16 First steps into Quantum Theory

A spaceship sinks slowly toward the surface of an unknown planet. Robots are sent out to collect on-site experimental data. People in the hull are discussing the incredible findings out there, they are coining concepts and thinking in images. Images, pictures, yes! They try to establish quantitative relations. Some equations are central, others lose importance as they result from the first ones. Five months later the theory "stands" and another five months later it has been tested by prediction and experimental validation of numerous processes.

In the following pages we attempt such an exploratory adventure. We cut down on the months above. Instead of explaining experiments in detail, we will start from their idealization and focus on their gist. And throughout we will draw the correct conclusions, introduce concepts and use pictures that work! We shall be ingenious inventors. On our way three hazards await us.

1. Dangerous copycat criminals. No guys, here we are only pretending to be geniuses. For all these apparently bold statements we already know they can be calculated. Quantum theory is 90 years old. It is correct. It is correct because it is in agreement with *all* experiments. Even today's theory of elementary particles is right (here we simply stick to the non-relativistic version). Do physicists adhere to a dogma? No: who ever finds a single reproducible anomaly receives the Nobel price. This is the unity of physics (the consistency of *the* mathematics which nature obeys) that "Einstein opponents" like to forget sometimes.

2. The dangers of impatience. It will take a while. Our new inventions need a mathematical formulation. The corresponding postulates must not contradict each other, should not overlap and they must not predict anything that cannot be measured (unfortunately the unnecessary "ideal measurement" postulate is to be found in certain textbooks). The new theory must explain why the old one works so well. Once its applications to a variety of situations work out (they are analytically possible, make sense and are verified by experiments) and do so repeatedly (reproducibility), then we gain confidence: this is how the world is.

3. The danger of persisting wrong education. "Quantum physics is the realm of models." said someone wearing a crooked smile. "You can think like this or like that. Here physicists are only describing. Now this des-

> cription changes again." Incredible! Where does this come from? Trashy novels, some schoolbooks that is. Novels and pictures may be excused as long as they are related to reality. But the "atomic model" belongs to the past. The Thomson model is wrong. The Bohr–Sommerfeld theory is false. Enumeration of historic mistakes miseducates, distracts and wastes precious time.

If relativity had shaken our representation of the world, well then quantum theory is an earthquake. It is caused by the following four words.

Particles are spread out

Our first finding is based on a wealth of experiments. Instead of choosing one, let us stress that in the end, *all* experiments show this extension property as long as one looks closely enough.

Barely has he grasped the above four words, does the human analog computer already call out a number of images. An object that is spread out, must have a shape. Maybe it is a spherical ball or a spherical shell, or a car tire, or a cloud or a mouse. Indeed, we will see that a talented experimentalist will be able to qualitatively imprint any desired shape on a particle. One thing, however, is impossible and that is to think of the shape as consisting of several particles. Overruled. *One* particle, *one* electron is spread out. Well then it should be a function on space, a field. The function shall be called $\psi(\vec{r}, t)$. At time t an electron is in state ψ.

If particles are indeed spread out, this means the end of the tale of point masses. Mechanics predicts the motion $\vec{r}(t)$ of points. But it cannot tell us anything about the evolution of the shape of a particle. Mechanics is wrong. Even conventional electrodynamics is wrong. However in our discussion we will mainly try to amend the "first part of the theory". This is what we call quantum mechanics. As in Mechanics we will assume that forces are known.

Figure 16–1. A mouse hiding under the carpet. Middle: Two level curves as seen from above. Right: The "mouse–particle" from 100 meters height.

How much or how little mechanics is wrong, will depend on the distance as shown in Figure 16–1. From a distance of 100 meters a mouse will look like a moving point. This yields the first explanation for the lack of evidence of the size of particles at our every day scale. Typically we are too far away from

them. In this limiting case quantum mechanics should reproduce the Newtonian equations of motion (see (16.43) below). For which $\vec{r}(t)$? — the center $*$ of the particle.

A second reason why we have not seen any evidence of particle interference or a particle rainbow is that the stones which we throw all contain roughly 10^{23} particles. They are bound together and we may consider the stone itself to be one particle with 10^{23} times larger mass. How the smaller spread (of the center) is related to an increased mass can be shown (but it won't fit into this introduction). Individual free particles slip through our eyes.

The aforementioned analog computer also recalls other pictures, those in the category "field". An electric field is also spread out over space. Such as the Coulomb field that sits in radial direction around a charge in the vacuum or the closed field lines around a magnet. Fields in space are out there! We know everything about electromagnetic fields. It all fits into the Maxwell's equations. But we know nothing yet about the field ψ. Quantum mechanics is so new and different, because it makes us ask so many questions right from the start. Let us number them for later reference:

1. Does ψ have inner structure? (it could be a vector like $\vec{E}$)
2. Does ψ evolve over time on its own?
3. Is there a state of rest and which shape does it have? (16.1)
4. How does the field react to exterior potentials?
5. What happens when an extended electron reaches a counting tube?

Questions two through four ask for an equation of motion for ψ. Would we not have asked those questions even if we had not known anything about electromagnetic fields?! We can no longer put off looking at experiments.

Experiment 1: a sleeping electron

Since about 1920 we have known that the heavy nucleus only occupies a small fraction of the volume of an atom. Think of a football field (electron) and a pea (nucleus). Moreover almost the entire mass of the atom is concentrated in the nucleus: proton mass = 1836 times electron mass. Rutherford's heavy α particles (2 protons, 2 neutrons) would barely have noticed the light envelope and happily diffracted from the nearly point-like gold nuclei.

The experiment, which we are going to investigate in highly idealized form (please neglect the Earth's gravitational pull) consists of a hydrogen atom (proton and electron) placed carefully in the middle of a big vacuum tube. Then we sit back and wait patiently. Experimental evidence: nothing. NOTHING happens!!! It does **not** radiate! The electron is sleeping. Its field is completely motionless.

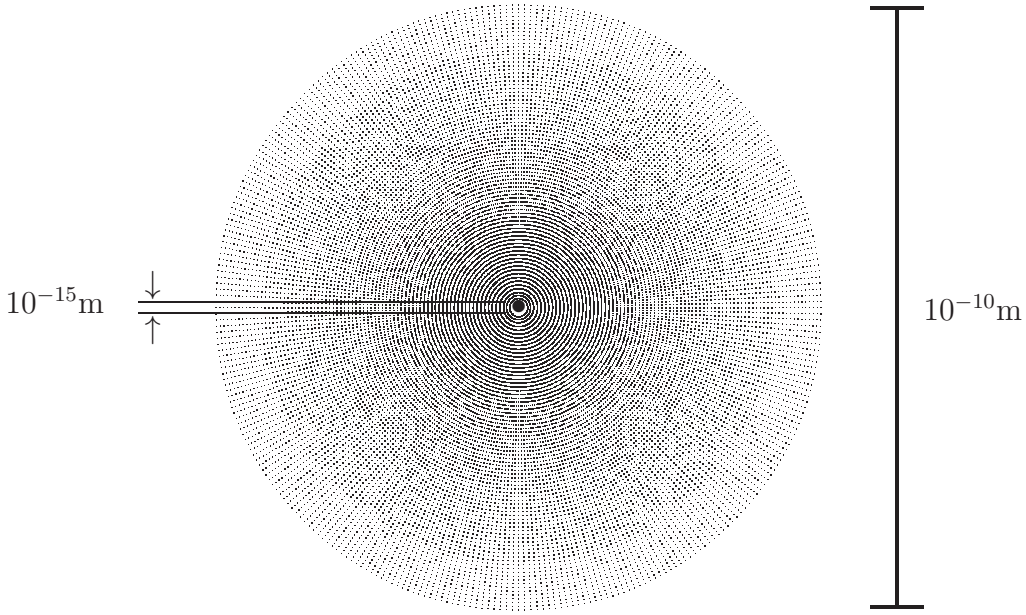

Figure 16–2 . The ground state of an H atom.

Were the field changing in time, then its charge most probably would act as antenna and emit electromagnetic waves. Now we are dangerously close to the "models", we were bashing earlier. Apart from those there are of course the insignia of Atomic Energy Commissions, which dutifully show electrons circling the nucleus. Well these too leave no doubt that the full blow of quantum mechanics has not reached their marketing departments.

But who would trust the scribblers working here? Let us see then what [*Landau/Lifshitz*] have to tell us about it in their 3rd volume, section 1, first paragraph :

> ... a model of an atom in which the electrons move round the nucleus in classical orbits. During such motion, as in any accelerated motion of charges, the electrons would have to emit electromagnetic waves continually. By this emission, the electrons would lose their energy, and this would eventually cause them to fall into the nucleus. Thus, according to classical electrodynamics, the atom would be unstable, which does not at all agree with reality.

The electron in Figure 16–2 is in the **ground state**. The H atom had to be placed carefully in empty space to avoid waking the electron (excited H atoms do radiate). Also, Rutherford's gold atoms, including their heavy nuclei, were sleeping soundly. We have been dishonest concerning the proton. Without further ado we had treated it as a classical small ball. However it is legitimate to proceed in steps on our way to quantum physics : heavy particles last. The fact that the atom as a whole is a field shall be left for later.

To sum it up, there are **stationary states**, those which do not change in time

and do not radiate. How the sleeping electron "looks" depends on its bedding, that is on the shape of the potential. Here the electron was sleeping in a bed having the shape of the Coulomb potential of the proton. In the old mechanics, statics means solving $0 = m\ddot{\vec{r}} = -\mathrm{grad}\, V(\vec{r})$, i. e. to search for the minimum of V. Quantum statics is a lot more complicated, since we need to solve a partial differential equation to determine the shape of a sleeping electron. But that was the same for electromagnetic fields: the arbitrarily difficult problems in electrostatics and magnetostatics are also of the shape determination kind. Nevertheless statics is merely a special case (of any theory). In nature (i. e. in physics) things are moving. Quantum dynamics is the main and much more interesting (but more difficult) matter.

