbar\omega)^2 \left({\hat{a}^\dagger}^2 \hat{a}^2 + 2\hat{a}^\dagger \hat{a} + \frac{1}{4}\right)$$

Compute $\langle E\rangle$ and $\sigma(E)$ for the state $|s\rangle$, and show that $\sigma(E)/\langle E\rangle$ goes like $1/|s|$ for large s.

5. Let ω be 1 per second. Consider a classical particle of mass 1 gram, swinging harmonically with angular frequencey ω, and let it have a velocity of 1 cm per second at the equilibrium position. Compute its classical energy $E_{\rm cl}$. Determine the value of $|s|$ necessary to have $\langle E\rangle = E_{\rm cl}$. Can we say that the energy is narrowly defined ?

6. We now investigate the time dependence of states like $|s\rangle$.

[181] NB : in this item, it is very useful to realize that $\hat{a}\,|s\rangle = s\,|s\rangle$, $\langle s|\,\hat{a}^\dagger = s^*\,\langle s|$, and $\hat{a}\,\hat{a}^\dagger = 1 + \hat{a}^\dagger \hat{a}$.

(a) Suppose that the state is $|s\rangle$ at time $t = 0$. Show that at later time t the coefficients c_n are to be replaced by

$$c_n = \exp\left(-\frac{1}{2}|s|^2\right) \frac{s^n}{\sqrt{n!}} \exp\left(-i\omega t(n + 1/2)\right)$$

(b) Show that this is, physically, completely æquivalent to replacing

$$s \quad \rightarrow \quad s\, e^{-i\omega t}$$

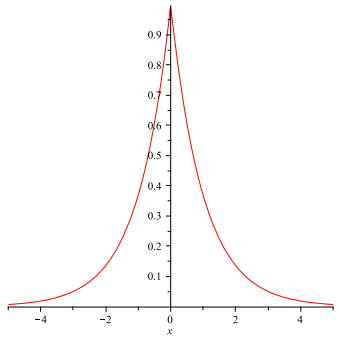
(c) Show that this means that $\langle x\rangle\,(t)$ and $\langle p\rangle\,(t)$ behave exactly like $x_{\rm cl}$ and $p_{\rm cl}$ would do for the *classical* harmonic oscillator.

8 Other one-dimensional systems

— L. Burke (1952). "On the tunnel effect" (Abstract).
The Quarterly Journal of Experimental Psychology.
Vol. 4. Issue 3. pp. 121-138.

8.1 A kinky function

In this section we shall study a few simple one-dimensional systems ; we shall use wave-function language throughout. As we have remarked before, the wave function must be continuous everywhere, but kinks are allowed : we will now have a closer look at this. A very simple function that is quadratically integrable and shows a kink is given by



$$\Lambda(\alpha;x) \equiv \exp(-\alpha|x|) \quad , \tag{229}$$

where α is a parameter with the dimension of inverse length[182]. The plot shows this function, with $\alpha = 1$. The derivative of $|x|$ is the *sign* of x, that is, it is -1 for $x < 0$ and $+1$ for $x > 0$. The derivative of $\Lambda(x)$ is therefore[183] **E 53**

$$\Lambda'(\alpha;x) = \Big(1 - 2\theta(x)\Big)\ \alpha\ \Lambda(x) \quad , \tag{230}$$

and its *second* derivative reads

$$\Lambda''(\alpha;x) = \alpha^2\,\Lambda(\alpha;x) - 2\alpha\,\delta(x)\,\Lambda(a;x) = \alpha^2\,\Lambda(\alpha;x) - 2\alpha\,\delta(x) \quad , \tag{231}$$

since $\Lambda(\alpha;0) = 1$. The Dirac delta crops up again ! Since the Schrödinger equation involves precisely this second derivative, we can immediately draw two conclusions. In the first place, the presence of a kink must be accounted for by a potential containing a Dirac delta[184] ; and where no such delta spikes occur, that is where the potential makes at most only *finite* jumps, the derivative of the wave function must also be continuous, just like the wave function itself.

[182] This is private notation : I use the symbol Λ because it displays exactly its own kink.
[183] For the *step* function $\theta(x)$, see appendix 11.
[184] More about this in the next section.

8.2 Pinning down a particle : the single pin

Let us assume, just for fun, that the wave function of a one-dimensional system of a particle of mass m is precisely given by the kink function[185] :

$$\psi(x) = \Lambda(\alpha; x) \quad . \tag{232}$$

We can now work our way towards the time-independent Schrödinger equation by writing

$$-\frac{\hbar^2}{2m}\,\psi''(x) = -\frac{\hbar^2\alpha^2}{2m}\,\psi(x) + \frac{\hbar^2\alpha}{m}\delta(x)\,\psi(x) \tag{233}$$

and by transferring the last term in the right to the left we obtain

$$-\frac{\hbar^2}{2m}\,\psi''(x) - \frac{\hbar^2\alpha}{m}\delta(x)\,\psi(x) = -\frac{\hbar^2\alpha^2}{2m}\,\psi(x) \quad . \tag{234}$$

This is precisely of the form of a time-independent Schrödinger equation :

$$-\frac{\hbar^2}{2m}\,\psi''(x) - V(x)\,\psi(x) = E\,\psi(x) \quad , \tag{235}$$

if we choose

$$V(x) = -\frac{\hbar^2\alpha}{m}\,\delta(x) \quad , \quad E = -\frac{\hbar^2\alpha^2}{2m} \quad . \tag{236}$$

We can deftly turn this argument upside down, and we are left with the conclusion that for a system with the following Hamiltonian :

$$\hat{H} = \frac{1}{2m}\,\hat{p}^2 - k\,\delta(\hat{x}) \tag{237}$$

for some given positive number k, there is precisely *one* bound state. It has the form of the kink function, with α given by

$$\alpha = \frac{m\,k}{\hbar^2} \quad , \tag{238}$$

and the (negative !) energy eigenvalue is

$$E = -\frac{k^2 m}{2\hbar^2} \quad . \tag{239}$$

we see that the presence of the negative potential pin at $x = 0$ serves to keep the particle localized in the neighbourhood of $x = 0$. This is of course completely different from the *classical* situation, where the particle is either completely free at $x \neq 0$, or is lurking down the abyssal potential hole. The difference arises from the fact that the particle's wave function is always spread out over all x values, and it must therefore always be 'aware' of the pin.

E 54 **E 55**

[185] It is quite usual not to worry about the normalization of the wave function. The understanding is that, provided that the wave function is quadratically integrable, we can always find the correct normalization *in principle*, should anyone insist.

8.3 The double-pin model

We can employ what we have learned even better. Let us assume that the (unnormalized) wave function is

$$\psi_+(x) \equiv \Lambda(\alpha; x) + \Lambda(\alpha; x - r) \ , \tag{240}$$

for some chosen, positive value of r, so that it shows *two* kinks, one at $x = 0$ and another one at $x = r$. This function is *symmetric* under reflection around the point $x = r/2$. Taking the second derivative gives us

$$\psi_+''(x) = \alpha^2 \, \psi_+(x) - 2\alpha \, \delta(x) - 2\alpha \, \delta(x - r) \ ; \tag{241}$$

and by exploiting the fact that

$$\psi_+(0) = \psi_+(r) = 1 + e^{-ar} \tag{242}$$

we can write Eq.(241) as

$$\psi_+''(x) = \alpha^2 \, \psi_+(x) - \frac{2\alpha}{1 + e^{-ar}} \Big(\delta(x) + \delta(x - r) \Big) \, \psi_+(x) \ . \tag{243}$$

We have again built something appropriate for a time-independent Schrödinger equation :

$$-\frac{\hbar^2}{2m} \, \psi_+''(x) + V(x) \, \psi_+(x) = E_+ \, \psi_+(x) \ , \tag{244}$$

with

$$E_+ = -\frac{\hbar^2 \alpha^2}{2m} \quad , \quad V(x) = -\lambda \Big(\delta(x) + \delta(x - r) \Big) \quad , \quad \lambda = \frac{\hbar^2}{m} \frac{\alpha}{1 + e^{-\alpha r}} \ . \tag{245}$$

The potential is that of *two* pins, sitting at $x = 0$ and $x = r$ as we might have expected. The relation between α and λ, the strength of the pins, is no longer as simple as in the single-pin model, but it can be solved numerically without any problem[186].

We can of course also look at an *antisymmetric* wave function :

$$\psi_-(x) = \Lambda(\beta; x) - \Lambda(\beta; x - r) \ . \tag{246}$$

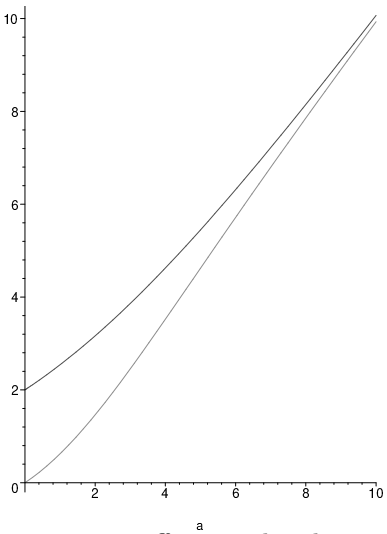
The time-independent Schrödinger equation now requires : **E 56**

$$E_- = -\frac{\hbar^2 \beta^2}{2m} \quad , \quad V(x) = -\lambda \Big(\delta(x) + \delta(x - r) \Big) \quad , \quad \lambda = \frac{\hbar^2}{m} \frac{\beta}{1 - e^{-\beta r}} \ . \tag{247}$$

Note that the relation between λ and β is different from that between λ and α. Since λ is a given number, α and β must be different. Let us write

$$f_+(\alpha) = \frac{\alpha}{1 + e^{-\alpha r}} \quad , \quad f_-(\beta) = \frac{\beta}{1 + e^{-\beta r}} \ . \tag{248}$$

[186] In practice, the value of λ is of course given, and α will just have to adapt itself so as to give the right answer.



Here we plot, for $r = 0.5$, $f_+(\alpha)$ (lower curve) and $f_-(\beta)$ (upper curve). The values for α and β must be such that $f_+(\alpha) = f_-(\beta)$. It follows that always $\alpha > \beta$, as can also be seen from the fact that the denominator in f_+ is always *larger* than unity, and that in f_- is always *smaller* than unity. This implies that $E_+ < E_-$: the symmetric state has the lowest energy. Also, it is possible that λ is so small that no β can be found, and only the symmetric state exists.

An interesting effect can be observed by changing r. Since λ is fixed, α and β must adapt themselves to the change in r, and so we may write $\alpha = \alpha(r)$ and $\beta = \beta(r)$. For α, we then have

$$\alpha(r) = \xi \left(1 + e^{-r\alpha(r)}\right) \tag{249}$$

for fixed ξ. Differentiating to r gives us

$$\alpha'(r) = \xi\, e^{-r\alpha(r)} \left(-\alpha(r) - r\alpha'(r)\right) \quad , \tag{250}$$

or

$$\alpha'(r) \left(1 + e^{-r\alpha(r)} r\xi\right) = -\alpha(r)\xi\, e^{-r\alpha(r)} \quad ; \tag{251}$$

which means that $\alpha'(r)$ is *negative*. In the same manner, we see that

$$\beta(r) = \xi \left(1 - e^{-r\beta(r)}\right) \tag{252}$$

leads to

$$\beta'(r) \left(1 - e^{-r\beta(r)} \xi r\right) = \beta\xi\, e^{-r\beta(r)} \quad . \tag{253}$$

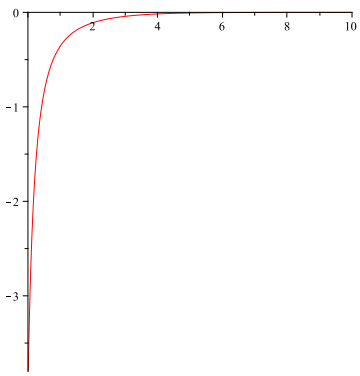
By using Eq.(252) we can rewrite

$$1 - e^{-r\beta(r)} \xi r = \frac{e^{r\beta(r)} - 1 - \beta(r) r}{e^{\beta(r) r} - 1} \tag{254}$$

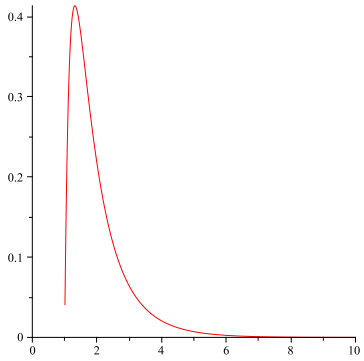
which is positive[187], therefore $\beta'(r)$ is positive. Physically, this means that in the state ψ_+ the energy *increases* when the pins are pushed apart, and in the state ψ_- it *decreases*. The state ψ_+ therefore represents an *attractive* force

[187] $e^x \geq 1 + x$.

between the pins, and ψ_- represents a *repulsive* force. Below, we give the force between the pins as a function of their distance, for $\hbar = \lambda = m = 1$.



The force for the state ψ_+. Since $\alpha \approx 2(1-r)$ for small r, the force approaches -4 as $r \downarrow 0$.



The force for the state ψ_-. For $r < 1$ there is no solution, and we see that the force vanishes as $r \downarrow 1$.

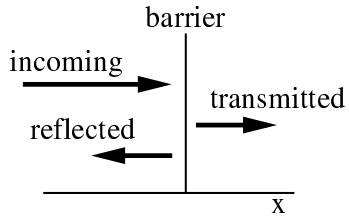
We see here the embryo of an idea that has become very important in physics : *forces can be represented by particles.* **E 57**

8.4 A single barrier

We can apply what we have learned to some other simple systems. The first of these is given by the potential

$$V(x) = \lambda\delta(x) \quad , \quad \lambda > 0 \ , \tag{255}$$

i.e. a positive delta-function-shaped barrier. Let us imagine how a physicist would study such a barrier. A reasonable approach would be to fire a stream of particles with given energy at the barrier, and then studying how many of them are reflected, how many make it through, and what they look like in that case ; all, of course, for various values of the energy. We therefore envisage the following quantum-mechanical setup. To the left of the barrier (*i.e.* for $x \leq 0$) the wave function consists of two components, a stream of particles moving to the right, towards the barrier, and a stream of reflected particles moving to the left ; on the right-hand side of the barrier we only have a single stream, that of the transmitted particles, moving to the right. Of course, no particles are supposed to be moving towards the barrier from the right. We take the wave number k of the particle waves to be positive ; because the energy eigenvalue $E = \hbar^2 k^2/2m$ must be unique, all three waves must have the same value of k. Waves moving to the right are represented by $\exp(ikx)$; those moving to the left are represented by $\exp(-ikx)$. In the two regions we **E 58**



thus have the following forms for the wavefunction : **E 59**

$$\psi(x) = \begin{cases} \psi_1(x) = e^{ikx} + \rho\, e^{-ikx} & , \quad x \leq 0 \\ \psi_2(x) = \tau\, e^{ikx} & , \quad x \geq 0 \end{cases} \tag{256}$$

We have arbitrarily (but reasonably, in our setup) assigned unit strength to the incoming wave : the strengths ρ of the refected wave and τ of the transmitted waves it is our job to figure out. The first matching condition states that $\psi_1(0) - \psi_2(0) = 0$ hence

$$\tau - 1 - \rho = 0 \ . \tag{257}$$

For the second matching condition, we have a look at the time-independent Schrödinger equation, this time integrated from an infinitesimal negative position $-\epsilon$ to an infinitesimal positive position $+\epsilon$:

$$\int_{-\epsilon}^{+\epsilon} dx \, \left(-\frac{\hbar^2}{2m} \psi''(x) + \lambda\, \delta(x)\, \psi(x) \right) = \int_{-\epsilon}^{+\epsilon} dx \, \left(E\, \psi(x) \right) \ . \tag{258}$$

The left-hand side evaluates to

$$-\frac{\hbar^2}{2m} \Big(\psi'(\epsilon) - \psi'(-\epsilon) \Big) + \lambda\, \psi(0) \ ;$$

and the right-hand side is, for very small ϵ, approximately equal to

$$2\epsilon\, E\, \psi(0) \ .$$

If we now take ϵ to zero, the right-hand vanishes, and we recover the second matching condition and for this system in the form

$$\psi_2'(0) - \psi_1'(0) = \frac{2m\lambda}{\hbar^2}\, \psi(0) \ , \tag{259}$$

so that

$$ik(\tau - 1 + \rho) = \frac{2m\lambda}{\hbar^2} \ . \tag{260}$$

The two matching conditions (257) and (260) can readily be solved, and we find

$$\tau = \frac{1}{1 + i\, \gamma/k} \quad , \quad \rho = \frac{-i\, \gamma/k}{1 + i\, \gamma/k} \quad , \quad \gamma \equiv \frac{m\lambda}{\hbar^2} \ . \tag{261}$$

The combined number γ can be interpreted as a measure of the *strength* of the barrier[188]. We are now in a position to inspect the physics of this simple experiment. The *transmission coefficient* T is the density of the transmitted particles as a fraction of the density of the stream of incoming particles. Since this latter was normalized to unity, we here simply have

$$T = |\tau|^2 = \frac{1}{1 + \gamma^2/k^2} \ . \tag{262}$$

[188]It would be less appropriate to talk of the *height* of the barrier, since the Dirac delta function is infinitely high anyway.

