Space, Time and Quanta

An Introduction to Contemporary Physics

Robert Mills

The Ohio State University

W. H. Freeman and Company
New York

Introduction

I.1 THE PICTURE GETS FUZZY

The world of nineteenth-century physics was a world of certainties. It consisted entirely of matter—solid, liquid, or gaseous. For philosophers, solid matter was the very epitome of the "real": A thing was real if you could grasp it or hurt your toe kicking it, and abstractions like love and beauty—and energy for that matter—had a most doubtful claim to reality. Matter was everything. The idea of a force field such as the electric or magnetic field, something existing at each point in space and causing forces on electric charges or currents, had been developed by Michael Faraday and James Clerk* Maxwell, but such a field was still understood very much in mechanical terms, as the displacement or movement of a

*Pronounced like "Clark."

material medium, the *ether.* This ether was supposed to fill all of space, even where it was completely empty of any ordinary kind of matter—what you'd call a perfect vacuum. To a present-day physicist, Maxwell's theory of the electromagnetic field seems to give a powerful description of something completely *non*material whose reality has nothing at all to do with matter as we know it, and yet Maxwell himself was apparently unable to understand his fields except in terms of a material ether.

A few other abstractions were beginning to enter physics—and were causing distress in the process. Two important examples are *energy* and *entropy,* which were key ingredients of the developing theory of heat, known as thermodynamics. It was becoming clear in the last century that the amount of energy in the universe is conserved—energy can't be created or destroyed—but it's very hard to hold it in your hands and say, "This is something real." Entropy, too, which measures the amount of disorder in the universe and plays an important role in the irreversible flow of time, was even more difficult than energy to understand as something real.

The behavior of matter was then believed to be governed with complete precision and complete predictability by various "laws of motion." The motion of physical objects was described by Newton's laws, which dated from the seventeenth century, whereas the electric and magnetic force fields, conceived as I said entirely in terms of a material ether, were elegantly described by Maxwell's equations, developed in the 1860s. Chemical and thermal processes were not so well understood, but here also scientists felt they were well on the way to an equally complete—and equally mechanical—description.

Space and-time were believed to be absolute, and the flow of time was thought to move the universe and the human race forward from one moment to the next at a universal and inexorable pace. Isaac Newton's view still held in the 1800s: "Absolute, true and mathematical time, of itself, flows equably without relation to anything external ...." The laws of physics, the very general statements that describe the behavior of matter and the forces between different bits of matter, were being refined and verified with ever-increasing precision, until it seemed that the only remaining task was to measure the constants of nature to more and more decimal places—and to tidy up a few loose ends, a few little nagging inconsistencies.

In this century we have tugged at the loose ends—and have seen all of nineteenth-century physics unravel.

The giants in this enterprise are clearly Albert Einstein and Niels Bohr, though many others played key roles in tearing apart the old picture and building a new one. Matter has been supplanted by fields as

the more fundamental concept, time has lost its unique and absolute status, and our faith in physics itself as an exact science has been badly shaken by the fuzziness and indeterminacy of quantum theory. The fact that physicists are still doing physics, however, is not simply a tribute to their pigheadedness but is rather due to the extremely high level of exactness and predictability that still remains, despite the fuzziness, when we probe the behavior of things in the proper fashion—when we ask the right kinds of question.

If you don't examine it too closely, the new picture looks just like the old one. When you do look closely, it's like a newspaper photograph, with the thousands of dots that make it up corresponding only roughly to the picture you thought you saw. In the case of physics, our close look reveals a structure that's radically different from the nineteenth-century picture, and that seriously violates nearly all our common-sense impressions of how things behave. These common-sense notions, of course, like the nineteenth-century picture of the physical world, correspond to the newspaper photograph viewed from a distance. One of things I want to do in this book is explain, in terms of the new physics, why the old-fashioned picture works so well—and I hope you will wonder, as I do, why in fact the little dots arrange themselves into a picture at all.

I.2 CENTRAL THEMES

What are the most distinctive characteristics of twentieth-century physics? I would say the loss of objective certainty, the fundamental role of symmetry, and the movement toward unification.

Conceptually, we've seen a loss of absolutes and an increasing appreciation of the important role of the observer in defining reality. We've seen the loss of complete predictability—a hallmark of classical physics—and we've seen a peculiar fuzziness in how well we can describe the state of a physical system. We have also seen a greatly increased level of mathematical abstraction in the theoretical description of nature: The mathematical structures we use often have no direct physical meaning, and sophisticated rules of interpretation are required before they can be related to experimental observations. We've had to deal with changes like these both in Einstein's theories of relativity and in quantum theory.