The proton carries a positive elementary charge e. The atom on the whole is electrically neutral. It is tempting to think of the electron cloud in Figure 16–2 as a charge density $\rho(\vec{r})$ with $\int \rho = -e$. To find any evidence for this one has to look into the details of measurements (see below) or one has to include other particles into the quantum mechanical description of the system. It would be fatal to consider one part of the cloud interacting with another part and try to calculate its electrostatic energy. That will be wrong.[1] One can talk about a charge probability density (see later). But "cloud" isn't half bad, it is made of completely new physics.

Experiment 2: an electron splits

Humans are extended, but we are unable to leave a leg at our grandma's and our head at the university. To cut onions with a knife, that works. Electrons can tear up without tears, and a light field can do it too. In 1927 Thomson shot single electrons through a foil: the atomic nuclei were the knives. Davisson and Germer let electrons collide obliquely with a washboard (the surface of a crystal is like an atomic washboard). From this experiment we can draw all necessary conclusions, even quantitative ones.

First of all, in order to be able to explain Figure 16–3, we need a realization that is best deduced by calculation from the Schrödinger equation. If a particle that does not feel any potentials, then its field will extend, the wave packet will melt away. After being ejected from the heated filament and accelerated to a certain (known) energy electron will be rather localized. But when it reaches the crystal "washboard" the cloud of the single electron is already widely extended.

[1] Anticipation. The energy of the cloud is $\int \psi^* H \psi = -1\,\text{Ry}$ ($1\,\text{Ry} = 13.6$ eV) and is the sum of kinetic $\int \psi^* T \psi = 1\,\text{Ry}$ and potential energy $\int |\psi|^2 V = -2\,\text{Ry}$. These go as ψ^2. But the electrostatic energy of the a charge density $\varrho = -e|\psi|^2$ would go as $|\psi|^4$ and give $-11/8\,\text{Ry}$ (after subtracting the proton's contribution). — By the way let Figure 16–2 show ψ and *not* $|\psi|^2$. Once we know more about its internal structure, we can add the ψ–phase, for instance by coloring the diagram at each point.

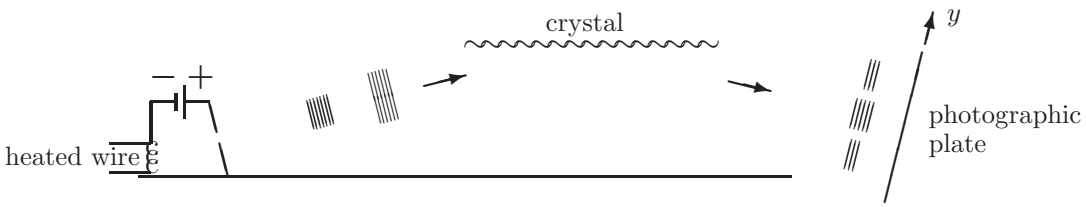

Figure 16–3 . A diffracted electron interfering with itself.

Upon leaving the crystal the electron acts like a water wave that is reflected on corrugated metal. The partial waves of the electron interfere. When it reaches the photographic plate the electron is in a camel–backed state. We may adjust this experiment to our needs. In Figure 16–3 the photographic plate replaces already the Geiger–Müller counter, used by Davisson and Germer at varying y–positions. Instead of the crystal lattice we rather think of a periodic potential. It is zero at the bottom of the wavy line and increases towards the top. Imagine that it does not vary in the z–direction (perpendicular to the page). This makes the problem two–dimensional. Do you expect stripes to appear on the plate ? Wait a minute. We shall postpone the discussion of the atrocious things that the horrible photographic grains are doing to our poor electron, until the section after the next one.

The structure of the particle field : $\psi \in \mathbb{C}$

If parts of the electron field can be superposed (added) to produce interference, then this means that the field is a wave in some sense. A wave is something that has a periodic dependence on $\vec{k}\,\vec{r} - \omega t$. Let us recall what sound and light are. Sound waves can interfere.

$$\vec{k}_1 = (\kappa, \eta, 0)\ ,\ \ \vec{k}_2 = (\kappa, -\eta, 0)\ ,\ \ \frac{\vec{k}_1 + \vec{k}_2}{2} = (\kappa, 0, 0)\ ,\ \ \frac{\vec{k}_1 - \vec{k}_2}{2} = (0, \eta, 0)\ \ ,$$

$$\cos\left(\vec{k}_1\vec{r} - \omega t\right) + \cos\left(\vec{k}_2\vec{r} - \omega t\right) \ = \ 2\,\cos\left(\kappa x - \omega t\right)\,\cos\left(\eta y\right) \quad . \tag{16.2}$$

The superposition of two tilted waves propagating toward the right at different angles, yields a wave going to the right with y-modulation. To the left of the tilting event the air density is assumed to be a plane wave $\sim \cos(\vec{k}\,\vec{r} - \omega t)$. At this point we run into trouble, because for a single electron field ψ flying toward the right in the direction of $\vec{k}$ such "waviness before" has never been observed. The following line ($c := \cos(\vec{k}\,\vec{r} - \omega t),\ s := \sin(\vec{k}\,\vec{r} - \omega t)$) shows us how this

can be:[2]

$$1 = c^2 + s^2 = |c + \mathrm{i}\, s|^2 = |\, \mathrm{e}^{\mathrm{i}\,(\,\vec{k}\,\vec{r} - \omega t\,)}\,|^2 = |\psi(\vec{r}, t)|^2 \quad . \tag{16.3}$$

The phase "waves", but the charge(–probability) density $-e|\psi|^2$ does not. From now on we stick to a general particle field that takes complex values: $\psi = |\psi| * \exp(\mathrm{i}\,\mathrm{phase})$. Should we allow the imaginary unit i to appear in theoretical physics? Why not![3] However it must not be contained in measurable quantities. By the way, electromagnetic fields can also be combined to $\vec{\psi} = \vec{E} + \mathrm{i}c\vec{B}$. We have now answered question 1 in (16.1).

Now we are bursting with curiosity to find out if and how these tilted *e*–to–the–i–waves interfere. For the two $\vec{k}$ vectors given in (16.2) we get

$$\mathrm{e}^{\mathrm{i}\vec{k}_1\vec{r} - \mathrm{i}\omega t} + \mathrm{e}^{\mathrm{i}\vec{k}_2\vec{r} - \mathrm{i}\omega t} = \mathrm{e}^{\mathrm{i}\kappa x - \mathrm{i}\omega t + \mathrm{i}\eta y} + \mathrm{e}^{\mathrm{i}\kappa x - \mathrm{i}\omega t - \mathrm{i}\eta y} = 2\,\mathrm{e}^{\mathrm{i}(\kappa x - \omega t)}\,\cos(\eta y) \quad . \tag{16.4}$$

Excellent. Due to the cosine modulation the field has zeros. The spacing between its zeros is π/η. Hence we can measure η. The ideal situation to which (16.4) applies is commonly called the **double slit experiment**. Consider two parallel lines (separation a) on a plane parallel to a screen (distance $\ell \gg a$). The two superposed waves could have been emitted from these two parallel lines and would become almost plane later on, far away from the plane (you can draw it!). The tangent of the tilting angle (with respect to the horizontal) is given on the one hand by the two components of $\vec{k}_1$ and on the other hand by the two lengths defined above. Hence $\eta/\kappa = (a/2)/\ell$. If we know a then $k = |\,\vec{k}\,| = \sqrt{\kappa^2 + \eta^2}$ is known as well. Nothing new happens when returning to the Davisson–Germer experiment, it is the passage to many "slits" on the washboard in Figure 16–3. In brief, from the lattice parameter of the crystal and the image on the screen we obtain $k = |\,\vec{k}\,|$. k is also $|\,\vec{k}\,|$ before the crystal is reached.

We know the energy E received by the particle during the acceleration phase in Figure 16–3, there it is an almost point–like (still classical) particle. $E = p^2/(2m)$ tells us about its linear momentum $\vec{p}$. What if we produced a variety of different E and made a series of measurements of $\vec{p}$ and $\vec{k}$? The experiment yields

$$\vec{p} \;=\; \hbar\,\vec{k} \quad , \qquad \hbar = 1.05\,\ldots \cdot 10^{-34}\mathrm{Js} \quad . \tag{16.5}$$

The higher the energy of the particle, the shorter its wavelength $\lambda = 2\pi/k$. Therefore people at accelerator facilities call for higher energies, because they want to "see" ever smaller objects in this "particle light". The relation (16.5)

[2] When a complex number z is written as $z = r\,\mathrm{e}^{\mathrm{i}\,\varphi}$ we call $\mathrm{e}^{\mathrm{i}\,\varphi}$ the "phase factor" and φ the "phase angle" or "phase" for brevity. $r = |z|$ is the "modulus".

[3] Aside, one can return to purely real quantities by writing i as a matrix: $\mathrm{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then $\psi = \Re e\,\psi + \mathrm{i}\,\Im m\,\psi = \begin{pmatrix} \Re e\,\psi & -\Im m\,\psi \\ \Im m\,\psi & \Re e\,\psi \end{pmatrix}$. Mostly this is rather unwieldy though.

is only valid for particles in some $e^{\mathrm{i}\vec{k}\,\vec{r}}$ state. We shall see that exactly in this case its linear momentum $\vec{p}$ is well defined (with value $\hbar\,\vec{k}$).

