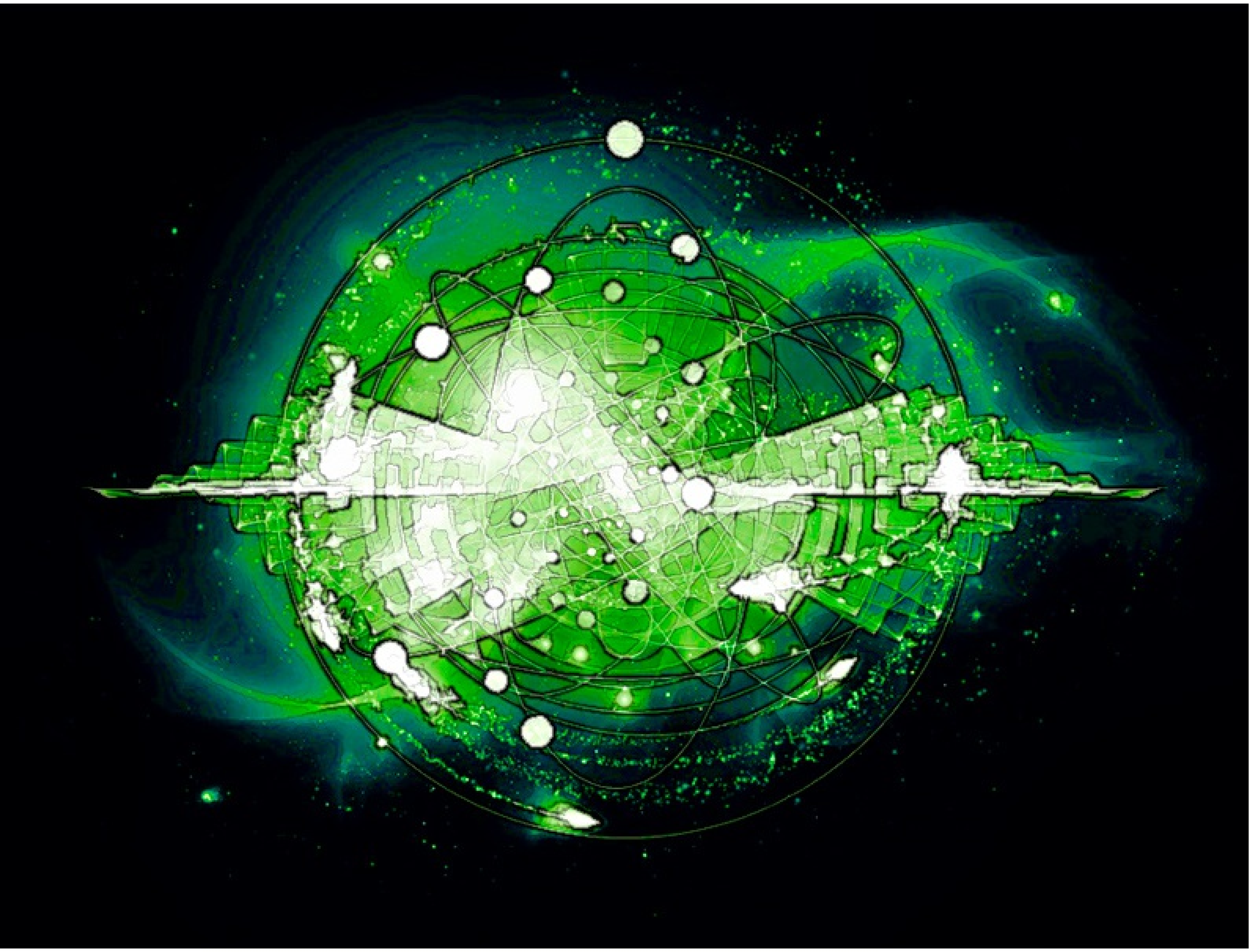
MIKE CORWIN

Remodeling Reality

RELATIVITY, QUANTUM MECHANICS, AND THE MODERN WORLDVIEW



bookboon
The eBook company

Mike Corwin

Remodeling Reality

The Impact of Relativity and Quantum Mechanics on Our Worldview

Remodeling Reality: The Impact of Relativity and Quantum Mechanics on Our Worldview
1st edition

ISBN 978-87-403-0958-4

Contents

1 The First 2000 Years

Your goals for this chapter are to know the following:

- The most significant factor that distinguishes Greek thought from earlier ways of thinking about the universe? Give some examples.
- The ancient (Greek) world-views of Thales, Democritus, Plato, Aristotle, and Ptolemy.
- The most significant discoveries and understandings about the Universe that were made by the Greeks.
- What most differentiates the philosophy of Socrates from that of the pre-socratics. What most differentiates the philosophy of Aristotle from that of Plato.
- What characterizes the following historical eras: the Greco-Roman (Classical) era, the Early Middle Ages (Dark Ages), and the Late Middle Ages.

The evolutionary origin of consciousness is not well understood. But it is clear that at some point in our evolutionary past, an inner world free from passive captivity in sensory impressions came to be – a life of the mind, if you like. The task of this inner world was to impose order on the content of our senses; that is, there came a point in our evolutionary past when our minds began to demand that the universe make sense.

The species to which all living humans belong, *Homo sapiens*, probably emerged somewhere around 100,000 years ago. We really don't know much concerning the specific worldviews of early humans. We have only some burial sites, figurines, and cave paintings to provide tantalizing hints. These suggest that their thinking was of a type that would now be called magical or superstitious. It seems clear that these stone-age people sought to employ magical rituals to influence the external world, hoping to positively affect hunting, fertility, and other survival-related aspects of their lives. Though such thinking may seem primitive to us, it clearly reflects a mind attempting to bring order into the universe.

With the coming of civilization some 10,000 years ago, the magical thinking of the earliest peoples had evolved into mythology. Myths, though differing in their local details, have some common threads running through them. Often powerful non-human, but anthropomorphic, figures create and control the world and its inhabitants. Myth arose from our need to make sense of the world as a whole, and, particularly, of our place as human beings in it. We see in myth attempts to find cause and effect explanations for the experienced world. Early people wove basic sensory knowledge of the world into a pattern that seems reasonable. For instance, the Mesopotamian creation myth used their knowledge of how silt deposits form land where fresh and salt water meet. Thus, although there are some obvious differences between the mytho-poetic approach and the scientific approach, we can also see connections. Myths are the first rungs on the ladder of discovery. Embedded within them are basic truths about both the universe and the human condition.

1.1 The Birth of Science

The next step in our understanding of the universe was taken in ancient Greece. Although it is probably an exaggeration to think in terms of “the Greek miracle” or of “motherless Athena,” as is frequently done, it is clear that about 600 BCE, a new approach to understanding the universe emerged. Although the Greeks had their myths, they went beyond the myths to search for physical explanations. Unlike earlier cultures, they were not content to explain the universe completely in terms of the actions of the gods; the Greeks insisted on thinking in terms of natural processes. This attitude is exemplified in the statement of a writer belonging to the Hippocratic school on the nature of epilepsy.

> “It seems to me that the disease is no more ‘divine’ than any other. It has a natural cause, just as other diseases have. Men think it divine merely because they do not understand it. But if they called everything divine which they do not understand, why, there would be no end of divine things!”

These proto-scientists made the remarkable assumption that an underlying rational unity and order existed within the flux and variety of the world. Nature was to be explained in terms of nature itself, not of something fundamentally beyond nature, and in impersonal terms rather than by means of personal gods and goddesses. Science was born here, not motherless, to be sure, but nonetheless a new and distinctly different way of looking at the world.