In terms of practical knowledge, we've probed both the infinitesimal and the infinite—the microscopic and the cosmic—and in both directions have found structures unlike anything that scientists of the nineteenth century could have imagined: exotic objects such as quarks and gluons in the small, and curved spacetime, black holes, and the Big Bang in the large. The theories we've been developing to describe these structures embody all the characteristics mentioned: loss of objectivity, laws of physics governed by symmetry requirements, and always the pursuit of unity.

I want to elaborate now on these central themes, which run through all the parts of this book.

Loss of Objective Certainty

The relativity of truth in Einstein's view of space and time and the uncertainties associated with quantum theory have led to some serious soul-searching about the nature of the *real.* We are no longer sure what it means for a thing to be real, and we don't know what it means to *know.* Indeed this epistemological need, this need to understand what it is to "know" a thing, has become more urgent as the observer has taken center stage in physics. Facts seem not to have an independent objective existence but rather are mediated in every case through an observer and are stated and known only relative to that observer. Absolutes remain, however. At the immediate level it's relationships that we deal with, but the description of these relationships (as far as we can tell) has absolute validity! Only relative velocities have meaning, for example, but we can make firm statements about the different relative velocities in some experimental situation—statements that are true from *every* point of view.

Einstein insisted, and we've come to believe him, that space itself—the space through which the earth and the stars move—cannot be thought of either as being at rest or as having any definite state of motion. Not only has it become impossible to answer the question "How fast is the earth moving through space?" but the very question no longer has any absolute meaning. Even the gravitational force, although it seems such a real part of our everyday life, appears now to depend on your point of view. For example, the pull of gravity seems to vanish when you're in a state of free fall—in an elevator whose cable has snapped, for example, or in an orbiting spacecraft. (Einstein claimed that the point of view of someone in the falling elevator, in which it appears that the vehicle is at rest and that there's no gravitational force at all, is just as valid as the point of view in which there *is* a force and the vehicle, along with its contents, is accelerating downward because of that force.) The sense that we have lost our solid foundations and that everything is *relative* was taken by many people as applying to other things besides physics, sending ripples of both apprehension and liberation through

our Western culture and elevating Einstein himself to a kind of mystical eminence. Note that this kind of cultural use of the idea of relativity is very vague and has no foundation in physics. And note too, as I said just now, that certain kinds of statements in physics *do* have absolute validity.

Quantum theory too has left a trail of insecurity and uncertainty as it has developed in the twentieth century. Even more than relativity, it has shaken our belief in the possibility of objective knowledge about reality—and our belief in that reality itself. It has robbed physics, at a very fundamental level, of its power to make precise predictions about future events. I think the shaking of our beliefs has taken place only at the philosophical level (though I think that's important enough). For one thing, these uncertainties and unpredictabilities become negligible in the world of big objects like people and machines. In addition, we find we can make precise and reliable statements *about* the uncertainties, which enable us to do high-precision physics as long as we ask only meaningful questions. For example, we've discovered that it's not meaningful to speak of the position of a particle in the absence of a direct observation but that it *is* meaningful to inquire into the *probability* of finding it at one location or another. The laws of quantum physics, in fact, cannot predict events; they are reduced to predicting the *odds* that different events might occur.

This replacement of certainties by probabilities is basic to our quantum world. It's not like a coin toss or a roll of dice, which is unpredictable simply because you don't have precise enough information about what is in fact a completely determined motion. In quantum physics there doesn't really seem to be *any* real objective state of affairs at a time when a system is not under observation. Different possible sequences of events, even though they're mutually exclusive according to classical physics, all contribute to the final outcome; that is, *all* the possible histories leading up to the moment of observation play a role in determining the probabilities for that observation. This interference among the different possible histories is in fact closely related to the interference of classical waves such as sound waves: There are "dead spots" and "live spots" in an auditorium because the sound waves from the performer follow different paths as they bounce off the walls of the room and can either cancel or reinforce each other as they arrive at the location of the listener.

Much of the early shaping of this new way of looking at the world and much of the battling for this new view were done by Niels Bohr and the group that he gathered around him at his institute in Copenhagen in the early years of this century.

Symmetry

Another thread that runs through the physics of this century is the notion of symmetry. From being an accidental property of certain things, nice but not of fundamental importance, symmetry has come to play a central role in physics, apparently controlling the very structure of the laws of physics and the number and character of the elementary particles of nature. I need to explain what I mean here by symmetry. You can speak of the symmetry of an object, or you can speak—as we do in physics and as I'll explain in a moment—of the symmetry of physical laws. It is the symmetry revealed in the *laws* of physics that has become so important in recent years.