In quantum mechanical processes it is important to define the **system** properly, occasionally this can be rather tricky business. The system is composed of the particles to be treated quantum mechanically, potentials (classical forces) and if need be initial conditions. The measurement device is $n\,o\,t$ to be included, it usually inflicts nasty treatment to particles (to be discussed soon). The system occupies a state ψ , which in general evolves over time. Even you are a system, and (be assured) your state will change. **Superposition** means that the pieces of a particle's field can simply be added, even if they overlap. The common saying that for two given states ψ_1 and ψ_2 , $\psi = \alpha\psi_1 + \beta\psi_2$ is also a state is a bit dangerous. Let us be clear, if ψ_1 and ψ_2 are possible states of a system then their linear combination is also a possible state, as long as it is **normalized**. Roughly, by normalization the requirement $\int \psi^*\psi = 1$ is meant. This is further discussed after (16.8), and in (16.30) and (16.31).

Question $\neq$ answer

We have to keep our promise and come back to Figure 16–3 . On the right an electron in cloudy tatters is whizzing along quite happily. Altogether this shredded cloud carriesthe elementary charge $-e$. Next it encounters a screen made of photographic grains, each of which is asking the rather unpleasant question "Electron, are you here ?" To blacken a grain needs the whole charge $-e$. Yes or No! We ought to feel sorry for the poor electron. It is neither here, nor there, but extends all over. It cannot talk its way out of it and explain its state. Well, it has to answer some grain, so that at least one of them is satisfied. Question $\neq$ answer. The question A corresponds to a macroscopic apparatus and this in turn will be associated with an operator A . But the answer is the result of a measurement and therefore a solid real number a . Actually the grain raises two questions : "Are you at my x–coordinate ?" and "Are you at my y–coordinate ?". But we may simplify things and let the electron be projected onto the plane of Figure 16–3 . Then the screen is a line and a is a y-value.

The source in the Davisson–Germer experiment shall send out one electron per minute. We started the experiment in the evening, but became impatient after ten minutes. So, we replaced the photographic plate by a new one and developed the old one. It shows 10 randomly distributed points. Crestfallen, we had gone to bed. But when we awake the next morning something amazing has happened. 500 points are organized into parallel dark stripes. Now we can read the value of η — or wait until next week to improve the precision. This is a first answer to question 5 in (16.1).

To sum up every subfield of physics in a few words, that is a worthwhile effort. For quantum mechanics we are now able to say "what it is", at least roughly. A

first turning point is reached along our journey into quantum land. The following is on the one hand a summary of what we know so far. On the other hand (16.6) incorporates a few new points. So, it is also an outlook, an appetizer, a longish heading. If the formulae are disregarded for the moment, then it is

Quantum mechanics in words :

(16.6)

A quantum mechanical system is in a state. An equation of motion, containing the given potentials, governs the time evolution of the state, but also allows to deduce its possible static forms.	$\psi(\text{Variable}\,, t) \in \mathbb{C}$ $\mathrm{i}\,\hbar\,\dot{\psi} = H\,\psi$ $H\,\varphi = E\varphi$
The question (classical macroscopic measurement device, observable A) is generally distinct from the answer (value a). For such questions (except those "fitting" to the state) one can only determine (calculate and measure by repeating the same experiment over and over again) the probability for obtaining a certain real value a .	$A\,\varphi_a = a\,\varphi_a$ $P_a = \lvert \int \varphi_a^* \psi \rvert^2$

Obviously the above propositions are organized into two very different structures. The first part (above the horizontal line) tells us how the future of a field is uniquely and complete prescribed, sort of like electrodynamics. Why is this not enough ?! — this question had been suppressed for over 50 years, but recently it has attracted attention again. Below the line we have THE mystery of quantum physics. Here physicists are divided into two camps. The first ones say that this is how it is, full stop (Copenhagen interpretation). For them, those who are still looking for "hidden parameters" have not understood a thing. The other camp suspects that the equation of motion also describes the behavior of the measurement device as a many particle system (Everett postulate). According to John Stewart Bell, the bottom half is only valid *for all practical purposes* (FAPP). For more on this topic see the quotes at the end of the chapter.

Curiously, the overwhelming success of quantum mechanics, in a long standing series of agreements between experiments and analytical results, is founded almost entirely on the first half, i. e. on solving the Schrödinger equation : atomic spectra, chemical bonds, electron bands in solids, tunneling, phonons, magnetism, statistical physics, nulcei and many more. Whether it is an auxiliary assumption or not, the lower half is going to help us to translate our first formulation of quantum mechanics into mathematical formulae.

Measured values are the eigenvalues of A

It is becoming unavoidable, to make inventions and only test them later. In the lower half of (16.6) we have mentioned special questions, but there are also many normal ones. Now we have the following idea : if we focus an electron to $y = a$

on the photographic plate by means of potentials, then this leads to the grain at $y = a$ blackening with certainty. Measurements of y yield a with probability 1. So the state and the question can fit. If it is compatible we call the state φ instead of ψ. In case a high school lecture gets this far, we encourage the teacher to adorn his course with all kinds of examples from every day life. Oh, student X has fallen into his ground state: no agitation what so ever. His neighbor Y is being noise. "Would you please adopt an eigenstate for my question?!". Every day life can be manipulated, for instance if the interviewee may only operate a device with Yes and No buttons. Or three buttons or a keyboard. The tatters of cloud in Figure 16–3 have a keyboard with infinitely many infinitesimally thin buttons at their disposal.

Let an operator A be given. We alter the state ψ (in our minds and experiments) until one of them fits. Which means that a repeated measurement of A produces a single value a. This state φ_a is called the **eigenstate of A for the measured value** a. Now let us see if we can remember a structure which associates values and corresponding functions to a given operator. Of course this is familiar. It is an eigenvalue problem. $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has eigenvalues $+1$ and -1, and their respective eigenvectors are $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$. For functional spaces we know for instance that $e^{\mathrm{i}\vec{k}\vec{r}}$ are the eigenfunctions of the Laplace operator Δ with eigenvalues $-k^2$ (see Fourier transform and diffusion equation).

Tentatively, we postulate that, in the general case, the possible values of a measurement of A are obtained by

$$\overbrace{A\,\varphi_{a\nu}(x)}^{\downarrow} = a\;\varphi_{a\nu}(x) \quad . \tag{16.7}$$

The arrow indicates on what A is acting upon (which dependency). The indices a, ν can only be attached after solving $A\varphi = a\varphi$. The index $\nu = 1...\eta$ is only needed if there is more than one solution for the eigenvalue a. η is the **degree of degeneracy** of a. Only normalizable functions are acceptable solutions and distinct solutions are to be orthogonal to the previous solutions. In the case when A is a matrix, all of this is well known from chapter 4 (main diagonal transformation). The variable x reminds us of the coordinate in 1D. However, it is highly convenient that we can also think of x as the index of vector components ($A\varphi_j = A_{jk}\varphi_k$) or as $\vec{r}$ in 3D (or as a set of variables $\vec{r}\,\sigma$ for a spinning particle).

Of course, eigenvectors with a finite number of components are always normalizable. In this case we have a discrete set of eigenvalues. Given an infinite number of components however, normalizability is a genuine (non trivial) condition. It excludes non-normalizable vectors. For eigenfunctions (with continuous component index) this holds as well. Eigenvalues may occur in any possibly imaginable way on the a axis. They can be discrete (angular momentum), be continuous on one half of the line (energy of a free particle), have holes (energy

bands) or fill the whole axis (position and linear momentum). The distribution of the eigenvalues along an axis is called the **spectrum** of the operator A.

Probability = modulus squared

In this subsection, we take it easy and consider only operators A, whose eigenvalues a are discrete and non–degenerate. We suppose that we have solved $A\varphi_a = a\varphi_a$, meaning that all a, φ_a are known. The probabilities of throwing dice have been discussed in chapter 14. If the system is an eigenstate $\psi = \varphi_a$ of A, then the probability of measuring a is 1. The case $\psi \neq \varphi_a$ is more interesting (ψ is none of the φ_a). For the Davisson–Germer experiment the stripes on the plate are darker where more of the cloud hits. Somehow the probability for a should depend on the "component" of ψ along the "direction" of φ_a. The component — for a vector that would mean its projection onto the basis vectors. Perhaps we should expand ψ along the A eigenfunctions φ_a. The latter form a CONS (complete orthonormal system), as we already know for the vectorial case at least. Generally, the a–th coefficient of this expansion will be complex and thus cannot be directly identified with the probability P_a. Moreover the sum of all P_a has to be 1. The coefficient's square modulus might work out though. Let's see.

$$\psi \;=\; \sum_a c_a\,\varphi_a \quad , \qquad P_a = |c_a|^2 \;\; . \tag{16.8}$$

If we define a scalar product "$\bullet$" on the functional space, then we could orthonormalize the basis vectors with $\varphi_b^* \bullet \varphi_a = \delta_{ba}$ *and also* normalize the state ψ according to $\psi^* \bullet \psi = 1$, then (16.8) actually works, because

$$1 = \psi^* \bullet \psi = \sum_{b,a} c_b^* c_a \varphi_b^* \bullet \varphi_a = \sum_a |c_a|^2 = \sum_a P_a \;\; . \tag{16.9}$$

But of course, nature has the last word.

Experiment 3: a photon

Newton supposedly thought, that light is a ray of small grains. Probably, people of the 19th century were amused by this, because for them it surely is a field! Nowadays we know that both electrons and photons are fields, of which fixed portions stay together during a measurement. Photons and electrons are close relatives.

This is something truly beautiful which has been unearthed by quantum measurements: it is a *unification*. There are only fields, that is to say quantum fields. The electron always arrives at the counting tube with its whole charge $-e$. The photon is an electromagnetic (e.m.) wave of frequency ω and always arrives with its whole energy $\hbar\omega$. It should simply be ascribed to the whims

of human history, that we have arrived at this conclusion from two completely different perspectives, from the idea of grains for particles and for light quanta from the picture of classical fields.

