

III. Probability and Uncertainty

1. Competing approaches: One year earlier than Schroedinger had published his development of de Broglie's matter wave ideas, i.e., in 1925, a young theorist, Werner Heisenberg¹, uninfluenced by the matter wave ideas, published a *completely* different approach to the problems of atomic structure and quantum phenomena. To most physicists Heisenberg's approach seemed extremely abstract and devoid of intuitive content.

It had been motivated by Heisenberg's growing conviction that the problems encountered in trying to get beyond the treatment of simple systems (from free electrons scattering off crystals or isolated bound electrons in simple hydrogenic atoms, etc. to electrons subject to electromagnetic fields and complicated atoms and molecules with several active outer electrons) were due to an excessive reliance on classical concepts like particle trajectories in circumstances where there seemed to be no possibility of observing them or measuring them. This attitude was due, in large part, to the influence on Heisenberg's thinking of Einstein's previous revolutionary analysis of space and time concepts in terms of what we actually measure and to the even earlier analysis of the concepts of classical physics by Ernst Mach², which also emphasized the procedures of measurement. The philosophical position that interpreted and attributed meaning solely on the basis of observation and procedures of measurement was called **positivism**, or sometimes **logical positivism**, and was gaining adherents in those days.

To apply this viewpoint Heisenberg substituted, in the standard classical equations for atomic problems, not the values of the standard classical quantities (which were un-measurable), but large rectangular sets of numbers more closely associated with the measurable frequencies and intensities of radiation emitted from atomic systems. He invented rules for adding and multiplying these large (often infinitely large!) sets of numbers and did not learn, until months later from Max Born, that such things had been studied by mathematicians for some time under the name **matrix algebra**.

Only his closest colleagues had any understanding of his method and most physicists were revolted by what seemed to be a meaningless formalism! So, when Schroedinger introduced his **Wave Mechanics** approach, based on deBroglie's ideas, it was welcomed as much more intuitively appealing! Strangely, however, whenever one knew how to use both Schroedinger's

Wave Mechanics and Heisenberg's **Matrix Mechanics** to attack the same problem, both methods gave the same answer.

In short order Schroedinger (wanting to show that one didn't *need* Heisenberg's unfamiliar matrices!) proved that the two approaches were completely equivalent and showed how to calculate Heisenberg's matrices from his wave functions and conversely. One result of this equivalence, however, was to bring to a head an open question.

Just what was the physical significance of Schroedinger's wave functions and/or Heisenberg's matrices? Did they, themselves, represent anything real?!

The wave mechanical camp (originally de Broglie, Schroedinger and Einstein, but soon joined by growing numbers of physicists) wanted to *identify* the matter waves with the electrons and atoms and molecules. By giving up the old idea that these things were particles, they hoped to retain classical continuity, determinism and a visualizable picture of the atomic world.

The matrix mechanical camp (Heisenberg, Bohr, Born, Pauli, Jordan, Dirac) were more in the grip of positivistic philosophy and thought the new developments meant that physics had to be purged of intuitive concepts that were not susceptible to observation and measurement. In one grueling encounter Heisenberg and Bohr subjected Schroedinger to such a merciless critique of realistic wave thinking that Schroedinger left physically ill and claiming he wished he'd never become involved with "this damned quantum jumping".

What turned out to be an essential step in advancing the problem of interpretation was proposed by Max Born³ and Wolfgang Pauli, again in 1926. Convinced by his study of collision processes and scattering experiments of the corpuscularity of interactions at the atomic scale, Born proposed first that the wave function (quickly corrected by Pauli to its **absolute value squared**) was proportional to the **probability** of finding the associated particle at the point where the wave function had been evaluated. Thus the waves of de Broglie and Schroedinger were, essentially, waves of probability!

Rather rapidly thereafter (especially as a consequence of Dirac's approach to general **Quantum Mechanics**, which treated both Schrodinger's and Heisenberg's formulations as special cases of a large class of possible formulations, all of which had equivalent status) the probability interpretation of most aspects of quantum theory became pervasive.

2. Radioactivity: the first utterly probabilistic phenomena: Although it was not recognized as such at the time, the first encounter with fundamentally probabilistic processes in physics occurred with the discovery, in 1896, of radioactivity. Following upon Roentgen's discovery of X-rays, Henri Becquerel⁴ was looking for X-ray emission from phosphorescent materials exposed to the sun. He would put the materials, various salts of Uranium, next to photographic plates that had been wrapped in black paper to protect them from ordinary light. At one point the weather turned overcast, he put the arrangement in a desk drawer for a few days and then, oddly, developed the film expecting to find a very faint exposure at best. To his astonishment the exposure was intense and he knew immediately that he had discovered a new penetrating radiation that did not require the stimulus of sunlight (Fig. 3.1).