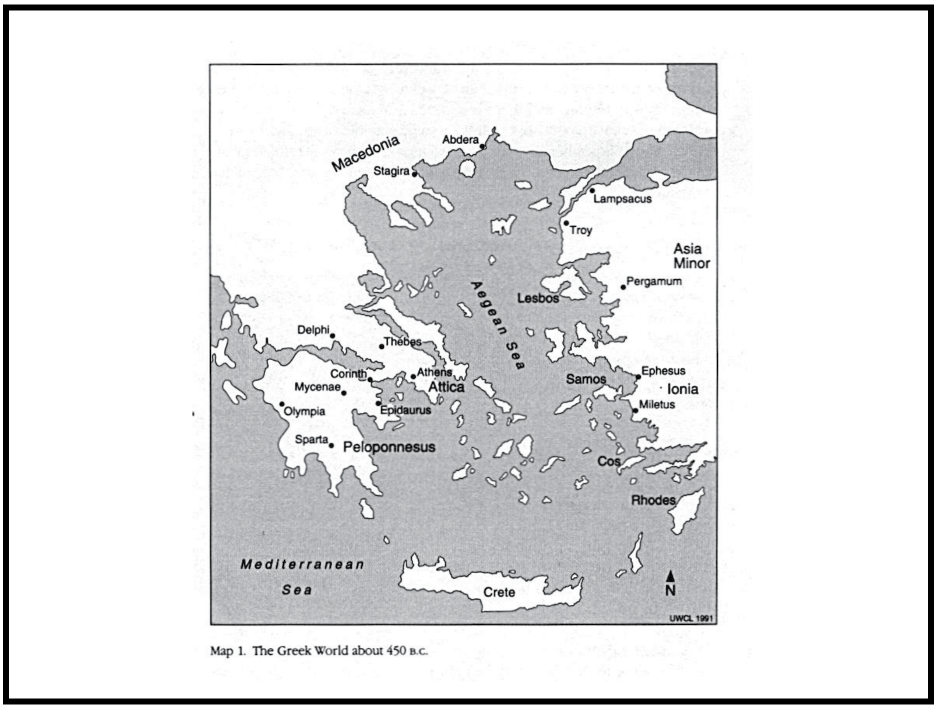


Figure 1.1 Map of Greece

Thales (624–547 BCE) was born in the Greek city of Miletus across the Aegean Sea from the Greek mainland. The inhabitants of this region were known as Ionians (Greeks who fled the Dorian invasion). Its location on the coast of Asia Minor provided Thales with exposure to the cultures of both the Babylonians and the Egyptians, and in fact, he visited both regions. It was his knowledge of Babylonian astronomy that gave rise to the story, probably apocryphal, that he predicted the solar eclipse of May 28, 585 BCE.

We consider Thales the first scientist because, as far as we can tell from the admittedly incomplete historical record, he was the first to approach the world from a scientific perspective. He wondered how the universe came to be and came up with an answer far different from that depicted in the creation of the gods myth of Hesiod's Theogony (8th century BCE). It seemed to him that all things either came from moisture or were sustained by moisture. He concluded that the universe grew from water. According to Thales the earth is a flat disc floating on a sea of water. The unique element in the cosmology of Thales is the idea that the universe developed over time through natural processes from some undifferentiated state. The first recorded use of a physical model in explaining a natural phenomenon is Thales belief that earthquakes are caused by disturbances in the water that supports the earth.

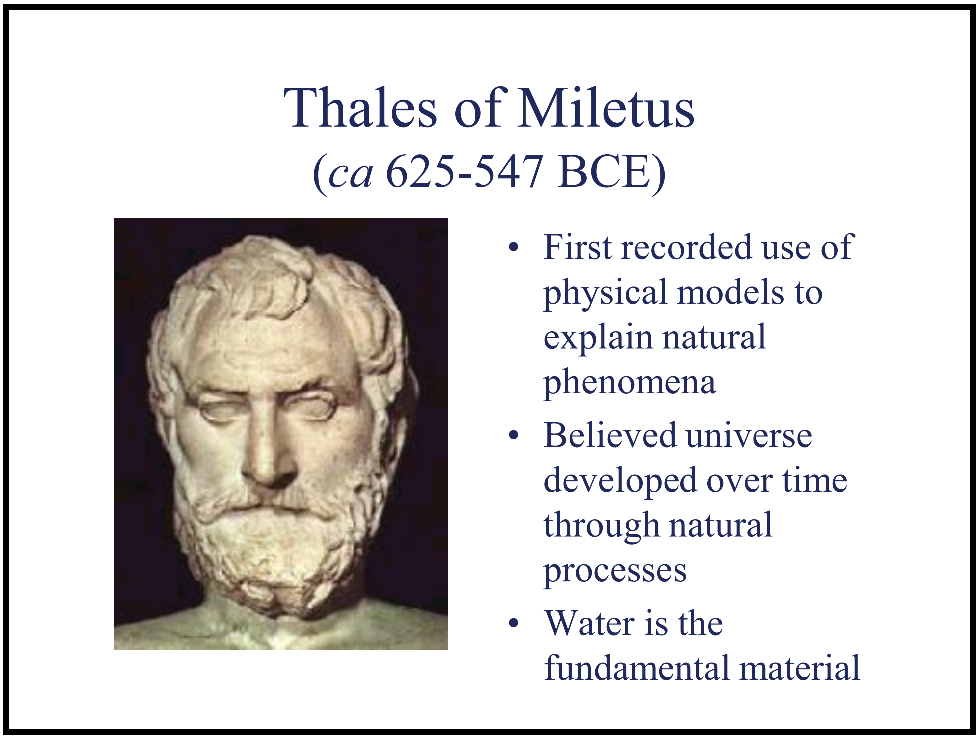


Figure 1.2 Thales

Thales had a student named Anaximander (610–546 BCE), who introduced the notion that the universe was spherical, an idea that survived for more than 2000 years. He saw the earth as suspended in space (rather that floating on water). He also believed that living creatures arose from the moist elements when it had been partially evaporated by the sun. According to Anaximander, humans in the remote past resembled fish, perhaps the first theory of biological evolution.

In the second half of the fifth century, the approach of Thales and Anaximander was adopted and extended by Leucippus of Miletus (*fl.* 440 BCE) and Democrutus of Abdera (*c.* 470–*c.* 400 BCE). Democritus constructed a complex explanation of all phenomena as the result of material interactions. He taught that the world was composed exclusively of uncaused and immutable material atoms. These invisibly minute and indivisible particles perpetually moved about in a boundless void and by their random collisions and varying combinations produced the phenomena of the visible world. This concept is known as materialism. In the words of Democritus, "nothing exists except atoms and the void; all else is mere opinion."

It is interesting to note that a central concept in the thinking of Thales, Anaximander, and Democritus is that there is no real distinction between the terrestrial and celestial realms. Only later did Greek thinking regress to needing a fifth essence (the quintessence) for celestial objects. These early Greek thinkers, known as the presocratics, were the first we know of to systematically seek natural explanations of natural phenomena. The Babylonians and ancient Hebrews had a literature embodying stories expressing awe about the heavens and the earth. However, they concerned themselves with the question 'why', and did not attempt to answer the question 'how.'

This earlier and simpler phase of Greek thought terminates in the fifth century with a thinker of an entirely different type, Socrates (470–399 BCE). With Socrates and his student Plato (427–347 BCE), we have a unique synthesis of Greek science and Greek religion. They taught that the visible world contains within it a deeper meaning, in some sense both rational and mythic in character, which is reflected in the material world but which emanates from an eternal dimension that is both source and goal of all existence.

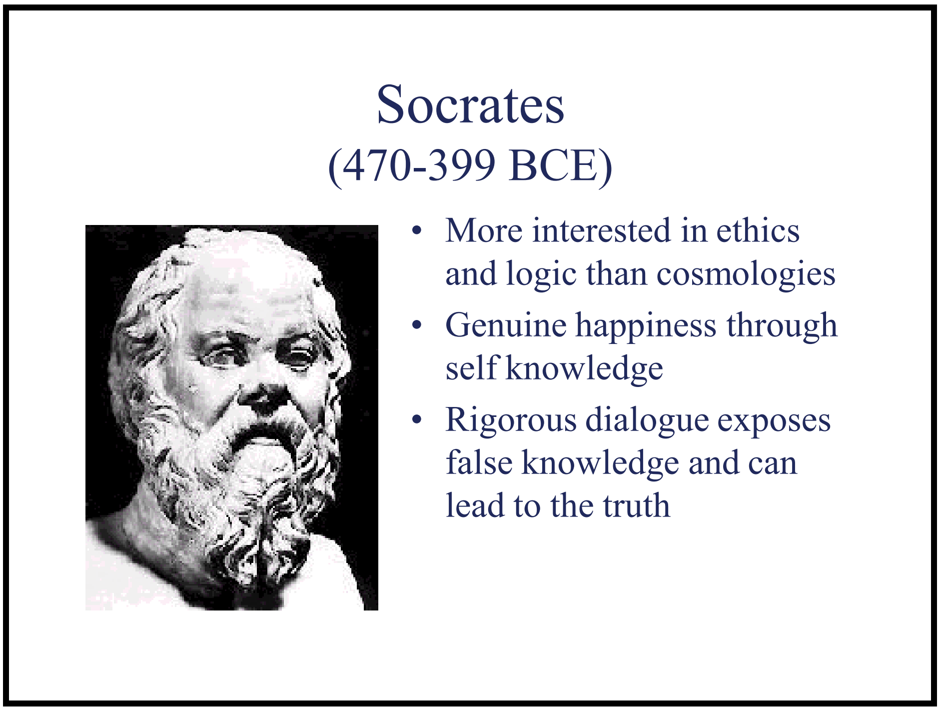


Figure 1.3 Socrates

With Aristotle (348–322 BCE), a student of Plato's and a teacher of Alexander the Great, the pendulum began to swing back toward the more down-to-earth perspective of the presocratics. Plato asserted the existence of archetypal Ideas or Forms as primary, while the visible objects of conventional reality are their direct derivatives. These Ideas, according to Plato, possess a quality of being, a degree of reality, that is superior to that of the concrete world. On the other hand, Aristotle assumed that true reality was the perceptible world of concrete objects, rather than the imperceptible world of Plato's eternal Ideas.