The *reflection coefficient* is the analogous fraction for the reflected wave :

$$R = |\rho|^2 = \frac{\gamma^2/k^2}{1+\gamma^2/k^2} \quad . \tag{263}$$

It is gratifying to establish that no particles have somehow got lost in our experiment, since

$$T + R = 1 \quad . \tag{264}$$

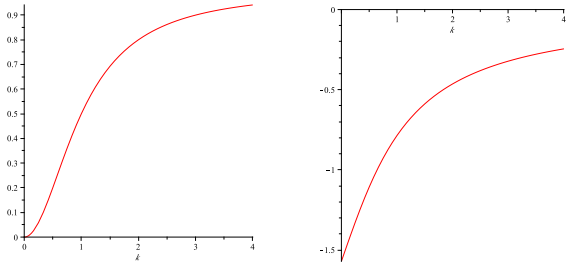
This 'conservation of probability', or *unitarity*, is a general feature of a consistent theory. A second topic of interest is the precise shape of the outgoing (transmitted) wave. We can write[189]

$$\tau = \frac{1}{1+\gamma^2/k^2}\sqrt{\frac{1-i\gamma/k}{1+i\gamma/k}} = \sqrt{T}\, e^{i\phi(k)} \quad , \quad \phi(k) = -\arctan(\gamma/k) \quad . \tag{265}$$

The outgoing wave can therefore be written as

$$\psi_2(x) = \sqrt{T}\, \exp(ikx + \phi(k)) \quad , \tag{266}$$

that is, as an (attenuated) version of the incoming wave, where the phase has been shifted by $\phi(k)$, which is therefore reasonably called the *phase shift*. Even if we had no idea of the existence of the barrier, the phase shift would tell us that, nevertheless, *something* happened to the particles that we observe at positive x.



We plot the transmission coefficient and the phase shift as a function of k. We have taken the strength $\gamma = 1$. Actually, in this case it is only the ration γ/k that counts. For low energies (small k) almost all particles are reflected, and the few that make it through the barrier have a phase shift of almost $-\pi/2$. At very high energies, more or less all particles zip through, and their phase shift becomes very small, so that you wouldn't notice that there was a barrier.

As a last remark, we may investigate what happens if we replace the barrier by a 'pin', by simply inverting the sign of λ. In fact, nothing very much happens ! The transmission coefficient T is unchanged, so that, surprisingly, particles may 'bounce off a hole in the ground'. The only difference that we can notice is in the sign change in the phase shift.

[189] To see this, we may employ the useful representation of the arc tangent

$$\arctan(z) = \frac{1}{2i}\log\left(\frac{1+iz}{1-iz}\right) \quad .$$

8.5 The double barrier : a wave filter

The reflection behaviour of the above model is simple, perhaps too simple to be interesting. We can enliven things by putting in a second barrier at some position $L > 0$:

$$V(x) = \lambda\,\delta(x) + \lambda\,\delta(x-L) \quad . \tag{267}$$

We now distinguish three regions, rather than two :

$$\psi(x) = \begin{cases} \psi_1(x) = e^{ikx} + \rho\, e^{-ikx} & , \quad x \le 0 \\ \psi_2(x) = c\, e^{ikx} + d\, e^{-ikx} & , \quad 0 \le x \le L \\ \psi_3(x) = \tau\, e^{ikx} & , \quad x \ge L \end{cases} \quad , \tag{268}$$

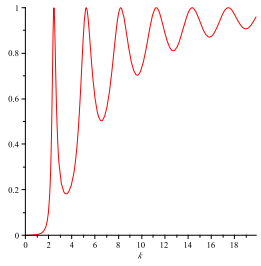
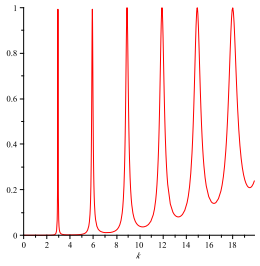
where ρ, c, d and τ again have to be obtained from the matching conditions, of which we now have four rather than two. Using the same γ as before, we can write them as

$$\begin{aligned} 0 &= c + d - 1 - \rho \ , \\ 0 &= ik(c - d - 1 + \rho) - 2\gamma(c + d) \ , \\ 0 &= e^{ikL}(\tau - c) - e^{-ikL} d \ , \\ 0 &= ik\left(e^{ikL}(\tau - c) + e^{-ikL} d\right) - 2\gamma\tau\, e^{ikL} \quad . \end{aligned} \tag{269}$$

Solving for our four coefficients is less easy than in the single-barrier case, but straightforward[190]. The most interesting of these is of course

$$\tau = \frac{1}{(1 + i\gamma/k)^2 + (\gamma^2/k^2)\, e^{2ikL}} \quad , \tag{270}$$

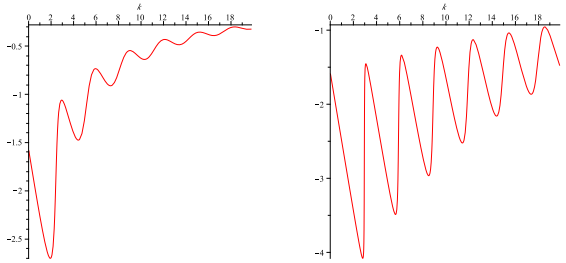
since it gives us the transmission coefficient (and, by unitarity, the reflection coefficient) as well as the phase shift of the transmitted wave. In contrast to the previous case, where the only relevant parameter was γ/k, we now also have kL as a dimensionless parameter. If we take $L = 1$ (in other words, adopt L as our unit of length), the physics depends on γ.

Here we plot T as a function of k, with $L = 1$ and $\gamma = 3$ (left-hand plot) and $\gamma = 15$ (right-hand plot). As γ increases, the T plot approaches a series of spikes for $k = n\pi$, $n = 1, 2, 3, \ldots$ By inserting yet *another* barrier, we see that for some k values the transmission becomes again perfect ! That, for infinitely strong barriers, waves with $kL = n\pi$ are transmitted

[190] Ominous words...

is of course not so surprising, since then an exact number of half wavelengths fits between the barriers. A similar interesting pattern can be observed for the phase shifts, which we give here. Again, we use $\gamma = 3$ (left-hand plot) and $\gamma = 15$ (right-hand plot). For very large k these approach zero, as you might expect. For large values of γ, the phase shift starts to follow a saw-tooth pattern that we can understand as follows. For $\gamma \to \infty$, we have



$$\tau \sim \frac{k^2}{\gamma^2\,(e^{2ikL}-1)} = \frac{k^2\,e^{-ikL}}{2i\gamma^2\sin(kL)} = \frac{k^2}{2\gamma^2|\sin(kL)|}\,\frac{-i\,e^{-ikL}}{\mathrm{sign}(\sin(kL))} \quad . \tag{271}$$

The first factor in the last expression does not, of course, influence the phase shift. The second factor describes a point on the unit circle in the complex plane that moves with increasing k from $-i$ clockwise to $+i$, and then jumps immediately back to $-i$ to start all over again. The jumps occur, of course, whenever $\sin(kL) = 0$, or $k = n\pi$, $n = 1, 2, 3 \ldots$.

This example shows that, even in simple cases, a surprisingly rich behavior can be obtained. This is what makes these simpleminded, one-dimensional models useful for the understanding of the much more complicated real life.

8.6 Tunneling

We now come to a model with a – very slightly – better claim to realism. We consider a stream of particles moving towards a potential barrier of finite height W and thickness L, that is

$$V(x) = \begin{cases} 0 & \text{for } x < 0 \\ W & \text{for } 0 \le x \le L \\ 0 & \text{for } x > L \end{cases} \quad . \tag{272}$$

We shall mostly concentrate on energy eigenvalues E that are 'below' the barrier : $0 < E < W$. Classically, a particle with such an energy cannot cross the barrier. Instead, when it comes in from the left[191] and hits the barrier, it simply bounces back and moves away to the left again. In quantum mechanics, the situation is somewhat different : as we shall see the particle has a *finite* probability of making it through the barrier !

Our system thus contains *three* distinct regions. To the left and to the right of the barrier, the time-independent Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\,\psi''(x) = E\,\psi(x) \quad , \tag{273}$$

[191] The 'left' is the region of negative x.

while inside the barrier

$$-\frac{\hbar^2}{2m}\,\psi''(x) = -(W-E)\,\psi(x) \quad . \tag{274}$$

The general solution is therefore also split up into three pieces :

$$\psi(x) = \begin{cases} e^{ikx} + rho\, e^{-ikx} & \text{for } x < 0 \\ c\, e^{\kappa x} + d\, e^{-\kappa x} & \text{for } 0 \le x \le L \\ \tau\, e^{ikx} & \text{for } x > L \end{cases} \quad , \tag{275}$$

with

$$k^2 = \frac{2mE}{\hbar^2} \quad , \quad \kappa^2 = \frac{2m(W-E)}{\hbar^2} \quad . \tag{276}$$

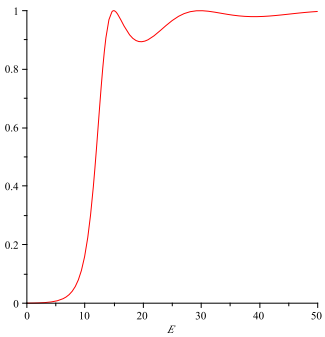
As long as $0 \le E < W$, κ is a real number ; for larger energies, $E > W$, it becomes imaginary. As we have learned from our investigation of the kink function, for a barrier of *finite* height both the wave function itself and its first derivative must be continuous everywhere ; and especially at the important points $x = 0$ and $x = L$. The four matching conditions therefore read

$$\begin{aligned} 0 &= c + d - 1 - \rho \ , \\ 0 &= \kappa(c-d) - ik(1-\rho) \ , \\ 0 &= \tau\, e^{ikL} - c\, e^{\kappa L} - d\, e^{-\kappa L} \ , \\ 0 &= ik\tau\, e^{ikL} - \kappa\left(c\, e^{\kappa L} - d\, e^{-\kappa L}\right) \ . \end{aligned} \tag{277}$$

As before, we can now solve for our four unknowns. Again, τ is of special interest :

$$\tau = \frac{4i\,\kappa k\, e^{-ikL}}{(k+i\kappa)^2\, e^{\kappa L} - (k-i\kappa)^2\, e^{-\kappa L}} \quad . \tag{278}$$

The transmission coefficient T is again given by $T = |\tau|^2$.The ingredients in this our problem are m, $\hbar$, L, W and k. We can therefore assume $m = \hbar = L = 1$, and the physics will then depend on the values of W, and of course k or rather the energy $E = \hbar^2 k^2/2m = k^2/2$. Here we plot T as a function of E, for $W = 10$. We see that even for $E = W$ the transmission is still rather poor, although *classical* particles would be perfectly transmitted for any $E > W$. As expected, for large energies the transmission becomes total. Another case, of actual practical interest, is that where W is very much larger than E. In that limit we may negelct k with respect to κ, and approximate



$$|\tau| \sim \frac{4k}{\kappa\left(e^{\kappa L} - e^{-\kappa L}\right)} \quad . \tag{279}$$

For large values of L, the transmission coefficient then takes the form

$$T \sim 16 \frac{E}{W} \exp\left(-\sqrt{\frac{8mW}{\hbar^2}}\, L \right) \ , \tag{280}$$

and is therefore *extremely* sensitive to small changes in L. Taking for m the value of the electron mass[192], we find for $W = 100$ electronvolts the value

$$\sqrt{\frac{8mW}{\hbar^2}} \sim 10^{11} \text{ per meter} \ , \tag{281}$$

so that a change of 10^{-11} meter in L leads to a drastic effect on the transmission : and such distances are on the atomic scale ! This effect (made possible, of course, because $\hbar$ is so very, *very* small) is used extensively in Scanning Tunneling Microscopy.

8.7 Extra ! Periodic potentials

8.7.1 General properties

Another class of models of interest is that of *periodic potentials*. A periodic potential $V(x)$ in one dimension, with period L, has the property that

$$\cdots = V(x-L) = V(x) = V(x+L) = V(x+2L) = \cdots$$

Note that we have here an *infinitely large* system, that looks unchanged if we jump from x to $x + L$. What follows in this paragraph involves some fairly abstract reasoning[193], but we shall end with a very useful result.

Let $\psi(x)$ be an energy eigenfunction (that is, a solution of the time-independent Schrödinger equation) :

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x) \ . \tag{282}$$

Then, of course, also

$$-\frac{\hbar^2}{2m}\psi''(x+L) + V(x+L)\psi(x+L) = E\psi(x+L) \ , \tag{283}$$

and since $V(x+L) = V(x)$, we see that if $\psi(x)$ is a solution, then so is $\psi(x+L)$. Let $\phi_1(x)$ and $\phi_2(x)$ be two orthogonal eigenfunctions with the same energy E. In matrix notation, we can write

$$\phi_1(x) \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \phi_2(x) \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ . \tag{284}$$

[192] About 10^{-30} kilogram.

[193] And, unsurprisingly, some linear algebra !

A general eigenfunction with energy E can then be written as

$$\psi(x) = a_1\,\phi_1(x) + a_2\,\phi_2(x) \leftrightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad . \tag{285}$$

Since $\phi_{1,2}(x+L)$ are also solutions (see above), we must have

$$\phi_1(x+L) \leftrightarrow \begin{pmatrix} m_{11} \\ m_{12} \end{pmatrix} \quad , \quad \phi_2(x+L) \leftrightarrow \begin{pmatrix} m_{21} \\ m_{22} \end{pmatrix} \quad . \tag{286}$$

And, of course, we also have

$$\begin{aligned} \psi(x+L) &= a_1\,\phi_1(x+L) + a_2\,\phi_2(x+L) \\ &= \begin{pmatrix} m_{11}\,a_1 + m_{21}\,a_2 \\ m_{12}\,a_1 + m_{22}\,a_2 \end{pmatrix} = \mathcal{M} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \ , \end{aligned} \tag{287}$$

where $\mathcal{M}$ is the 2×2 matrix

$$\mathcal{M} = \begin{pmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{pmatrix} \quad . \tag{288}$$

We see that, in matrix language, going from x to $x+L$ is a matrix multiplication. If we shift by kL (with integer k), we have

$$\psi(x+kL) \leftrightarrow \mathcal{M}^k \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad . \tag{289}$$

We can find out something about the matrix $\mathcal{M}$. From excercise 35 we know that the Wronskian of ϕ_1 and ϕ_2 is independent of x :

$$\phi_1(x+L)\phi_2'(x+L) - \phi_1'(x+L)\phi_2(x+L) = \phi_1(x)\phi_2'(x) - \phi_1'(x)\phi_2(x) \ , \tag{290}$$

and if we insert Eq.(286) in this expression, it is easily seen that we necessarily have

$$m_{11}m_{22} - m_{12}m_{21} = \det(\mathcal{M}) = 1 \quad . \tag{291}$$

The two eigenvalues σ_1 and σ_2 of the matrix $\mathcal{M}$ must therefore obey

$$\sigma_1\,\sigma_2 = 1 \quad . \tag{292}$$

Now imagine that we like to jump from x to $x + kL$, with very large k. Then we would have to multiply by a large power of $\mathcal{M}$. If one of the eigenvalues, σ_1, say, has absolute value larger than one, $|\sigma_1| > 1$, the components of $\mathcal{M}^k$ become huge and the wave function will therefore not be normalizable. We must therefore have that both σ_1 and σ_2 have absolute value precisely one, or in other words

$$\sigma_1 = e^{iw} \quad , \quad \sigma_2 = e^{-iw} \ , \tag{293}$$

for some phase w that we may take to run from $-\pi$ to π. We see that there are in our infinite, periodic system, energy eigenfunctions, which we shall denote by $\phi_w(x)$, with the property

$$\phi_w(x+L) = e^{iw}\phi_w(x) \ , \tag{294}$$

where we may take w to be positive and negative to distinguish the two cases.