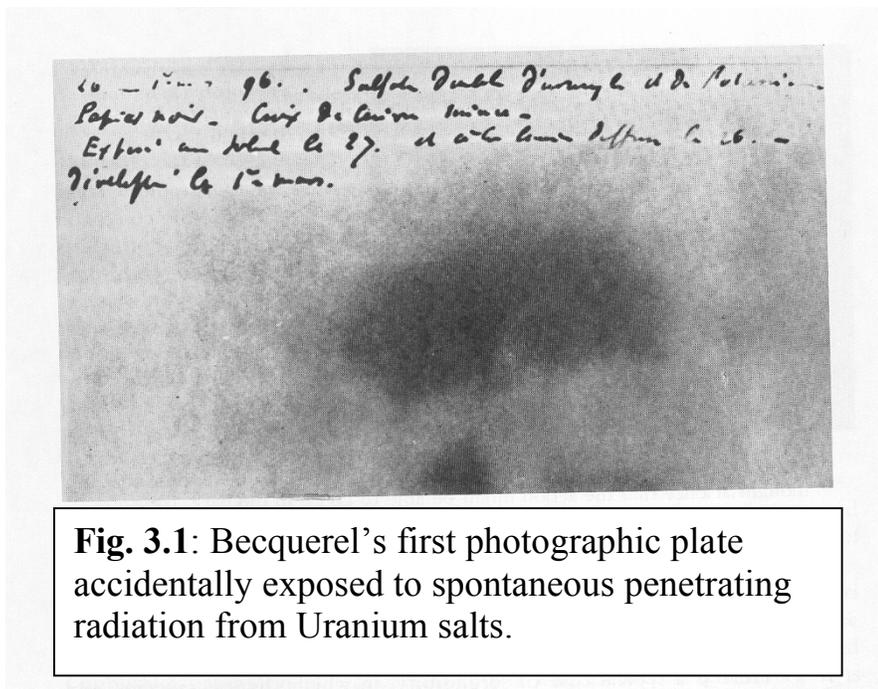


Fig. 3.1: Becquerel's first photographic plate accidentally exposed to spontaneous penetrating radiation from Uranium salts.

The next stage in the story was the heroic effort of the Curies, particularly Marie⁵, after the early accidental death of her husband, to find and classify

the most intensely radioactive substances that existed. Among other things this led to the discovery of Radium. Although Marie lived to the age of 67 she probably died from radiation induced disease. It was reported that her cookbooks from her home were still radioactive fifty years after her death!

The final stage in the unraveling of the nature of radioactivity came with the work of Ernest Rutherford and Frederick Soddy⁶ who, in the first decade of the new century, were the principal workers in disentangling the baffling complex of phenomena associated with radioactivity. It was determined first that the radiations emitted by radioactive substances were three in kind. Alpha (α) radiation, consisting of the positively charged nuclei of Helium atoms, beta (β) radiation, consisting of electrons and gamma (γ) radiation, consisting of very high energy photons (**Fig. 3.2**).

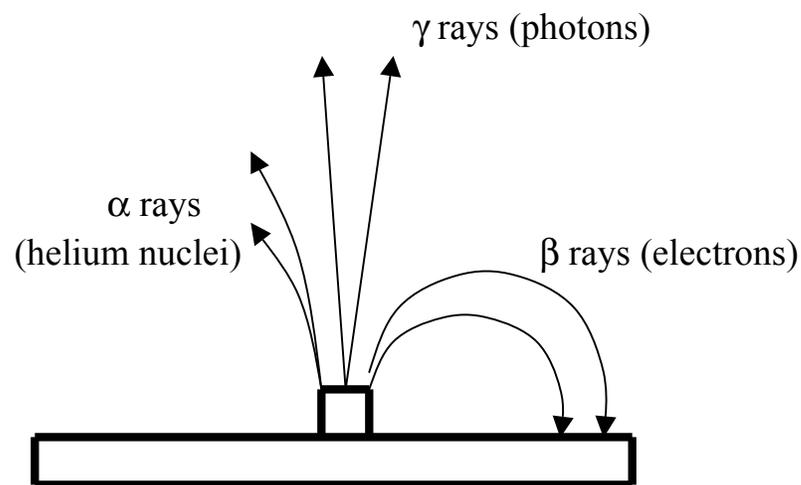


Fig. 3.2: The three kinds of spontaneous radiations from radioactive materials. Emitted vertically into a magnetic field pointing into the page, α particles (positively charged Helium nuclei) curve slightly to the left, β particles (negatively charged electrons) curve strongly to the right and γ rays (uncharged photons) do not curve. A version of this figure appeared in Marie Curie's doctoral thesis.

Second, it was determined that the radiation came from the nucleus of an atom and did not fundamentally involve the orbital electrons. Indeed, when

the radiation was of the α or β kind the nucleus involved *changed into a nucleus of a different kind!* In a sense the holy grail of the old alchemists had been realized; the transmutation of the elements!

Third, and most strangely, it was determined that the only thing that seemed to govern the time of radioactive emission from, or decay of a nucleus was a *probability!* Of any given kind of radioactive nucleus one could say that if it was intact at one moment, then at a given later time there was a definite probability that it would still be intact. The probability decreased with increasing time interval and if one doubled the time interval then one squared the probability. More precisely, if (prob. to remain intact by t) = $p(t)$, then $p(nt) = p(t)^n$. (**Fig. 3.3**)

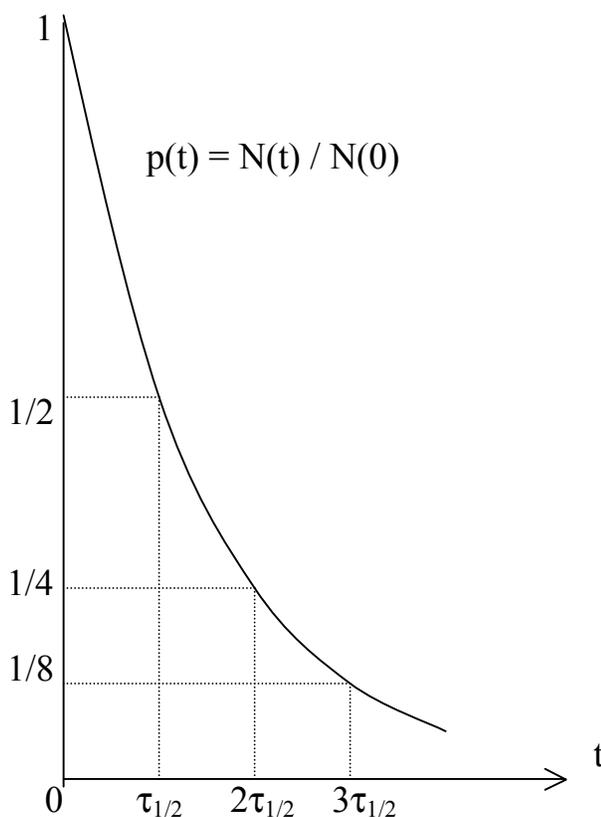


Fig. 3.3: The probability law for radioactive decay. With $N(t)$ as the number of undecayed nuclei left after time t when $N(0)$ is the initial number of undecayed nuclei, the non-decay probability is, $p(t) = N(t)/N(0)$. We then find that, $p(nt) = p(t)^n$. In particular, $p(2\tau_{1/2}) = p(\tau_{1/2})^2$, and $p(3\tau_{1/2}) = p(\tau_{1/2})^3$, where $\tau_{1/2}$ is the **half life** such that $p(\tau_{1/2}) = 1/2$.