Aristotle placed a new and fruitful stress on the value of observation and classification. Aristotle's writings were the first to create a comprehensive system of philosophy, encompassing morality and aesthetics, logic and science, politics and metaphysics. He provided a language and logic, a foundation and structure, and, not least, a formidably authoritative figure without which the philosophy, theology, and science of the West could not have developed as they did.

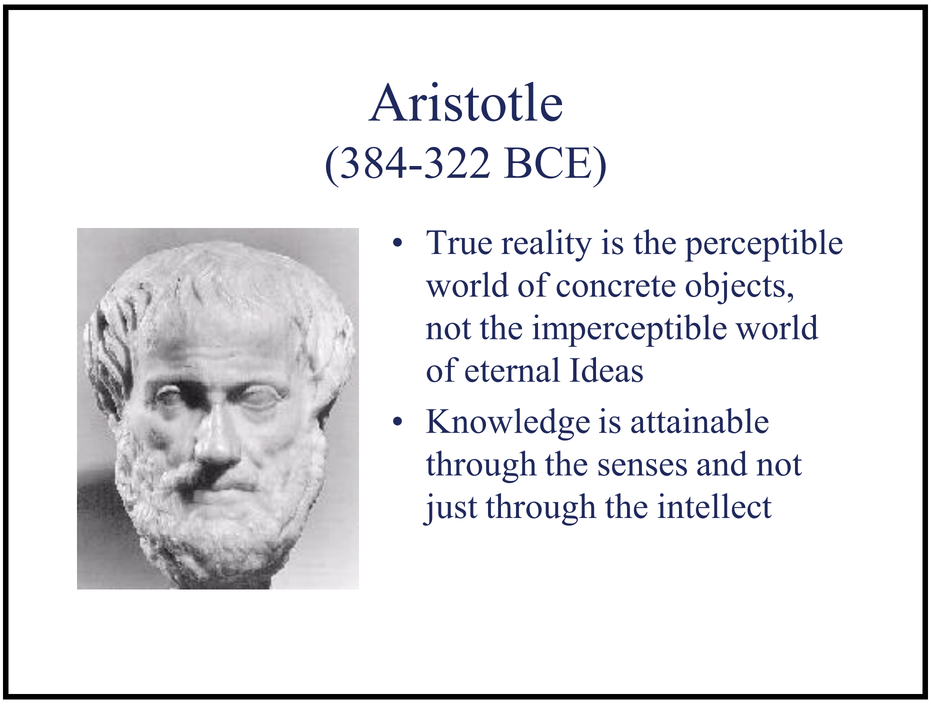


Figure 1.4 Aristotle

1.2 The Greek Worldview

The Greek worldview is based primarily on the teachings of Aristotle and its defining characteristic is that it is geocentric. It is the most long-lived cosmological model in history, lasting into the 17th century.

Aristotle taught that the universe was spherical and finite, with rotating spheres carrying the moon, sun, planets, and stars around a stationary earth at its center. To support the fact that the earth did not move, he pointed out that if the earth were in motion, an observer on it would see the fixed stars as shifting their positions with respect to one another, a phenomenon known as parallax. However, parallax was not observed.

Aristotle offered several proofs that the earth was a sphere. One such proof involved lunar eclipses. At the time, it was known that lunar eclipses were caused by the shadow of the earth falling on the moon. The fact that the shadow was always circular showed that the earth was a sphere. He also pointed out that when one travels northward or southward, the position of the North Star changes with respect to the horizon. To further bolster his argument that the earth was the immovable center of the universe, he proposed a comprehensive theory of motion that required all earthly substances to naturally move toward the center of the earth.

Aristotle accepted Empedocles' view that there are four earthly elements, earth, air, fire, and water. Aristotle added an additional element called aether or quintessence (fifth element), which he believed composed the celestial bodies. However, Aristotle rejected, on grounds of logic, Democritus' view of atoms. Because atoms have extension in space, they could not be indivisible. Extension in space also implies composition and therefore atoms could not be elementary. Aristotle proposed that matter is continuous and infinitely divisible.

A Greek thinker living after Aristotle, Aristarchus (*c.* 310–230 BCE), used geometric arguments based on eclipses and the phases of the moon to estimate the relative sizes of earth, moon, and sun. He showed that the sun was many times larger than the earth and that the moon was much smaller. Reasoning that the smaller object should orbit the larger object, he concluded that the sun not the earth was the center of the universe, and that the earth orbited the sun once a year while spinning on its axis once every 24 hours.

This idea was taken seriously by Greek thinkers but ultimately rejected. If the earth were to orbit the sun, the positions of the stars must show parallax, which was not observed. Aristarchus argued that it was not observed because the nearest stars are at extremely great distances compared to the distance from the earth to the sun. In fact, the correct explanation. The nearest stars are so far away that parallax was not observed until the early 1800s, long after it had been established that the earth did orbit the sun.

Eratosthenes (*c.* 276–195 BCE), a younger contemporary of Aristarchus, used geometry to determine the circumference of the earth. It was known that at noon on a particular day of the year, the sun shone directly down a vertical well in Syene in southern Egypt. At that same time in Alexandria, north of Syene, the sun was not directly overhead, Eratosthenes was able to measure the angle that the sun made with the vertical as one-fiftieth of a circle (about 7°). From this, he was able to conclude that the earth's circumference is 50 times the distance between Alexandria and Syene.

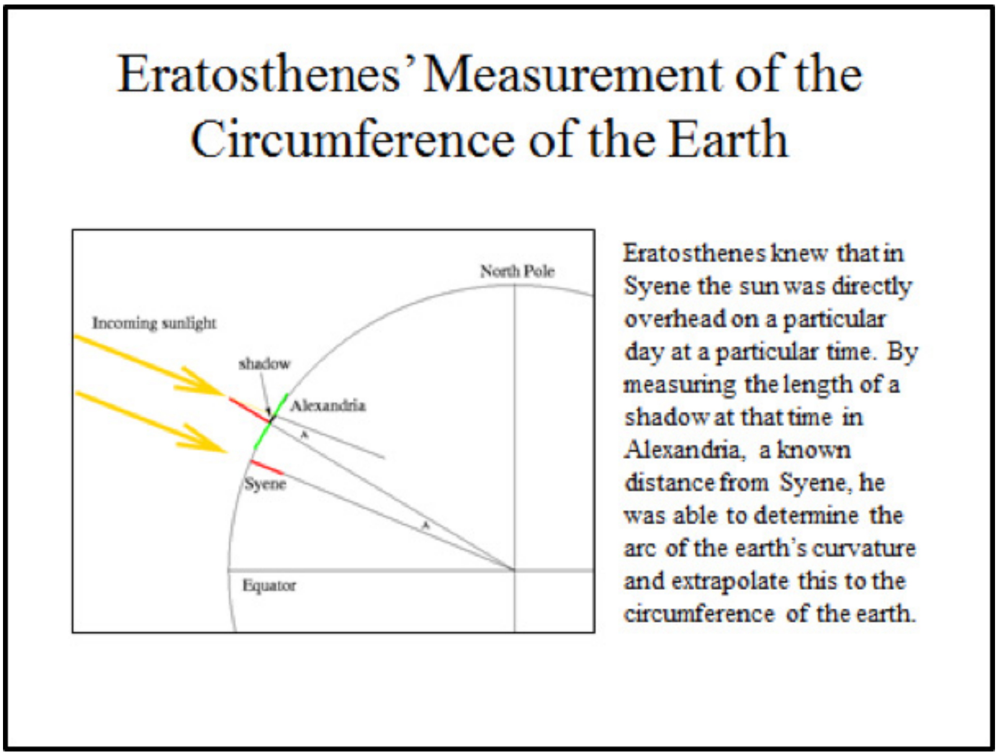


Figure 1.5 Eratosthenes

It is not possible to evaluate precisely the accuracy of Eratosthenes' solution because there is some uncertainty about the length of the unit he used. However, it is certain that his value was within 20% of the correct answer and may have been as close as 1%. (Eratosthenes' value was much closer to correct than the one used by Columbus almost 2000 years later. Columbus thought the earth was much smaller than it is. Had he known and accepted Eratosthenes' value, it would have been obvious that he could not reach China by sailing west.)

Hipparchus (*c.* 190 BC–*c.* 120 BC), refined the method of Aristarchus for measuring the relative distances and sizes for the earth, moon, and sun, and obtained better results. He compiled an accurate catalog of the positions of over 850 stars. Hipparchus had access to 1000 years of Babylonian records of star positions. By comparing these with his own, he was able to determine that the earth's axis of rotation sweeps around in a cone much like the motion of a top inclined with respect to the vertical. The earth's axis takes about 26,000 years for one complete sweep. This motion is referred to as the precession of the equinoxes.