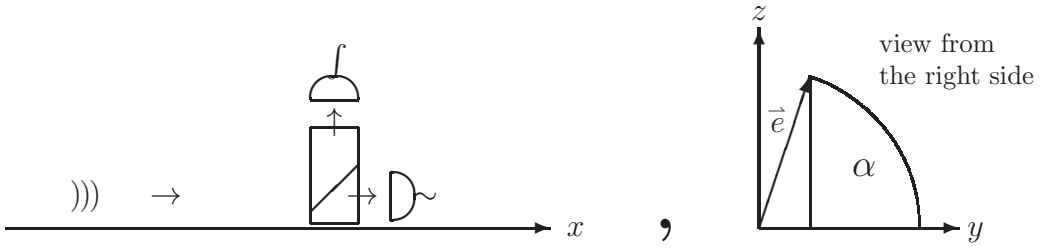

Figure 16–4 . How to coerce a photon to answer

We will only briefly venture into this topic, because the polarization of e.m. waves very nicely allows us to verify our rather audacious invention (16.8). In Figure 16–4 an e.m. wave is flying toward the right (x–axis) into a Nicol prism. It is polarized along a given direction $\vec{e}$. This is easy to achieve experimentally. $\vec{e}$ lies in the yz–plane. In the y and z directions the wave is confined by appropriate boundary conditions. The prism splits the beam into one part polarized along z and another part polarized along y (i. e. perpendicular to the figure plane). The first one goes through the prism, and the second one is directed toward the top. Let the incoming wave have exactly the energy $\hbar\omega$ (see field energy in chapter 11). Who are the guys asking those unfitting questions ? They are *photomultiplier tubes*, who are only satisfied if they get a whole piece of $\hbar\omega$. If the counter on the right gives a signal, then the one on top does not and conversely.

As for the Davisson–Germer experiment, it becomes interesting only after 500 fold reiteration. The unit vector $\vec{e} = (\,0\,,\,\cos(\alpha)\,,\,\sin(\alpha)\,)$ is known. Now we realize that in the case at hand, the expansion (16.8) has only two coefficients. They are $c_{\rm top} = \cos(\alpha)$ and $c_{\rm right} = \sin(\alpha)$. Can you guess what the experimental evidence shows ?

$$P_{\rm top} = \cos^2(\alpha) \;\;,\quad P_{\rm right} = \sin^2(\alpha) \;\;,\quad P_{\rm top} + P_{\rm right} \;=\; 1 \;\;. \tag{16.10}$$

Excellent! So our investigation of the behavior of photons validates (16.8). However, for those who see this for the first time, the claim "electrons are like photons" could be a tough monster. The same is valid for the temptation to begin with the spin variable of the electron. Indeed, we are at odds with the typical introductory literature on quantum mechanics. There you can also find a lot of philosophical, historical and long discussions of the uncertainty principle. We already know it from chapter 12 (Fourier): $\Delta k\,\Delta x \,\geq\, \frac{1}{2}$. So what ? As an inequality it is of little use for calculations in both electrodynamics and quantum mechanics.

"Derivation" of the Schrödinger Equation

The equation of motion of quantum mechanics can not be deduced directly (where would you start?). There is however an astonishing heuristic way. We combine six ideas and they will lead us to Schrödinger's equation.

1st We already know some equations of motion for fields (the diffusion equation, the second and the fourth of Maxwell's equations). These only contain the first derivative with respect to time. Hence the value of a field at time $t = 0$ completely determines its future. A snapshot is enough. We expect this to hold for ψ as well (here we consider a single particle). So we write

$$\psi(\vec{r}, dt) = \psi(\vec{r}, 0) + dt \cdot //// = \psi(\vec{r}, 0) + dt\, \mathcal{H}\, \psi(\vec{r}, 0) \quad . \tag{16.11}$$

//// is a function of $\vec{r}$, which we should be able to obtain from $\psi(\vec{r}, 0)$. This is why we can write on the right hand side of (16.11) that //// results from applying an operator to $\psi(\vec{r}, 0)$. $\mathcal{H}$ acts on the $\vec{r}$ dependence.

2nd Interference between pieces ψ_1, ψ_2 of a particle field tells us that each piece has its own future: $\psi_j(\vec{r}, dt) = \psi_j(\vec{r}, 0) + dt\, \mathcal{H}\, \psi_j(\vec{r}, 0)$ $(j = 1, 2)$. Since $\psi = \psi_1 + \psi_2$ the relation $\mathcal{H}\psi = \mathcal{H}\psi_1 + \mathcal{H}\psi_2$ must also be valid. Thus $\mathcal{H}$ must be a linear operator. In (16.11) we move the term $\psi(\vec{r}, 0)$ to the left and divide by dt. Hence

$$\mathrm{i}\,\hbar\,\dot{\psi} = \mathcal{H}\,\psi \quad , \quad \mathcal{H} \text{ is a linear operator} \quad . \tag{16.12}$$

ψ has complex values, as does the unknown $\mathcal{H}$. Due to this we may include the factor $\mathrm{i}\hbar$ on the left hand side, which means that we have redefined H a bit.

3rd The answer predicted by the equation of motion also depends on the question. Inspired by the picture of the hydrogen atom in Figure 16–2, we ask (just the same as for the diffusion equation) whether there are solutions to (16.12), whose shape does not change. To keep its $\vec{r}$ dependence over time, such a field ψ must be the product of two functions, only one of which can be a function of time (*separation ansatz*). The product *ansatz* inserted into (16.12) gives

$$\begin{aligned} \psi(\vec{r}, t) &= \chi(t)\,\varphi(\vec{r}) \quad \curvearrowright \quad \mathrm{i}\hbar\dot{\chi}\varphi = \chi\,\mathcal{H}\varphi \quad , \quad \frac{\mathrm{i}\hbar\dot{\chi}(t)}{\chi(t)} = \frac{\mathcal{H}\varphi(\vec{r})}{\varphi(\vec{r})} =: \mathcal{E} \\ \mathrm{i}\hbar\dot{\chi} &= \mathcal{E}\chi \quad \left(\curvearrowright\ \chi = e^{-\frac{\mathrm{i}}{\hbar}\mathcal{E}t}\right) \quad , \quad \mathcal{H}\varphi = \mathcal{E}\varphi \quad . \end{aligned} \tag{16.13}$$

The two quotients in the upper line show that the left quotient (and the right one as well) cannot depend on either $\vec{r}$ or t (as a consequence of our question). We call this constant $\mathcal{E}$. States of stable shape do exist, if $\mathcal{H}\varphi = \mathcal{E}\varphi$ admits solutions.

4th A quantum system (in a time independent potential) is radiating energy until it falls asleep. It can remain for quite a while in a stationary state of

higher energy. However, this is a metastable equilibrium (saddle point). Minimal perturbations from outside will cause the system to leave its excited state and thereby it will loose energy. This means that, sooner or later, the system will reach its ground state for one reason or another. For any physical system such a state of lowest energy must exist (otherwise we would not need power plants and all that, only a bit of shaking would be needed).

The ground state has to be an eigenstate for the energy question. Energy can be measured, it is an observable, hence there is an operator H associated to it. Traditionally we call it **Hamilton operator** (or Hamiltonian): $H\varphi_0 = E_0\varphi_0$. Since it provides the ground state, H plays a special part amongst the operators.

The ground state φ_0 does not radiate, its shape is stable and it is one of the solutions of (16.13). Now we have exhausted our wits, so we have a look at a series of experiments and conclude that an unperturbed system in any eigenstate φ_n of H does not radiate and conserves its shape.[4] With $H\varphi_n = E_n\varphi_n$ we also get $\mathcal{H}\varphi_n = \mathcal{E}_n\varphi_n$[5]. The correspondence between E_n and $\mathcal{E}_n$ should be system independent. Hence there is a universal function $\mathcal{E}$ such that

$$\mathcal{E}_n = \mathcal{E}(E_n) \quad , \quad \mathcal{H} = \mathcal{E}(H) = \sum_\nu \frac{1}{\nu!}\,\mathcal{E}^{(\nu)}(0)\; H^\nu \quad . \tag{16.14}$$

This Taylor series shows how to define a function of an operator.

5th We consider a system consisting of two parts which are separated by such a large distance that they do not take notice of each other. Thus, on the one hand, $H = H_1 + H_2$ (H_1 is the Hamiltonian of the left part and acts there only, and H_2 does the same for the right part). On the other hand, the operator $\mathcal{H}$ must also be additive, $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$. It follows that

$$\mathcal{E}(H_1 + H_2) = \mathcal{E}(H_1) + \mathcal{E}(H_2) \quad \curvearrowright \quad \mathcal{H} = \mathcal{E}(H) = \lambda \cdot H \quad . \tag{16.15}$$

6th For the value of the numerical factor λ in (16.15), well, only nature can tell us. Actually, we have already asked nature quite frequently, in classical mechanics that is. Quantum mechanics must include Newtonian mechanics as a limit. Carry out the corresponding calculation with an unspecified λ, then the limit provides the constraint $\lambda = 1$. There it is, the **Schrödinger equation**:

$$\mathrm{i}\,\hbar\,\dot{\psi} \;=\; H\;\psi \quad . \tag{16.16}$$

We have now answered questions 2 and 3 in (16.1). This essentially completes our invention. This is the core of quantum mechanics (see also (16.6)).

[4] This is true as long as only quantum *mechanics* is studied. However, once the e.m. fields are quantized as well transitions to lower φ_n become possible. Such a transition is a slow change on the time axis. The tale of the "quantum leap" is another one of those common lies.

[5] A linear operator is determined by its action on all elements of a basis. When you know $\mathcal{E}_n$ you know $\mathcal{H}$, more on this can be found in the section on Hilbert space.