In particular there was, for each radioactive nucleus, a **half life**, i.e. a time interval after which the probability was $1/2$ that the nucleus would still be intact. If you started with Zillions of such nuclei, after one half life, half of them would have decayed and half would be left. There was *no telling* which ones would go and which ones would not!

For physicists reared on classical determinism this was utterly astounding! In 1920 Max Planck wrote:

“How is it that a definite Uranium atom, after having remained completely unaltered and passive for untold millions of years, suddenly, in an immeasurably short space of time, without any determinable cause, explodes with a violence, compared with which our most powerful explosives are like toy pistols? It sends off fragments of itself with velocities of thousands of kilometers per second, and at the same time emits electro-magnetic rays of greater intensity than the hardest Roentgen rays, while another in its immediate neighborhood, and to all appearances exactly similar, remains in a passive state for still more millions of years until finally it meets the same fate. In fact, all attempts to affect the course of radio-active phenomena by external means, such as raising or lowering the temperature, have ended in complete failure. It appears, therefore, at present hopeless even to guess at dynamical laws which would account for this.”

In fact, some would argue, quantum mechanics does not *account* for radioactivity. It just incorporates it into a coherent general scheme for describing phenomena at the atomic and sub-atomic level. For a technological application of radioactivity see **Appendix A**.

3. Probability amplitudes: The application of Born’s probability interpretation of Schroedinger’s wave functions changed rather drastically the picture of the outer electronic structure of atoms from the earlier picture of Bohr orbits or even de Broglie waves. Now one had fuzzy/cloudy probability distributions that gradually tailed off to zero as one got further and further from the atom (Fig. 3.4). And, for the atomic states of definite energy, the clouds didn’t move! The distributions told you where you would be most likely to find the electrons if you looked closely, but the distributions didn’t change with time. So from position measurements alone

there was no indication that the electrons were moving at all. For this reason they came to be called stationary states.

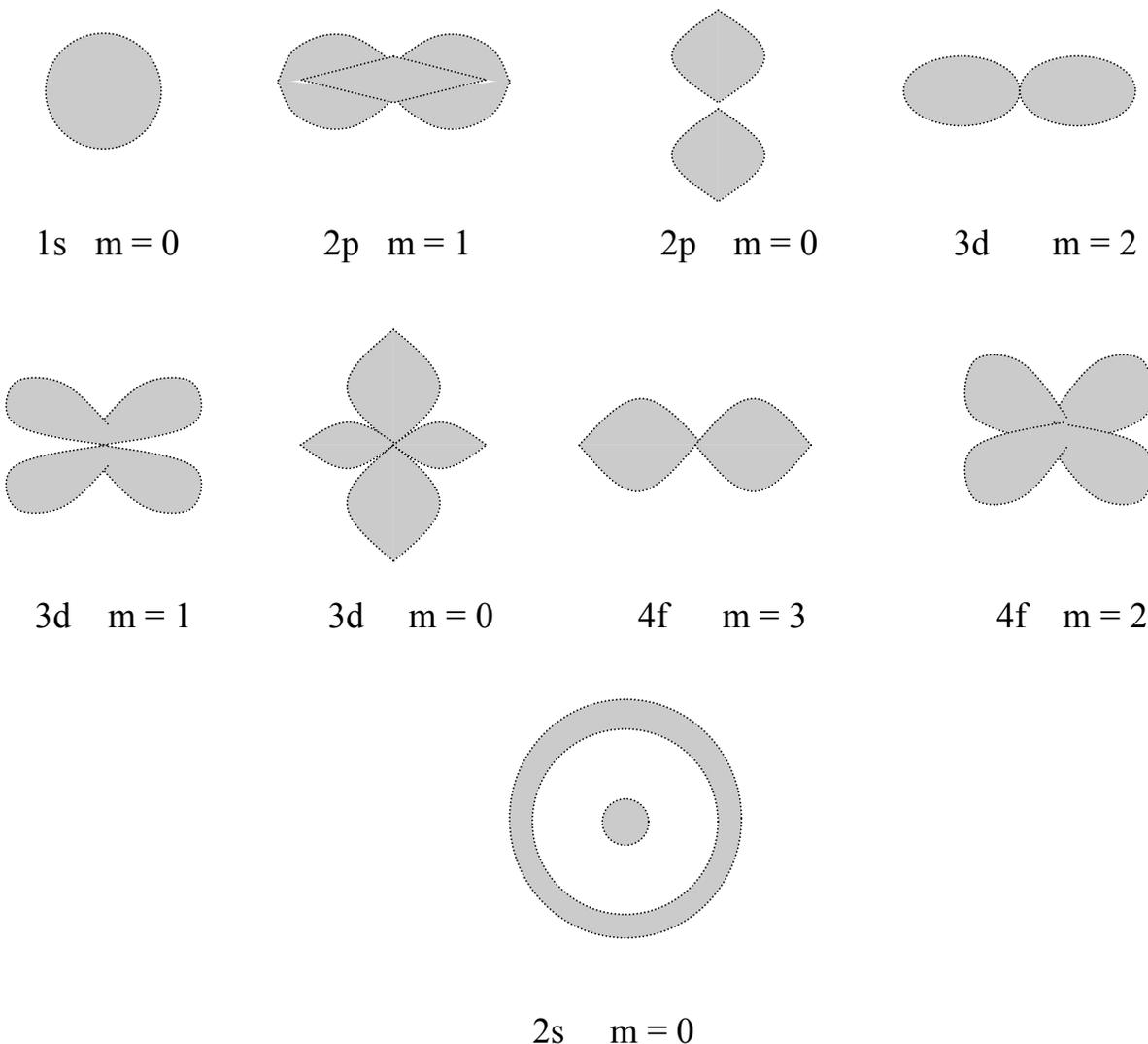


Fig. 3.4: Shapes of the dominant portions of the probability distributions for an electron in various Hydrogen bound states of definite energy, so-called stationary states. The integers, 1, 2, 3, 4, - - - indicate increasing energy levels with 1 denoting the ground state. The letters s, p, d, f, - - - indicate increasing orbital angular momentum magnitudes with s denoting magnitude zero. The m values refer to the orientation of the angular momentum.