Claudius Ptolemaeus, better known as Ptolemy, lived in Alexandria, Egypt in the middle of the second century CE. Relying heavily on the work of Hipparchus, Ptolemy compiled 13 volumes containing all known astronomical knowledge. Some of the material was original with Ptolemy.

Ptolemy is best known for his detailed geocentric model of the universe. The notion that celestial motions must always be represented by constant speeds and circular orbits was well established in Greek thought. Ptolemy retained this concept but added additional circles within circles to more accurately represent the observed locations of the planets in the sky. Although the model was very cumbersome, it was sufficiently accurate to remain the accepted view of the universe until the 17th century.

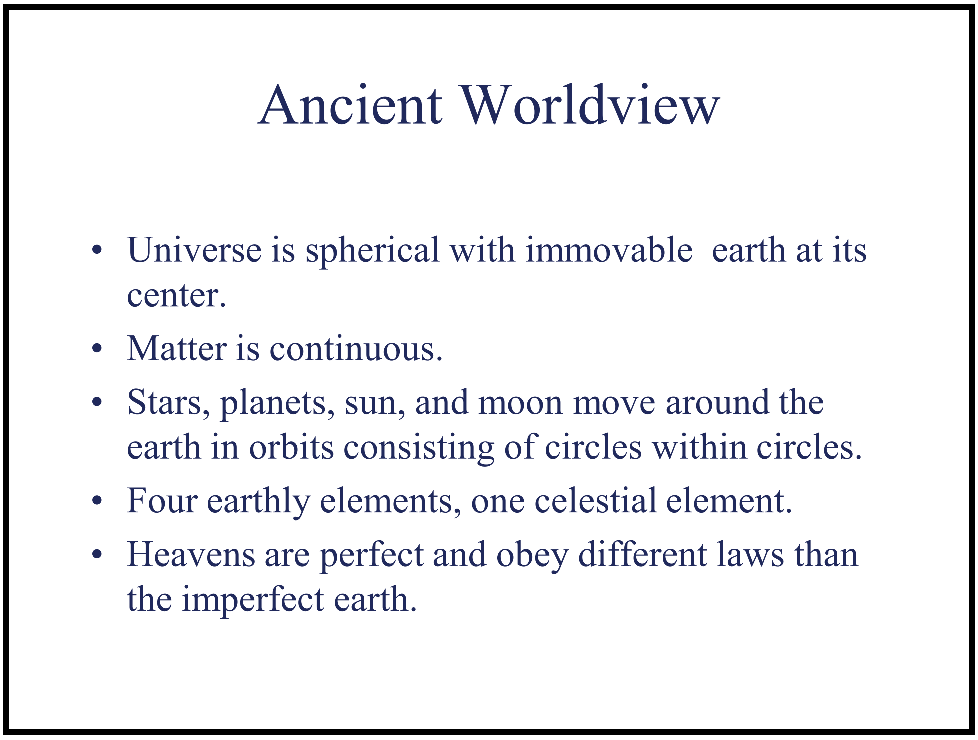


Figure 1.6 Ancient Worldview

The Romans began building their empire in the 4th century BCE and eventually controlled the entire Mediterranean world. They conquered Greece in the 2nd century BCE. However, the Romans recognized the superiority of much of Greek thought and culture and incorporated it into their culture. The resulting civilization is known as Greco-Roman.

In 313 CE the Roman emperor Constantine issued the Edit of Milan recognizing the right of Christians to practice their religion. Gradually Christianity became the official religion within the Empire. By the 4th century, the Empire was beginning to collapse. Commerce, particularly seaborne commerce, was in decline as was the population, especially in urban areas. The area under the control of Rome was shrinking and Rome itself was being subjected to barbarian invasions. This period of time is usually referred to as Late Antiquity. This is a transitional period, during which the ancient world slid slowly into the medieval.

1.3 The Middle Ages

The Middle Ages is a derogatory term coined after the fact to reflect the view that this period was a time of intellectual stagnation between that of classical (Greco-Roman) culture and its later reestablishment during the Renaissance (rebirth). At the time it first appeared in the 15th century, it referred to the period from the fall of the Roman Empire (~500 CE) to the beginning of the Renaissance (~1350). Later, the term referred to the division of history into Classical, Medieval, and Modern. More recent scholarship has somewhat modified the earlier divisions. Recognizing that the Roman Empire was in serious decline before 500 and that some aspects of Greco-Roman culture continued after 500, the era Late Antiquity was added.

Although the term Dark Ages is no longer used to characterize Late Antiquity and the Early Middle Ages, these eras were none the less a time of wide spread political insecurity, famine, and illiteracy. As the authority of secular Rome declined, the vacuum was slowly filled by the Roman Church. By 500 the Church was essentially the only authority in Western Europe.

Almost from the beginning of Christianity, safeguarding the faith was the Church's almost exclusive priority. Dialogue was often curtailed all together least faith be undermined and the authority of the Church questioned. There was a reaction against the pagan philosophers. Tertullian, a 3rd century Church leader, said "What has Athens to do with Jerusalem?" One consequence is that the value of observing, analyzing or understanding the natural world was greatly diminished.

Needless to say, not much progress in the understanding of the universe was made during this period. In fact, just the opposite. Much of what was known in the classical era was lost. Books were burned. For the Church authorities, direct study of the natural world was seen as a threat to the integrity of religious faith and thus to salvation.

One of the more absurd results of this was a rejection of the idea of a spherical earth, in part because the idea was supported by the pagan philosophers. Instead, the universe was seen as shaped like the Holy Tabernacle; the earth enclosed within a domed rectangular box. Based on the story in Genesis that the firmament is enclosed by water, the model included super-celestial waters resting on top of the tabernacle. Not until about the end of the 9th century was the idea of a spherical earth reinstated.

In the later Middle Ages, Christianity's earlier need to distinguish and strengthen itself by the more or less rigid exclusion of pagan culture lost some of its urgency, and a more relaxed attitude toward secular learning developed. "It seems to me a case of negligence if, after becoming firm in our faith, we do not strive to understand what we believe." Saint Anselm of Canterbury, (1033–1109). Contrast this with the attitude of Tertullian in the 3rd century, "All curiosity is at an end after Jesus, all research after the Gospel. Let us have Faith and wish for nothing more."

The change in the Church's attitude was accelerated by the rediscovery, mostly as a result of the Crusades, of a large body of Aristotle's writings that had been preserved by Islamic and Byzantine cultures in the East. The Arabic and Greek text were translated into Latin and widely circulated in the new universities of the West. At first the Church tried to suppress the teachings of Aristotle as they were in conflict with those of neo-Platonism which Saint Augustine, in the early 5th century, had established as the philosophical foundation of Christianity. This turned out to be impossible, and finally in the mid-13th century, Saint Thomas Aquinas was able to integrate Aristotle with Christian theology, in much the way that Augustine had before with Plato.

Aquinas blended Aristotelian philosophy and Christian doctrine by suggesting that rational thinking and the study of nature, like revelation, were valid ways to understand truths pertaining to God. According to Aquinas, God reveals himself through nature, so to study nature is to study God. The writings of Aquinas became the new philosophical foundation for the Christian religion and Aquinas' view of the relationship between God and the universe became the cosmological worldview of the Late Middle.

Through the work of Aquinas and other Scholastics, the Aristotelian-Ptolemaic cosmology was reintroduced to Europeans and, at the same time, permeated with Christian meaning. In retrospect, the blending of Aristotelian thought with theology may have been a mistake. Because the cosmology of Aristotle was so fundamentally a part of his philosophy, the synthesis had the (perhaps unintended) consequence of making the Greek view of the physical universe a part of Christian dogma. The Church became locked into a geocentric model of the universe. During the Scientific Revolution this would bring the Church into direct conflict with science, one factor that eventually eroded its authority.

2 The Scientific Revolution and the Modern Worldview

Your goals for this chapter are to know the following:

- What is meant by the Scientific Revolution that took place between the times of Copernicus and Newton. In what ways the modern world-view that resulted from the Scientific Revolution differd from the ancient world-view.
- How the science of Galileo differd from that of the Greeks (Aristotle). What Galileo's most significant contributions were.
- The most significant contributions of Copernicus, Brahe, Kepler, and Newton.
- What the scientific method is. What it means to say that in order for a statement to be scientific it must be falsifiable.
- What is meant by a mechanistic-deterministic worldview. Why a clockwork is a good analogy for the mechanistic-deterministic worldview. What Diests are.

2.1 Nicolaus Copernicus

History is generally divided into eras each of which can be characterized in a way that differentiates it from all others. But, in fact, these eras blend, more or less gradually, into one another. For example, there is a sense in which the Scientific Revolution is simply a continuation of the Renaissance. History also has a tendency to date the various eras in round numbers. For instance the Late Middle Ages is usually said to have occurred between the years 1000 and 1350 and the Renaissance between 1350 and 1600. The Scientific Revolution is usually said to have occurred in the 17^{th} century. However, all treatments of the Scientific Revolution inevitably begin with Nicolaus Copernicus, a Polish canon in a Catholic cathedral in the first half of the 15^{th} century – a thoroughly Renaissance man in both time and spirit.