8.7.2 Infinitely many Dirac pins

The simplest example of a periodic system as dsecribed above has an infinite series of negative Dirac deltas, spaced at distances L :

$$V(x) = -\lambda \sum_{n=-\infty}^{\infty} \delta(x - nL) \qquad (\lambda > 0) \ . \tag{295}$$

This can be seen as a primitive description of an infinite array of identical ions, that each attract an electron : a model of a one-dimensional crystal lattice. For simplicity we take our units of length, mass and time such that $\hbar = m = L = 1$, and then the only remaining physical parameter in the model is λ, or rather the potential strength

$$\gamma = \frac{\hbar^2 \lambda}{m} \tag{296}$$

that we have already seen before. As usual, we have to investigate the matching conditions. If we can solve them at one potential point, we can of course solve them everywhere, so we only have to study the effects of requiring

$$\begin{aligned} &\lim_{\epsilon \to 0} \quad \phi_w(L+\epsilon) - \phi_w(L-\epsilon) = 0 \ , \\ &\lim_{\epsilon \to 0} \quad \phi'_w(L+\epsilon) - \phi'_w(L-\epsilon) = -2\gamma\, \phi_w(L-\epsilon) \ . \end{aligned} \tag{297}$$

Since we have at our disposal Eq.(294), we can write these two matching conditions in our case as

$$\begin{aligned} &e^{iw} \phi_w(0) - \phi_w(L) = 0 \ , \\ &e^{iw} \phi'_w(0) - \phi' w(L) = -2\gamma\, \phi_w(L) \ . \end{aligned} \tag{298}$$

It is useful to start with states with *negative* energy. Since the particle is free between the Dirac pins, we decide that we may write

$$\phi_w(x) = a\, e^{zx} + b\, e^{z(L-x)} \quad , \quad E = -\frac{\hbar^2 z^2}{2m} \ , \tag{299}$$

for $0 \le x \le L$. The matching conditions then take the form (with $L = 1$) :

$$\begin{aligned} &e^{iw} \left(a + b\, e^z \right) - \left(a\, e^z + b \right) = 0 \ , \\ &e^{iw} \left(a - b\, e^z \right) - \left(a\, e^z - b \right) = -\frac{2\gamma}{z} \left(a\, e^z + b \right) \ . \end{aligned} \tag{300}$$

The first of these establish the ratio of a and b :

$$b = a\, \frac{e^z - e^{iw}}{e^{z+iw} - 1} \ , \tag{301}$$

and then adding the two conditions results in

$$2a\, e^{iw} - 2a\, e^z = -\frac{2\gamma}{z}\, a\, \frac{e^{2z+iw} - e^{iw}}{e^{z+iw} - 1} = \frac{2a\gamma}{z}\, \frac{e^{2z} - 1}{e^{-iw} - e^z} \ . \tag{302}$$

Unsurprisingly, a drops out since it is related to the normalization about which we do not care, and we are left with

$$1 - e^z \left(e^{iw} + e^{-iw} \right) + e^{2z} = \frac{\gamma}{z} \left(e^{2z} - 1 \right) \quad . \tag{303}$$

Dividing by e^z we finally arrive at the result, valid for $E < 0$ [194]:

$$\cos(w) = \cosh(z) - \gamma \, \frac{\sinh(z)}{z} \quad , \quad E = -\frac{\hbar^2 z^2}{2m} \quad . \tag{304}$$

What about positive energies ? Well, we have nowhere used that z has to be real. So we can simply exchange z for $z = ik$, and find immediately the condition

$$\cos(w) = \cos(k) - \gamma \, \frac{\sin(k)}{k} \quad , \quad E = +\frac{\hbar^2 k^2}{2m} \quad . \tag{305}$$

8.7.3 Band structure

We have seen that, for the model above, Eq.(304) or Eq.(305) have to be satisfied, for negative or positive energies, respectively. But this may not always be possible ! After all, $\cos(w)$ is bound to be between -1 and 1, and the right-hand-sides may exceed that for some values of z and k. Those values, therefore, are *not physically allowed*, and the corresponding energies are not in the spectrum of the Hamiltonian. Such energy values are called *forbidden*, and if they form some interval that is called an *empty energy band*. Finding out what are the forbidden and allowed (*i.e.* non-forbidden) energy bands, and for what values of z and k they occur, forms a major part of condensed-matter physics. We shall now investigate the *band strucure* (as it is called) of our model.

Let us start with negative energies, for which we have (remember the remarks on choosing our units)

$$E = -\frac{z^2}{2} \quad , \quad \cos(w) = \cosh(z) - \frac{\gamma}{z} \sinh(z) \quad . \tag{306}$$

We employ the facts that

$$\cosh(z) = \cosh(z/2)^2 + \sinh(z/2)^2 \quad , \quad \sinh(z) = 2 \cosh(z/2) \sinh(z/2) \tag{307}$$

and

$$\cosh(z/2)^2 - \sinh(z/2)^2 = 1 \tag{308}$$

to notice that, for $\cos(w) = 1$ for some $z = z_1$, we must have

$$\sinh(z_1/2) \left(\sinh(z_1/2) - \frac{\gamma}{z_1} \cosh(z_1/2) \right) = 0 \quad , \tag{309}$$

[194] Since $\sinh(z) \geq 0$ and $\cosh(z) \geq 1$ for $z \geq 0$, we also immediately see that there is no negative-energy solution for *negative* γ. This is only to be expected, of course, since whatever influence Dirac *spikes* may exert on our particle, it will certainly *not* be to attract them into bound states.

so that the value of z_1 that gives us $\cos(w) = 1$ must obey

$$\sinh(z_1/2) = 0 \quad \text{or} \quad \frac{\gamma}{z_1} = \tanh(z_1/2) \ . \tag{310}$$

Of these, the first $z_1 = 0$ we consider to be spurious since $0/0$ is not well-defined ; the second requirement always gives us a solution, with $z_1 > \gamma$ (since tanh is always smaller than one). Therefore, all energies smaller than $-{z_1}^2/2$ are forbidden — a good thing since there must be *some* lowest energy, after all ! For $\cos(w) = -1$ we find, using the same machinery of hyperbolo-trigonometrics,

$$\cosh(z_2/2) \left(\cosh(z_2/2) - \frac{\gamma}{z_2} \sinh(z_2/2) \right) = 0 \tag{311}$$

from which (since $\cosh(z_2/2) = 0$ is impossible) we have

$$\frac{z_2}{\gamma} = \tanh(z_2/2) \ . \tag{312}$$

The energy value $-{z_2}^2/2$ gives the *upper* value of the allowed negative-energy band – but only for $\gamma > 2$ since $\tanh(z/2) < z/2$. For smaller values of γ the lowest allowed-energy band extends from the negative values into the positive ones, which we shall now look at.

For positive energies we must have

$$E = \frac{k^2}{2} \quad , \quad \cos(w) = \cos(k) - \frac{\gamma}{k} \sin(k) \ . \tag{313}$$

Again, we may use

$$\cos(k) = \cos(k/2)^2 - \sin(k/2)^2 \quad , \quad \sin(k) = 2 \sin(k/2) \cos(k/2) \tag{314}$$

to see that, for $\cos(w) = 1$, we must have

$$\sin(k/2) \left(\sin(k/2) + \frac{\gamma}{k} \cos(k/2) \right) = 0 \ , \tag{315}$$

so that

$$\sin(k/2) = 0 \quad \text{or} \quad \frac{\gamma}{k} = -\tan(k/2) \ . \tag{316}$$

Similarly, for $\cos(w) = -1$ we have

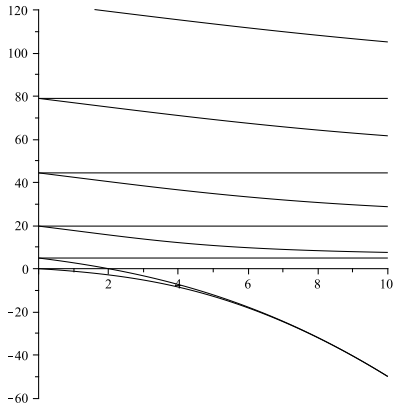
$$\cos(k/2) \left(\cos(k/2) - \frac{\gamma}{k} \sin(k/2) \right) = 0 \tag{317}$$

with solutions

$$\cos(k/2) = 0 \quad \text{or} \quad \frac{k}{\gamma} = \tan(k/2) \ . \tag{318}$$

Since $\tan(z/2) > z/2$, the latter solution only falls in the interval between 0 and π if $\gamma < 2$, which ties in neatly with what we saw for $E < 0$. We can now

put everything together. In the plot, we give the band boundaries for γ op to 10.



For small values of γ we see that the forbidden bands shrink to nothing, as you would expect since for $\gamma = 0$ we are simply talking about a *free* particle. As γ increases there are increasingly large empty energy bands, and an allowed band of increasingly negative energy (corresponding to the bound state of the 'single Dirac pin' discussed earlier). For larg*ish* values of γ the allowed energy bands shrink, and for very large values of γ the only possible values are (for negative energies) $z = \gamma$ or $E = -\gamma^2/2$, and $k = \pi, 2\pi, 3\pi, \ldots$ for which the (positive) energies are $E = n^2\pi^2/2$ with $n = 1, 2, 3, \ldots$ — precisely those of the 'particle in a box' that we have seen before ! A final remark is in order on the movement of our particle. For negative energies, the wave functions have no constant overall complex phase since there is a phase shift of size w at every pin ; but in between pins its phase *is* constant and hence the probability current vanishes (*cf.* excercise 39), so we may say that the particles are actually sitting still for negative energies. For positive energies, the particles *do* move with momentum k, to the right, or to the left.

8.8 Exercises 53 to 60

Exercise 53 : **Integrating the kink**
Show that

$$\int dx \ \ |\Lambda(x)|^2 = \frac{1}{\alpha}$$

Exercise 54 : **Dimensional arguments once more**
For the 'single-pin' model of Eq.(237), show that the parameter k must have the dimension of energy×length. Use this to show that α as given in Eq.(238) and E as given in Eq.(239) have the correct dimensions of inverse length and of energy, respectively.

Exercise 55 : **Checking Heisenberg – again**
Compute $\langle x \rangle$, $\langle p \rangle$, $\sigma(x)$ and $\sigma(p)$ for the bound state in the 'single-pin' model, and show that the uncertainty relations hold.

Exercise 56 : **Antisymmetry with two pins**
Verify Eq.(247) by following the same steps as for the symmetric wave function in the two-pin model.

Exercise 57 : **Another solvable case**
In this exercise we make heavy use of hyperbolic trigonometric functions. We recall that

$$\frac{d}{dx}\cosh(x) = \sinh(x) \quad , \quad \frac{d}{dx}\sinh(x) = \cosh(x) \quad , \quad \cosh(x)^2 - \sinh(x)^2 = 1 \ .$$

Furthermore you may use

$$\int_{-\infty}^{\infty} \frac{1}{\cosh(x)^2} dx = 2 \quad , \quad \int_{-\infty}^{\infty} \frac{\sinh(x)^2}{\cosh(x)^4} dx = \frac{2}{3} \quad , \quad \int_{-\infty}^{\infty} \frac{x^2}{\cosh(x)^2} dx = \frac{\pi^2}{6}$$

1. We consider a one-dimensional system of a particle with mass m. Its wave function is given by
$$\psi(x) = \frac{N}{\cosh(x/L)} \ ,$$
with $L > 0$ and N the appropriate normalization factor. Show that this particle is in an eigenstate of an Hamiltonian with the classical potential
$$V(x) = -\frac{\hbar^2}{mL^2 \cosh(x/L)^2}$$
with energy eigenvalue $E = -\hbar^2/(2mL^2)$.
2. Determine N (which you may take real and positive).
3. Show that in this state $\langle x \rangle = 0$ and $\langle p \rangle = 0$.
4. Compute $\langle x^2 \rangle$ and $\langle p^2 \rangle$.
5. Show, by explicit computation, that the Heisenberg inequality is obeyed.

Exercise 58 : **Streaming to the left, streaming to the right**
Consider a free particle in one dimension, and assume that at time $t = 0$ its (unnormalized) wave function is given by $\exp(ikx)$, with $k > 0$. Show that this is a momentum eigenstate with momentum eigenvalue $\hbar\, k$. Write down the time-dependent form of this wave function and show that this is a plane wave moving to the *right*, that is it moves towards larger x values as time increases. Also show that a wave function of the form $\exp(-ikx)$ describes a wave moving to the *left*.

Exercise 59 : **A typing error ?**
In Eq.(256) we distinguish the regions $x \leq 0$ and $x \geq 0$. A purist may complain that these regions overlap, namely at $x = 0$. Show that this is not necessarily a typing error !

Exercise 60 : A potential step
We consider a one-dimensional system of a particle of mass m in a potential of the form

$$V(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ W & \text{for } x > 0 \end{cases}$$

We shall start by taking $W > 0$. A stream of particles comes in from the left, and impinges on the potential step.

1. Give the time-independent Schrödinger equation for negative and for positive x.

2. Give the general form of the wave function in the two regions. For $x > 0$, make a distinction between $E > W$ and $E < W$. Why are there two independent solutions for $x \leq 0$, but only one for $x > 0$?

3. Show that the wave function and its first derivative must be continuous for all x.

4. Use these continuity conditions to find de coefficient of the incoming wave and the reflected wave in terms of the coefficient of the wave function for $x > 0$ (which we may take to unity for simplicity).

5. Compute the reflection coefficient R and show that it only depends on the ratio $u = E/W$.

6. Sketch the function $R(u)$ for u from 0 to 5.

7. Now assume that $W < 0$. Show that the case $E < W$ is impossible. Repeat the previous steps to compute the reflection coefficient.

8. Compare your results with what you would expect for a classical system.

Exercise 61 : Your author is lazy again
Verify that the matching conditions of Eq.(**??**) are indeed the correct ones, and check that the solutions given for A, B, C and D are indeed the correct solutions.

9 Two-particle systems

— *Tertius Interveniens, das ist Warnung an etliche Theologos, Medicos vnd Philosophos, sonderlich D. Philippum Feselium, dass sie bey billicher Verwerffung der Sternguckerischen Aberglauben nict das Kindt mit dem Badt aussschütten vnd hiermit jhrer Profession vnwissendt zuwider handlen*
by Johannes Kepler (1571-1630)

9.1 Two particles in one dimension

So far the quantum systems we have considered contain only a single particle ; but, of course, the more interesting systems are those of more particles in some kind of mutual interaction. For simplicity we shall restrict ourselves here to two particles, labelled 1 and 2, that can move about in a single dimension. The states are then described by the characteristics of both particles, for instance $|x_1, x_2\rangle$ are position eigenstates for both particles. A reasonable choice for the Hamiltonian is then

$$\hat{H} = \frac{1}{2m_1}\hat{p}_1^2 + \frac{1}{2m_2}\hat{p}_2^2 + V(\hat{x}_1, \hat{x}_2) \quad , \tag{319}$$

where $V(x_1, x_2)$ is the combined potential energy. We assume that the two particles have their own position and momentum operators, that are independent of each other, and therefore the commutators read

$$[\hat{x}_1, \hat{p}_1] = [\hat{x}_2, \hat{p}_2] = i\hbar \quad , \quad [\hat{x}_1, \hat{p}_2] = [\hat{x}_2, \hat{p}_1] = 0 \quad . \tag{320}$$

9.2 Translation invariance and momentum conservation

It frequently happens that the potential V is *translation invariant*, that is, it depends only on the relative position of the particles :

$$V(\hat{x}_1, \hat{x}_2) = V(\hat{x}) \quad , \quad \hat{x} \equiv \hat{x}_1 - \hat{x}_2 \quad . \tag{321}$$

In addition to the operator $\hat{x}$ for the relative position, we can also define an operator $\hat{P}$ for the total momentum of the system :

$$\hat{P} \equiv \hat{p}_1 + \hat{p}_2 \quad . \tag{322}$$

We can easily see that

$$[\hat{x}, \hat{P}] = 0 \quad , \quad [\hat{P}, \hat{p}_1] = [\hat{P}, \hat{p}_2] = 0 \quad , \tag{323}$$

and this implies that also

$$[\hat{H}, \hat{P}] = 0 \quad : \tag{324}$$

the total momentum is conserved, as a direct consequence of the translation invariance (321). This is a general phenomenon : if the Hamiltonian has an invariance, there is a corresponding conserved quantity[195].

[195] The same holds in classical physics, but the *proof* is quite different there.

9.3 Global and internal observables

Consider the hydrogen atom, a system of a proton and an electron bound together by their mutual Coulomb attraction. In talking about it, we can of course keep track of the position (or momentum) of the proton and electron separately ; but it is much more useful to distinguish the atom's *global* position/momentum and other observables that describe its *internal* structure. That is, we want to define in addition to $\hat{x}$ and $\hat{P}$ also a global position operator $\hat{X}$ and and internal momentum observable $\hat{p}$, whose form we have to derive. We shall write

$$\hat{X} = \alpha\,\hat{x}_1 + \beta\,\hat{x}_2 \quad , \quad \hat{p} = \gamma\,\hat{p}_1 + \delta\,\hat{p}_2 \quad , \tag{325}$$

and we want to figure out what we must take for the constants α, β, γ and δ. The commutators $[\hat{P}, \hat{p}]$ and $[\hat{X}, \hat{x}]$ vanish automatically, but we also want

$$[\hat{X}, \hat{P}] = i\hbar\ (\alpha + \beta) \quad \text{and} \quad [\hat{x}, \hat{p}] = i\hbar\ (\gamma - \delta) \tag{326}$$

to have the standard value $i\hbar$: we therefore require

$$\alpha + \beta = 1 \quad , \quad \gamma - \delta = 1 \quad . \tag{327}$$

It is now easy to express $\hat{p}_{1,2}$ in terms of $\hat{P}$ and $\hat{p}$:

$$\hat{p}_1 = \hat{p} - \delta\,\hat{P} \quad , \quad \hat{p}_2 = \gamma\,\hat{P} - \hat{p} \quad . \tag{328}$$

The kinetic part of the Hamiltonian is then

$$\frac{1}{2m_1}\hat{p}_1^2 + \frac{1}{2m_2}\hat{p}_2^2 = \left(\frac{\delta^2}{2m_1} + \frac{\gamma^2}{2m_2}\right)\hat{P}^2 + \left(\frac{1}{2m_1} + \frac{1}{2m_2}\right)\hat{p}^2 - \left(\frac{\delta}{m_1} + \frac{\gamma}{m_2}\right)\hat{P}\hat{p} \quad . \tag{329}$$

If the global and internal observables are to be truly independent, the last term in this expression must vanish, and together with Eq.(327) we can then work out the four constants :

$$\alpha = -\delta = \frac{m_1}{M} \quad , \quad \beta = \gamma = \frac{m_2}{M} \quad , \quad M = m_1 + m_2 \quad . \tag{330}$$

We can also define the so-called *reduced mass* m_r of the system as

$$\frac{1}{m_r} \equiv \frac{1}{m_1} + \frac{1}{m_2} \quad \Rightarrow \quad m_r = \frac{m_1 m_2}{M} \quad , \tag{331}$$

so that the two looked-for operators are seen to be

$$\hat{X} = \frac{m_1}{M}\hat{x}_1 + \frac{m_2}{M}\hat{x}_2 \quad , \quad \hat{p} = \frac{m_2}{M}\hat{p}_1 - \frac{m_1}{M}\hat{p}_2 \quad . \tag{332}$$

We see that $\hat{X}$ corresponds precisely to the *centre-of-mass position* of the system ; since we can write

$$\hat{p} = \frac{m_1 m_2}{M}\left(\frac{1}{m_1}\hat{p}_1 - \frac{1}{m_2}\hat{p}_2\right) \tag{333}$$

we can recognize that $\hat{p}$ decribes, very reasonably, the *relative velocity* of the particles multiplied by the reduced mass. The Hamiltonian now takes the form

$$\hat{H} = \frac{1}{2M}\hat{P}^2 + \frac{1}{2m_r}\hat{p}^2 + V(\hat{x}) \ . \tag{334}$$

The total energy is now nicely split up into two parts : the global part (the first term) shows that the atom, as a whole, is a free particle of mass M, and the other terms describe the energy due to the atom's internal structure in terms of a single particle with mass m_r.