Schroedinger's equations did permit one to form wave functions for atomic electrons that would have time dependent probability distributions so that the electrons would clearly be moving. But this required the superposition of amplitudes corresponding to distinct energy values for the atom. In other words putting the probability distributions in motion seemed to require *denying the atom any precise energy*, whatever that would mean!

Not only that, the probability distributions for the stationary states tended to divide the total volume of the atom up into disjoint regions of space with zero probabilities on the boundaries (**Fig. 3.5**). It seemed the electrons had to be confined to one or another of the compartments or, if they *moved* between the compartments, they had to cross the boundaries with impossibly high velocities so they wouldn't spend time near the boundaries.

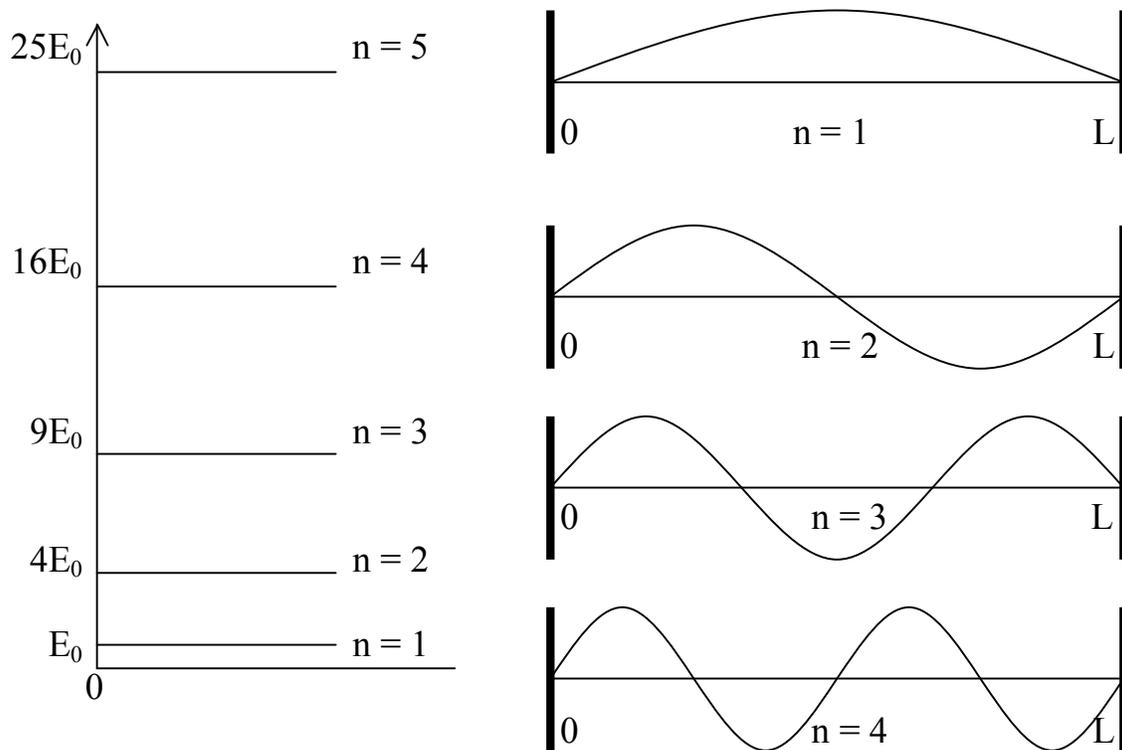


Fig. 3.5: Schroedinger energy levels and wave functions for a quantum particle trapped in a one-dimensional box. Note how the wave functions divide up the box into separate domains for the particle.

Not only that, the probability distributions trailed off to zero *gradually* as one moved further and further from the atom (**Fig. 3.6**). This permitted the electrons to be found in regions where, classically, they would have to have negative kinetic energy to satisfy the precise total energy that the distributions corresponded to. In fact the real penetration of these classically