We say that an object is symmetrical if it has a certain *invariance* property—that is, if the object is left unchanged by some operation on it. A cylinder, for example, is invariant under rotations about its axis, and a sphere, showing a greater degree of symmetry, is invariant under rotations about any axis through its center. (Each of these examples has also a *reflection symmetry:* It looks the same in a mirror.) [What other examples can you think of? Give an example of an object that has no symmetry at all.]

The symmetries that I need to talk about are the symmetries of the laws of physics themselves. For example, you may repeat some experiment in different locations or in different orientations, and you'll find that everything behaves in exactly the same way—according to the same rules. You can even look at things from a moving frame of reference. For example, if you get on a high-speed train and try to walk or play ping-pong, you find that the same rules apply as if the train were at rest (assuming no vibration, of course). The reflexes and muscular skills that you learned on the ground work just as well as ever. An even more striking example is the fact that we have no sense at all that (in old-fashioned terms) the earth is moving through space at a prodigious speed (about 20 miles per second relative to the sun and around three times faster than that relative to the center of our galaxy). The laws that govern the motion of our bodies and the objects around us work exactly the same as if we were at rest. In physicists' language, the laws of physics are invariant with respect to a variety of possible choices of reference frame.

The first indication of the power of symmetry to dictate physical laws was presented in 1905 by Einstein, who overthrew the nineteenth-century understanding of space and time and rewrote the laws of physics on the strength of a symmetry principle, the *principle of relativity.* His vision was so clear, and so beautiful, that after the initial shock scarcely anyone doubted that he had it right. This principle of relativity is the requirement referred to earlier:

> *The laws of physics—including the behavior of light—must be exactly the same for any two observers moving with constant velocity relative to each other.*

It's called the principle of relativity because it makes the idea of any absolute stationary reference frame meaningless in practice and says, rather, that we can describe events only relative to some observer and that all uniformly moving observers are equivalent. We continue to say that things are "stationary" or "moving," and it's usually clear what we mean, but to be precise we have to specify the reference frame *relative to which* a thing is stationary or moving. This principle leads to drastic changes in the description of nature devised by Newton. The new description goes by the name of the *special theory of relativity,* or just *special relativity.*

It turns out, surprisingly, that when we relate the views of two observers moving relative to each other, we find that the dimensions of space and time get mixed—that what looks like pure space to one observer looks like an admixture of time and space when viewed by the other. This mixing is closely analogous to what happens when the coordinate axes are rotated on a plane or in ordinary space. As a consequence, we have to think of space and time together as a single four-dimensional space (three dimensions of space, plus time), for which we use the word *spacetime.*

Einstein's vision of the deep equivalence of different points of view led him to pursue even further the idea that the laws of physics should appear exactly the same to all possible observers: He now insisted on including observers who are accelerating or rotating. This further step was suggested by the fact that the inertial effects of acceleration or rotation—like the forces experienced by astronauts during blast-off and like the centrifugal force that throws you to one side during a sharp turn in a car—are indistinguishable from gravitational effects. That is, you can attribute such effects to your own motion or, equally well, to the presence of different gravitational forces.

Einstein's inspiration was to extend this simple kind of equivalence to all kinds of motion and to state it as a universal rule, referred to as the *principle of equivalence:*

The laws of gravity must be such that the apparent forces due to any possible kind of motion are indistinguishable from gravitational forces.

This would not make sense if it weren't possible for gravitational forces to change with time and (as you'll see) to depend on the velocity of the object they're acting on. This vision led Einstein, by a difficult and tortuous path, to a general theory of gravity that not only far transcends Newton's in power and sophistication but also leads us to a new view of space and time as *curved*. We live apparently in a curved universe, and the forces that we attribute to gravity turn out not to be forces at all but simply a necessary consequence of that curvature. It's as if you were a two-dimensional creature sliding smoothly along some irregular surface with dimples distributed here and there and were being deflected in your path every time you passed near one of the dimples.

This curved universe can be said to have a shape, either open and infinitely extended or closed like the surface of a balloon—we don't know which. Every star or planet sits in the middle of one of the dimples, and the dimple can become infinitely deep in the case of a star that has collapsed and become a black hole.

The theory that describes this curved spacetime is called the *general theory of relativity,* or just *general relativity.* It gives to spacetime itself an active, dynamical role, and it foreshadows, in a curious way, the modern concept of gauge fields (see below). Note that general relativity is again based on a symmetry principle—a requirement of invariance: the insistence that the laws should be the same for all observers, including even those in accelerating and rotating frames of reference.