E 62

9.4 Separable Hamiltonians

The above Hamiltonian is seen to be the sum of two parts that can, in principle, be considered separately, the global and the internal part. Such systems, in which the energy is a sum of contributions from 'sub'systems, is called *separable*. That this is a useful property can be seen if we go to the wave-function representation of the time-independent Schrödinger equation. Let us write the two-particle wave function in terms of X and x :

$$\psi(x_1, x_2) \quad \rightarrow \quad \psi(X, x) \ .$$

The time-independent Schrödinger equation now reads

$$-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial X^2}\psi(X,x) - \frac{\hbar^2}{2m_r}\frac{\partial^2}{\partial x^2}\psi(X,x) + V(x)\psi(X,x) = E\psi(X,x) \ . \tag{335}$$

Now, consider wave functions that *factorize*, that is they can be written as a product of two factors, each dependent only either only X or x :

$$\psi(X,x) = \phi(X)\,\chi(x) \ . \tag{336}$$

The Schrödinger equation then takes on the form

$$-\frac{\hbar^2}{2M}\phi''(X)\chi(x) - \frac{\hbar^2}{2m_r}\phi(X)\chi''(x) + V(x)\phi(X)\chi(x) = E\phi(X)\chi(x) \ . \tag{337}$$

Dividing by $\phi(X)\chi(x)$ we obtain

$$\left[-\frac{\hbar^2}{2M}\,\frac{\phi''(X)}{\phi(X)}\right] + \left[-\frac{\hbar^2}{2m_r}\,\frac{\chi''(x)}{\chi(x)} + V(x)\right] = E \ . \tag{338}$$

The first term in square brackets depends *only* on X, and the second one *only* on x — while their sum is the constant E. This is of course only possible if both square brackets are, in fact, themselves constant ! The Schrödinger equation therefore *separates* into two independent ones :

$$\begin{aligned} &-\frac{\hbar^2}{2M}\phi''(X) = E_X\phi(X) \ , \\ &-\frac{\hbar^2}{2m_r}\chi''(x) + V(x)\chi(x) = E_x\chi(x) \ , \\ &E_X + E_x = E \ . \end{aligned} \tag{339}$$

These two equations, both of the form of a one-particle time-independent Schrödinger equation, can be solved separately, and the total energy is the sum of the two energies. There are two important observations to be made. In the first place, we do not *have* to solve the system using wave functions : we may as well use the operator form of Eq.(339) :

$$\frac{1}{2M}\hat{P}^2\ |\psi\rangle = E_X\ |\psi\rangle \quad , \quad \frac{1}{2m_r}\hat{p}^2\ |\psi\rangle + V(\hat{x})\ |\psi\rangle = E_x\ |\psi\rangle \tag{340}$$

if we decide that that is easier. In the second place, do *not* make the mistake of concluding that all solutions to the Schrödinger equation have the factorization property ! The factorizing solutions provide a *basis* of states, but a generic superposition of such factorizing states is, in general, not itself factorizing.

E 63 **E 64**

It is clear that the separation trick will always work provided the Hamiltonian can indeed be written as a sum of terms, each term containing operators that do not occur in the other terms.

9.5 Three dimensions

The treatment given so far carries over into more dimensions without any problems. Suppose the two-particle Hamiltonian is

E 65

$$\hat{H} = \frac{1}{2m_1}\hat{\vec{p}}_1^{\,2} + \frac{1}{2m_1}\hat{\vec{p}}_2^{\,2} + V(\hat{\vec{x}}_1 - \hat{\vec{x}}_2) \ . \tag{341}$$

We have introduced a new notation here: a *vector of operators*. For instance

$$\hat{\vec{x}} = \left(\hat{x}, \hat{y}, \hat{z}\right) \ ; \tag{342}$$

so that

$$\hat{\vec{x}}^{\,2} = \hat{x}^2 + \hat{y}^2 + \hat{z}^2 \ . \tag{343}$$

This usually does not lead to confusion, but *if* it does, we can always go back to writing out all components explicitly. We can define the global and internal observable-operators by

$$\begin{aligned} \hat{\vec{P}} &= \hat{\vec{p}}_1 + \hat{\vec{p}}_2 \ , \ \hat{\vec{p}} = \frac{1}{M}\left(m_2\hat{\vec{p}}_1 - m_1\hat{\vec{p}}_2\right) \ , \\ \hat{X} &= \frac{1}{M}\left(m_1\hat{\vec{x}}_1 + m_2\hat{\vec{x}}_2\right) \ , \ \hat{\vec{x}} = \hat{\vec{x}}_1 - \hat{\vec{x}}_2 \ . \end{aligned} \tag{344}$$

The Hamiltonian can then be written as

E 67

$$\hat{H} = \frac{1}{2M}\hat{\vec{P}}^{\,2} + \frac{1}{2m_r}\hat{\vec{p}}^{\,2} + V(\hat{\vec{x}}) \ . \tag{345}$$

A potential that depends not on $\hat{\vec{x}}_1$ and $\hat{\vec{x}}_2$ separately but only on their difference $\hat{\vec{x}}$ leads to translation invariance, and hence to conservation of the total momentum :

$$[\hat{\vec{P}}, \hat{H}] = 0 \ . \tag{346}$$

If the potential is also rotationally invariant, that is, if it depends not on the complete $\hat{\vec{x}}$ but only on its length $\hat{\vec{x}}^2$, it is called a central potential, with concomitant conservation of *angular momentum*. **E 66**

9.6 Exercises 62 to 69

Exercise 62 : **A one-dimensional harmonic molecule**
We consider a one-dimensional system consisting of two particles of equal mass m, that attract each other due to a potential of the form

$$V(x_1, x_2) = \frac{1}{2} m\omega^2 \left(x_1 - x_2\right)^2$$

Compute M and m_r, and write down the Hamiltonian using global and internal observabels.

Exercise 63 : **Energies of the harmonic atom**
Show that the total-energy eigenvalues of the one-dimensional harmonic atom can be written as

$$E = E(P, n) = \frac{1}{4m} P^2 + \hbar\omega\sqrt{2}\left(n + \frac{1}{2}\right)$$

where P is a continuous, real number and n a non-negative integer. Explain the occurrence of the numbers '4' and '$\sqrt{2}$' in this result. Also, determine the degeneracy of the energy eigenvalues.

Exercise 64 : **Two harmonic particles**
Consider a one-dimensional system of two particles of mass m_1 and m_2, respectively, with a Hamiltonian given by

$$\hat{H} = \frac{1}{2m_1}\hat{p}_1^2 + \frac{1}{2m_2}\hat{p}_2^2 + \frac{1}{2}m_1\omega^2\hat{x}_1^2 + \frac{1}{2}m_2\omega^2\hat{x}_2^2 \ ,$$

with obvious notation $\hat{p}_{1,2}$ and $\hat{x}_{1,2}$.

1. Show that the system is separable, and that the energy eigenvalues are given by the expression

$$E_{n_1,n_2} = \hbar\omega\left(n + 1 + n_2 + 1\right) \ .$$

2. Go over to global and internal variables as discussed in the text, and show that the Hamiltonian can also be written as

$$\hat{H} = \frac{1}{2M}\hat{P}^2 + \frac{1}{2m_r}\hat{p}^2 + \frac{1}{2}M\omega^2\hat{X}^2 + \frac{1}{2}m_r\omega^2\hat{x}^2 \ ,$$

with again obvious notation.

3. Show that the system is still separable, and that the energy eigenvalues can *also* be written as

$$E_{N,n} = \hbar\omega^2 \left(N + n + 1\right) \quad .$$

4. An energy eigenstate with given N and n is, in general, *not* and energy eigenstate with given n_1 and n_2, *i.e.* it is in general a superposition. Determine which $n_{1,2}$ can be in that superposition, and figure out when the energy eigenstate is *not* a superposition.

Exercise 65 : **Three-dimensional harmonic oscillator**
We consider a system of a single particle in a three-dimensional harmonic oscillator potential :

$$\hat{H} = \frac{1}{2m}\hat{\vec{p}}^2 + \frac{1}{2}m\omega^2\hat{\vec{x}}^2 \quad .$$

1. Write the Hamiltonian in terms of the three Cartesian coordinates, and show that it is separable.

2. Show that the energy of this system depends on three nonnegative integers :

$$E = E_{n_1,n_2,n_3} = \hbar\omega \left(n_1 + n_2 + n_3 + \frac{3}{2}\right)$$

3. Determine the degeneracy, that is, compute how many combinations of n_1, n_2 and n_3 sum to a given value n.

Exercise 66 : **Angular momentum**
In classical physics, the angular momentum of a particle with respect to some origin is given by

$$\vec{L} = \vec{x}_{\rm cl} \times \vec{p}_{\rm cl}$$

In quantum mechanics, we have a similar definition. For a single particle with position and momentum operators $\hat{\vec{x}}$ and $\hat{\vec{p}}$, respectively :

$$\begin{aligned}
\hat{L}_x &= \hat{y}_y\hat{p}_z - \hat{z}\hat{p}_y \\
\hat{L}_y &= \hat{z}_y\hat{p}_x - \hat{x}\hat{p}_z \\
\hat{L}_z &= \hat{x}_y\hat{p}_y - \hat{y}\hat{p}_z
\end{aligned}$$

which we might indeed write as

$$\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$$

1. Show that

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z \ , \ [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x \ , \ [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$$

which we could write as the weird-looking

$$\hat{\vec{L}} \times \hat{\vec{L}} = i\hbar\,\hat{\vec{L}}$$

2. Show that
$$[\hat{\vec{L}}, \hat{\vec{x}}^2] = [\hat{\vec{L}}, \hat{\vec{p}}^2] = 0$$
Do this componentwise, for instance by computing $[\hat{L}_x, \hat{x}^2 + \hat{y}^2 + \hat{z}^2]$.

3. Show that, for a central potential,
$$[\hat{\vec{L}}, \hat{H}] = 0$$
and that therefore also *angular momentum* is conserved.

4. Show that the components of angular momentum are *in*commensurable (this is *easy*!). Show that, on the other hand,
$$[\hat{\vec{L}}, \hat{\vec{L}}^2] = 0$$

5. Show the following : if a particle moves under the influence of a central potential, and we have measured the values of $\vec{L}^2$ and L_z, say, these values will be preserved throughout the motion of the particle.

6. Show that for such a particle a maximal commensurable set of observables is formed by E, $\vec{L}^2$ and $\vec{L}$.

Exercise 67 : **Three-dimensional harmonic atom**
Generalize the model of Exercises 62 and 63 to the case of three dimensions.. Be careful ! Both the result for the total energy and that of the degeneracy changes.

Exercise 68 : **Degeneracies of non-isotropic harmonic oscillators**
Consider a three-dimensional harmonic oscillator that is *non-isotropic*, which means that the oscillation frequencies are not the same in all directions. The Hamiltonian is
$$\hat{H} = \frac{1}{2m}\hat{\vec{p}}^2 + \frac{m}{2}\left(\omega_x \hat{x}^2 + \omega_y \hat{y}^2 + \omega_z \hat{z}^2\right)$$

1. Show that this system is separable, and give the form of the energy eigenvalues.

2. Suppose that the frequencies are in the proportion
$$\omega_x : \omega_y : \omega_z = 1 : \sqrt{2} : \sqrt{3}$$
Show that there is no degeneracy. (Hint : $\sqrt{6}$ is irrational).

3. Let the frequencies be in the proportion
$$\omega_x : \omega_y : \omega_z = 1 : 1 : 2$$
Compute the degeneracies of the five lowest energies. Try to find the general formula.

Exercise 69 : Stuck with one another
We consider a one-dimensional system of two particles. The particles can move freely except that they can not be further apart than a distance L (imagine the particles tied together by a chain of length L).

1. Find the potential that describes this situation, and show that it is translation-invariant.

2. Let the particles have mass m and $2m$ respectively. Give the Hamiltonian for this system in terms of the particles' position and momentum.

3. Go over to global and internal position and momenta, and give the Hamiltonian in terms of these.

4. Derive the form of the total energy of the system.

10 Discrete probability

Consider a situation where an observation of a quantity A can result in several possible outcomes a, which we shall denote by $a_1, a_2, a_3, \ldots, a_K$. Think of throwing a dice, or a coin : or, indeed, of a quantum measurement. We say that the probability (or *probability density*) $P(a_j)$ is the *fraction of times* that the result of the observation $a = a_j$ comes up if the experiment is repeated over and over again[196]. Obviously we must have

$$P(a_j) \geq 0 \quad , \quad \sum_{j=1}^{K} P(a_j) = 1 \ . \tag{347}$$

The *expectation value* of a is the average value that will come out of a large number of repetitions of the experiment. By the very definition of probability, the number of times that a_k comes up if we repeat the experiment N times is expected to be $N \times P(a_j)$, and so the expectation value of a (denoted by brackets) is

$$\langle a \rangle = \frac{1}{N} \sum_{j=1}^{K} N \times P(a_j) \times a_j = \sum_{j=1}^{K} a_j \ P(a_j) \ . \tag{348}$$

The expectation value is not necessarily one of the possible outcomes ! For instance if we throw dice (with possible outcomes 1,2,3,4,5,6) the expectation value is 3.5 provided the dice are not loaded. Also note that expectation values are *linear* :

$$\begin{aligned} \langle a + b \rangle &= \langle a \rangle + \langle b \rangle \ , \\ \langle ca \rangle &= c \ \langle a \rangle \ \text{ if } c \text{ is a fixed constant } . \end{aligned} \tag{349}$$

By the same reasoning we can find the expectation value of a^2, of a^3, and so on : we have[197]

$$\langle a^p \rangle = \sum_{j=1}^{K} {a_j}^p \ P(a_j) \ . \tag{350}$$

Of particular importance is the *variance*, that is the spread of outcomes around the expectation value : we define it as the average of the difference between outcome and expectation value, where this difference is squared to get rid of its sign. It is denoted by $\sigma(a)^2$:

$$\sigma(a)^2 \quad = \quad \left\langle \Big(a - \langle a \rangle \Big)^2 \right\rangle$$

[196] This is the so-called *frequentist* approach to probability. Other approaches are also possible, but we shall not discuss them in these notes. Note that we must invoke somehow the *limit* that the nuber of repeats becomes infinite, since real randomness implies statistical fluctuations : for instance, a series of 10000 throws with a true coin that gives exactly 5000 heads and 5000 tails is itself not very likely, having a probability of slightly less that 0.8%.

[197] Note that this rule gives correctly that $\left\langle a^0 \right\rangle = 1$.

$$= \left\langle a^2 - 2a\,\langle a\rangle + \langle a\rangle^2 \right\rangle = \left\langle a^2\right\rangle - 2\,\langle a\rangle\,\langle a\rangle + \langle a\rangle^2$$
$$= \left\langle a^2\right\rangle - \langle a\rangle^2$$
$$= \sum_{j=1}^{K} {a_j}^2\, P(a_j) \;-\; \left(\sum_{j=1}^{K} a_j\, P(a_j)\right)^2 \quad . \tag{351}$$

Its square root[198], $\sigma(a)$, is called the *standard deviation*, or *uncertainty* of a: it is *not* any uncertainty in the determination of a, since A may be observed as accurately as we please ; rather it is the precision with which we know something about the outcome *beforehand*.

In this context the following comes in useful : suppose that $\sigma(a) = 0$. We then know that the outcome a is always equal to its expectation value $\langle a\rangle$: which simply means that there actually is only *one* possible outcome of the observation! We find that in that case

$$P(a_j) = 1 \quad , \quad P(a_n) = 0 \;\text{ if } a_n \neq a_j \tag{352}$$

for some given j ; and then of course also $\langle a\rangle = a_j$. The whole probability density is then concentrated at the single value a_j. So, indeed, if the uncertainty in a is zero, then we know exactly what observation of A will give us.