Copernicus was a scholar rather than a scientist in the modern sense of the word. As scholars of the time did, he immersed himself in the newly translated classical literature, not with the intention of making new discoveries, but of recovering old discoveries. Copernicus is sometimes credited with discovering the heliocentric model of the solar system. In fact he read about it in a book. Greek thinkers had proposed it centuries before the common era, principle among them, Aristarchus.

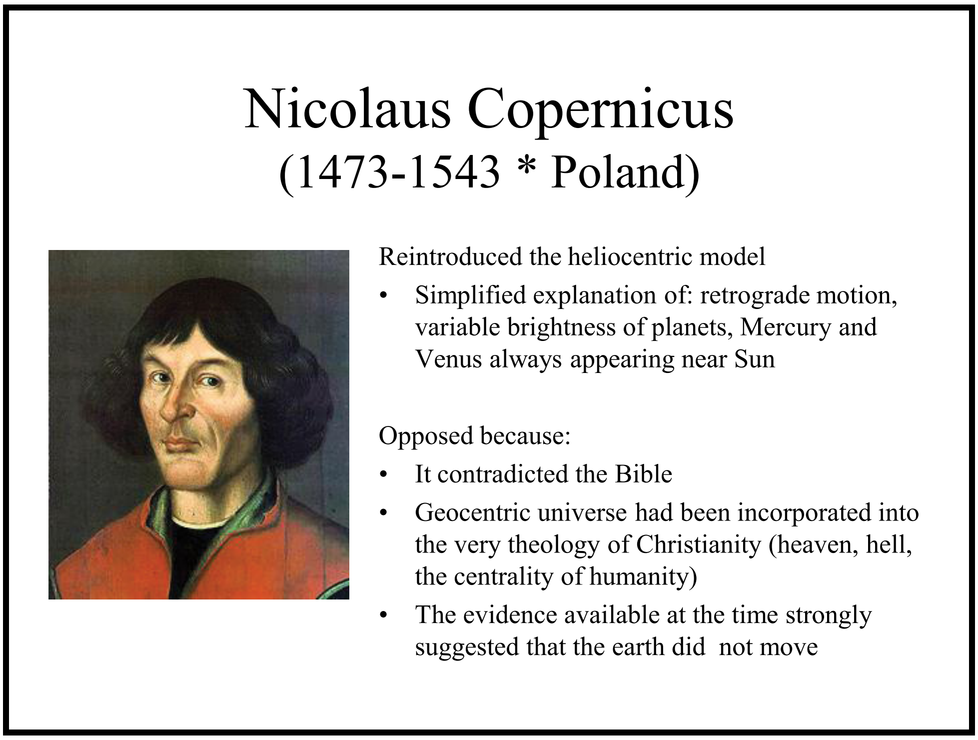


Figure 2.1 Nicolaus Copernicus

Copernicus read of this suggestion and realized that it explained in a simple manner many things about the motion of the planets that had complex, implausible explanations using Ptolemy's geocentric model. Copernicus felt that a satisfactory representation of the solar system should be coherent and physically plausible, not requiring a different construction for each phenomenon, as Ptolemy's system did. To him, Ptolemy's system was ugly and therefore could not represent the work of the Creator.

(Upon hearing in the late 13th century of Ptolemy's model of the universe and of the extremely complicated mathematics it required, Alfonso X, Spanish monarch and astronomer, is said to have replied, "If the Lord Almighty had consulted me before embarking on creation, I should have recommended something simpler.)

As early as 1514, Copernicus circulated among his friends a short manuscript describing his heliocentric views. However, he was reluctant to publish. Most scholars believe that it was not fear of the Church that caused his reluctance. The Church did not take a hard line on the issue at the time. In fact, it was in general supportive of Copernicus. It was only later, during the counter-Reformation, that people such as Giordano Bruno and Galileo Galilei suffered retribution for their views on the nature of the universe.

2.2 Tycho Brahe and Johannes Kepler

Copernicus got it right about the earth going around the sun, but he got the orbits in which the earth and other planets orbit the sun wrong. Copernicus' model continued to use the circles-within-circles orbits of Ptolemy's model. His model consisted of a moving earth in a cosmos otherwise ruled by Aristotelian and Ptolemaic assumptions. Observations soon showed that Copernicus' model was somewhat better at predicting the exact locations of the planets in the sky at some future date, but still not completely accurate. Both models had to be wrong. The Danish astronomer, Tycho Brahe (1546–1601), set himself the task of coming up with the correct model.

Brahe realized that progress in astronomy required systematic, rigorous observation, night after night, using the most accurate instruments available. This program became his life's work. Brahe improved and enlarged existing instruments, and built entirely new ones. The telescope had not been inverted yet, so all of Brahe's instruments were naked-eye. Brahe began making observations and recording data in 1572 and continued to do so until his death in 1601.

Brahe's model of the solar system was a hybrid of the geocentric and the heliocentric. He accepted the arguments of Copernicus that having the planets orbit the sun rather than the earth provided a simpler explanation for the observations, but was convinced that the earth did not move. In Brahe's model, the sun, with its orbiting planets, orbited a stationary earth.

Brahe planned to use his extensive data on the locations of the planets to demonstrate the correctness of his model. The idea was to use the model to calculate the position in the sky of a planet at some point in the past. He would then go back to his data to show that the calculation produced the actual observed location. If it could consistently do this, as Ptolemy's and Copernicus' could not, then his model would be shown to be correct. However, these calculations were extremely difficult. Brahe could not do them himself, so, in 1600 he hired the German mathematician, Johannes Kepler to do them for him.

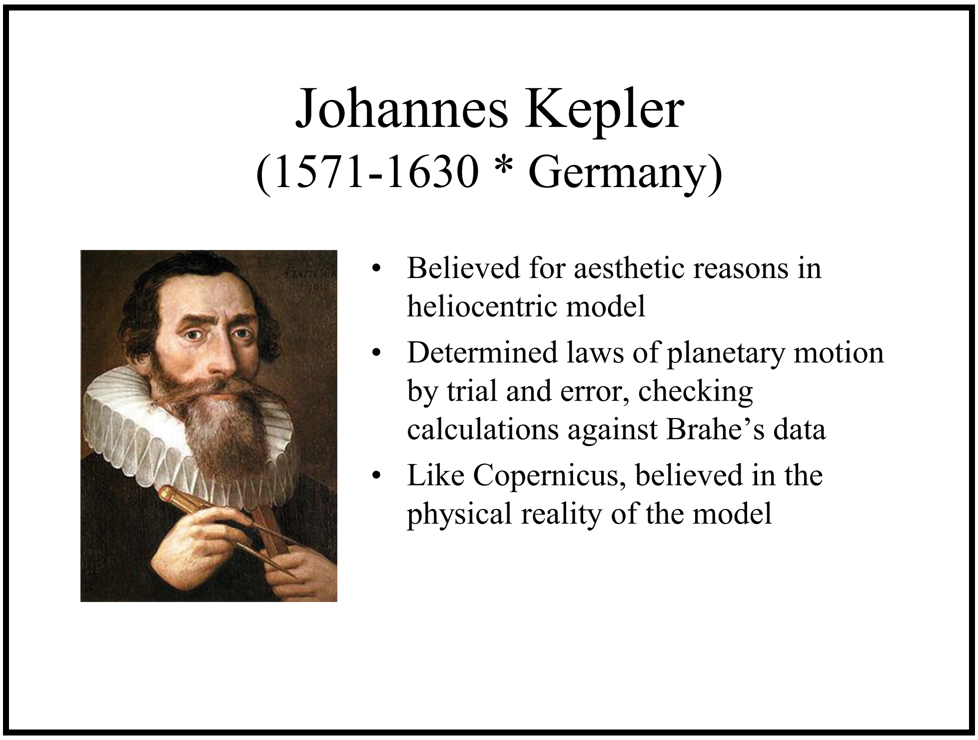


Figure 2.2 Johannes Kepler

Although he was hired to make the calculations necessary to demonstrate the correctness of Brahe's geocentric model, Johannes Kepler had for some time been a convinced Copernican. Not that he believed that Copernicus' model was correct in all its details; he knew that its slight inaccuracies meant that it was wrong. But the aesthetic superiority of a heliocentric view was compelling to him.

Brahe died shortly after Kepler was hired. Kepler succeeded him as imperial mathematician and astrologer to the Holy Roman Emperor, with the responsibility of completing Brahe's unfinished work. Kepler now had access to Brahe's decades of unprecedentedly accurate astronomical observations.

Kepler had entered Brahe's employment with a specific heliocentric model of his own. Now he had the opportunity to check his model against the data, and soon found that his model was wrong. He did not give up, however. For four years he repeatedly devised new models, checked them against the data, and found out that they were wrong. In these attempts he focused on the planet Mars. He reasoned that the Creator would not have created a different orbit for each planet; that would be unaesthetic, something incompatible with his concept of God. If he could figure out the orbit of Mars, he was sure it would be the orbit of all the other planets as well.