At first it may look like a messy patchwork of arguments, reasoning and not yet closely examined formulae. Examples and applications are missing. We cannot leave it that way. Firstly, we take this dozen of pages to a patent clerk. He accepts our request and tells us to return in a week. In the postulates section you can find the legal jargon into which he has translated our work.

For the time being, let us keep in mind a few of the unanswered questions. Are there acceptable/unacceptable operators? How to define the scalar product "$\bullet$"? (see the following section). Will the norm $\psi^* \bullet \psi = 1$ be conserved over time? What does the Hamilton operator of specific systems look like? What about the operators for location, momentum and any other observable?

One specific question shall be answered right away. It is sometimes raised by students who are already in the midst of studying quantum theory. How do we get ψ? How do we get the initial condition $\psi(\vec{r}, 0)$ for which the equation of motion (16.16) tells us the future? Let us remember how this is done in mechanics. How do we get $\vec{r}(0)$ and $\dot{\vec{r}}(0)$ for which Newton knows what is going on? We grab a mass point, wait until it stops fidgeting and throw it (applying forces onto it by hand, using Newton), at $\vec{r}(0)$ with $\dot{\vec{r}}(0)$. Now it is in the desired state of uniform motion. In quantum mechanics we proceed by direct analogy

> Whence ψ? — Define the system, wait until it reaches its ground state, adapt potentials s u c h that at a given time t_1 the Schrödinger equation leads to the desired state $\psi(t_1)$. (16.17)

Hence a state ψ is *reproducible* (at least in principle), i. e. we can make it again as many times as we want. Therefore we can measure its probability content. Moreover the information contained in ψ can be reconstructed from measurements (from several different, possibly non commuting operators). At least in this sense ψ really has an objective existence — as much as an electromagnetic field.

Hilbert space

This section can be skipped in the first reading. We take a break from our previous speculations about nature's deepest secrets. Those who know Parts I an II (chapters 1, 4, 12) will be happy to learn "what this stuff is used for". It is not surprising that the mathematical structures at work in quantum mechanics also appear elsewhere in the mathematics of nature. Nature is unique.

For a given quantum system the set of all its states forms a space. It is called **Hilbert space**, HS for short. It has elements (the states), a scalar product and a norm are defined on it, and operators rotate (and dilate) the states in HS. A finite dimensional HS is a vector space, whose vectors however have complex valued components. Aside from this detail we only need to recollect what we

know about vectors and matrices. And afterward we can simply transfer the main results to the infinite dimensional Hilbert space. That is really nice.

Finite number of components: the discrete case. Let the vectors $\vec{\varphi}$, $\vec{\chi}$, $\vec{\psi}$, ... have n complex components. In order for the scalar product to yield a real valued norm, we introduce complex conjugation:

$$\vec{\varphi}^* \vec{\chi} \;:=\; \sum_{j=1}^{n} \varphi_j^* \chi_j = \sum_{\sigma=1}^{n} \varphi^*(\sigma)\chi(\sigma) = \varphi^* \bullet \chi = \chi \bullet \varphi^* \quad , \tag{16.18}$$

where we have replaced the index $_j$ by the functional argument σ (a discrete variable), and thus the arrow on vectors can be dropped. As usual, the known argument of the functions is omitted at the right end of (16.18). ψ is called normed if and only if $\psi^* \bullet \psi = 1$. An operator B is linear if and only if $B(\alpha\varphi + \beta\chi) = \alpha\, B\varphi + \beta\, B\chi$ is true for any two elements φ, χ of HS. Hence a linear operator B is a matrix of $n \times n$ components.

Observables correspond only to matrices A, for which the solution of the eigenvalue equation $A\varphi = a\varphi$ yields eigenvalues a that are real and eigenfunctions φ_a that are mutually orthogonal (due to (16.9)) and thus form a CONS. We assume all of the above and ask what kind of matrix A should be. The following reads as a single line

$$\begin{aligned} \varphi_b^* \bullet A\varphi_a \;&=\; \varphi_b^* \bullet a\varphi_a = a\delta_{ab} = (a\delta_{ab})^* = (b\delta_{ba})^* = (\varphi_a^* \bullet A\varphi_b)^* \\ &= \big(\big[\, A^T \varphi_a^* \,\big] \bullet \varphi_b \big)^* = \varphi_b^* \bullet A^{T*} \varphi_a \qquad \curvearrowright \quad A \stackrel{!}{=} A^{T*} =: A^\dagger \quad . \end{aligned} \tag{16.19}$$

Such a matrix is commonly called **hermitian**. In the final step we compare the left and the right end of this chain of equations and use the fact that it must be valid for all a, b in order to conclude.[6] The dagger index † stands for $*T$ or $T*$ since the order of transposition and complex conjugation can be reversed. For a non hermitian matrix B (i. e. $B \neq B^\dagger$) $B^\dagger$ called its hermitian adjoint matrix (or hermitian adjoint operator). In (16.19) we have used the bullet because this form is directly valid in the continuous case.

The converse is also a rather neat reasoning. If A is hermitian, then the above conditions are derived easily.

$$\begin{aligned} & a\, \varphi_a^* \bullet \varphi_a = \varphi_a^* \bullet A\varphi_a = (A^\dagger \varphi_a)^* \bullet \varphi_a = (A\varphi_a)^* \bullet \varphi_a = a^* \varphi_a^* \bullet \varphi_a \\ \curvearrowright \quad & a \text{ is real} \qquad , \qquad 0 = \varphi_b^* \bullet (A\varphi_a - a\varphi_a) = (b-a)\,(\varphi_b^* \bullet \varphi_a) \\ \curvearrowright \quad & \text{if } a \neq b \;\; \varphi_a \text{ is orthogonal to } \varphi_b \quad . \end{aligned} \tag{16.20}$$

In the degenerate case, i. e. for $a = b$, the φ's are orthogonalisable (as in chapter 4, section 4.3). Hence the $\varphi_{a\nu}$ form a CONS. The "ON" herein stands

[6] Two operators are identical if for all elements of the space they produce identical results. Linearity entails that we only need to check this on a basis.

for the

Orthonormalization relation $$\varphi^*_{a\nu} \bullet \varphi_{b\mu} \;=\; \delta_{ab}\,\delta_{\nu\mu} \quad . \tag{16.21}$$

Every hermitian matrix generates its own basis of HS. Given a basis one can expand with respect to it. That is the

Expansion theorem $$\psi = \sum_{a\,\nu} c_{a\,\nu}\,\varphi_{a\,\nu} \quad \curvearrowright \quad c_{a\,\nu} = \varphi^*_{a\nu} \bullet \psi \quad , \tag{16.22}$$

where we have taken the scalar product with $\varphi^*_{a\nu}$ on the left and ψ on the right. If ψ has norm 1, then the set of coefficients also has norm 1 :

$$1 = \psi^* \bullet \psi = \sum_{a\nu\,b\mu} c^*_{a\nu}\,c_{b\mu}\,\varphi^*_{a\nu} \bullet \varphi_{b\mu} = \sum_{a\nu} |c_{a\nu}|^2 \quad . \tag{16.23}$$

To the orthonormalization relation (16.21) (sum over the variables yields Kronecker delta of the indices) there is a converse (sum over the indices yields Kronecker delta of the variables),

$$\psi(\sigma) = \sum_{a\nu} c_{a\nu}\,\varphi_{a\nu}(\sigma) = \sum_{\sigma'} \left[\sum_{a\nu} \varphi^*_{a\nu}(\sigma')\,\varphi_{a\nu}(\sigma) \right] \psi(\sigma') \quad \curvearrowright \quad \Big[\quad \Big] = \delta_{\sigma'\sigma} \;,$$

i. e. the

Completeness relation $$\sum_{a\nu} \varphi^*_{a\nu}(\sigma')\,\varphi_{a\nu}(\sigma) \;=\; \delta_{\sigma'\,\sigma} \quad . \tag{16.24}$$

Apart from a few stars we have only used vector algebra. The simplest CONS is $\varphi_a(\sigma) = \delta_{a\sigma}$ (these are the $\vec{e}_a$ with $a = 1, 2, 3$). (16.21) : $\varphi^*_a \bullet \varphi_b = \sum_\sigma \delta_{a\sigma}\delta_{b\sigma} = \delta_{ab}$ $(= \vec{e}_a\,\vec{e}_b)$. (16.22) : $\psi(\sigma) = \sum_a c_a \delta_{a\sigma}$ $(\vec{a} = a_j\,\vec{e}_{\,j})$, $\psi(\sigma) = \sum_a c'_a f_a(\sigma)$ $(\vec{a} = a'_j\,\vec{f}_{\,j}\,)$. A Hilbert vector h with components $\psi(\sigma) =: \langle\sigma|h\rangle$ along the Kronecker basis has the components $c_{a\nu} =: \langle a\nu|h\rangle$ along another basis $\varphi_{a\nu}(\sigma) =: \langle\sigma|a\nu\rangle$. Using these neat definitions $c_{a\nu} = \varphi^* \bullet \psi$ (unusual) can be rewritten elegantly as $\langle a\nu|h\rangle = \langle a\nu|\sigma\rangle\,\langle\sigma|h\rangle$ (usual).