A final result, which we shall not prove here, is the following : if we know all the expectation values $\langle a^q\rangle$, $q = 1, 2, 3, \ldots$, then the probability density $P(a_n)$ is completely determined.

11 The Dirac delta function

The Dirac delta function (or, more exactly, delta *distribution*) can be seen as the continuum analogue of the Kronecker delta. Whereas the Kronecker delta can be defined as follows :

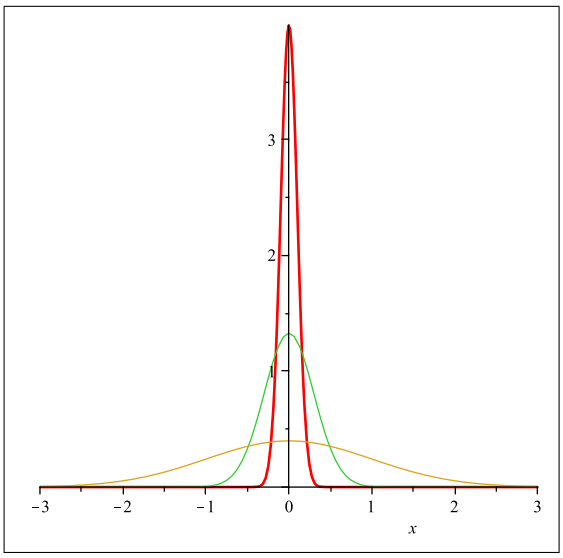
$$\delta_{m,n} = 0 \quad \text{if } m \neq n \quad , \quad \sum_n \delta_{m,n}\, f_n = f_m \tag{353}$$

for a sequence of numbers f_n, the Dirac delta obeys

$$\delta(x-y) = 0 \quad \text{for } x \neq y \quad , \quad \int_{-\infty}^{\infty} dy\, \delta(x-y)\, f(y) = f(x) \tag{354}$$

for a function $f(x)$. The dirac delta is most easily visualized as an infinitely narrow, infinitely tall *spike* sitting at argument 0.

[198] $\sigma(a)^2$ has the *dimension* not of a but of a^2 : for purposes of discussion the preference is therefore usually given to $\sigma(a)$ rather than $\sigma(a)^2$.



The Dirac delta $\delta(x)$ can be seen as the limit $a \to 0$ of

$$\delta_a(x) = \frac{\exp(-x^2/2a^2)}{\sqrt{2\pi a^2}} \quad ,$$

a Gaussian that becomes more and more narrow as a decreases. In the picture we show the shapes for $a = 1, a = 0.3$ and $a = 0.1$.

Another useful fact to know about the Dirac delta is the following. Let us define the *step function* by

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{355}$$

This function, then, is a simple step upwards by one unit, situated at $x = 0$. The *derivative* of this function is our Dirac delta :

$$\frac{d}{dx}\,\theta(x) = \delta(x) \quad . \tag{356}$$

For quantum states labeled by *discrete* numbers n, the orthonormality condition reads

$$\langle n|\, n'\rangle = \delta_{n,n'} \quad ; \tag{357}$$

for quantum states labeled by *continuous* numbers, such as the position eigenstates $|x\rangle$, the orthonormality condition ought rather to be written as[199]

$$\langle x|\, x'\rangle = \delta(x - x') \quad . \tag{358}$$

A result that we present here without proof is a useful relation between complex exponents and delta functions :

$$\sum_{n=-\infty}^{\infty} e^{2i\pi nx} = \sum_{k=-\infty}^{\infty} \delta(x - k) \quad . \tag{359}$$

[199] A technical detail : this orthonormality condition would seem to be clumsy since it gives $\langle x|\, x\rangle = \infty$. In fact, for continuous variables such as position, a *physical* system can never be in a pure eigenstate since we cannot measure the position with infinite accuracy, and the only states we really encounter are therefore always superpositions. This makes the mathematics work out ! At the level of *these* lecture notes, the main message is *not to worry too much about the distinction between discrete and continuous.*

12 The Gaussian integral

Let us consider computing the following integral, which occurs a lot in physical problems :

$$J_0 = \int_{-\infty}^{\infty} dx\ e^{-cx^2} \quad . \tag{360}$$

This integral seems very difficult. But there exist the beautiful 'doubling trick' : instead of J_0 we shall compute ${J_0}^2$! First we write

$${J_0}^2 = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\ e^{-c(x^2+y^2)} \quad . \tag{361}$$

Next we transform to *polar coordiniates* :

$$x = r\ \sin(\phi) \quad , \quad y = r\ \cos(\phi) \ , \tag{362}$$

so that

$$x^2 + y^2 = r^2 \quad , \quad dx\ dy = r\ dr\ d\phi \ . \tag{363}$$

The integral now becomes

$${J_0}^2 = \int_0^{2\pi} d\phi \int_0^{\infty} dr\ r\ e^{-cr^2} = 2\pi \int_0^{\infty} dr\ r\ e^{-cr^2} \quad . \tag{364}$$

Finally, we go over from r to $s = r^2$, so that $r\ dr = ds/2$:

$${J_0}^2 = \pi \int_0^{\infty} ds\ e^{-cs} = \frac{\pi}{c} \ ; \tag{365}$$

and this gives us the final result

$$J_0 = \int_{-\infty}^{\infty} dx\ e^{-cx^2} = \sqrt{\pi/c} \ . \tag{366}$$

Once you've seen this trick, it is easy to remember ! We can also get extra results, once we know J_0. Let us consider the more general expression

$$J_n = \int_{-\infty}^{\infty} dx\ e^{-cx^2}\ x^n \quad , \quad n = 0, 1, 2, 3, \ldots \tag{367}$$

In a very compact way, we can do *all* these integrals at once, by considering their so-called *generating function*

$$\begin{aligned} F(z) &\equiv \sum_{n\geq 0} \frac{z^n}{n!} J_n \\ &= \int_{-\infty}^{\infty} dx\ e^{-cx^2} \sum_{n\geq 0} \frac{z^n}{n!} x^n = \int_{-\infty}^{\infty} dx\ e^{-cx^2+zx} \quad . \end{aligned} \tag{368}$$

Since we have the simple identity

$$-cx^2 + zx = -c\left(x - \frac{z}{2c}\right)^2 + \frac{z^2}{4c} \quad , \tag{369}$$

the integral $F(z)$ can be immediately computed by shifting the integration variable x to $x' = x - z/2c$:

$$F(z) = \int_{-\infty}^{\infty} dx\ e^{-c(x-z/2c)^2 + z^2/4c} = \int_{-\infty}^{\infty} dx'\ e^{-cx'^2}\ e^{z^2/4c} = \sqrt{\frac{\pi}{c}}\, e^{z^2/4c} \ , \tag{370}$$

and again expanding the exponential in the last lemma we arrive at

$$F(z) = \sum_{n\geq 0} \frac{z^n}{n!}\, J_n = \sum_{n\geq 0} \sqrt{\frac{\pi}{c}}\, \frac{z^{2n}}{n!(4c)^n} \quad . \tag{371}$$

We can now read off the results for all J_n :

$$J_{2k+1} = 0 \quad , \quad J_{2k} = \sqrt{\frac{\pi}{c}}\, \frac{(2k)!}{k!\,(4c)^k} \quad . \tag{372}$$

For instance, the mean ($2k+1 = 1$) of the Gaussian distribution $\exp(-x^2/(2b^2))$ is zero, while its variance ($2k = 2$) is seen to be[200] $\sigma(x)^2 = b^2$.

13 The Gamma function

This very important function is defined by its integral representation :

$$\Gamma(x) = \int_0^{\infty} dz\ z^{x-1}\ e^{-z} \quad . \tag{373}$$

This definition only works, strictly speaking, as long as $x > 0$ (although this can be extended by so-called analytic continuation). A few simple values are

$$\Gamma(1) = \int_0^{\infty} dz\ e^{-z} = 1 \tag{374}$$

and

$$\begin{aligned} \Gamma(1/2) &= \int_0^{\infty} dz z^{-1/2}\ e^{-z} \\ &= 2\int_0^{\infty} dt\ e^{-t^2} = \int_{-\infty}^{\infty} dt\ e^{-t^2} = \sqrt{\pi} \ , \end{aligned} \tag{375}$$

[200] By writing $c = 1/(2b^2)$.

where we have used $z = t^2$ and the Gaussian integral of the previous section. A very important property is obtained by partial integration :

$$\begin{aligned} \Gamma(x+1) &= \int_0^\infty dz\; z^x\; e^{-z} \\ &= \big[- z^x\; e^{-z} \big]_{z=0}^{z=\infty} + \int_0^\infty dz\; x\, z^{x-1}\; e^{-z} \quad , \end{aligned} \tag{376}$$

so that we find

$$\Gamma(x+1) = x\,\Gamma(x) \quad . \tag{377}$$

This is very reminiscent of the factorial function ; and, indeed, for integer values of x we see that

$$\Gamma(n+1) = n! \quad . \tag{378}$$

A few results for not-too-large arguments are now simply found :

$$\begin{aligned} &\Gamma(2) = 1 \;\;,\;\; \Gamma(3) = 2 \;\;,\;\; \Gamma(4) = 6 \;\;,\;\; \Gamma(5) = 24 \;\;,\;\; \Gamma(6) = 120 \;\;, \\ &\Gamma(3/2) = \frac{\sqrt{\pi}}{2} \;\;,\;\; \Gamma(5/2) = \frac{3\sqrt{\pi}}{4} \;\;,\;\; \Gamma(7/2) = \frac{15\sqrt{\pi}}{8} \;\;,\;\; \Gamma(9/2) = \frac{105\sqrt{\pi}}{16} \quad . \end{aligned} \tag{379}$$

A very useful formula (which we shall not prove in these notes, although the proof is not particularly deep) that involves the Gamma function is *Euler's integral* :

$$\int_0^1 dz \;\; z^a\,(1-z)^b = \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)} \quad . \tag{380}$$

14 Solutions to the Exercises

14.1 Excercises 1 to 6

Solution to Exercise 1
If there are three possible positions, the matrix notation should have 3 entries :

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \quad , \quad \langle\psi| = \begin{pmatrix} \alpha^* & \beta^* & \gamma^* \end{pmatrix}$$

The probability for a given position is the absolute value squared of the corresponding component of the ket : if all three outcomes are equally probable we must therefore have

$$|\alpha|^2 = |\beta|^2 = |\gamma|^2 = 1/3$$

The most general such state is therefore

$$|\psi\rangle = \begin{pmatrix} e^{i\phi_1}/\sqrt{3} \\ e^{i\phi_2}/\sqrt{3} \\ e^{i\phi_3}/\sqrt{3} \end{pmatrix}$$

where $\phi_{1,2,3}$ are arbitrary.

Solution to Exercise 2
First, some notation : define real and imaginary parts of the product by

$$\langle\chi|\,\psi\rangle = A + iB$$

The absolute value mentioned in the excercise is then

$$|\,\langle\chi|\,\psi\rangle\,| = \sqrt{A^2 + B^2}$$

And now to work ! Consider the valid quantum state

$$|\phi\rangle = |\psi\rangle - e^{i\phi}\,|\chi\rangle$$

where ϕ is some angle. No matter what is the *interpretation* of this state, it exists and therefore has also a bra

$$\langle\phi| = \langle\psi| - e^{-i\phi}\,\langle\chi|$$

and the product of the state with itself must be nonnegative :

$$\begin{aligned} 0 &\leq \langle\phi|\,\phi\rangle \\ &= \langle\psi|\,\psi\rangle + \langle\chi|\,\chi\rangle - e^{-i\phi}\,\langle\chi|\,\psi\rangle - e^{i\phi}\,\langle\psi|\,\chi\rangle \\ &= 2 - 2A\,\cos\phi - 2B\,\sin\phi \end{aligned}$$

so that we find that in all cases

$$F(\phi) \leq 1 \quad , \quad F(\phi) = A\,\cos\phi + B\,\sin\phi$$

The extremum value of $F(\phi)$ is found from its derivative

$$F'(\phi) = -A\,\sin\phi + B\,\cos\phi$$

and we find the extremizing value

$$F'(\tilde{\phi}) = 0 \;\;\Rightarrow\;\; \tan\tilde{\phi} = B/A \;\;\Rightarrow\;\; \sin\tilde{\phi} = \frac{B}{\sqrt{A^2+B^2}}\;,\; \cos\tilde{\phi} = \frac{A}{\sqrt{A^2+B^2}}$$

so that F is bounded by

$$F(\tilde{\phi}) = \sqrt{A^2+B^2} \leq 1$$

and this proves the first assertion : notice, valid for *any* two states !

Now the second part : suppose that $|\,\langle\chi|\,\psi\rangle\,| = 1$. we then have $F(\tilde{\phi}) = 1$, or in other words, $\langle\phi|\,\phi\rangle = 0$. But this means that $|\phi\rangle$ itself must be zero, in other words

$$|\psi\rangle - e^{i\tilde{\phi}}\,|\chi\rangle = 0 \tag{381}$$

so that the two states differ, indeed, by only a complex phase[201].

Solution to Exercise 3

1. This is simply a matter of doing the multiplication.
2. Since $\langle\chi_j|\,\chi_j\rangle = 4$, it suffices to divide each coefficient by 2 to normalize these states to unity.
3. By simple inspection, or solving the system of linear equations, we find

$$\begin{aligned}
|\psi_1\rangle &= (|\chi'_1\rangle + |\chi'_2\rangle + |\chi'_3\rangle + |\chi'_4\rangle)/2 \;\;, \\
|\psi_2\rangle &= (|\chi'_1\rangle + |\chi'_2\rangle - |\chi'_3\rangle - |\chi'_4\rangle)/2 \;\;, \\
|\psi_3\rangle &= (|\chi'_1\rangle - |\chi'_2\rangle + |\chi'_3\rangle - |\chi'_4\rangle)/2 \;\;, \\
|\psi_4\rangle &= (|\chi'_1\rangle - |\chi'_2\rangle - |\chi'_3\rangle + |\chi'_4\rangle)/2 \;\;,
\end{aligned}$$

   where the $|\chi'_j\rangle$ are the normalized versions of the $||chi_j\rangle$.

Solution to Exercise 4
We have $\langle\psi_j|\,\psi_j\rangle = 8$ for each j. In addition,

$$\langle\psi_1|\,\psi_2\rangle = 0 \;\;,\;\; \langle\psi_1|\,\psi_3\rangle = 3 + 5i \;\;,\;\; \langle\psi_1|\,\psi_4\rangle = 0 \;\;,$$

[201] I do not remember seeng this simple result anywhere in the books. However, it is very useful !

$$\langle\psi_2|\,\psi_3\rangle = 3-3i \;\;,\;\; \langle\psi_2|\,\psi_4\rangle = -4\sqrt{2} \;\;,\;\; \langle\psi_3|\,\psi_4\rangle = (1+i)(1-2\sqrt{2}) \;\;.$$

Solution to Exercise 5

1. The given time evolution is only a complex phase, which by definition does not influence the states' physical contents.
2. The states' product becomes $e^{i(\beta-\alpha)t}\,\langle 1|\,2\rangle$ and is still zero. This ought to be expected since the states' physical contents didn't change and therefore the states' overlap must remain zero.
3. It is sufficient to treat the states *pairwise*, just like above.

Solution to Exercise ??
For the normalization we have

$$\begin{aligned}\langle\theta|\,\theta\rangle &= (\cos\theta)^2\,\langle 0|\,0\rangle + \cos\theta\,\sin\theta\,\Big(\,\langle 0|\,90\rangle + \langle 90|\,0\rangle\Big) + (\sin\theta)^2\,\langle 90|\,90\rangle \\ &= (\cos\theta)^2 + (\sin\theta)^2 = 1 \;\;.\end{aligned}$$

Now, assume that the polaroid is held up at 0 degrees, say. The probability of transmission is then

$$|\,\langle 0|\,\theta\rangle\,|^2 = (\cos\theta)^2 \;\;,$$

and the average value of this proability is

$$\frac{1}{2\pi}\int_0^{2\pi}(\cos\theta)^2\;d\theta = \frac{1}{2} \;\;.$$

Since all θ are equally likely, the orientation of the polaroid does *not* matter.

Solution to Exercise 6

1.

$$\begin{aligned}\langle\nu_e|\,\nu_m u\rangle &= \Big(\cos\theta\,\langle\nu_1| + \sin\theta\,\langle\nu_2|\Big)\Big(-\sin\theta\,|\nu_1\rangle + \cos\theta\,|\nu_2\rangle\Big) \\ &= -\sin\theta\,\cos\theta\,\langle\nu_1|\,\nu_1\rangle + \sin\theta\,\cos\theta\,\langle\nu_2|\,\nu_2\rangle \\ &\quad -\sin\theta^2\,\langle\nu_2|\,\nu_1\rangle + \cos\theta^2\,\langle\nu_1|\,\nu_2\rangle \\ &\quad -\sin\theta\,\cos\theta\;+\;\sin\theta\,\cos\theta \;\;=\;\; 0\end{aligned}$$

2. In the previous result, only the terms $\langle\nu_1|\,\nu_1\rangle$ and $\langle\nu_2|\,\nu_2\rangle$ survive : and these products do not change under the time evolution.