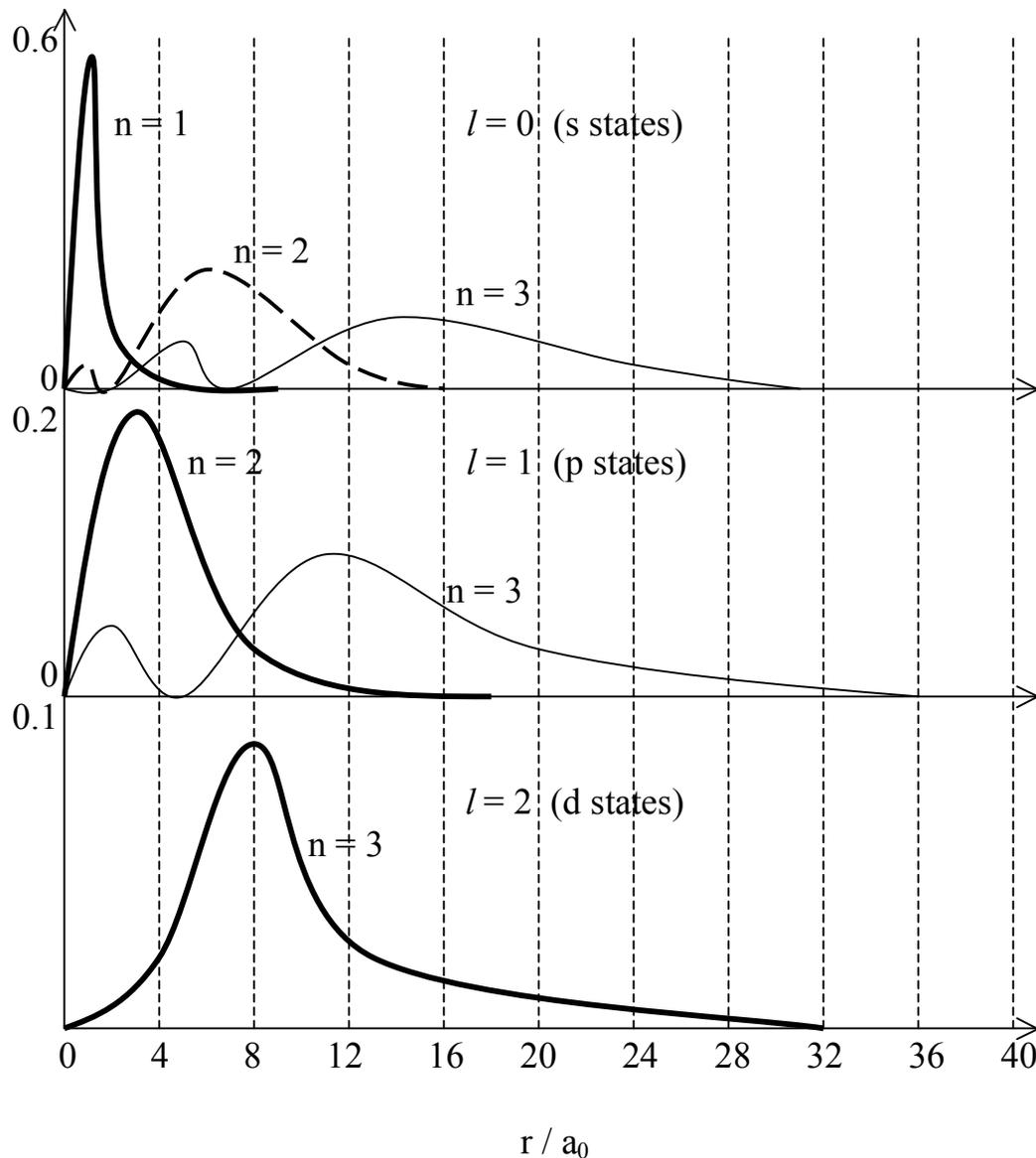


Fig. 3.6: Radial probability distributions for atomic stationary states. The distributions tail off to zero at infinity, thus allowing electrons to be found where their classical kinetic energy would be negative. The zeros of the distributions divide space into isolated spherical shells.

forbidden regions, the so-called **barrier penetration** by the quantum waves, was a crucial feature of the quantum account of several phenomena, e.g. molecular oscillations and α decay of nuclei (**Fig. 3.7**).

One reaction to these puzzles was to reassert the electron to be a tiny moving piece of charged matter and to deny that the probability distributions gave an adequate account of the details. If so, then quantum mechanics was an

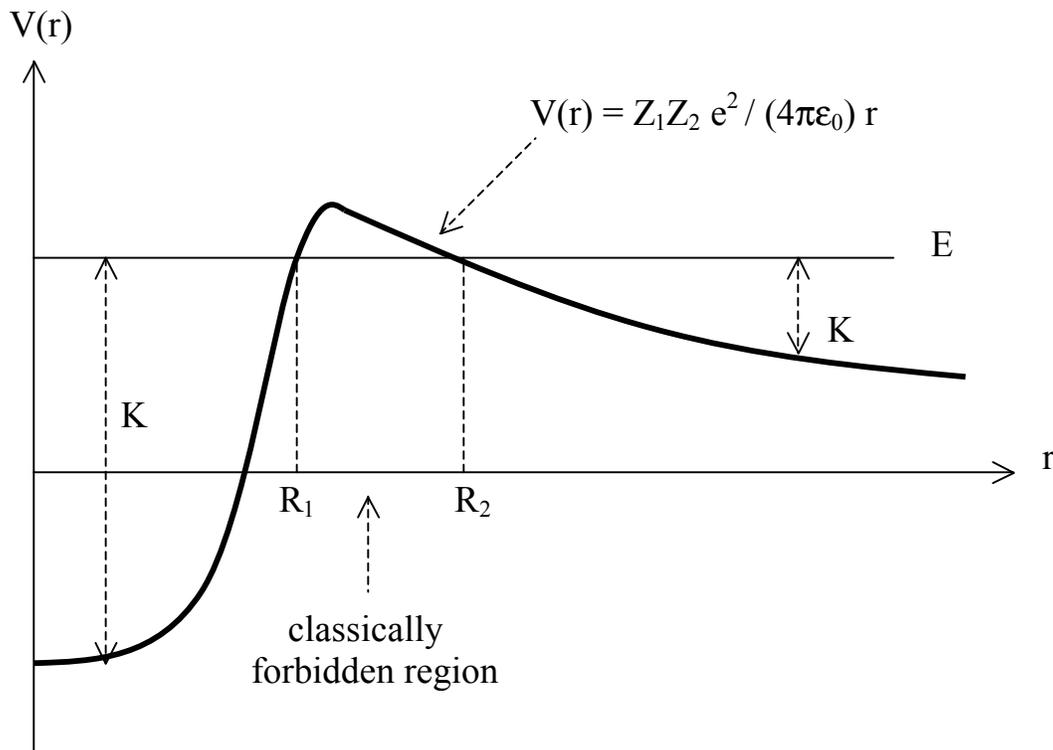


Fig. 3.7: Approximate potential energy curve for a particle of positive charge, $Z_1 e$, in the presence of a nucleus of positive charge, $Z_2 e$. Inside the nucleus, $r < R_1$, and with total energy, E , higher than the potential energy, the particle has positive kinetic energy, K , and is classically trapped within the potential energy wall. Outside the nucleus, $r > R_2$, the particle is repelled by the nucleus and with total energy, E , can not, classically, get closer than the potential barrier permits, i.e., R_2 . But quantum mechanically the particle, originally inside the nucleus, can penetrate the barrier and slip through to the outside with non-zero probability. This appears to be the mechanism for α radioactivity.

incomplete theory of what was going on (Einstein, de Broglie)! This view came to define a vigorous dissenting position, forcefully articulated by David Bohm⁷ in 1951, against the emerging dominant view and is still today active among some physicists and many philosophers of physics.