After years of unsuccessful attempts using various combinations of circles, he finally gave up on that approach. Finally, in 1605 he hit upon the correct combination of path and speed that would match the calculations to Brahe's observations. Mars moves in an elliptical path with varying speeds depending of the distance between it and the sun. Mars speeds up as it approaches the sun and slows as it recedes. It does this in such a way that an imaginary line drawn between Mars and the sun sweeps out equal areas in equal time intervals. As he had suspected, this orbit worked for the other planets as well. Although the manuscript presenting this information was completed in 1605, it was not published until 1609 due to legal disputes with Brahe's heirs over the use of Brahe's observations.

This correct orbit was arrived at strictly by trial and error. Kepler had no model in mind that allowed him to predict it and no clear explanation for why the planets moved in this way. The explanation would have to wait more than 50 years for Isaac Newton to figure it out. However, the accuracy with which the model was able to predict the past locations of the planets in the sky, as verified by Brahe's observations, left little doubt that the orbit was correct.

2.3 Galileo Galilei

Galileo Galilei was a contemporary of Kepler, and, like Kepler was a convinced Copernican long before there was anything other than aesthetic reasons for supporting the heliocentric model. Other than that, the two men had little in common. Kepler was very mild mannered, somewhat sickly, and modest. Galileo was the opposite. Kepler was a Protestant and Galileo a Catholic, both strong in their faith. Galileo dismissed much of Kepler's work as useless fiction and refused to accept elliptical orbits for the planets, continuing to believe they had to be circular in some way.

Galileo is significant in science for two distinct reasons. First of all, he was the first, in 1609, to use a telescope to study the heavens and in this way made several important discoveries that undermined the Ptolemaic model accepted by most scholars and the Christian churches, both Catholic and Protestant. However, these discoveries did not prove that the earth itself orbited the sun, as Galileo liked to claim. Secondly, he is generally credited with inventing the scientific method as we understand it today, or at the very least, with being the first to apply it systematically.

Although Galileo did not invent the telescope, he was the first to use it to gain knowledge of the heavens. Among his discoveries were the mountains and craters on the moon. Because the moon was part of the celestial realm, Aristotle and religious dogma required it to be perfect. Well, almost perfect. It was clearly blemished, perhaps signifying that as the closest celestial object to the earth, it was a transitional object between the imperfect earth and the absolutely perfect heavens beyond. In any case, the scholars and churches taught that the moon is a perfectly smooth and spherical object. If you already believe this and look at the moon with the naked eye, it is easy to believe this is true.

Even through Galileo's relatively low power telescope, it clearly is not true. Galileo had trouble convincing others of this. They either refused to look through the telescope or claimed that the irregularities were an artifact of the telescope itself rather than a true image of the moon. The resemblance of the moon's features to those on the earth misled Galileo somewhat. He thought that the dark relatively smooth surfaces on the moon were oceans and named them seas. Today we call them maria, the Latin word for seas.

Galileo discovered that the planet Venus went through phases just as the moon does. This was important because it proved that Venus orbited the sun rather than the earth, thus proving the Ptolemaic model wrong. He was also able to demonstrate what some others had suspected, that the Milky Way, the band of diffuse light that arcs across the night sky from horizon to horizon, is actually composed of hundreds of thousands of stars. He observed sunspots and used them to calculate the speed of rotation of the sun, about one revolution every 25 days.

Perhaps his most important discovery was the four (now called Galilean) moons of Jupiter. One of the strongest arguments in favor of the geocentric model was the fact that our moon orbits the earth. No one disputed this. The argument went that the earth could not possibly move because if it did, it would leave the moon behind. In the days before the discovery of gravity, this was a very powerful argument. However, whether one believed in a geocentric or a heliocentric universe, it was clear that Jupiter moved; it had to orbit something whether the earth or the sun. The fact that Jupiter was somehow able to move without leaving its moons behind destroyed the argument.

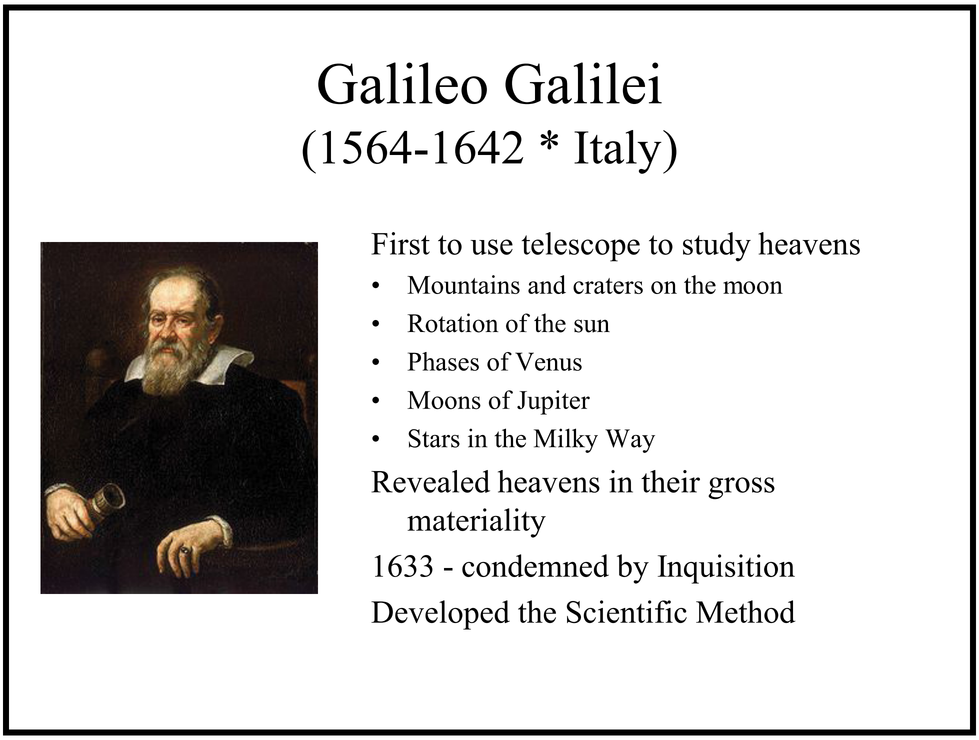


Figure 2.3 Galileo Galilei

As most everyone knows, Galileo got into serious trouble with the Church later in his life. In 1616, Galileo had been instructed by the Church "not to hold or defend" the heliocentric model, though he was allowed to discuss it as a hypothesis. However, in his book, Dialogue Concerning the Two Chief World Systems, published in 1632, he not only defended the heliocentric model, he ridiculed the then reigning Pope's arguments in favor of the geocentric model. For this, he was called before the Inquisition in 1633.

He was threatened with torture if he did not publicly recant, which he did. He avoided torture, but was found "vehemently suspect of heresy" and sentenced to house arrest, where he remained for the rest of his life. In spite of his troubles with the Church, he remained a devout Catholic throughout his life. His justification for proposing theories contrary to the Bible is summarized in his statement "The Bible tells you how to go to heaven, not how the heavens go."

His trial before the Inquisition ended Galileo's work as an astronomer. Fortunately for science, it did not end his work as a physicist. During his almost decade of house arrest, Galileo made original contributions to the science of motion through an innovative combination of experiment and mathematics. Galileo is perhaps the first to clearly state that the laws of nature are mathematical. His studies of motion laid the groundwork for Isaac Newton's formulation of his three laws of motion. The first of these laws, logically just a special case of the second, is simply a restatement of work done by Galileo, and was included specifically to recognize Galileo's contribution. Galileo's empirical approach in his studies of motion is what we now know as the scientific method.

The Scientific Revolution is not yet complete. Kepler has shown the heliocentric model to be correct by determining the orbits of the planets around sun, but he has no explanation for why they move in those particular orbits. Galileo has made important discoveries in mechanics, completely destroying Aristotle's theory of motion, but was not able to replace it with a similarly comprehensive theory. The unfinished work of Kepler and Galileo would have to wait another 25 years for the genius of Isaac Newton to complete their work.

2.4 Isaac Newton

It is no exaggeration to say that Isaac Newton is the single most important contributor to the development of modern science. The Latin inscription on Newton's tomb, despite its bombastic language, is thus justified in proclaiming, "Mortals! Rejoice at so great an ornament to the human race!" It is perhaps a slight exaggeration to say, as Alexander Pope did as an epitaph for Newton:

> "Nature and Nature's laws lay hid in night; God said, Let Newton be! and all was light."