The next step, the beatification of symmetry, you might say, was Noether's theorem (1918), which gives a relation between symmetries and conservation laws in physics. A conservation law is any rule, derived from the basic laws of physics, that says the total amount of some quantity is constant and doesn't change with time. A notable example is energy, which can be neither created nor destroyed but only transformed from one form to another. What Emmy Noether* showed (with some exceptions that don't concern us here) was that for every symmetry of the laws of physics there is a corresponding conservation law. We now know also, though it wasn't part of her original theorem, that the converse is true: Every conservation law must be associated with a corresponding symmetry. The proof of the theorem depended on the fact that all the laws then known were of a certain type—what we call derivable from a Lagrangian. It is astonishing to me, given the tremendous changes that have been wrought in physics since 1918, that Noether's theorem still applies and has become, in fact, much more general than it was to begin with. It's so easy to invent universes where it doesn't apply that I take it to be a fundamental *principle* of our universe, rather than just a theorem, that symmetries and conservation laws go hand in hand. For this reason I would like to promote Noether's theorem to *Noether's principle:*

The laws of physics must be such that every symmetry of nature corresponds to a conservation law, and vice versa.

The next stage—what could be called the canonization of symmetry—was the discovery during the last 35 years that many, and maybe all, of the laws of physics themselves can be generated from symmetry principles. It now seems that all the interactions of physics are caused by a special kind of field called a *gauge field,* whose structure and behavior are completely dictated by a new symmetry requirement: the requirement of local symmetry. What's a local symmetry? Recall that symmetry, as I've been talking about it, is the equivalence of the laws of physics from different related points of view—broadly speaking, different frames of reference. The laws show a *local symmetry* if the equivalence persists even when you choose a different point of view at every point in space and at every possible time.

It seems likely that every continuous symmetry of nature is associated in this way with a gauge field, and in each case the conserved quantity corresponding to that symmetry (by Noether's principle) is exactly the thing that interacts with the gauge field. The familiar fact that gravity interacts with mass is an important example; mass is a form of energy, and energy is one of the conserved quantities related to the symmetries of space and time. (In fact, gravity interacts with all forms of energy, though often in a less obvious way than we're used to.) The electromagnetic field interacting with electric charge is a second example, and finally we have several new kinds of gauge field, interacting with exotic new kinds of "charge," that extend the idea to include all the interactions we know of.

Unification

The third characteristic—not new to this century, to be sure, but representing more and more a major theme of physics—is *unification.* Unifica-

*German, Nöther: pronounced like "neater" but with the lips pursed as for "ooh."

tion means showing that what may look like very different areas of physics are in fact governed by the same fundamental laws, or showing that what look like a large number of different particles are in fact made up of a small number of fundamental ones. We have seen unification in previous centuries: when sound and heat, for example, were identified as mechanical motion at a microscopic scale, governed by the same mechanical laws that Newton had devised for larger objects; and again when light was found to be an electromagnetic wave, governed by exactly the same equations that describe static or slowly varying electric and magnetic fields.

In this century, we've seen the unification of space and time, which are understood now as being both of the same fundamental character. We have seen the further unification of spacetime with gravity, which is no longer thought to be an arbitrarily added force but is believed to be the result of spacetime itself being curved. We have seen all the matter and forces that make up the world as we know it unified through the ideas of *quantum field theory:* All of the forces are directly due to fields like the gravitational and electromagnetic fields, whereas each of the particles, both familiar and exotic, that populate the universe is seen as a kind of quantum "ripple" of one of the fundamental fields. In the case of the electromagnetic field, these ripples, or *quanta* (the singular is *quantum*), of the field are called *photons.* Everything is now fields. (Gravity is still the troublemaker—like a field but somehow different.) I have to admit that it's hard to describe just what a field is; I'll try to do so in §6.2 and §11.5. This idea of a field is just one example of the increased abstractness of the mathematics that we need to use to describe the physics.

Gravity, incidentally, is just one example of our constant struggle with semantics. The words we use to describe phenomena reflect the mode of thinking that we're in at the moment. Many times it's completely natural to speak of a "gravitational field," a force field pervading space and producing the familiar gravitational forces, but at other times we need to use the geometrical picture of a curved spacetime, in which the idea of a gravitational field has no place. The actual physical effects are the same, but the language is very different.