Another reaction, based on the fact that the quantum mechanical probability distributions seemed to be all one needed to solve any (solvable) problem that could be unambiguously formulated, was to *identify* the electrons with the complex of all the probability distributions they could be associated with. This branded the former material, corpuscular, deterministic conceptions as *illusions* historically fostered by a preoccupation with probabilities very close to 1 or 0. This view did not receive celebrity endorsement.

The emerging dominant middle ground (Bohr, Heisenberg, Born, Pauli, Dirac⁸) did not identify the electron with its associated probabilities but did regard the probabilities as determining everything one could ever know, in principle, concerning the electron. This was an important aspect of what came to be called the **Copenhagen Interpretation**.

As time went on quantum probability amplitudes were applied ever more widely and successfully to the nuclei of atoms, to collections of atoms in molecules and solids, to macroscopic super-cooled liquids and even to the interior of stars. But the increasing successes of the theory only intensified the seriousness of the interpretational puzzles.

From early on all these efforts to understand the quantum world were strongly influenced and guided by the **Uncertainty Relations** discovered by Heisenberg.

4. Heisenberg's uncertainty relations: In late 1926 Heisenberg undertook a systematic study of the brand new, but now essentially completely formulated, Quantum Mechanics, for the purpose of clarifying just what limitations, if any, were really placed on the assignment of values, the measurement of values and the change with time of values of physical quantities in the new theory. It was understood that there were many pairs of physical quantities such that one could not attribute *precise* values to both

members of a pair at the same time. But what about approximate values? Aren't those what we always really use anyhow?! In short order Heisenberg discovered the basic uncertainty relations that bear his name and which quickly came to be regarded as providing the key to the internal consistency of the probabilistic Copenhagen interpretation.

Heisenberg later wrote that fate had prepared him for this task by his having failed a question in an oral exam as a graduate student. The question concerned the resolving capacity of optical instruments. Humbled into studying the subject he came to understand the need to use short wavelength light and large aperture lenses to obtain high resolution images of tiny objects, a version of high precision position measurement (**Fig. 3.8**).

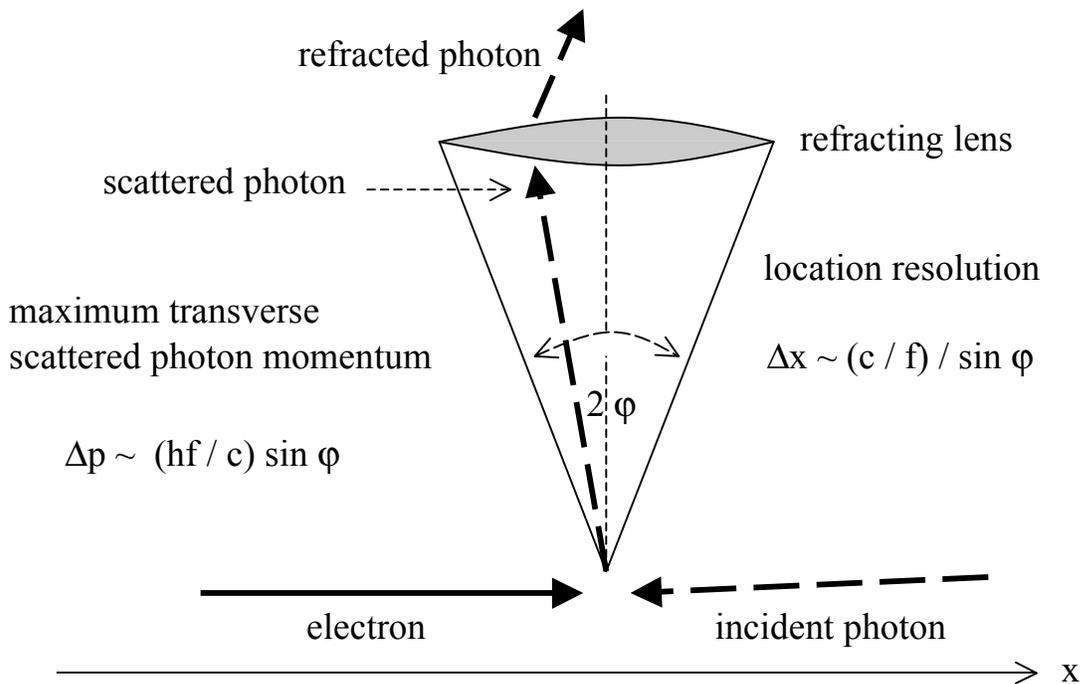


Fig. 3.8: Heisenberg's analysis of using an idealized X-ray microscope to localize an electron. To produce a sharp image indicating where the electron is (was) one must employ high frequency radiation and a large aperture lens. The photon structure of the radiation would then seriously disrupt the electrons motion uncontrollably. To reduce the disturbance to the motion one must settle for poor resolution of the position.

But according to Einstein and de Broglie the short wavelength light was composed of high momentum photons. So the large aperture lens would make it impossible to assess the momentum that was transferred to the tiny object when the photons bounced off it into the lens. So the sharp image of the tiny object would be gained at the expense of losing knowledge about the momentum of the object. One could soften the momentum disturbances by using long wavelength light and small apertures. But then the image indicating where the object was would be fuzzy! This would be the first of many illustrations of how the quantum world seemed to conspire to prevent us from determining the precise values of all the quantities needed to describe the present and predict the future.