Infinitely many components, continuously indexed by x , functional case. Sums over indices are replaced by integrals over x . The scalar bullet now refers to a preceding integral. All previous equations remain valid. Rewriting (16.18) to (16.24) without bullets and using integrals is a useful exercise for the reader. Matrix operations also contain a sum over σ . In the continuous case we can collect the elements of a matrix A as $K(x, y)$ and write altogether

$$\varphi^* \bullet \chi = \int dx\;\varphi^*(x)\,\chi(x) \quad , \quad A\psi = \int dy\;K(x, y)\,\psi(y) \quad . \tag{16.25}$$

If $\langle x|\hat{A}|y\rangle$ is written instead of $K(x,y)$ the corresponding operator remains visible (see also (16.28) below, Dirac notation). We now encounter two new situations. First of all, the integrals in (16.25) must exist. Normalization is now an actual restriction. The completeness relation now reads

$$\sum_{a\nu} \varphi^*_{a\nu}(y)\, \varphi_{a\nu}(x) \;=\; \delta(y-x) \quad . \tag{16.26}$$

Secondly, eigenvalues and degeneracy indices are now allowed to range over continuous values. The corresponding sum is to be replaced by an integral — in which case (16.26) has to be deduced anew.

In the functional case $\delta(x-a)$ is the trivial basis (and a its index). For this basis (16.22) yields $c_a = \psi(a)$, and (16.21) and (16.24) go as $\int dx\; \delta(x-a)\,\delta(x-b) = \delta(a-b)$ and $\int da\; \delta(x'-a)\,\delta(x-a) = \delta(x'-x)$, respectively.

We are done. An operator is hermitian if and only if its kernel fulfills $K(x,y) = K(x,y)^{T*} = K(y,x)^*$. Its eigenfunctions form a CONS. Each linear operator has its kernel. The kernel of ∂_x is $K(x,y) = \partial_x \delta(x-y)$. The kernel of the translation operator T_a is $K(x,y) = \delta(x+a-y)$. The kernel of the parity operator P is $K(x,y) = \delta(x+y)$. How does P act on a function of x? It reverses the sign of its argument x. Is ∂_x hermitian? It is not, because $\partial_x \delta(x-y)$ changes sign under transposition (but $\mathrm{i}\,\partial_x$ is hermitian). Is P hermitian? Yes. Is T_a hermitian? It is not, but $T_a + T_{-a}$ and $\mathrm{i}(T_a - T_{-a})$ are. How come all of this sounds so familiar? Because we have seen it in the chapter on the Fourier Transform (especially Figure 12–5. (16.23) is Parseval's theorem (12.19), (12.50)). Once the German book (of which the present volume is essentially the translation) was criticized for lacking quantum mechanics. Hey, chapter 12 is teeming with it.

One last question. Given a B with known effect on functions, how do we obtain its kernel? We also know the result $\chi_{a\nu}$ of applying B on any basis element $\varphi_{a\nu}$:

$$\chi_{a\nu}(x) \;:=\; B\,\varphi_{a\nu}(x) = \int dz\; K(x,z)\,\varphi_{a\nu}(z) \quad .$$

We multiply by $\varphi^*_{a\nu}(y)$, sum over a,ν, use the completeness relation and obtain

$$K(x,y) = \sum_{a\nu} \chi_{a\nu}(x)\,\varphi^*_{a\nu}(y) \;=\; \Big\langle x\big|\hat{B}\big|a\nu\Big\rangle\,\langle a\nu|y\rangle \;=\; \Big\langle x\big|\hat{B}\big|y\Big\rangle \quad . \tag{16.27}$$

In the following three lines, we have collected all there is to know about the Dirac's shorthand notation:

$$\begin{aligned}
\psi_{\,\text{index}}(\text{variable}) \;&=:\; \big\langle \text{variable}\,\big|\,\text{index} \;:=\; \text{all specifications of the state } \psi\,\big\rangle \\
A\,\big\langle \text{variable}\,\big|\,\text{index}\big\rangle \;&=:\; \Big\langle \text{variable}\,\Big|\,\text{“abstract operator” } \hat{A}\;\Big|\,\text{index}\Big\rangle \\
\text{and} \qquad & \quad |\,any\rangle\langle any\,| \;=\; 1 \quad ,
\end{aligned} \tag{16.28}$$

where one has to sum (or integrate respectively) over *any* in the last equation.

Finally, to write down the abstract and general **spectral representation** of a hermitian operator, we go back to (4.57), i. e. to $A = \sum_\ell \vec{f}_\ell \circ \lambda_\ell \vec{f}_\ell^{\,*}$, and replace ℓ by a, ν. Then λ_ℓ becomes a. The result is

$$\hat{A} = \sum_{a,\nu} |\, a\nu \,\rangle\, a\, \langle a\nu \,| \quad . \tag{16.29}$$

Use the Dirac notation only when it makes sense, please. Otherwise notations like $\psi(\vec{r}, t)$ are preferable. We are living in 3D.

Seven postulates

I

Any quantum system can be completely described by a single valued function $\psi(\, x\,,\, t\,) \in \mathbb{C}$ (the information carrier). x is a set of variables, one for each degree of freedom. Generally we may write $x = 1, 2, \ldots$ where $1 :=$ is the set of variables for particle 1 and so on.

II

To each observable corresponds a linear hermitian operator A. The following table of correspondence is part of this postulate.

classical quantity	name in Q.	letter	space	mode of operation
position (1D)	position	X	$\psi(x)$	$X = x$
momentum (1D)	momentum	p	$\psi(x)$	$p = \frac{\hbar}{\mathrm{i}}\,\partial_x$
momentum (3D)	momentum	$\vec{p}$	$\psi(\vec{r})$	$\vec{p} = \frac{\hbar}{\mathrm{i}}\nabla$
angular momentum	angular momentum	$\vec{L}$	$\psi(\vec{r})$	$\vec{L} = \vec{r} \times \vec{p} = \vec{r} \times \frac{\hbar}{\mathrm{i}}\nabla$
—	parity (3D)	P	$\psi(\vec{r})$	$P\,\psi(\vec{r}) = \psi(-\vec{r}\,)$
—	spin (1^{st} comp.)	σ^x	two–com–ponent	$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
energy	Hamiltonian	H	e.g.	e.g.
.	.	.	$\psi(\vec{r})$	$H = -\frac{\hbar^2}{2m}\Delta + V(\vec{r}, t)$
⋮	⋮	⋮	⋮	⋮

III

Possible measurement values are the eigenvalues of A, which are obtained when solving $A\varphi_{a\nu} = a\varphi_{a\nu}$ while requiring single valuedness (see **I**) and normalizability (see **IV**).[7]

[7] Warning: After a measurement whose result is a, generally the system is $n\,o\,t$ in the state φ_a see *[Landau/Lifshitz] III*, section 7. Position measurements are an exception. Trajectories do exist in cloud chambers.

IV

The eigenstates of A are to be normed according to

$$\int dx\ \varphi^*_{a\nu}(x)\varphi_{b\mu}(x) \;=\; \delta_{ab}\,\delta_{\mu\nu}\ ,\ \ \delta(a-b)\,\delta_{\mu\nu}$$
$$\text{or} \qquad \delta_{ab}\,\delta(\mu-\nu)\ ,\ \ \delta(a-b)\,\delta(\mu-\nu)\ , \tag{16.30}$$

where the delta functions apply to continuously indexed parts of the spectrum of A. Let such φ's be called "delta–normed".

But the actual state of a system must always have norm one,

$$\int dx\ |\,\psi(x)\,|^2 \;=\; 1\ . \tag{16.31}$$

Note: The state of a system ψ may be arbitrarily close to a delta–normed eigenstate φ_a (a continuous). Only its strict delta–normalization would be at odds with (16.32) (e.g. giving $|\delta|^2$ in P). In such a case (16.31) can always be satisfied by suitable embedding such as e.g. by adding ∞–high potential walls far away from the system.

V

The probability of obtaining one of a set of discrete measurement results a and the probability density for continuously distributed measurement values a follow the same formula

$$P(a,t) = \sum_\nu |c_{a\nu}|^2 = \sum_\nu \left| \int dx\ \varphi^*_{a\nu}(x)\psi(x,t) \right|^2 \ . \tag{16.32}$$

If ν is a continuous variable, then $\sum_\nu$ in (16.32) has to be replaced by $\int d\nu$.

VI

The equation of motion of quantum mechanics is

$$\mathrm{i}\,\hbar\,\dot{\psi} \;=\; H\ \psi \quad , \qquad\qquad \equiv\ (16.16)$$

where the operator H is to be found in table **II**. (16.16) is also valid when H depends on time (for instance because of $V(\vec{r},t)$, see the above table).

VII

Pauli principle: When the sets of variables of two identical particles are exchanged the following relation must hold

$$\psi(1,2,\ldots) \;=\; \mp\,\psi(2,1,\ldots) \quad . \tag{16.33}$$

The minus sign corresponds to fermions, the plus sign to bosons.

We are rather satisfied upon first skimming of the patent clerk's words. Apparently, several of our formulations were already alright. The table in postulate **II** answers one of our remaining questions. Wherefrom did those details come? Surely, they must arise from heuristic considerations (see the next section). There are operators who lack a classical equivalent. How nice that the vectorial case (σ^x) is among them. Postulate **V** also takes degeneracy into account. The probability to obtain a is added up from η parts: correct, good job. If all eigenvalues a are distinct, then there is no $\sum_\nu$. When two eigenvalues move together then the combined a is obtained with probability = sum of the original P values. It remains to be shown, whether the normalization in **VI** works in all cases. Our distinction of two structures in (16.6) is no where to be seen. Well, it seems the clerk did not understand it. But what about postulate **VII**? Did he (cheekily!) call a certain Mister Pauli? This must again involve some experimental results. Sometimes judiciary people do not like to take reality into account.

When you see Maxwell's equations for the first time, you start to check their coherence. That is what we did in chapter 11, and now we are going to do it with these postulates.