We have seen the different elementary particles and fields of nature fall into families, forming patterns whose symmetries reflect the fundamental symmetries I mentioned. The electric and magnetic fields are related in this way to each other, and the combined "electromagnetic field" is in turn related, by further symmetries, to the fields responsible for the weak interactions—the interactions that produce certain kinds of radioactivity. The proton, the neutron, and the pi meson, members of

the family of strongly interacting particles called *hadrons,* are now known to be composed of particles more elementary still, known as *quarks,* fewer in number and related again by appropriate symmetries. We believe now that there are six types of quark and six types of *lepton* (the electron and its family of related particles), along with three or four types of gauge field that are responsible for all the forces of nature. Physicists dream of removing the apparent arbitrariness of these different families and of deriving them all from a single grand group of symmetries. Many schemes have been proposed for performing this ultimate unification, but we're not there yet. I believe myself that a radical revolution at the deepest conceptual level may still be necessary before the pieces all fit together.

Part I of this book belongs to Albert Einstein. A separate introduction to Part I would make too many introductions before we get down to work, so I hope you will glance again at the various references in this introduction to Einstein and to his vision of a universe with no fixed framework and no meaningful distinction between "moving" and "stationary" as applied to both objects and observers. I hope to make clear in what follows what fueled this vision and how far it has led us.

CHAPTER 13

The Historic Principles

It's important to remember that in the early years of quantum theory, there was a lot of confusion and argument among physicists. The universe at the microscopic level was behaving in an impossible way, and there seemed to many to be no hope of reconciling the contradictory features that were emerging. Light could *not* consist of particles and still show the familiar interference effects that signal the presence of waves; neither could electrons be waves and still have the sharply defined momentum and energy that characterize particles of definite mass; and surely the laws of physics could not have given up their allotted function of predicting the future with arbitrary precision once initial conditions are known. Yet all these things appeared to be happening.

During this period, people were trying very hard to make some kind of sense of this new quantum world, and they achieved many deep and valid insights even before they managed to develop a proper mathematical quantum theory. Some of these insights took the form of "principles," which served as a guide to living with the contradictions; they did not in themselves constitute anything like a true theory, but they enabled people to do physics in many of the situations ordinarily encountered. In a sense these principles represent a distillation of the most distinctive features of the new physics: characteristics that would have to be—and indeed became—part of the proper theory when it finally appeared.

In this short chapter I want to introduce the most notable of these historic principles, reminding you though of their *preliminary* character. From our present perspective, these are no longer fundamental principles or axioms but are rather consequences or features of a fully developed theory.

13.1 CORRESPONDENCE

The *correspondence principle* is the requirement that, when a physical theory is overthrown, the new theory must correspond to the old theory—that is, the new and the old must agree—*in those realms of physics in which the old theory has been tested.* Bohr formulated this idea as a guide in the development of quantum theory, but it is pretty obvious that it's much more general in scope. You saw this logic in Part I when we talked about relativity, which clearly has to agree with nonrelativistic physics when velocities are small compared to c. The point now is that the new quantum theory—whatever theory that turns out to be—must agree with classical physics for large-scale events (baseballs, planets, and so on).

How can this be? Is a baseball supposed to have a wavelength and to exhibit interference fringes when you throw it through two holes in a wall? Are sound waves supposed to be quantized and hit the ear in a little bang as each quantum of energy arrives? When you think about it, you find that there's no real contradiction of our experience, basically because Planck's constant, the fundamental constant that describes all quantum effects, is so very small. In fact, $\hbar$ is about 10^{-34} J · s (eqs. 10.3.5 and 12.3.26), and its smallness relative to everyday units is much more extreme, for example, than the bigness of c, 3×10^8 m/s. That is to say, we expect quantum effects to be *much* less noticeable in our everyday experience even than relativistic effects.

To see how this works in practice, consider the baseball. Its wavelength is given by the de Broglie relation (eq. 10.4.1), $\lambda = h/p$, which means that if $p = 1$ in standard units (kg · m/s) then the wavelength is 10^{-34} m, which is far smaller than an atom (10^{-10} m) or even a proton

(10^{-15} m). [A thrown baseball might have $p = mv \approx (0.1\text{ kg}) \times (30\text{ m/s}) = 3\text{ kg} \cdot \text{m/s}$, which doesn't affect the discussion—a factor of 10 more or less hardly matters.] What's the meaning of a wavelength this small? It can produce no observable effect. For one thing, it sets the scale of any possible interference pattern, which is therefore far too tiny to see even with the fanciest instruments. For another, it means that the fuzziness of a wave pulse, which is the fuzziness in the location of the ball at any given time, is also unobservably small. The trajectory of the ball has to be correctly predicted also—this issue is addressed in more detail in §14.3.