The *quantitative* statement of this first uncertainty relation (**Fig. 3.9**) is, that for any Cartesian coordinate axis (call it the x axis),

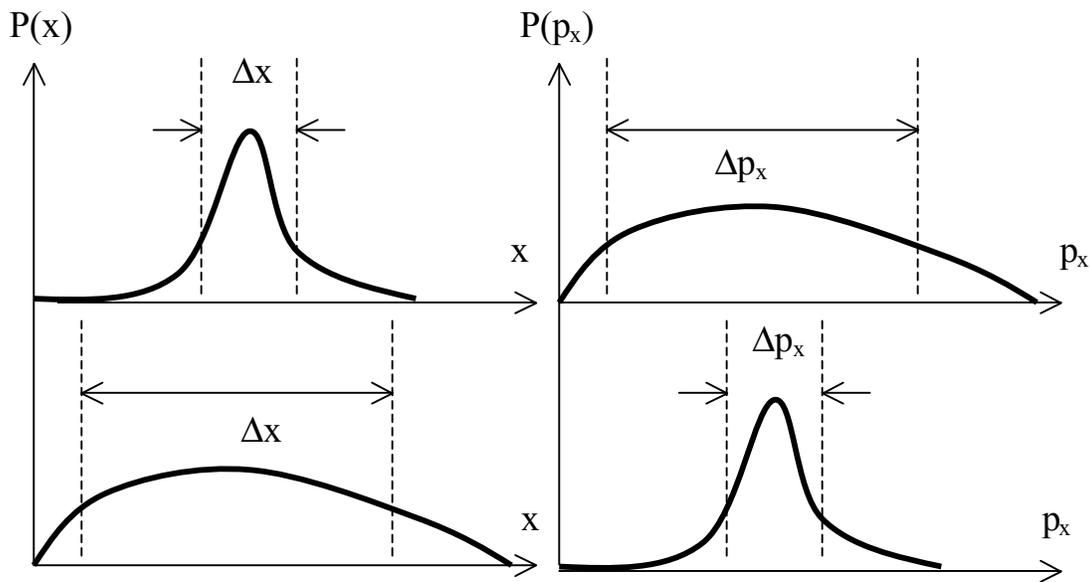


Fig. 3.9: Sample probability distributions for the position and momentum of a quantum particle compatible with Heisenberg's uncertainty relation. In the top distributions, the position is relatively well defined and the momentum is very uncertain. In the bottom distributions the momentum is relatively well defined and the position is very uncertain. The uncertainty relation prohibits both distributions from being narrowly peaked.

$$\Delta x \Delta p_x \geq \hbar / 2 ,$$

where Δx measures the uncertainty in the x-coordinate of the position of interest and Δp_x measures the uncertainty in the corresponding momentum variable. So if Δx is 'small', Δp_x must be 'large', and if Δp_x is 'small', Δx must be 'large'.

There's nothing special about x, of course, and so we also have

$$\Delta y \Delta p_y \geq \hbar / 2, \quad \text{and} \quad \Delta z \Delta p_z \geq \hbar / 2.$$

Now the punch in these relationships did not come from Heisenberg's quaint commentary about using microscopes! No matter how many accounts like that one presented there would always be people who regarded them merely as challenges to be overcome. Most physicists were convinced instead by the fact that Heisenberg had shown that there simply *did not exist in the quantum world any quantum states* (described by de Broglie waves or Heisenberg matrices or Schroedinger waves or Dirac states, etc.) *in which these relationships were violated!*

Nevertheless, for a number of years after these relationships were published many who refused to be intimidated by the mathematics tried to conjure ways to overcome them. Einstein's was the most famous (and clever) effort. Niels Bohr rose to the challenge of trying to show Einstein the error of his ways. After many trips back to the drawing board Einstein finally quit trying, convinced only that he hadn't been clever enough.

Other uncertainty relations (there are an infinite number of them) quickly followed Heisenberg's original ones. Probably the most important ones for our purposes are the angular momentum uncertainty relations and the energy-time uncertainty relations.

The angular momentum relations are,

$$\Delta J_x \Delta J_y \geq (\hbar / 2) | \langle J_z \rangle |, \quad \text{and cyclic permutations thereof}$$

(where $| \langle J_z \rangle |$ is the absolute value of the **average value** of J_z).

This tells us, among other things, that if a system has non-zero average angular momentum component about some specified axis, then it can not have *precise* angular momentum components about *any* axis perpendicular

to the specified one. Conversely, if a system has precise angular momentum component about some specified axis, then it must have zero average angular momentum components about any axis perpendicular to the specified one.

The energy-time relations are obtained from

$$\Delta A \Delta E \geq (\hbar/2) | \langle \dot{A} \rangle |,$$

(for *any* quantity, A and the total energy, E, and where $\langle \dot{A} \rangle$ is the rate of change with time of the average value of A).

This tells us that any quantum state in which a quantity, A, has its average value changing with time must have $\Delta E > 0$, i.e. it can not have precise total energy. Furthermore, it tells us that the length of time, Δt , one must wait before the average value of A changes by an amount equal to the uncertainty in A, ΔA , satisfies

$$\Delta t \Delta E \geq \hbar / 2.$$

Today, while no one challenges the mathematical derivation of the uncertainty relations (which requires a precise definition of the symbol Δ) from the quantum mechanical theory, many people still question the physical interpretation that is placed upon them. It is frequently asserted that physical ‘explanations’ of them, like Heisenberg’s original microscope account, are internally inconsistent since they start by assuming things which they end up proving can not exist. While I agree with this criticism, I do not think it renders the ‘explanatory’ accounts useless. What they really do, I would claim, is to show that the classical *definitions* of the physical quantities lead to internal contradictions in the quantum world. So the old familiar symbols really mean different things now. Closely related things, but different!

Pairs of quantities subject to uncertainty relations which prohibit both quantities from having precise values simultaneously except, possibly, in special circumstances are called **incompatible quantities**.