Newton entered Trinity College of Cambridge University in 1661. The Cambridge curriculum at that time was still strongly classical, but Newton preferred to read the more advanced ideas of modern philosophers such as Descartes and astronomers such as Copernicus, Galileo, and Kepler. In private studies he had begun to master the field of mathematics as shown by notebooks he kept at the time. No one at Cambridge apparently recognized his genius. Newton obtained his degree from Cambridge in August 1665 without honors or distinction.

The university temporarily closed for the next two years as a precaution against the Great Plague, and Newton returned to his mother's farm. During these two years (his 23rd and 24th) Newton filled notebook after notebook with ideas and experimental observations. These may be the two most productive years in the entire history of science.

In that relatively short period of time, Newton made brilliant and important discoveries regarding light and color. He continued his studies of mathematics and invented the calculus, which he then used to describe the motion of objects. Finally, and perhaps most significant of all, he developed a mathematical equation describing gravity. Thus Newton not only knew how the planets moved, he knew why they moved that way. What Kepler had laboriously determined through trial and error, Newton using his laws of motion and the law of gravity could calculate on the back of an envelope.

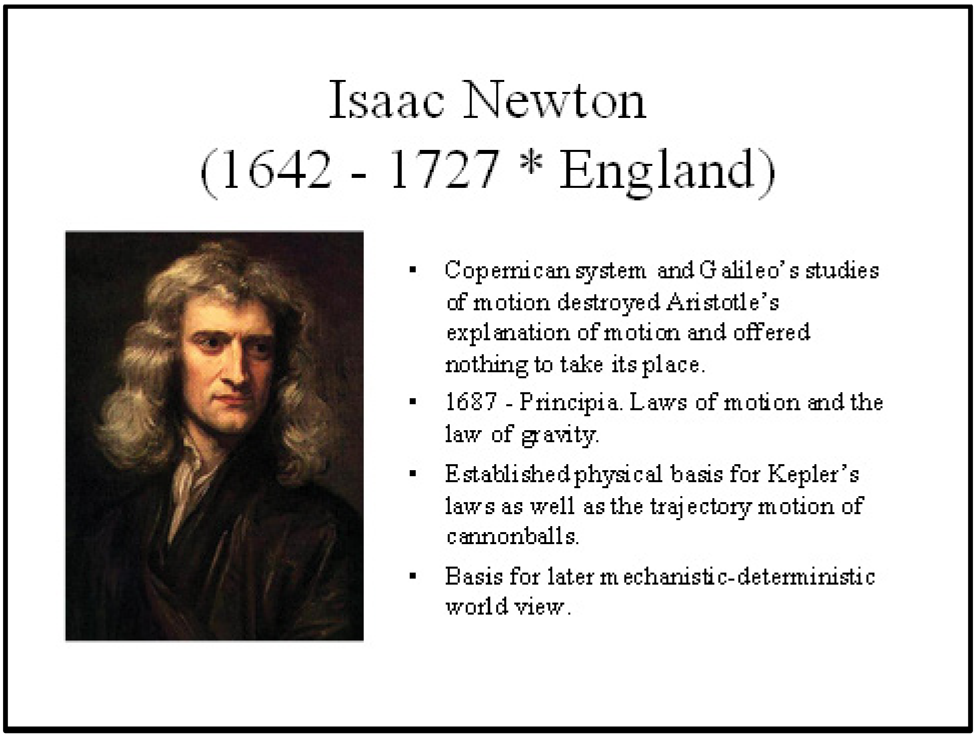


Figure 2.4 Isaac Newton

Although Galileo never dropped two balls from the Leaning Tower of Pisa, the story that Newton arrived at his theory of gravity after watching an apple fall from a tree in his mother's orchard appears to be true. At least, Newton said it was.

Newton returned to Cambridge in 1667, and was elected a minor fellow at Trinity. Finally, his talents were beginning to be recognized. The next year he became a senior fellow upon taking his master of arts degree, and in 1669, before he had reached his 27th birthday, he became Lucasian Professor of Mathematics. The duties of this appointment offered Newton the opportunity to organize the results of his earlier optical researches, and in 1672 he published his work on light and color. This work established his reputation as a scientist of the first magnitude.

Newton was, however, a highly secretive and suspicious person who found it extremely difficult to submit his ideas to the scrutiny of others. Many of the great discoveries made during the two years on the farm remained unpublished for decades. It was not until 1684 that Edmond Halley finally persuaded Newton to make his work on motion and gravity known.

Halley was intensely interested in planetary orbits, and also those of comets. He and fellow scientist Robert Hooke suspected that an inverse-square relationship produced the orbits, but were not able to deduce from this hypothesis a theoretical orbit that would match the observed planetary motions. Halley traveled to Cambridge to seek the advice of Newton. What would be the orbit of a body subjected to such a force? Newton told him he had already solved the problem – the orbit would be an ellipse – but that he had mislaid his calculations to prove it.

Shortly afterwards Newton sent Halley a copy of his demonstration. Realizing the significance of what Newton had done, Halley, utilizing great skill and tact, persuaded the reluctant Newton to develop and publish his ideas on celestial mechanics. Newton's *Mathematical Principles of Natural Philosophy* (commonly known as the *Principia* from its Latin title), containing Newton's three laws of motion and his law of gravity, was published in 1687. Halley read the manuscript, corrected the proofs, and paid the publication costs out of his own pocket.

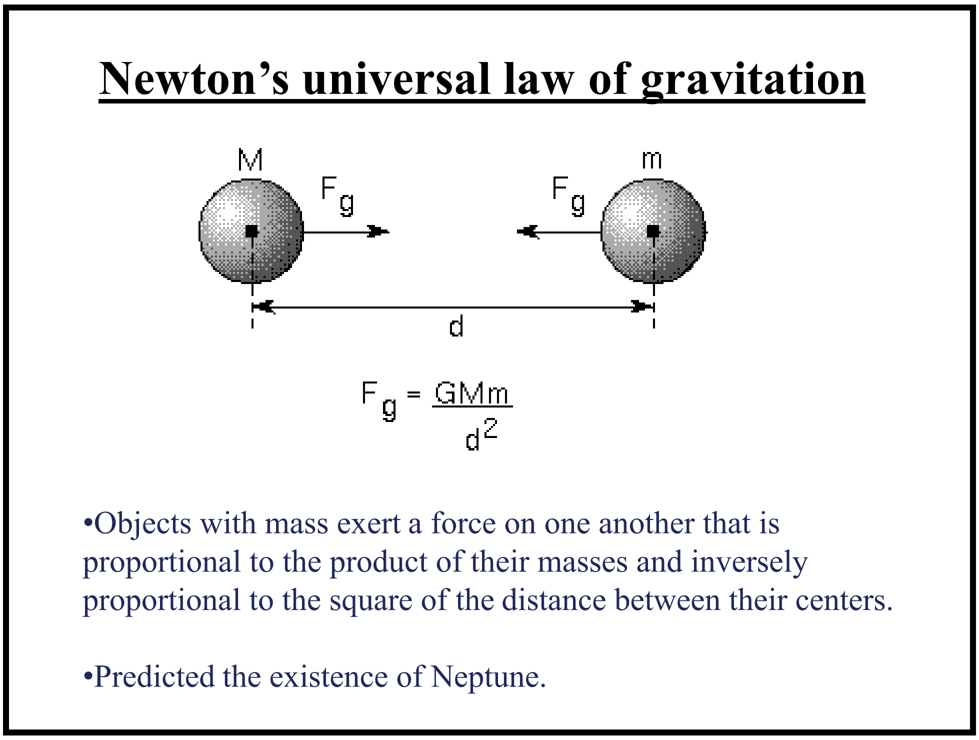


Figure 2.5 Newton's law of gravity

With the publication of the *Principia*, it seemed as if the science of mechanics was complete. The relationship between applied force and subsequent motion was now firmly established. The single cosmological force, gravity, had been completely described. Objects moved in accordance with strict natural laws that could be understood mathematically. Some continental scientists and philosophers were at first skeptical. They felt that Newton's concept of gravity as a force acting through a distance was insufficiently mechanical to be correct. Newton was also bothered by this. In a letter to a fellow scientist, he said,

> "That gravity should be innate, inherent, and essential to matter, so that one body may act on another body at a distance through a vacuum,…is to me so great an absurdity that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it."

However the spectacular success of Newton's mathematical description of the motion of both earthly and heavenly objects soon overcame the philosophical objections and Newton was celebrated as the greatest scientist who had ever lived. (We will return to the problem of action-at-a-distance in a later chapter.)

Newton viewed the material universe as consisting of atoms whose motion is determined by precise mathematical laws. Newton and virtually all of his contemporaries took the existence of a Creator as an obvious fact. However, this mechanical universe brought into question the role of the Creator with respect to the universe. Does the Creator interfere with the mechanical cause and effect from time to time? Or, did the Creator create the universe and the laws governing it and then allow the universe to evolve in accordance with those laws?