Now, what about classical waves? If you're dealing with a wave at the everyday level, with a wavelength of the order of 1 m (or 1 mm or 1 km—it doesn't matter), then it may consist of quanta, but the momentum and energy of the individual quanta are at the scale of 10^{-34} in everyday units and far too tiny to be observable. Any ordinary wave would have to comprise ~ 10^{34} quanta per second, and only the collective effect of all of them, purely wavelike, would be observed.

This discussion shows that the quantum nature of matter doesn't cause any problem in principle with our everyday experience, but it doesn't show in detail how quantum theory can reproduce the detailed *laws* of macroscopic physics, Newton's laws of motion and Maxwell's laws for EM fields in particular. This is one of the things you have to look out for when you start considering an actual mathematical quantum theory. I think it's amazing that a theory can be so radically different in its basic character from our familiar classical physics and still reproduce that classical behavior in all its detail.

PROBLEMS

1. **(a)** Explain why, in looking for photons in EM waves at the everyday level, it doesn't make much difference whether you use wavelengths of millimeters or kilometers.
   (b) Find the number of photons per second in a 1-mW beam of **(i)** 600-nm visible light waves, **(ii)** 500-m radio waves, and **(iii)** 1-nm x-rays.
2. Estimate the mass of a dust particle (say the density is comparable with water, 1 g/cm^3), and find the velocity at which its de Broglie wavelength would be around the diameter of a proton. Does the wavelength get bigger or smaller as the velocity increases?

13.2 UNCERTAINTY

Werner Heisenberg, one of the pioneers in the development of quantum theory, devised the *uncertainty principle* as an explicit example of the loss of classical certainty. The "uncertainty" referred to is not just human uncertainty but intrinsic uncertainty built into the laws of physics: There are physical variables that simply cannot be known simultaneously with complete precision. In its original form, the uncertainty principle refers to a particular example of this kind of incompatibility, that of the position and momentum of a particle. The reason for this incompatibility is simple: The momentum of a particle is related to an associated wavelength by the de Broglie relation, eq. 10.4.1, and your knowledge of that wavelength gives a measure of the fuzziness in your knowledge of the particle's position.

The principle is stated (in one dimension for now) in terms of Δx and Δp, which are defined as the indefiniteness of the particle's position and of its momentum, respectively. The Heisenberg relation is an inequality that says that Δx and Δp can't both be made arbitrarily small:

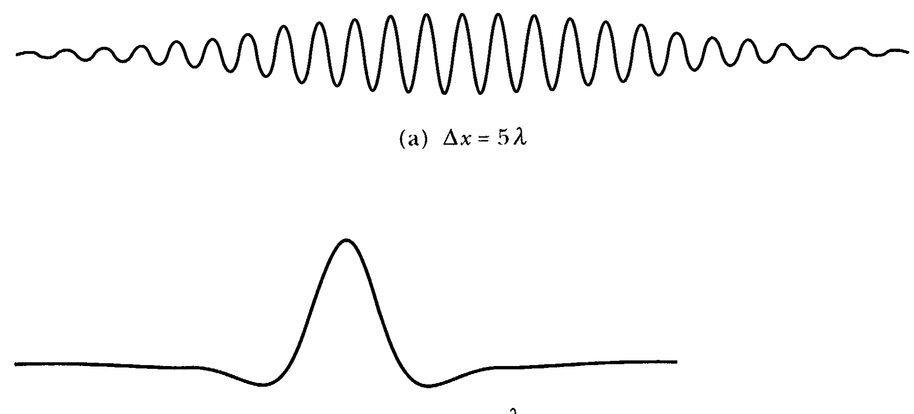
$$\Delta x \Delta p \gtrsim h \qquad (1)$$

That is, if x is known—or measured—to a precision Δx, then p cannot be known—or predicted—to a precision better than about $h/\Delta x$. Equivalently, if the momentum is known to a precision Δp, then the uncertainty in position can be no smaller than about $h/\Delta p$.

When Heisenberg presented this result in 1927, he made no use of the wave character of matter (which he opposed at that time) but looked at the problem entirely in terms of the process of observation itself. His approach exploits the dual nature of light, using the Planck relationship, but avoids admitting the dual nature of electrons. A rough version of the argument is this: Suppose you use light to observe the position of an electron. The picture is fuzzy because the wavelength of the light sets a limit on how precisely you can locate the electron, so the uncertainty in the position of the electron, Δx, will be of order λ. The photons have momentum h/λ, or $h/\Delta x$, and the direction of the photon after it bounces off the electron is uncertain (because of the aperture of the observing lens). The photon therefore gives the electron an uncontrollable impulse (transfer of momentum) of that order, making the electron's momentum uncertain by that amount. Soft, gentle photons (photons of very low energy) have long wavelengths and produce very

fuzzy pictures, while short-wavelength photons, the kind you need for sharp pictures, are very energetic and shake the electron in an uncontrollable way.