5. Resolving quantum world paradoxes: The reader may remember that in the first chapter of the readings, **Quantization**, we encountered the seeming

paradox that the possible precise values for the components of angular momentum were mostly incompatible with the Pythagorean theorem applied to the relation between the squared magnitude of the angular momentum and the squares of the mutually perpendicular components. Then again, in the second chapter, **Wave-particle Duality**, we encountered the paradox of states of precise total energy allowing the bound particle to be found where the intrinsically positive kinetic energy would have to be negative; a paradox made initially more virulent by the Born-Pauli probability interpretation discussed in this chapter.

But with our new understanding of uncertainty in the quantum world, in particular, with our recognition of Heisenberg-like uncertainty relations (UR) for incompatible quantities we are finally in a position to resolve these paradoxes and show them to be a consequence of combining *properties* of the quantum world with *presuppositions* of the classical world. Those presuppositions must be given up, however counterintuitive that may seem!

First we consider the angular momentum paradox. According to the quantization rules for angular momentum, if the squared magnitude of an angular momentum is precisely $(3/2)(5/2) \hbar^2$ (which *is* one of the possible, precise values), then the only precise values allowed for any component of the angular momentum are, $\pm (1/2) \hbar$ or $\pm (3/2) \hbar$. Pythagoras's theorem then, applied to these constraints would seem to require that,

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2,$$

or

$$(15/4) = \begin{matrix} (1/4) & (1/4) & (1/4) \\ \text{or} & + & \text{or} & + & \text{or} \\ (9/4) & (9/4) & (9/4) \end{matrix} \quad \text{in some combination or another.}$$

But no combination works.

The angular momentum URs,

$$\Delta J_x \Delta J_y \geq (\hbar / 2) | \langle J_z \rangle |, \quad \text{and cyclic permutations thereof,}$$

comes to our aid by telling us, e.g., that if J_z has one of its possible, precise values, $\pm (1/2) \hbar$ or $\pm (3/2) \hbar$, so that $| \langle J_z \rangle | > 0$, then J_x and J_y can not possess any precise values at all! They are intrinsically indefinite. This means the

Pythagorean theorem can not hold for precise values for everything, but only, in our example, for precise values of \mathbf{J}^2 and J_z^2 , but *average* values for J_x^2 and J_y^2 . Those average values can be *anything* between $(1/4) \hbar^2$ and $(9/4) \hbar^2$, and Pythagoras's theorem is saved!

For the energy paradox - the possibility of finding the particle outside the region where the kinetic energy is positive, as it must *always* be - the resolution is again suggested by URs. Remember that the inference of a negative value for the kinetic energy depended upon the total energy having a precise value. This time the relevant URs are members of a family involving the energy. For any quantity, A, we have,

$$\Delta A \Delta E \geq (\hbar/2) |\langle \dot{A} \rangle|,$$

where \dot{A} denotes the time rate of change of A. In other words, neither the total energy, E, nor A, itself, can be definite and precise if the value of A is, on average, changing with time. For the particular case in which $A = x$, the position of the particle, we have,

$$\Delta x \Delta E \geq (\hbar/2) |\langle v \rangle|.$$

where v is the velocity of the particle. So if we localize the particle under circumstances where its moving, on average, the total energy can not be definite and the kinetic energy need not be negative. Again if $A = p$, the momentum of the particle, we have,

$$\Delta p \Delta E \geq (\hbar/2) |\langle F \rangle|,$$

or the energy and momentum must both be indefinite if the quanton is localized where the average force, $\langle F \rangle$, is non-zero, since the force is the time rate of change of the momentum. The force usually depends on the position and if we localize the particle *where* the force is non-zero, on average, the energy must be indefinite.

There is a loophole in all this. Might we not localize the particle in a manner and in a location where both the average velocity and the average force are zero? If so, our URs don't require an indefinite energy. This is a possibility in some situations. Nevertheless, our URs suggest that localizing the particle

tends to render the energy indefinite and that holds generally even though we can't seem to capture it securely in any single UR.

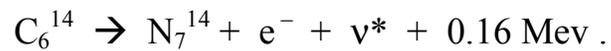
Footnotes

1. Werner Heisenberg (1901 – 1976)
2. Ernst Mach (1838 – 1916)
3. Max Born (1882 – 1970)
4. Henri Becquerel (1852 – 1908)
5. Marie Curie (1876 – 1934)
6. Frederick Soddy (1877 – 1956)
7. David Bohm (1917 – 1992)
8. Paul Dirac (1902 – 1984)

Appendix A: Carbon dating

As we will see in more detail later, the chemical properties of the atoms of an element are determined by the number and arrangement of the outer electrons surrounding the nuclei of the electrically neutral atoms. That number and arrangement is, in turn, determined by the number of positively charged protons in the nuclei which is fixed for each element. But the nuclei also contain neutrons which have next to no influence on the chemical properties and the number of which can vary within limits. For a given element, the atoms with different numbers of neutrons in their nuclei are called **Isotopes** of the element.

Many elements have isotopes that are spontaneously radioactive. One such biologically important element is Carbon and the isotope is C_6^{14} with 8 neutrons and 6 protons in the nucleus. The most common and stable isotope is C_6^{12} with 6 neutrons and 6 protons in the nucleus. In the atmosphere the ratio of the occurrence of these two isotopes is about 1 to 10^{12} and this is the ratio found in living organisms. But once the organism dies and is no longer metabolically exchanging Carbon with the environment the C_6^{14} in the remains starts to disappear via the decay process,



This decay converts the Carbon nucleus to a Nitrogen nucleus and emits an electron, e^- , and an anti-neutrino, ν^* , with kinetic energy of 0.16 million electron volts (Mev). It has a half life of 5730 yrs, so if one finds remains with half as much ratio of C_6^{14} to C_6^{12} as a living organism would have, one infers that the remains come from an organism that died about 5730 yrs ago.

Carbon dating is accurate to about 10 half lives, i.e., about 58,000 yrs. Beyond that the concentration of C_6^{14} is rendered unreliable by non-metabolic deposition of C_6^{14} from the environment.