3.

$$\begin{aligned}
\langle \nu_\mu | \nu_e; t \rangle &= \Big(-\sin\theta \, \langle \nu_1| + \cos\theta \, \langle \nu_2| \Big) \Big(e^{i\alpha\, t} \cos\theta \, |\nu_1\rangle + e^{i\beta\, t} \sin\theta \, |\nu_2\rangle \Big) \\
&= \sin\theta \, \cos\theta \, \left(-e^{i\alpha\, t} + e^{i\beta\, t} \right) \\
&= -\frac{1}{2} \sin(2\theta) \, \left(e^{i(\alpha-\beta)t/2} - e^{-i(\alpha-\beta)t/2} \right) \, e^{i(\alpha+\beta)t/2} \\
&= -i \, \sin(2\theta) \, \sin\left(\frac{1}{2}(\alpha-\beta)t \right) \, e^{i(\alpha+\beta)t/2}
\end{aligned}$$

4. $P(\nu_e \to \nu_\mu)$ is equal to $|\, \langle \nu_\mu | \nu_e; t \rangle \,|^2$. The probability of persistence $P(\nu_e \to \nu_e)$ is of course simply $1 - P(\nu_e \to \nu_\mu)$.

5. Any *overall* complex phase is irrelevant, and therefore any physical effects must be related to the *relative* phase $\alpha - \beta$. Whenever θ is a multiple of $\pi/2$, either $\sin\theta$ or $\cos\theta$ vanishes so that the states $|\nu_e\rangle$ and $|\nu_\mu\rangle$ become *identical* to $|\nu_1\rangle$ and $|\nu_2\rangle$, respectively (if $\sin\theta = 0$), or to $|\nu_2\rangle$ and $|\nu_1\rangle$, respectively (if $\cos\theta = 0$). In those cases there can be no oscillations since $|\nu_e\rangle$ and $|\nu_\mu\rangle$ evolve independently.

14.2 Excercises 7 to 16

Solution to Exercise 7

$$\begin{aligned}
\hat{1} \, |\psi\rangle &= \left(\sum_j |j\rangle \, \langle j| \right) \left(\sum_{j'} c_{j'} \, |j'\rangle \right) = \sum_{j,j'} c_{j'} \, |j\rangle \, \langle j|\, j'\rangle \\
&= \sum_{j,j'} c_{j'} \, |j\rangle \, \delta_{j,j'} = \sum_j c_j \, |j\rangle = |\psi\rangle
\end{aligned}$$

Solution to Exercise 8

$$\begin{aligned}
\sum_j P(x_j) &= \sum_j |c_j|^2 = \sum_j \langle \psi | \, j\rangle \, \langle j | \, \psi\rangle \\
&= \langle \psi | \left(\sum_j |j\rangle \, \langle j| \right) |\psi\rangle = \langle \psi | \, \hat{1} \, |\psi\rangle = \langle \psi | \, \psi\rangle = 1
\end{aligned}$$

Solution to Exercise 9
If a_j are the possible outcomes for X, then the possible outcomes for $X^p + X^q$ are ${a_j}^p + {a_j}^q$; and the possible outcomes for $\exp(zX)$ are $\exp(z\, a_j)$ as seen from the definition of the exponent as given in the excercise.

Solution to Exercise 10

$$\begin{aligned}\langle x^p \rangle &= \sum_j P(x_j)\, {x_j}^p = \sum_j \langle \psi | j \rangle \; {x_j}^p \; \langle j | \psi \rangle \\ &= \langle \psi | \left(\sum_j {x_j}^p \; |j\rangle \langle j| \right) |\psi\rangle = \langle \psi | \, \widehat{X^p} \, |\psi\rangle = \langle \psi | \, \hat{X}^p \, |\psi\rangle\end{aligned}$$

Solution to Exercise 11
Since $\langle X \rangle$ is just a *number*, not an operator, we can immediately write

$$\begin{aligned}\left\langle (X - \langle X \rangle)^2 \right\rangle &= \left\langle X^2 - 2X \langle X \rangle + \langle X \rangle^2 \right\rangle \\ &= \langle X \rangle^2 - 2 \langle X \rangle \langle X \rangle + \langle X \rangle^2 = \left\langle X^2 \right\rangle - \langle X \rangle^2\end{aligned}$$

which is enough for the first lemma of the two equations : by putting in the notion of Eq.(39), we obtain the last lemma.

Solution to Exercise 12

1. On the one hand,

$$\langle n | \, \hat{Q} \, |n\rangle = \langle n | \left(\hat{Q} \, |n\rangle \right) = \langle n | \left(\lambda_n \; |n\rangle \right) = \lambda_n \; \langle n | n \rangle$$

   On the other hand,

$$\langle n | \, \hat{Q} \, |n\rangle = \left(\hat{Q}^\dagger \, |n\rangle \right)^\dagger |n\rangle = \left(\hat{Q} \, |n\rangle \right)^\dagger |n\rangle = \left(\lambda_n \; |n\rangle \right)^\dagger |n\rangle = {\lambda_n}^* \; \langle n | n \rangle$$

   Therefore $\lambda_n = {\lambda_n}^*$ since $\langle n | n \rangle$ is never zero.

2. This goes in the same way. On the one hand,

$$\langle k | \, \hat{Q} \, |n\rangle = \langle k | \left(\hat{Q} \, |n\rangle \right) = \lambda_n \; \langle k | n \rangle$$

   and on the other hand

$$\langle k | \, \hat{Q} \, |n\rangle = \left(\langle k | \, \hat{Q} \right) |n\rangle = \lambda_k \; \langle k | n \rangle$$

   If $\lambda_k \neq \lambda_n$ we must therefore have $\langle k | n \rangle = 0$.

Solution to Exercise 13

1. By its definition,

$$\langle k \rangle = \sum_{k=1}^{6} P(k)\, k = \frac{1}{6}\Big(1+2+3+4+5+6\Big) = \frac{7}{2}$$

2. By its definition,

$$\langle k \rangle^2 = \sum_{k=1}^{6} P(k)\, k^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$$

Then by its own definition

$$\sigma(k) = \sqrt{\langle k^2 \rangle - \langle k \rangle^2} = \sqrt{\frac{91}{6} - \frac{49}{4}} = \sqrt{\frac{35}{12}}$$

1. We know the possible outcomes for K, therefore

$$\hat{K} = \sum_{k=1}^{6} k\ |k\rangle \langle k|$$

2. The state $|\psi\rangle$ is of the form

$$|\psi\rangle = \sum_{k=1}^{6} c_k\ |k\rangle \quad , \quad c_k = \frac{1}{\sqrt{6}}$$

and therefore

$$\langle \psi | \psi \rangle = \sum_{k=1}^{6} |c_k|^2 = 6 \times \frac{1}{6} = 1$$

The probability to find $K = n$ ($n = 1, \ldots, 6$) is therefore equal to $|c_n|^2 = 1/6$ in all cases, just like the classical case.

3. The states $|k\rangle$ are now simply multiplied by a complex phase, which cannot change their physical properties ; or, æquivalently, the coefficients c_k are multiplied by a complex phase, which cannot change their absolute value.

4. The state is normalized since

$$\sum_{k=1}^{6} \left(\frac{\sqrt{k}}{\sqrt{21}}\right)^2 = \sum_{k=1}^{6} \frac{k}{21} = 1$$

The various expectations and such are

$$\langle k \rangle = \sum_{k=1}^{6} \frac{k}{21}\, k = \frac{13}{3}$$

$$\langle k^2 \rangle = \sum_{k=1}^{6} \frac{k}{21}\, k^2 = 21$$

$$\sigma(k) = \sqrt{21 - \left(\frac{13}{3}\right)^2} = \frac{2}{3}\sqrt{5}$$

These values are *not* those for an unloaded classical die.

5. A possible answer is
$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|1\rangle + |6\rangle\Big)$$
which gives $\langle k\rangle = 7/2$ but $\sigma(k) = 5/2$. This is also the largest possible value for $\sigma(k)$.

Solution to Exercise 14
Even if the kinetic energy is given, there is an *infinite* number of *directions* in which the particle can move, and the energy eigenvalues are infinitely degenerate. The single exception is when the particle *doesn't* move : then the energy and velocity are both exactly *zero*.

Solution to Exercise 15
The Hermiticity is trivial : the eigenvalues are all ± 1.

Solution to Exercise 16
The Hermiticity is again trivially established. The eigenvalue equation reads
$$-\lambda(1-\lambda)(2-\lambda) = 0$$
so the eigenvalues are 0,1 and 2, no degeneracies.

14.3 Excercises 17 to 21

Solution to Exercise 17

1. These are a matter of simply writing out the commutators explicitly.
2. The first term is proven by writing $\hat{C} = \hat{B}_2 \cdots \hat{B}_N$ and applying Eq.(63), and so on.
3. This is again trivial, it suffices to take $\hat{B}_1 = \hat{B}_2 = \cdots = \hat{B}_N$.

Solution to Exercise 18
This is a matter of simply writing out the matrix multiplications.

Solution to Exercise 19

1. The three matrices are
$$\hat{M}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \hat{M}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \hat{M}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and each of them has eigenvalues 1 and -1 by standard linear algebra : the equation for the eigenvalues λ read $\lambda^2 - 1 = 0$ in all three cases. Let the eigenvectors be given by

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

The eigenvector equations (eigenvalues are ± 1 !) are then

$$\begin{aligned} \hat{M}_x \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= \pm \begin{pmatrix} \beta \\ \alpha \end{pmatrix} \\ \hat{M}_y \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= \pm \begin{pmatrix} -i\,\beta \\ i\,\alpha \end{pmatrix} \\ \hat{M}_z \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= \pm \begin{pmatrix} \alpha \\ -\beta \end{pmatrix} \end{aligned}$$

Solutions for the normalized eigenvectors are therefore

$$\begin{aligned} \hat{M}_x &: \lambda = 1 : \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} , \lambda = -1 : \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \\ \hat{M}_y &: \lambda = 1 : \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix} , \lambda = -1 : \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{pmatrix} \\ \hat{M}_z &: \lambda = 1 : \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \lambda = -1 : \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

2. Suppose the result of measuring M_z was 1. The system is then in the state with vector notation

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

This state has an overlap of precisely $\pm 1/\sqrt{2}$ with the eigenvectors of $\hat{M}_x$, and therefore both outcomes for the measurement of M_x have probability 1/2. Exactly the same goes for $\hat{M}_y$, apart from some factors i that drop out in the probability.

3. Goes the same way. All the eigenvectors have overlap with the eigenvectors of the *other* operators that have absolute values equal to $1/\sqrt{2}$.

Solution to Exercise 20

1.

$$\langle \chi | \chi \rangle = \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b} \end{pmatrix} \begin{pmatrix} a\,e^{i\phi_a} \\ b\,e^{i\phi_b} \end{pmatrix} = a^2 + b^2$$

So we shall from now on assume $a^2 + b^2 = 1$.

2.

$$\begin{aligned}
\langle M_z\rangle &= \langle\chi|\,\hat{M}_z\,|\chi\rangle = \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b}\end{pmatrix}\begin{pmatrix} 1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix} a\,e^{i\phi_a}\\ b\,e^{i\phi_b}\end{pmatrix}\\
&= \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b}\end{pmatrix}\begin{pmatrix} a\,e^{i\phi_a}\\ -b\,e^{i\phi_b}\end{pmatrix} \;=\; a^2-b^2\\
\langle M_x\rangle &= \langle\chi|\,\hat{M}_z\,|\chi\rangle = \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b}\end{pmatrix}\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} a\,e^{i\phi_a}\\ b\,e^{i\phi_b}\end{pmatrix}\\
&= \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b}\end{pmatrix}\begin{pmatrix} b\,e^{i\phi_b}\\ a\,e^{i\phi_a}\end{pmatrix}\\
&= ab\left(e^{i(\phi_a-\phi_b)}+e^{-i(\phi_a-\phi_b)}\right) = 2ab\,\cos(\phi_a-\phi_b)\\
\langle M_y\rangle &= \langle\chi|\,\hat{M}_z\,|\chi\rangle = \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b}\end{pmatrix}\begin{pmatrix} 0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix} a\,e^{i\phi_a}\\ b\,e^{i\phi_b}\end{pmatrix}\\
&= \begin{pmatrix} a\,e^{-i\phi_a} & b\,e^{-i\phi_b}\end{pmatrix}\begin{pmatrix} -ib\,e^{i\phi_b}\\ ia\,e^{i\phi_a}\end{pmatrix}\\
&= iab\left(e^{i(\phi_a-\phi_b)}-e^{-i(\phi_a-\phi_b)}\right) = -2ab\,\sin(\phi_a-\phi_b)
\end{aligned}$$

3. Since $\hat{M}_j^2 = \hat{1}$ $(j = x, y, z)$ as proven before, we arrive immediatelly at

$$\sigma(M_x)^2 = 1 - 4a^2b^2\cos(\phi_a-\phi_b)^2 \quad , \quad \sigma(M_y)^2 = 1 - 4a^2b^2\sin(\phi_a-\phi_b)^2$$

4. The *form* of the inequality arises immediately once we remember that $[\hat{M}_x, \hat{M}_y] = 2i\hat{M}_z$. Furthermore,

$$\sigma(M_x)^2\sigma(M_y)^2 = 1 - 4a^2b^2 + 16a^4b^4\cos(\phi_a-\phi_b)^2\sin(\phi_a-\phi_b)^2$$

The *minimum* value of this expression as a function of the ϕ's is of course obtained by dropping the last term, and then we find

$$\begin{aligned}
\sigma(M_x)^2\sigma(M_y)^2 &\geq 1 - 4a^2b^2 = 1 - 4(1-b^2)b^2\\
&= (1-2b^2)^2 = (a^2-b^2)^2 = \langle M_z\rangle^2
\end{aligned}$$

which clinches it. The three special cases follow by inspection.

Solution to Exercise 21
This is just a matter of perseverance.

14.4 Excercises 22 to 29

Solution to Exercise 22
If

$$|\psi_C\rangle = e^{-iCt/\hbar}\,|\psi\rangle$$

then

$$i\hbar\frac{d}{dt}\left|\psi_C\right\rangle = Ce^{-iCt/\hbar}\left|\psi\right\rangle + e^{-iCt/\hbar}i\hbar\frac{d}{dt}\left|\psi\right\rangle = Ce^{-iCt/\hbar}\left|\psi\right\rangle + e^{-iCt/\hbar}\hat{H}\left|\psi\right\rangle$$

which proves the assertion if $\hat{H}$ and $e^{-iCt/\hbar}$ can be interchanged, for which we made the assumption on $\hat{H}$.

Solution to Exercise 23
For the state $|\psi\rangle$ as given by Eq.(115) we implement the Schrödinger equation directly :

$$\begin{aligned} i\hbar\frac{d}{dt}\left|\psi\right\rangle &= i\hbar\frac{d}{dt}\sum_n e^{-iE_nt/\hbar}c_n\left|E_n\right\rangle(0) \\ &= \sum_n e^{-iE_nt/\hbar}c_nE_n\left|E_n\right\rangle(0) \\ &= \sum_n e^{-iE_nt/\hbar}c_n\hat{H}\left|E_n\right\rangle(0) \\ &= \hat{H}\sum_n e^{-iE_nt/\hbar}c_n\left|E_n\right\rangle(0) = \hat{H}\left|\psi\right\rangle \end{aligned}$$

Solution to Exercise 24

1. If $\hat{H}\left|\psi\right\rangle = E\left|\psi\right\rangle$ then

   $$\left\langle\psi\right|\left[\hat{H},\hat{A}\right]\left|\psi\right\rangle = \left\langle\psi\right|\hat{H}\hat{A}-\hat{A}\hat{H}\left|\psi\right\rangle = \left\langle\psi\right|E\hat{A}-\hat{A}E\left|\psi\right\rangle = (E-E)\left\langle\psi\right|\hat{A}\left|\psi\right\rangle = 0$$

2. This is a matter of applying the Heisenberg equation *twice*. Extensions are trivial, *e.g.*

   $$\frac{d^3}{dt^3}\left\langle A\right\rangle = \frac{d}{dt}\left(\frac{d^2}{dt^2}\left\langle A\right\rangle\right) = -\frac{i}{\hbar^3}\left\langle\left[\hat{H},\left[\hat{H},\left[\hat{H},\hat{A}\right]\right]\right]\right\rangle$$

3. This comes straightaway from the definition of the uncertainty :

   $$\frac{d}{dt}\sigma(A)^2 = \frac{d}{dt}\left\langle A^2\right\rangle - 2\left\langle A\right\rangle\frac{d}{dt}\left\langle A\right\rangle$$

   and $[\hat{H},\hat{A}^2] = [\hat{H},\hat{A}]\hat{A}+\hat{A}[\hat{H},\hat{A}]$. The result as written has been prettified : note that $\left\langle A\right\rangle\left\langle\left[\hat{H},\hat{A}\right]\right\rangle = \left\langle\left[\hat{H},\hat{A}\right]\right\rangle\left\langle A\right\rangle$ since expectation values are just *numbers* and therefore commute.

Solution to Exercise 25
This done by expressing the commutators for $\hat{x}'$, etcetera, in terms of the original

ones. By simple goniometry we then arrive at the desired result. Note that here it is important that **(a)** the commutators are diagonal, and **(b)** have the same value in all three directions.