It turns out that eq. 1 can be made precise by defining the uncertainties Δx and Δp more accurately and by making use of basic properties of matter waves, along the lines of the argument mentioned in the first paragraph of this section. The momentum of a particle is associated with the wavelength of the associated wave, so a particle of well-defined momentum has to have a wave of definite wavelength. On the other hand, the wave associated with a well-localized particle, whose position is sharply determined, must be a localized wave—a wave packet—spread out only over a region of length Δx, as shown in the examples in fig. 1. Note that how well the wavelength is determined depends on how large Δx is compared to the wavelength λ. In fig. 1(a), Δx is much larger than λ, and you can see that λ is quite definite, though not completely so. The indefiniteness in λ is associated with the overall shape of the pulse: To get a localized packet, you have to superpose waves of slightly different wavelength, in the spirit of §11.3, in order to get them to cancel outside the region of the packet. In fig. 1(b), Δx is small, and it's much harder to make a good estimate of λ. The relation between Δx and the uncertainty in λ is a theorem in Fourier transforms (the theory of the analysis of arbitrary wave shapes in terms of pure sine waves), which is expressed



(a) $\Delta x = 5\lambda$

(b) $\Delta x = \frac{\lambda}{5}$

Fig. 1 Wave pulses.

most conveniently in terms of the angular wave number $k = 2\pi/\lambda$ (eq. 11.1.2). If Δx and Δk are defined as the standard (or root mean square) deviations* of x and k, respectively, from their mean values, then they must satisfy the exact inequality

$$\Delta x \, \Delta k \geq \frac{1}{2} \tag{3}$$

It is then a simple consequence of the de Broglie relation (10.4.1),

$$p = \frac{h}{\lambda} = \hbar k \tag{4}$$

that

$$\Delta x \, \Delta p \geq \frac{\hbar}{2} \tag{5}$$

the exact statement of the uncertainty principle. Note that this is an *inequality.* You can be *more* uncertain than this about both x and p but not less. In fact, in most normal experimental situations there is much greater uncertainty than this; remember that $\hbar$ is a very small quantity.

The variables x and p of the preceding discussion are referred to as *complementary variables.* In fact, every variable in physics is found to be one of a pair of complementary variables, related to each other by an uncertainty relation similar to eq. 5. (One example is angular momentum and angle of orientation in the case of rotational motion.) In a three-dimensional world, the position and momentum of a particle are vectors. The uncertainty principle applies to only one component at a time, because it is only p_x that is complementary to x, and so forth. There is no such restriction relating Δx and Δp_y for instance.

*The standard deviation of a variable is the square root of the mean of the squares of the deviations. That is, if y_i, $i = 1, 2, 3, \ldots$, are a set of measured values of y, and the true average value is $\bar{y}$, then the standard deviation is given by

$$\Delta y = [\langle (y_i - \bar{y})^2 \rangle_{av}]^{1/2} \tag{2}$$

You should examine carefully how the formula reflects the verbal definition.

PROBLEM

1. (a) Find the minimum momentum uncertainty Δp for an electron confined to a region of length 1Å (about the size of the hydrogen atom), and compare it with the momentum of an electron in the first Bohr orbit (§12.3).
   (b) Make crude estimates of experimentally reasonable limits on Δx and Δp (work from Δv, the uncertainty in speed) for a thrown baseball, and compare it with the limit imposed by the uncertainty principle. Work out a few additional examples.

13.3 COMPLEMENTARITY

The principle of *complementarity,* which, like the correspondence principle, we owe to Bohr, is much more qualitative in character than the others discussed here and more difficult to use in a systematic way. A generalization of the ideas leading to the uncertainty principle, the complementarity principle attempts to confront the contradictions of quantum theory head-on and, in a sense, to formulate the fact of contradictions as a principle. The essence of the principle is that two different kinds of experiment, or two different kinds of physical variable, can be *complementary** in that they are mutually exclusive (incompatible), but *both are needed for a complete picture.*

The prime example of this is wave–particle duality itself. The wave properties and the particle properties of light are complementary, then, because you need both to give a complete description but you can't observe both at the same time. When one photon leaves its mark on a photographic film or produces a photoelectron, there is no trace of the wave properties—no extension in space, no observation of wavelength or frequency. When you tune your radio to a particular station, on the other hand, you are sorting out from all the waves in the atmosphere just those of a particular frequency, by allowing them to produce a resonance in an electrical circuit that will oscillate at just that frequency and no other—a process that seems entirely meaningless in terms of localized particles of radiation.