1st test: On the requirement of single-valuedness in postulate **I**. Let us consider an ad hoc operator $L_z := (\hbar/\mathrm{i})\,\partial_\varphi$ ($\varphi :=$ polar angle of a plane system)[8], and ask for its eigenfunctions $\chi(\varphi)$. The obvious guess is $\chi \sim \mathrm{e}^{\mathrm{i}m\varphi}$ ($m \in \mathbb{C}$, unknown). If the angle φ goes over 2π, then a single valued χ has to fall back to the values it had from $\varphi = 0$ onwards. $e^0 \stackrel{!}{=} e^{\mathrm{i}m2\pi} \;\curvearrowright\; m = 0, \pm1, \pm2, \ldots$. Due to the required single-valuedness, the angular momentum $A = L_z$ has a discrete spectrum $a = \hbar\, m$.

2nd test: On the requirement of normalizability in postulate **III**. The simplest Hamilton operator in the table is $H = -\partial_x^2$ (1D, $V = 0$, prefactor set to 1). The

[8] L_z is the third component of the angular momentum operator in the table of postulate **II**, expressed in polar coordinates. For an electron confined to a circular ring (R) the Hamiltonian is $H = L_z^2/(2\theta)$ and the energy eigenvalues are $E = \hbar^2 m^2/(2\theta)$.

ansatz $\varphi \sim e^{\lambda x}$ yields $E = -\lambda^2$. But the norm integral is divergent for real λ. Hence $\lambda = \mathrm{i}k$, $\varphi \sim e^{\mathrm{i}kx}$, $E = k^2$. The continuous spectrum of a free particle is bounded from below (otherwise power plants would not be needed).

3rd test: On the probability (density) in postulate **V**. When (16.32) is summed (or integrated) over, it has to add up to one. Let us work it out for continuous a:

$$\int da\, P(a,t) \begin{array}{l} = \int da \sum_\nu \int dx\, dy\, \psi^*(x)\varphi_{a\nu}(x)\varphi^*_{a\nu}(y)\psi(y) \\ = \int da \sum_\nu \int db \sum_\mu c^*_{a\nu}\, c_{b\mu} \int dx\, \varphi^*_{a\nu}(x)\varphi_{b\mu}(x) \end{array} = \int dx\, |\psi|^2 = 1\,. \quad (16.34)$$

In the first line we use the completeness relation. In the second line we start from the right side, expand ψ and use the orthonormalization relation.

4th test: On the equation of motion in postulate **VI**. How does it guarantee that the norm (16.31) remains unity when the system is left alone with (16.16)?

$$\partial_t \int dx\, |\psi|^2 = \int dx\, \left(\dot\psi^*\psi + \psi^*\dot\psi\right) = \frac{1}{\mathrm{i}\hbar}\int dx\, \left(-[H\psi]^*\psi + \psi^* H\psi\right) = 0\ , \quad (16.35)$$

because H is hermitian.

5th test: Phase factors. They are the subject of a question to the postulates as a whole. The eigenvalue equation $A\varphi_{a\nu} = a\varphi_{a\nu}$ does not specify a complex prefactor of the solution $\varphi_{a\nu}$. It may even depend on time t (but not on $\vec r$, because A acts on the $\vec r$ dependence). Normalization only determines its modulus. Hence a phase factor $e^{\mathrm{i}\vartheta_{a\nu}(t)}$ for each of the $\varphi_{a\nu}$ can be chosen arbitrarily. Woe betide us if a physical prediction were to depend on it! Due to (16.22) the coefficients $c_{a\nu}$ still do depend on the phase choice, but the probabilities (16.32) are clearly real, measurable quantities. When it is searching for stationary states (as in (16.13) with $\mathcal{H} = H$) ψ keeps the prefactor $e^{-\frac{\mathrm{i}}{\hbar}Et}$, but again this disappears in a, P. Solving $H\varphi(\vec r) = E\varphi(\vec r)$ also leaves a constant phase factor. To draw the conclusion from all this, one may say that states which differ only by a constant phase factor are the *same*. For calculations a specific phase should be chosen and conserved.

The most important operators

Once again, we use a kind of heuristic arguments that are (probably) absent in textbooks. The following are derivations for the contents of the table in postulate II.

Position X (1D): The measured values a of a 1D position may be distributed continuously along the real axis. (The photographic "line"is asking questions all over). Utmost certainty (of measuring a) should be a consequence of utmost localization of the state $\varphi_a(x) = \delta(x-a)$ (which is already normed). Remember that a linear operator is defined by its action on all elements of a basis, here this is $X\,\delta(x-a) = a\,\delta(x-a)$. Now we integrate this equation with an arbitrary weight function $\psi(a)$:

$$\int da\ \psi(a)\, X\, \delta(x-a) = \int da\ \psi(a)\, a\, \delta(x-a) \ \curvearrowright\ X\ \psi(x) \ = \ x\,\psi(x) \ . \quad (16.36)$$

This tells us how X acts. Hence the eigenvalue equation reads $(x-a)\,\varphi_a(x) = 0$ and yields $\varphi_a(x) = C\cdot\delta(x-a)$ with $|C| = 1$ due to (16.30). To obtain the **probability density for position measurements**, we equip (16.32) with $\varphi_a(x) = \delta(x-a)$ and obtain

$$\begin{aligned} P(a) = \left| \int dx\ \delta(x-a)\psi(x) \right|^2 \ &= \ |\psi(a)|^2 \quad , \qquad (16.37) \\ \curvearrowright \quad \begin{array}{l}\text{the probability to find}\\ \text{a particle in } (a,b)\end{array} \ &= \ \int_a^b dx\ |\psi(x)|^2 \ . \end{aligned}$$

(16.37) typically stands at the beginning of treatises on quantum physics, but note that here it comes as a result. The generalization to 3D is harmless and looks like this: $\vec{R}\,\psi = \vec{r}\,\psi$, $P(\vec{r}) = |\psi(\vec{r})|^2$ and $\varphi_{\vec{a}}(\vec{r}) = \delta(\vec{r}-\vec{a})$.

Momentum p (1D):

Thinking about the left half of Figure 16–3 (left of the corrugated metal, left of the double slit) we had argued that the state of the electron was similar to $\mathrm{e}^{\mathrm{i}kx}$ (with the x axis being parallel to the arrow), and this state corresponds to a well defined momentum $q = \hbar k$ (we write q instead of p, because p is already in use for the operator we are looking for). For the system to have a definite momentum, means that q is an eigenvalue of p. Hence we know the eigenstates for each eigenvalue, i. e. $p\,\mathrm{e}^{\mathrm{i}kx} = \hbar k\,\mathrm{e}^{\mathrm{i}kx}$. Again this equation may be integrated over with a suitable weight function. Any idea? This time an arbitrary Fourier transformed state is the proper weight:

$$\begin{aligned} \int \frac{dk}{2\pi}\,\widetilde{\psi}(k)\,p\,\mathrm{e}^{\mathrm{i}kx} \ &= \ \int \frac{dk}{2\pi}\,\widetilde{\psi}(k)\,\hbar k\,\mathrm{e}^{\mathrm{i}kx} \ = \ \frac{\hbar}{\mathrm{i}}\partial_x \int \frac{dk}{2\pi}\,\widetilde{\psi}(k)\mathrm{e}^{\mathrm{i}kx} \\ \curvearrowright \qquad p\ \psi(x) &= \frac{\hbar}{\mathrm{i}}\,\partial_x\,\psi(x) \ . \qquad (16.38) \end{aligned}$$

Since it is arbitrary, $\psi(x)$ can be omitted on both sides. This leaves us with the operator identity given in the table. Of course you can again transpose this result to 3D, norm the momentum eigenfunctions correctly to $\delta(\vec{q}\,' - \vec{q})$, and

write down the probability density $P(\vec{q})$.

Potential $V(\vec{r})$: This is a function of an operator $\vec{R}$, into which we have already included how it acts ($\vec{R} \to \vec{r}$). Hence the eigenfunctions of V are $\delta(\vec{r}-\vec{a})$ with eigenvalues $V(\vec{a})$. The operator V acts on any ψ, by multiplying it with $V(\vec{r})$. We admit that this is a rather cheap (but common) step.

Kinetic energy T (1D): This is also a function of an operator, namely $p^2/(2m)$. At least, this view is a rather obvious invention (again think of Davisson–Germer):

$$T \;=\; \frac{p^2}{2m} \;=\; -\,\frac{\hbar^2}{2m}\,\partial_x^2 \quad , \qquad \text{eigenvalues} \;=\; \frac{\hbar^2 k^2}{2m} \quad . \tag{16.39}$$

Of course, the letter k on the right in (16.39) reminds us, that it had been used as an indexing label for the T eigenfunctions via $e^{\mathrm{i}kx}$. T has a positive (i. e. half–sided) continuous spectrum, in 3D as well.

Concerning angular momentum the reader is referred back to the table. Let the statement that T and V can be added together to form the Hamilton operator rest on more or less plausible arguments (V increases the energy at $\vec{r}$ independently of the kinetic part) and everything works well through all applications. In reality V mostly has an electrostatic origin, $V = q\phi$ with ϕ the electrostatic potential. Even a general V can be thought of as electrostatic.

THE Hamilton operator H with which all of non relativistic quantum mechanics gets along, contains both the electromagnetic potentials, ϕ (in place of V/q) and the vector potential $\vec{A}$ of a magnetic field. **Quote:**

$$H = \frac{1}{2m}\left(\vec{p} - q\,\vec{A}(\vec{r},t)\right)^2 \;+\; q\,\phi(\vec{r},t) \quad . \tag{16.40}$$

Why bother with "quantization rules", if there is only one application of them? For systems containing multiple particles (numbered), the above H is labeled with the particle index (same for m, q, $\vec{r}$). The Hamilton operator of the system is the sum of the indexed H-es and the potentials of the particle interactions $V(i,j)$.