Solution to Exercise 26
we simply take into account that position, momentum, and *derivatives* now have three components. The Ehrenfest theorem in three dimensions is therefore

$$\frac{d}{dt}\left\langle \vec{p} \right\rangle = -\left\langle \vec{\nabla} V(\vec{x}) \right\rangle$$

Solution to Exercise 27

1. Since the potential goes all the way up to infinity, the particle can never 'escape' and will always be localized somewhere around $x = 0$.
2. $\hat{T} = \hat{p}^2/(2m)$ hence $\hat{H} = \hat{p}^2/(2m) + k\,\hat{x}^n$.
3. It is easy to see that $\hat{W}^\dagger = \hat{W}$, which is sufficient for its being the operator of an observable[202]. The Heisenberg equation that is asked for is simply the well-known expression $d\left\langle W \right\rangle/dt = i\left\langle [\hat{H}, \hat{W}] \right\rangle/\hbar$.
4. The integral can be trivially worked out :

$$\{d\left\langle W \right\rangle/dt\} = \lim_{\tau\to\infty} \int_0^\tau \frac{d\left\langle W \right\rangle}{dt}\, dt = \lim_{\tau\to\infty} \frac{1}{\tau}(\left\langle W \right\rangle(\tau) - \left\langle W \right\rangle(0)) = 0$$

   by the assumption that neither $\left\langle W \right\rangle(\tau)$ nor $\left\langle W \right\rangle(0)$ is infinite.
5. By simple algebra we find that

$$[\hat{p}^2, \hat{W}] = -4i\hbar\hat{p}^2 \quad , \quad [\hat{x}^n, \hat{W}] = 2ni\hbar\hat{x}^n \quad .$$

   Therefore,

$$\hat{H}, \hat{W}] = -4i\hbar\hat{T} + 2ni\hbar\hat{V} \quad .$$

6. The Ehrenfest theorem now says that

$$\frac{d}{dt}\left\langle W \right\rangle = 4\left\langle T \right\rangle - 2n\left\langle V \right\rangle$$

   and the desired result follows immediately.

[202] Sometimes people simply use $\hat{W} = \hat{x}\hat{p}$ but in my view that is not justified since that combination is not Hermitean.

Solution to Exercise 28
Eq.(137) can be written as the statement

$$f(x) = \sin(x) + \frac{\pi}{2}\cos(x) + x - \frac{\pi}{2} \geq 0 \quad .$$

Indeed, for $x = \pi$ we have $f(x) = 0$ since $\sin(\pi) = 0$ and $\cos(\pi) = -1$. Furthermore,

$$f'(x) = \cos(x) - \frac{\pi}{2}\sin(x) + 1 \quad ,$$

so that $f'(\pi) = 0$ as well. The curve for $f(x)$ therefore touches, but does not intersect, the real axis at $x = \pi$.

Solution to Exercise 29

1. This is trivial.
2. Both E_1 and E_2 have a 50% chance of being observed in both cases.
3. The time-evolving state that is $|+\rangle$ at $t = 0$ is

$$|+\rangle(t) \quad = \quad \frac{1}{\sqrt{2}}\left(\exp\left(-i\frac{E_1 t}{\hbar}\right)|1\rangle + \exp\left(-i\frac{E_2 t}{\hbar}\right)|2\rangle\right)$$

Therefore the time necessary to arrive at state $|-\rangle$ is given by

$$\frac{\pi\hbar}{|E_2 - E_1|}$$

and we indeed have

$$\frac{1}{2}|E_2 - E_1| = \frac{1}{2}(E_1 + E_2) - \min(E_1, E_2)$$

14.5 Excercises 30 to 45

Solution to Exercise 30
By the definitions given, we have

$$\langle\psi|\,\phi\rangle = \int dx\ \psi(x)^*\phi(x) \quad , \quad \langle\phi|\,\psi\rangle = \int dx\ \phi(x)^*\psi(x)$$

and these are one another's complex conjugates.

Solution to Exercise 31
If the wave function representation of $\hat{p}\,|\phi\rangle$ is $-i\hbar\frac{\partial}{\partial x}\phi(x)$, then

$$\left(\hat{p}\,|\phi\rangle\right)^\dagger |\psi\rangle = \int dx \left(i\hbar\frac{\partial}{\partial x}\phi(x)^*\right)\psi(x)$$

and partial integration leads to the original result, if we can neglect boundary terms.

Solution to Exercise 32
If the wave function has constant complex phase we may as well assume it to be real, so that $\psi(x)^* = \psi(x)$. Then

$$\psi(x)^* \frac{\partial}{\partial x}\psi(x) = \psi(x) \frac{\partial}{\partial x}\psi(x) = \frac{1}{2}\frac{\partial}{\partial x}\left(\psi(x)\right)^2$$

so that $\langle p \rangle$ becomes the integral of a total derivative. If the wave function vanishes at infinity, $\langle p \rangle$ therefore vanishes.

Solution to Exercise 33
We can use partial integration to make the replacement

$$\psi(x)^* \left(-\frac{\partial^2}{\partial x^2}\psi(x)\right) \quad \rightarrow \quad \left(\frac{\partial}{\partial x}\psi(x)^*\right)\left(\frac{\partial}{\partial x}\psi(x)\right)$$

which proves the formula. The only way in which $\langle E_{\rm kin} \rangle$ may possibly vanish is when $\psi'(x)$ is zero *everywhere*, in other words $\psi(x)$ is constant : but then it is either zero or not quadratically integrable. Both cases are inadmissible.

Solution to Exercise 34
The generalizations are

$$i\hbar \, \frac{d}{dt} \, \psi(\vec{x},t) = -\frac{\hbar^2}{2m}\vec{\nabla}^2 \, \psi(\vec{x},t) + V(\vec{x}) \, \psi(\vec{x},t)$$

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2 \, \psi(\vec{x}) + V(\vec{x}) \, \psi(\vec{x}) = E \, \psi(\vec{x})$$

$$\int d^3x \, |\psi(\vec{x})|^2 = 1 \ .$$

Solution to Exercise 36
The three-dimensional plane wave must of course read

$$\psi(\vec{x},t) \ \propto \ \exp\left(-\frac{i}{\hbar}(Et - \vec{x}\cdot\vec{p})\right)$$

Solution to Exercise 37

1. By combining the two exponentials we can write the wave function as

$$\psi(x,t) \propto \int_{-\infty}^{\infty} dx \, e^A$$

with

$$\begin{aligned} A &= -\frac{(p-p_0)^2}{4\sigma^2} - \frac{i}{\hbar}\left(\frac{p^2 t}{2m} - px\right) \\ &= -\left(\frac{it}{2m\hbar} + \frac{1}{4\sigma^2}\right)p^2 + \left(\frac{p_0}{2\sigma^2} + \frac{ix}{\hbar}\right)p - \frac{{p_0}^2}{4\sigma^2} \\ &= -\frac{\hbar m + 2it\sigma^2}{4\sigma^2\hbar m}\left(p - \frac{m(p_0\hbar + 2ix\sigma^2}{\hbar m + 2it\sigma^2}\right)^2 + B \quad , \\ B &= \frac{-m\sigma^2x^2 + i\hbar m p_0 x - i\hbar {p_0}^2 t/2}{\hbar(\hbar m + 2it\sigma^2)} \end{aligned}$$

and the integral over p leaves as a result

$$\psi(x,t) \propto e^B$$

2. By taking the real parto of B, since

$$|\psi(x,t)|^2 = \psi(x,t)\psi(x,t)^* = e^{B+B^*}$$

$$\begin{aligned} B + B^* &= \frac{-m\sigma^2x^2 + i\hbar m p_0 x - i\hbar {p_0}^2 t/2}{\hbar(\hbar m + 2it\sigma^2)} \\ &\quad + \frac{-m\sigma^2x^2 - i\hbar m p_0 x + i\hbar {p_0}^2 t/2}{\hbar(\hbar m - 2it\sigma^2)} \\ &= -2\sigma^2 \frac{(mx - p_o t)^2}{m^2\hbar^2 + 4\sigma^4 t^2} \end{aligned}$$

The value of N is obtained by applying Eq.(366) iwth

$$c = 2\sigma^2 \frac{(mx - p_o t)^2}{m^2\hbar^2 + 4\sigma^2 t^2}$$

3. The peak of the lump is that combination of x and t for which $|\psi|^2$ is maximal, *i.e.* c vanishes. This happens when $mx = p_0 t$, in other words when

$$x = t\,\frac{p_0}{m}$$

4. $|\psi(x,t)|^2$ is a Gaussian distribution in terms of x, so that the expected value of x coincides with the peak value (the Gaussian is symmetrical around its peak). From the last remark made in section 12 we see that $\sigma(x)^2$ must be given by

$$\sigma(x)^2 = \frac{1}{2c} = \frac{\hbar^2}{4\sigma^2} + \frac{\sigma^2 t^2}{m^2}$$

5. $\sigma(x)^2$ grows with t^2 so increaszes all the time. This is to be expected since the wave function is a combination of plane waves all moving with different velocity. At $t = 0$ we find $\sigma(x) = \hbar/(2\sigma)$ which is precisely the minimum value possible under the Heisenberg uncertainty relations. If σ is *small*, the spread in momenta is and the original speard in position is *large* ; on the other hand, since most momenta are then close together the wave spreads out *slowly*. For *large* σ the situation is reversed, and the wave starts as a narrow peak but spreads out rapidly.

Solution to Exercise 40
Since $\sin(-x) = -\sin(x)$, changing the sign of n is æquivalent to changing the sing of the wave function, and therefore does not give a physically different state. Since $\sin(0) = 0$, choosing $n = 0$ gives a wavefunction that is identically zero. In a sense, this is indeed a solution : the case of an *empty* box. But we do serious physics here, not some nerdish game.

Solution to Exercise 41
Put $x = Ly/n$. Then

$$\begin{aligned}\int_0^L dx\, \sin\left(\frac{n\pi x}{L}\right)^2 &= \frac{L}{n}\int_0^n dy\, \sin(\pi y)^2 = \frac{L}{2n}\int_0^n dy\, \Big(1 - \cos(2\pi y)\Big) \\ &= \frac{L}{2n}\left[y - \frac{1}{2\pi}\sin(2\pi y)\right]_{y=0}^{y=n} = L/2\end{aligned}$$

Solution to Exercise 42
The possible measurement outcomes for x are always between 0 and L. Therefore $\langle x\rangle$, $\langle x^2\rangle$ and $\sigma(x)$ are finite. Since $\sigma(x)\sigma(p) \geq \hbar/2$, $\sigma(p)$ can never be zero, and the system is never in an eigenstate of $\hat{p}$. The energy, however, is proportional to p^2, and the energy eigenstates are therefore a combination of waves with momentum $+p$ and $-p$. In classical physics this goes the same : a standing wave (fixed at its endpoints) is a combination of two travelling waves, travelling with *opposite* speed.

Solution to Exercise 43
The requirement on $\sum |c_n|^2$ arises from the fact that the energy eigenfunctions have absolute value squared that integrat to $L/2$. If $c_1 = c_2 = 1/\sqrt{L}$, $c_3 = c_4 = \cdots = 0$ the wave function takes the form

$$\begin{aligned}\psi(x,t) &= \frac{1}{\sqrt{L}}\left[\sin\left(\frac{\pi x}{L}\right)\exp\left(-\frac{i\hbar\pi^2 t}{2mL^2}\right) + \sin\left(\frac{2\pi x}{L}\right)\exp\left(-\frac{2i\hbar\pi^2 t}{mL^2}\right)\right] \\ &\propto \frac{1}{\sqrt{L}}\left[\sin\left(\frac{\pi x}{L}\right) + \sin\left(\frac{2\pi x}{L}\right)\exp\left(-\frac{3i\hbar\pi^2 t}{2mL^2}\right)\right]\end{aligned}$$

if we drop an overall complex phase. The probability density for position is then

$$\begin{aligned} |\psi(x,t)^2|^2 &= \frac{1}{L}\left[\sin\left(\frac{\pi x}{L}\right)^2 + \sin\left(\frac{2\pi x}{L}\right)^2 \right. \\ &\quad \left. + 2\sin\left(\frac{\pi x}{L}\right)\sin\left(\frac{2\pi x}{L}\right)\cos\left(\frac{3\hbar\pi^2 t}{2mL^2}\right)\right] \end{aligned}$$

We now employ the following elementary results:

$$\sin\left(\frac{n_1\pi x}{L}\right)\sin\left(\frac{n_2\pi x}{L}\right) = \frac{1}{2}\left(\cos\left(\frac{(n_1-n_2)\pi x}{L}\right) - \cos\left(\frac{(n_1+n_2)\pi x}{L}\right)\right)$$

and

$$\begin{aligned} \int_0^L dx\ x\ \cos\left(\frac{n\pi x}{L}\right) &= \left[\frac{L}{n\pi}x\sin\left(\frac{n\pi x}{L}\right)\right]_{x=0}^{x=L} - \frac{L}{n\pi}\int_0^L dx\ \sin\left(\frac{n\pi x}{L}\right) \\ &= \frac{L^2}{\pi^2 n^2}\left[\cos\left(\frac{n\pi x}{L}\right)\right]_{x=0}^{x=L} = \frac{L^2}{\pi^2 n^2}\left((-1)^n - 1\right) \end{aligned}$$

Putting everything together, we find

$$\langle x\rangle = \frac{L}{2}\left(1 - \frac{8}{9\pi^2}\cos\left(\frac{3\pi^2\hbar t}{2mL^2}\right)\right)$$

The position expectation value oscillates between about $0.32L$ and $0.68L$: it never reaches[203] the edges of the box.

Solution to Exercise 44

1. The classical kinetic energy is

$$E = \frac{1}{2}mv^2 = 0.5\times 10^{-3}\text{kg}\times\left(10^{-2}\frac{\text{m}}{\text{sec}}\right)^2 = 5\times 10^{-8}\text{Joule}$$

   The quantum mechanical energy is of course given by

$$E = \frac{\pi^2\hbar^2}{2mL^2}n^2 \approx 5.488\times 10^{-61}\text{Joule}\ n^2$$

   Equating these gives $n \approx 3\times 10^{26}$.

2. For the values given for the electron mass and the size of an atom, we find

$$E_1 \approx 1.4\times 10^{-18}\text{Joule}$$

   which is about 10 electronvolts. This is *roughly* of the order of atomic energy scales which are in the ballpark of a few electronvolts.

[203] For *this* state !

Solution to Exercise 45

1. We have the following time derivatives :

$$\dot{x}_{\rm cl} = \dot{\phi}_{\rm cl}\, R \cos\phi_{\rm cl} \quad , \quad \dot{y}_{\rm cl} = -\dot{\phi}_{\rm cl}\, R \sin\phi_{\rm cl}$$

   The angular momentum is

$$m\,(x_{\rm cl}\dot{y}_{\rm cl} - y_{\rm cl}\dot{x}_{\rm cl})$$

   with the result as given. We also have

$$\dot{x}_{\rm cl}^2 + \dot{y}_{\rm cl}^2 = R^2\dot{\phi}_{\rm cl}^2$$

   from which the expression for the kinetic energy follows.

2. This is simple algebra.

3. The commutation relation

$$[\hat{\phi}, \hat{p}_\phi] = i\hbar$$

   follows directly from the definition of $\hat{p}_\phi$. The commutator in Cartesian coordinates relies on the fact that

$$\frac{\partial}{\partial x}\arctan\left(\frac{x}{y}\right) = \frac{y}{x^2+y^2} \quad , \quad \frac{\partial}{\partial y}\arctan\left(\frac{x}{y}\right) = -\frac{x}{x^2+y^2}$$

4. The time-independent Schrödinger equation reads, quite trivially,

$$-\frac{\hbar^2}{2mR^2}\psi''(\phi) = E\,\psi(\phi)$$

5. The solutions to the time-independent Schrödinger equation are all of the form

$$\psi(\phi) = N\, e^{iA\phi} \quad , \quad \frac{\hbar^2 A^2}{2mR^2} = E$$

   with some normalization factor N. The periodicity implies that $A\phi$ and $A(\phi + 2\pi)$ must differ by $2n\pi$ for some integers n, so that $A = n\pi$. This gives the result for the energy. The normalization is obtained by realizing that we have to integrate along a circle with circumference $2\pi R$: hence

$$1 \equiv R\int_0^{2\pi} d\phi\, |\psi(\phi)|^2 = R\int_0^{2\pi} d\phi\, N^2 = 2\pi R N^2$$

   which gives us N as indicated.

14.6 Excercises 46 to 52

Solution to Exercise 46
All energies have the dimension of Joules. Since $x_{\rm cl}$ is expressed in meters and m in kilograms, ω^2 must have dimensionality Joule/(meter)^2/kilogram, in other words (second)$^{-2}$. Since m and $\hbar$ are the only other quantities in the problem there is no other possibility of forming a frequency.

Solution to Exercise 47
This is a matter of srtaightforward matrix multiplication, where you shouldn't be scared of the dots running off to the right and the bottom.

Solution to Exercise 48

1. $\hat{a}$ transforms a quantum state into another quantum states. Since these states must have the same dimensionality, $\hat{a}$ must be dimensionless.
2. The only quantities we are allowed to use are $\hbar$, m and ω ; and $\hbar\omega$ is easily seen to be the only possibility. We can consider the 'most general' combination $\hbar^\alpha \omega^\beta m^\gamma$ with exponents α, β and γ : the only solution for ending up with an energy is $\alpha = \beta = 1$, $\gamma = 0$.
3. This goes in the same way as the previous item.