*Not *complimentary.* Two things *complement* each other when each supplies something that the other lacks.

The kind of observation you make determines which aspect of the radiation is relevant, and, in fact, complementary observations invariably obstruct each other. An example of this can be seen in the two-slit interference experiment of §10.5. You saw there that you could observe the photons at the wall where their distribution shows the effect of interference between the two slits, *or* you could observe them close to the slits where what you're determining is which slit the photon goes through. These again are complementary observations, and each is incompatible with the other. No matter how gently you do it, if you succeed in determining which slit the photon went through, you destroy the possibility of interference; conversely, if you do see interference, you have to have no knowledge at all of which slit the photon went through.

The early practitioners of the new quantum theory became masters of exploiting complementarity to get at physical results during the period when there seemed to be no satisfactory interpretation of the theory. The scattering of x-rays from electrons (the Compton effect) is an example of this: You employ wave interference to select x-rays of a definite wavelength and frequency; you convert this information into momentum and energy for a single quantum; you use relativistic particle dynamics to analyze the scattering of the two particles, the photon and the electron, and deduce the momentum of the scattered photon; and finally you shift gears again and convert this back to a wavelength, which is measured by means of wave diffraction from a crystal lattice as described in §10.4. The exact wavelength shift of the scattered x-rays is impossible to explain in terms of classical waves, but it comes out naturally as a momentum transfer in the photon picture.

Again, we must watch and see how the proper theory, as it is now interpreted, deals with this—how it can preserve the *effects* of complementarity while removing the contradictions.

13.4 SUPERPOSITION

My final example of an historic principle, the *superposition principle,* extends the phenomenon of interference to a broad generalization about the quantum states of any system. You've seen (§10.5 and §10.6) how interference is a consequence of the way waves can be combined by simple addition (§11.3), so that where two waves have the same sign they reinforce each other and where they have opposite signs they cancel. The key here is the word *addition,* and it runs very deep—deeper even than the wave concept itself.

It works like this: Whenever a system can be in one or the other of two states, call them A and B, then it is possible for it to be in a state that is a combination of A and B, called a *superposition,* with properties analogous to those of a combination of waves. In such a combined state, if you do an experiment that tells whether the system is in state A or in state B, you will definitely get one answer or the other, with a certain probability for each, just like measuring which slit the photon went through in the two-slit experiment. On the other hand, if you do some other experiment complementary to the A-versus-B experiment (such as looking at the fringes in the two-slit experiment), then A and B can interfere—constructively, destructively, or in a variety of other ways—just as if A and B could somehow be added together to get the resulting state.

What this is giving us is an idea of the *state* of a physical system that is radically different from anything we've seen in classical physics. To describe the state of a classical particle at a certain moment of time, you give its position, $\underline{r}$ and its momentum $\underline{p}$, and a different state would have a different set, $\underline{r}'$ and $\underline{p}'$. How could you combine these states to get a third? If you tried, for example, to add the variables $(\underline{r} + \underline{r}', \underline{p} + \underline{p}')$, you'd have a state that has nothing in particular to do with the two states you started with, a state in which there's no possibility at all of the particle's being in one of the original states—and certainly no possibility of interference.

In the form of quantum theory developed by Schrödinger in 1925–1926, the superposition principle is satisfied by using waves to represent all the possible states of any physical system, in contrast to the classical way of giving positions and momenta of particles. The wave is called the Schrödinger wave function, and its behavior is governed by a wave equation, the *Schrödinger equation.* You'll learn more about the wave function and about the problem of what it means in Chapter 14.

Actually, waves are only one example of things that show superposition—that can be added and subtracted in a meaningful way. The most general sort of entity that shows superposition is a vector, which can be one of our familiar three-dimensional vectors or an abstract vector in a vector space of arbitrary—or even infinite—dimension. In the development of quantum theory, it was Dirac who took this step and made such abstract *state vectors* the building blocks of a beautifully general formal structure (known as a *linear algebra*) for quantum mechanics. For a system consisting only of nonrelativistic particles, this turns out to be equivalent to using a Schrödinger wave function to represent the state, but for more general systems, especially relativistic theories involving fields such as the EM field, the wave function picture is no longer adequate, and Dirac's way of thinking about the physics seems to represent its essential nature much better.

Each of the four principles I've introduced in this chapter is somewhat vague at this stage, but they become much more solidly based when a proper theory is available. Indeed, they become logical consequences of such a theory. I've given some indications here of how that theory will develop and include these principles as features, but it's important to remember that the principles were set down before quantum theory was properly formulated.