Classical limit

We remember Figure 16–1 and that Newton's equation of motion might come about in a suitable limit for the center of the particle. This center ought to be the mean of the position measurement value. For the mean measurement value a (of a system in the state ψ) a very general statement can be obtained. Again

we assume the case of continuous a values.

$$\begin{aligned}\langle a\rangle_\psi &= \int da\ a\,P(a,t) = \int da \sum_\nu \int dy\,dx\ \varphi_{a\nu}(y)\,\psi^*(y,t)\,\varphi^*_{a\nu}(x)\ A\,\psi(x,t)\\ &= \int dx\ \psi^*(x,t)\,A\,\psi(x,t)\ =:\ \langle A\rangle_\psi\quad .\end{aligned}\tag{16.41}$$

The real factor a was moved next to $\varphi^*_{a\nu}$, then transformed into A using $a\varphi^*_{a\nu} = (A\varphi_{a\nu})^*$ and in the x–integral we have used the hermiticity of A. The *sandwich* appearing in the second line is called the **expectation value** of A. Mostly the index ψ on $\langle\ \ \rangle$ is omitted, but beware and keep it in mind.

Time is simply a parameter in (16.41). However, we will now differentiate with respect to it, use the equation of motion and obtain the **Ehrenfest's theorem**:

$$\begin{aligned}\partial_t\,\langle A\rangle_\psi &= \int dx\ \dot\psi^* A\psi + \int dx\ \psi^* A\dot\psi + \int dx\ \psi^*\dot A\,\psi\\ &= \frac{\mathrm{i}}{\hbar}\,\langle HA - AH\rangle_\psi + \langle\dot A\rangle_\psi\quad .\end{aligned}\tag{16.42}$$

The last term is required in the rare case that the operator explicitely depends on time, such as for instance $V(\vec r,t)$. X however does not feel t: $\dot X = \dot x = 0$. The operator expression $AB - BA =:\ [\,A,B\,]$ is called **commutator**. To transform it ("calculate" it), it is best to write (or think of) a general element ψ of Hilbert space to the right of it.

And this is how we deduce Newton. In order for the calculations to be clearer, we will stay in one dimension and study the system described by $H = -\frac{\hbar^2}{2m}\,\partial_x^2 + V(x)$. A little algebra gives the commutators $[\,H,x\,] = Hx - xH = \frac{\hbar}{\mathrm{i}m}\,p$ and $[\,H,p\,] = Hp - pH = \mathrm{i}\hbar V'$. Using these we arrive at

$$\begin{aligned}m\,\langle x\rangle^\bullet &= \frac{m\mathrm{i}}{\hbar}\,\langle Hx - xH\rangle\ =\ \langle p\rangle\\ m\,\langle x\rangle^{\bullet\bullet} &= \langle p\rangle^\bullet = \langle -V'(x)\rangle\ =\ \langle F(x)\rangle\ \neq\ F(\langle x\rangle)\quad .\end{aligned}\tag{16.43}$$

The center is $\langle x\rangle$. Newton for the center would have been $m\,\langle x\rangle^{\bullet\bullet} = F(\langle x\rangle)$. That is not what we ended up with. The result (16.43) is very welcome nonetheless. It shows in which limiting case everything works out, that is the limit of localization. A measure for the extension of ψ is $\Delta x = \sqrt{\langle(x-\langle x\rangle)^2\rangle}$, see (12.101). If Δx becomes small compared to the distances on which F varies noticeably then the expressions on either side of $\neq$ in (16.43) become equal. There is a special case for which the $\neq$ sign in (16.43) disappears. The harmonic oscillator has $F(x) = -\kappa x$ and therefore $\langle -\kappa x\rangle = -\kappa\,\langle x\rangle$. The wavefunction may be going crazy, but the center $\langle x\rangle$ is moving harmonically to and fro.

The milestone (16.43) had to be reached without fail in this "first half of an introduction". What remains now is clear: exercises, applications, thinking about

and working out the related physics. Hannover is purely non-relativistic quantum mechanics (since it is lacking a particle accelerator). But before Hannover is completely understood, long hours will leave their footprints in the sands of time.

— — —

During the writing of these pages, from the outset they were addressed to readers with widely varying backgrounds. The only prerequisite is a burning interest in nature as it is. Dear housewives and retirees, the many pictures, colloquial speech and references to every day life were meant for you. The equations only got nasty in Hilbert space and in the last two sections.

A high school teacher should have easily found his way through this chapter (if not: sit down and work it out). Perhaps it will inspire beautiful $true$ tales. Young people are not as impenetrable as they may want to appear. They want to know things, if possible right away (just as "the impossible can be done, miracles take some time").

There is also the first quantum phyiscs at university, where people like to disagree.[9] Quite often the students deftly manipulate the formalisms, calculate all their exercises and everything seems alright. Really? Bear in mind the following important (not at all malicious) saying: If I have truly understood something, then I can also explain it to myself, my little sister (if she wants to listen) and — and here we are again at our dear housewives.

[9] Books on quantum mechanics are ten a penny. [*Landau/Lifshitz*, III] may be too difficult for beginners. A nice short introduction is to be found in [*Margenau/Murphy*]. [*Schiff*], *quantum mechanics* may be old but is still a good read. An example for a more modern textbook is [*Schwabl*, 3]. In German there is e.g. (slightly older) [*Becker/Sauter*, II]. — "But that is a *theory of electricity*" ?! Indeed and in volume II they do things right, with electrons that (a) behave quantum mechanically and (b) have the Coulomb interaction in their Hamiltonian.

W. Heisenberg [*Der Teil und das Ganze* {the part and the whole}
(dtv pocketbook No. 903, 9th ed. München 1985)] :

... then Einstein was disturbed, but the next morning at breakfast he had already come up with a new thought experiment, more complex than the first, which was now to show the invalidity of the uncertainty relation. By the evening this attempt had suffered the same fate as the first one and after this game had continued for several days, Einstein's friend Paul Ehrenfest, physicist from Leyden in Holland, said "Einstein, I am ashamed of you ; you argue against the new quantum theory in the same way, as your opponents have argued against the theory of relativity." But even this friendly admonition could not convince Einstein.

Again I realized, how infinitely difficult it is, to abandon a belief, that so far has been the foundation of our thinking and of our scientific work. Einstein had made it the work of his life, to explore the objective world of physical processes, which happen out there in space and time, according to permanent physical laws. The mathematical symbols of theoretical physics were to represent this objective world and to allow predictions about its future behavior. Now it was claimed, that such an objective world in space and time does not even to exist, when one descends to the scale of atoms, and that the mathematical symbols represent only the possible and not the factual. Einstein was not prepared to let – as he felt – the rug be pulled from under him.

Even later in his life, when quantum physics had already become a widely accepted field of physics, Einstein could not change his point of view. He would accept quantum theory as a preliminary, but not as a final explanation for atomic phenomena. "God does not play with dice", that was a principle, which for Einstein was an unshakable certainty, and which he would not let go of. Bohr could only respond : "But it cannot be our duty to tell God how to rule the world." [10]

R. Penrose [*The Emperor's New Mind*
(Oxford University Press; New Ed. March 18 1999, p. 324)] :

Regarding ψ as describing the "reality" of the world, we have none of this indeterminism that is supposed to be a feature inherent in quantum theory — so long as ψ is governed by the deterministic Schrödinger evolution. Let us call this the evolution process **U** . However, whenever we "make a measurement" magnifying quantum effects to the classical level, we change the rules. Now we do *not* use **U** , but instead adopt the completely different procedure, which I refer to as **R** , of forming the squared moduli of quantum amplitudes to obtain classical probabilities ! It is the procedure **R** and *only* **R** , that introduces uncertainties and probabilities into quantum theory.

The deterministic process **U** seems to be the part of quantum theory of main concern to working physicists; yet philosophers are more intrigued by the non–deterministic *state vector reduction* **R** (or, as it is sometimes graphically described : *collapse of the wavefunction*) . Whether we regard **R** as simply a change in the "knowledge" available about a system, or whether we take it (as I do) to be something "real", we are indeed provided with two completely *different* mathematical ways in which the state–vector of a physical system is described as changing with time. For **U** is totally deterministic, whereas **R** is a probabilistic law; **U** maintains quantum complex superposition, but **R** grossly violates it; **U** acts in a continuous way, but **R** is blatantly discontinuous. According to the standard procedures of quantum mechanics there is no implication that there be any way to "deduce" **R** as a complicated instance of **U** . It is simply a *different* procedure from **U** , providing the other "half" of the interpretation of the quantum formalism. All the non–determinism of the theory comes from

[10] We could not find an English translation of Heisenberg's reminiscence, the above is due to A.–A. Ludl.

R and not from **U** . *Both* **U** and **R** are needed for all the marvellous agreements that quantum theory has with observational facts.

M. Tegmark und J. A. Wheeler [*100 Years of Quantum Mysteries* (*Scientific American*, February 2001)] :

... The Copenhagen interpretation provided a strikingly successful recipe for doing calculations that accurately described the outcomes of experiments, but the suspicion lingered that some equation ought to describe when and how this collapse occurred.
(underlined by the present author)

... the Schrödinger equation itself gives rise to a type of censorship. This effect became known as decoherence ... coherent superpositions persist only as long as they remain secret from the rest of the world. Our fallen quantum card is constantly bumped by snooping air molecules and photons, which thereby find out whether it has fallen to the left or to the right, destroying ("decohering") the superposition and making it unobservable.

... Even though in the Everett view the wave function never collapses, decoherence researchers generally agree that decoherence produces an effect that looks and smells like a collapse.

... it is time to update the quantum textbooks : although these books, in an early chapter, infallibly list explicit nonunitary collapse as a fundamental postulate, the poll indicates that today many physicists – at least in the burgeoning field of quantum computation – do not take this seriously. The notion of collapse will undoubtedly retain great utility as a calculational recipe, but an added caveat clarifying that it is probably not a fundamental process violating the Schrödinger equation could save astute students many hours of confusion.