Solution to Exercise 49

1. We have
$$\{\hat{a}, \hat{a}^\dagger\} + 1 = \sum_n (2n+2) \left|n\right\rangle \left\langle n\right|$$
and therefore the desired Hamiltonian can be written as
$$\hat{H} = \frac{E_0}{4} \left(\{\hat{a}, \hat{a}^\dagger\}^2 + 2\{\hat{a}, \hat{a}^\dagger\} + 1 \right)$$
2. Putting $m = \hbar = \omega = 1$ for simplicity, we have
$$\hat{a} = \frac{1}{\sqrt{2}} (\hat{x} + i\hat{p}) \quad , \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}} (\hat{x} - i\hat{p})$$
so that the anticommutator reads, of course,
$$\begin{aligned} \{\hat{a}, \hat{a}^\dagger\} &= \frac{1}{2} \Big((\hat{x} + i\hat{p})(\hat{x} - i\hat{p}) + (\hat{x} - i\hat{p})(\hat{x} + i\hat{p}) \Big) \\ &= \hat{x}^2 + \hat{p}^2 \end{aligned}$$

The Hamiltonian is therefore given by

$$\hat{H} = \frac{E_0}{4}\left(\hat{x}^4 + \hat{p}^4 + \hat{x}^2\hat{p}^2 + \hat{p}^2\hat{x}^2 + 2\hat{x}^2 + 2\hat{p}^2 + 1\right)$$

Especially the occurrence of $\hat{p}^4$ makes this an extremele unlikely system.

Solution to Exercise 50
This is a matter of applying the Gaussian integral result of section 12.

Solution to Exercise 51
If $x = a\sin(\omega t)$ then

$$\frac{dx}{dt} = \omega\, a\, \cos(\omega t)$$

so that

$$\frac{dt}{dx} = \frac{1}{\omega\, a\, cos(\omega t)} = \frac{1}{\omega\sqrt{a^2 - x^2}}$$

Solution to Exercise 52

1. Suppose $\sum \alpha_n |n\rangle$ is an eigenstate of $\hat{a}^\dagger$. Now, there is *no* state $|\psi\rangle$ such that $\hat{a}^\dagger |\psi\rangle = |0\rangle$, therefore $\alpha_0 = 0$. But in that case $\hat{a}^\dagger |\psi\rangle$ cannot contain $|1\rangle$ since $|1\rangle = \hat{a}^\dagger |0\rangle$. Therefore, $\alpha_1 = 0$. And on, and on.

2. The recursion relation follows from

$$\hat{a}\,|s\rangle = \sum_{n,k\geq 0} \sqrt{n+1} c_k\, |n\rangle\, \langle n+1|\, k\rangle = \sum_{n\geq 0} \sqrt{n+1} c_{n+1}\, |n\rangle = \sum_{n\geq 0} s\, c_n\, |n\rangle$$

   The solution is checked by direct computation. The normalization of c_0 follows from the condition

$$1 \equiv \langle s|\, s\rangle = \sum_{n\geq 0} |c_n|^2 = \sum_{n\geq 0} \frac{|c_0|^2 |s|^{2n}}{n!} = |c_0|^2\, \exp(|s|^2)$$

3.

$$\begin{aligned}
\langle x\rangle &= \frac{1}{\sqrt{2}}\, \langle s|\, (\hat{a} + \hat{a}^\dagger)\, |s\rangle = \sqrt{2}\, \mathrm{Re}(s) \\
\langle p\rangle &= \frac{1}{i\sqrt{2}}\, \langle s|\, (\hat{a} - \hat{a}^\dagger)\, |s\rangle = \sqrt{2}\, \mathrm{Im}(s) \\
\langle x^2\rangle &= \frac{1}{2}\, \langle s|\, (\hat{a} + \hat{a}^\dagger)^2\, |s\rangle \\
&= \frac{1}{2}\, \langle s|\, \hat{a}^2 + (\hat{a}^\dagger)^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\, |s\rangle \\
&= \frac{1}{2}\, \langle s|\, \hat{a}^2 + (\hat{a}^\dagger)^2 + 2\hat{a}^\dagger\hat{a} + 1\, |s\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\left(s^2 + (s^*)^2 + 2s^*s + 1\right) = \langle x\rangle^2 + \frac{1}{2} \\
\langle p\rangle^2 &= \frac{-1}{2}\langle s|\,(\hat{a} - \hat{a}^\dagger)^2\,|s\rangle \\
&= \frac{-1}{2}\langle s|\,\hat{a}^2 + (\hat{a}^\dagger)^2 - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\,|s\rangle \\
&= \langle p\rangle^2 + \frac{1}{2}
\end{aligned}
$$

From these it follows that $\sigma(x) = \sigma(p) = 1/\sqrt{2}$ and the Heisenberg inequality is saturated.

4. The forms of $\hat{H}$ and $\hat{H}^2$ are trivial. With the same techniques as above we immediately find

$$
\langle E\rangle = \hbar\omega\left(|s|^2 + \frac{1}{2}\right)
$$

$$
\left\langle E^2\right\rangle = \hbar^2\omega^2\left(|s|^4 + 2|s|^2 + \frac{1}{4}\right)
$$

so that

$$
\sigma(E)^2 = \left\langle E^2\right\rangle - \langle E\rangle^2 = \hbar^2\omega^2|s|^2
$$

and

$$
\frac{\sigma(E)}{\langle E\rangle} = \frac{|s|}{|s|^2 + 1/2} \sim \frac{1}{|s|}
$$

5. In this example $\hbar\omega \approx 10^{-34}$ Joule. At the equilibrium position the particle has a mass of 5×10^{-8} Joule, so that $|s|$ is of the order of 10^{13}. The spread in energy is therefore better than one ppb.

6. The time-dependent form of the coefficients c_n is obtained by simply using Eq.(115). We may drop the overall complex phase factor $\exp(-i\omega t/2)$ since that cannot influence the physics ; and $\exp(-i\omega nt) = (\exp(-i\omega t))^n$. The replacement $s \to s\,exp(-i\omega t)$ is therefore justified. This means, of course that both $\mathrm{Re}(s)$ and $\mathrm{Im}(s)$ oscillate in the correct manner.

14.7 Excercises 53 to 61

Solution to Exercise 53

$$
\int dx|\Lambda(x)|^2 = \int_{-\infty}^{\infty} dx\; e^{-2\alpha|x|} = 2\int_0^{\infty} dx e^{-2\alpha x} = \frac{-1}{\alpha}\left[e^{-2\alpha x}\right]_{x=0}^{x=\infty} = \frac{1}{\alpha}
$$

Solution to Exercise 54

From the identity

$$\int dx\ \delta(x) = 1$$

we see that the dimensionality of the Dirac delta is the *inverse* of the dimensionality of its *argument*. Since $k\,\delta(x)$ must have dimension energy, k has dimension energy×length. We therefore have

$$\alpha = \frac{m\,k}{\hbar^2} = \frac{\text{kg Joule meter}}{\text{Joule}^2\ \text{sec}^2} = \frac{1}{\text{meter}}$$

$$E = k\,\frac{m\,k}{\hbar^2} = \text{Joule meter}\ \frac{1}{\text{meter}} = \text{Joule}$$

Solution to Exercise 55
This is a matter of straightforwardly applying the rules from the wave-function section. The normalized wave function is (see excercise 53)

$$\psi(x) = \sqrt{\alpha}\ \Lambda(x)$$

Since $\psi(x)$ is symmetric around zero, both $x\psi(x)$ and $\psi'(x)$ are *anti*symmetric, so that $\langle x\rangle = \langle p\rangle = 0$. For the higher moments we have

$$\begin{aligned}
\left\langle x^2\right\rangle &= 2\alpha \int_0^\infty dx\ x^2\ e^{-2\alpha\, x} = \frac{1}{2\alpha^2} \\
\left\langle p^2\right\rangle &= -\hbar^2 \int_{-\infty}^{\infty} dx\ \psi(x)^*\ \psi''(x) \\
&= -\hbar^2 \int_{-\infty}^{\infty} dx\ \psi(x)\ \left(\alpha^2\psi(x) - 2\alpha\delta(x)\psi(x)\right) \\
&= -\hbar^2\alpha^2 \int_{-\infty}^{\infty} dx\ \psi(x)^2 + 2\hbar^2\alpha \int_{-\infty}^{\infty} dx\ \psi(x)^2\,\delta(x) \\
&= -\hbar^2\alpha^2 + 2\hbar^2\alpha\psi(0)^2 = \hbar^2\alpha^2
\end{aligned}$$

We therefore find $\sigma(x)^2 = \left\langle x^2\right\rangle$, $\sigma(p)^2 = \left\langle p^2\right\rangle$, and $\sigma(x)\sigma(p) = \hbar/\sqrt{2}$, which is well above the lower bound $\hbar/2$.

Solution to Exercise 56
We take

$$\psi(x) = \Lambda(x) - \Lambda(x - r)$$

so that

$$\psi''(x) = \alpha^2\psi(x) - 2\alpha\delta(x) + 2\alpha\delta(x - r)$$

Also, we have

$$\psi(0) = 1 - e^{-\alpha r} \quad , \quad \psi(r) = -1 + e^{-\alpha r}$$

Since $\delta(x)\psi(x) = \delta(x)\psi(0)$ and $\delta(x-r)\psi(x) = \delta(x-r)\psi(r)$ we can write

$$\psi''(x) = \alpha^2\psi - 2\alpha\left(\delta(x) + \delta(x-r)\right)\frac{\psi(x)}{1-e^{-\alpha r}}$$

Multiplying by $-\hbar^2/2m$ and moving the potential term to the left-hand side we find the time-independent Schrödinger equation with Eq.(247).

Solution to Exercise 57

1. By differentiating to x we first find

   $$\psi'(x) = -\frac{N}{L}\,\frac{\sinh(x/L)}{\cosh(x/L)^2}$$

   and then

   $$\psi''(x) = \frac{N}{L^2}\,\frac{1}{\cosh(x/L)} - \frac{2N}{L^2}\,\frac{1}{\cosh(x/L)^3} \quad .$$

   This means that, multiplying by $-\hbar^2/(2m)$ we can write

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

   with the potential $V(x)$ and the energy eigenvalue E as given.

2. The normalization requirement can be written as

   $$1 = \int_{-\infty}^{\infty} \frac{N^2}{\cosh(x/L)^2}dx = N^2L\int_{-\infty}^{\infty}\frac{dy}{\cosh(y)^2} = 2N^2L \quad ,$$

   hence $N = 1/\sqrt{2L}$. Note that N must *anyhow* have the dimension of $L^{-1/2}$, the real 'effort' is the factor 2 ; and note the variable transformation $x = Ly$!

3. Since the wavefunction is symmetric and vanishes very fast for large $|x|$, the function $x|\psi(x)|^2$ is odd and integrates to zero ; hence $\langle x\rangle = 0$. Since $\langle x\rangle$ is therefore constant, $\langle p\rangle = m\,d\langle x\rangle/dt$ must also vanish.

4. For $\langle x^2\rangle$ the integral in straightforward :

   $$\langle x^2\rangle = N^2\int_{-\infty}^{\infty}\frac{x^2}{\cosh(x/L)^2}dx = N^2L^3\int_{-\infty}^{\infty}\frac{y^2\,dy}{\cosh(y)^2} = \frac{2N^2L^3}{3} = \frac{L^2}{3}$$

For $\langle p^2 \rangle$ we can use the fact that ψ is real, and partial integration, to see that

$$\langle p^2 \rangle = +\hbar^2 \int_{-\infty}^{\infty} \psi'(x)^2 \, dx = \frac{\hbar^2 N^2}{L^2} \int_{-\infty}^{\infty} \frac{\sinh(x/L)^2}{\cosh(x/L)^4} dx = \frac{\hbar^2 \pi^2}{12 L^2}$$

5. Since $\langle x \rangle = \langle p \rangle = 0$, we have

$$\sigma(x)^2 \sigma(p)^2 = \langle x^2 \rangle \langle p^2 \rangle = \frac{\hbar^2 \pi^2}{36}$$

hence $\sigma(x)\sigma(p) = \hbar\pi/6$ and this obeys Heisenberg since $\pi > 3$. Note that the length scale L drops out completely, as it should !

Solution to Exercise 58
The wave function

$$\psi(x) \propto e^{ikx}$$

obeys

$$-i\hbar \, \psi'(x) = \hbar \, k \, \psi(x)$$

and is therefore a momentum eigenstate with momentum $\hbar k$. The (kinetic) energy is therefore $E = \hbar^2 k^2 / 2m$, and the time-dependent form of the wave function is

$$\psi(x,t) \propto \exp\Big(-i(Kt - kx) \Big) \quad , \quad K = E/\hbar = \hbar k^2 / 2m$$

We have here a wave profile. As t increases, Kt increases and the same point of the wave profile (*i.e.* the same value for $Kt - kx$ is then found at a *larger* value of kx. If $k > 0$, this implies a *larger* value of x, so that the wave profile moves towards larger and larger x values as time goes on. Similarly, for $k < 0$ the time evolution moves the profile to smaller values of x.

Solution to Exercise ??
Outside the barrier, the solutions to the time-independent Schrödinger equation must be complex exponentials (or 'sines and cosines'). To the left of the barrier we allow for incoming and reflected waves : by assumption, we accept only outgoing wave to the right of the barrier. The fact that there we have no normalizing parameter reflects the fact that we do not care about the overall normalization so we can take *one* of the normalizing factor to unity. Inside the barrier the solution to the time-independent Schrödinger equation must be *real* exponentials. Insertion of the Ansätze into the Schrödinger equation proves the values of κ and k.

Solution to Exercise ??
This is a matter of insertion of the solutions into the equations, and then verifying them. Note : often students will assume that they have to *solve* the

equations themselves, and arrive at the given answer. This is of course very commendable, but *not* necessary.

Solution to Exercise 60

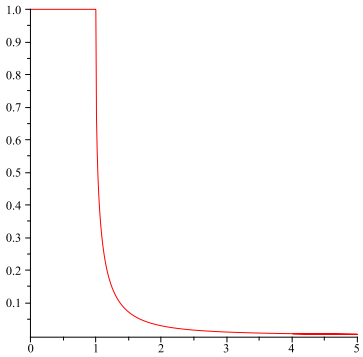
1.
$$x \leq 0 \quad : \quad -\frac{\hbar^2}{2m}\psi"(x) = E\,\psi(x)$$
$$x > 0 \quad : \quad -\frac{\hbar^2}{2m}\psi"(x) = (E-W)\,\psi(x)$$

2. For $x < 0$ the general form is
$$\psi(x) = A\,\exp(ikx) + B\,\exp(-ikx) \quad , \quad k = \sqrt{\frac{2mE}{\hbar^2}}$$
For $x > 0$ and $E > W$, we can choose
$$\psi(x) = \exp(i\hat{k}x) \quad , \quad \hat{k} = \sqrt{\frac{2m(E-W)}{\hbar^2}}$$
There is no term with $\exp(-i\hat{k}x)$ since there is no left-moving wave there. For $E < W$, we can take
$$\psi(x) = \exp(-\kappa x) \quad , \quad \kappa = \sqrt{\frac{2m(W-E)}{\hbar^2}}$$
No term with $\exp(\kappa x)$ can occur since then the wave function would not be normalizeable.

3. The wave function must always be continuous ; its derivative can only be discontinuous if there is an *infinite* potential step, but W is finite.

4. The continuity of ψ implies $A + B = 1$ in both cases. The continuity of ψ' implies $k(A-B) = \hat{k}$ for $E > W$ and $k(A-B) = -i\kappa$ for $E < W$.

5. In both cases $R = |B|^2/|A|^2$. If $E < W$, then we have $A = (1 + i\kappa/k)/2$ and $B = (1 - i\kappa/k)/2$ so that $|A| = |B|$ and $R = 1$. For $E > W$ we have $A = (1 + \hat{k}/k)/2$ and $B = (1 - \hat{k}/k)/2$ so that
$$R = \left(\frac{k-\hat{k}}{k+\hat{k}}\right)^2 = \left(\frac{\sqrt{E}-\sqrt{E-W}}{\sqrt{E}+\sqrt{E-W}}\right)^2 = \left(\frac{\sqrt{u}-\sqrt{u-1}}{\sqrt{u}+\sqrt{u-1}}\right)^2$$
and this depends, indeed, only on $u = E/W > 1$. The case $E = W$ is difficult to analyse. That is why we consider either $E < W$ or $E > W$.

6. 

At $u \leftarrow 1$, the slope of R is $-\infty$; for large u it approaches $u^{-2}/16$.

7. If $W < 0$ then $E < W$ would imply $E < 0$ which would make an incoming wave from the left an impossibility. For the rest, it suffices to take the case $E > W$ and replace $E - W$ by $E + |W|$. For instance, in the result for R, $\sqrt{u-1}$ is then replaced by $\sqrt{u+1}$.

8. In the classical analogue, R would always be either 1 (for $0 < E < W$) or 0 (in the other cases).

Solution to Exercise 59
Since the wave function is continuous, the various intervals may indeed overlap at exactly the points they meet — even if it would make mathematicians sad.

Solution to Exercise 61
A matter of inserting and working out.