Newton believed the former. He felt that divine intervention was necessary for the creation of the solar system and was also necessary to keep it operating smoothly. Most of the scientists and philosophers that succeeded Newton rejected his theistic arguments. Newton's rival Leibniz, for example, thought that God created the universal machine, set it in motion, and then had no need to intervene further in its operation. The universe unfolded according to mathematical laws with all the precision and inevitability of a well-made clock. This religious perspective is known as Deism.

Later, the French physicist and mathematician, Pierre Simon Laplace, developed this idea further. In a famous quote, he said:

> We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.

Laplace's perspective is known as the mechanistic-deterministic worldview.

The Enlightenment refers to a particular historic period that occurred primarily in Western Europe and the United States in roughly the 18^{th} and 19^{th} centuries. It is the time when the accomplishments of the Scientific Revolution were incorporated into a new worldview now called the modern worldview. Enlightenment also refers to the ideal of putting the whole of human life under the rule of reason. What is accepted as truth, to be believed or as a principle to put into practice, should not be based on authority, but on reasons that are judged to be sufficient in and of themselves. The Enlightenment was also a time of increased awareness of human rights, a time when more liberal political systems were established. The founding fathers of the American Revolution acted in the spirit of the Enlightenment.

3 Nineteenth Century Physics

Your goals for this chapter are to know the following:

- The concept of atoms from the time of the Greeks to the end of the nineteenth century.
- What evidence pointed to the wave nature of light in the nineteenth century.
- How the concept of energy differs from matter and light.
- How a gas was visualized in the nineteenth century. What changes occur in the gas as the temperature increases.
- How to solve problems involving the mathematical relationship for the velocity, frequency, and wavelength of a wave.

By the beginning of the 19th century, the mechanistic-deterministic model of the universe was well established. Although there was still no direct evidence for the existence of atoms and no clear model for them, most scientists accepted atoms as the fundamental components of the material universe. Newton's laws of motion described the behavior of objects acted upon by forces, and atoms were considered no more than very tiny objects. The nature of light was still unknown, but scientists were confident that it was just a matter of time before they determined the laws governing its behavior with the same clarity and certainty that Newton had established for mechanics. By the end of the 19th century, this goal had been accomplished and it appeared as if the laws of physics were complete.

3.1 Atoms in the 19th Century

Empedocles was a Greek philosopher who sometime around 450 BCE introduced the concept that all matter is made up, in differing proportions, of four elemental substances, earth, air, fire and water, and that it is the ratios of these substances within a particular object that determine its properties. Empedocles' theory was an important development in scientific thinking because it was the first to suggest that some substances that looked like pure materials, like stone, were actually made up of a combination of different elements.

Shortly afterwards, Leucippus introduced the idea of atoms, and his student, Democritus, developed the concept in more detail. Democritus taught that "nothing exists except atoms and the void: all else is mere opinion." He conceived of atoms as indivisible and eternal, and in fact the word atom is derived from the Greek word meaning "that which cannot be divided." The atoms of Democritus were small, discrete, and identical in composition, though they might differ in size and shape with differences in size and shape determining the specific properties of the substance. According to Democritus, perceived changes in the world were produced by changes in the groupings of atoms.

Aristotle accepted the theory of Empedocles but rejected that of Democritus. He argued that logic ruled out the concept of discrete atoms. The atoms of Democritus had extension in space and were identical in composition, but these properties are not compatible with indivisibility: extension in space implies divisibility and composition implies yet smaller parts. Aristotle taught that matter is continuous rather than discrete in nature and that "everything continuous is divisible into divisibles that are infinitely divisible." During the Middle Ages, it was considered heresy to believe in atoms, as the concept appeared to be in conflict with the transubstantiation of bread and wine into the body and blood of Christ.

Despite philosophical difficulties, the concept of atoms was too persuasive to be dismissed and by the 17th century most scientists, including Galileo and Newton, were atomists. In the early 19th century, the English chemist John Dalton used the concept to account for chemical reactions, placing atomic theory in a scientifically respectable context for the first time. By the end of the 19th century, the concept of the atom was well established, though some influential scientists at the time did not accept the atom as a real constituent of nature.

3.2 The Chemical Elements

A chemical element is any substance that cannot be broken down into simpler materials by chemical means. Hydrogen, oxygen, and iron are examples of common elements. An atom is the smallest possible quantity of an element. Such substances as salt and water, on the other hand, can be broken down chemically, and are called compounds. Compounds are composed of molecules, which are chemical combinations of atoms. For example, a water molecule is composed to two hydrogen atoms and one oxygen atom, and written H_2O.

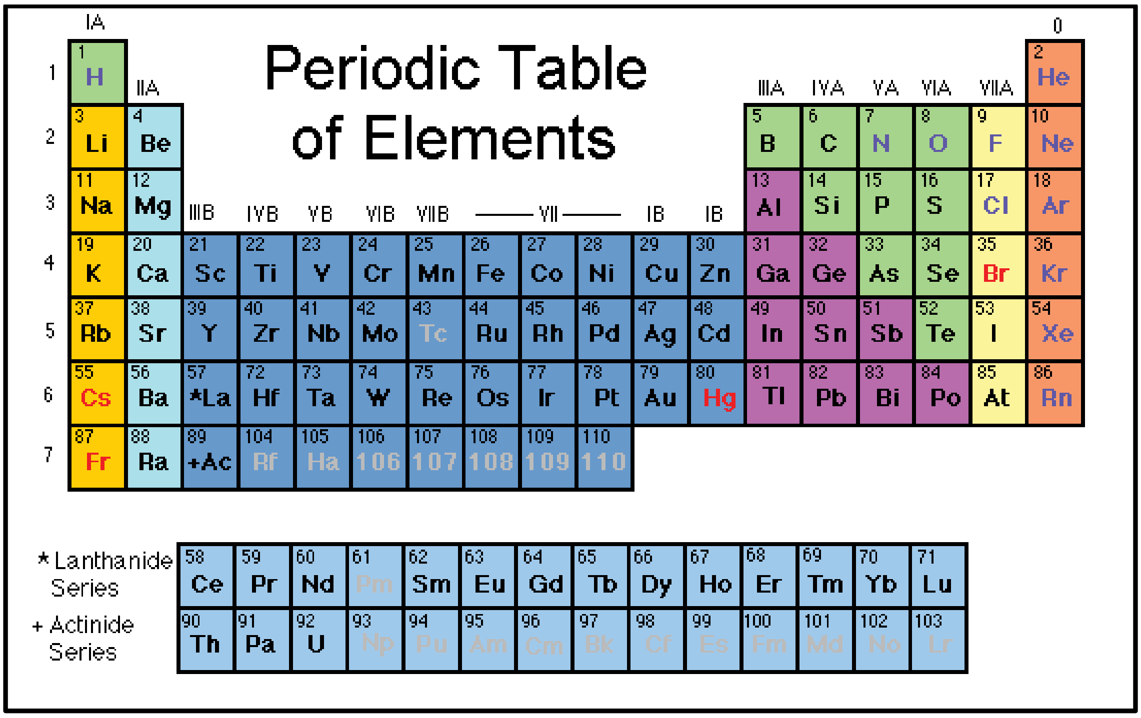


Figure 3.1 Periodic Table

3.3 The Nature of Light

Questions about the nature of light began at least as far back as the ancient Greeks. Various models were suggested, but it wasn't until the second half of the 17^{th} century that the nature of light began to be formally debated among scientists. The one thing that was clear was that light could transfer energy from one place to another. For example, sunlight can heat water. Two schools of thought arose as to how this was accomplished.

Robert Hooke and Christian Huygens, contemporaries of Newton, proposed a wave model. According to this view, light was like sound. When one speaks, the vocal cords cause the air in the throat to oscillate, which causes the air near it to oscillate, and so on until the air in the vicinity of the listener's eardrum oscillates. Each particle in the medium merely oscillates about some fixed position; there is no net transfer of air, just energy.

Newton preferred the idea of particles moving through space from one material object to another. When one throws a baseball at a target, energy is transferred. In this same manner, Newton viewed light as a stream of tiny particles moving at high speeds. Like bullets fired from a gun.

Although each model made different predictions regarding the behavior of light, at the time the technology required to distinguish clearly between the two did not exist. This began to change in the early part of the 19th century when evidence began to accumulate in favor of the wave model. In the first decade of the century, Thomas Young showed through his famous double-slit experiment that light could bend around obstacles and could interfere with itself. This behavior was predicted by the wave model. Also in mid-century the speed of light was measured, both in air and in water. As the wave model predicted, the speed of light in water was less than in air. The particle model predicted just the opposite.

The apparent *coup de grace* to the particle theory was delivered in 1865 by James Clerk Maxwell, an English theoretical physicist. Electric and magnetic phenomena, considered from ancient times to be separate and distinct, were discovered in 1820 to be closely related to one another. In the decades that followed, electromagnetism – as the science came to be known – was experimentally studied in detail and a tremendous body of knowledge was accumulated. Maxwell was able to represent all of the then known electromagnetic phenomena with a single physical theory consisting of four equations.

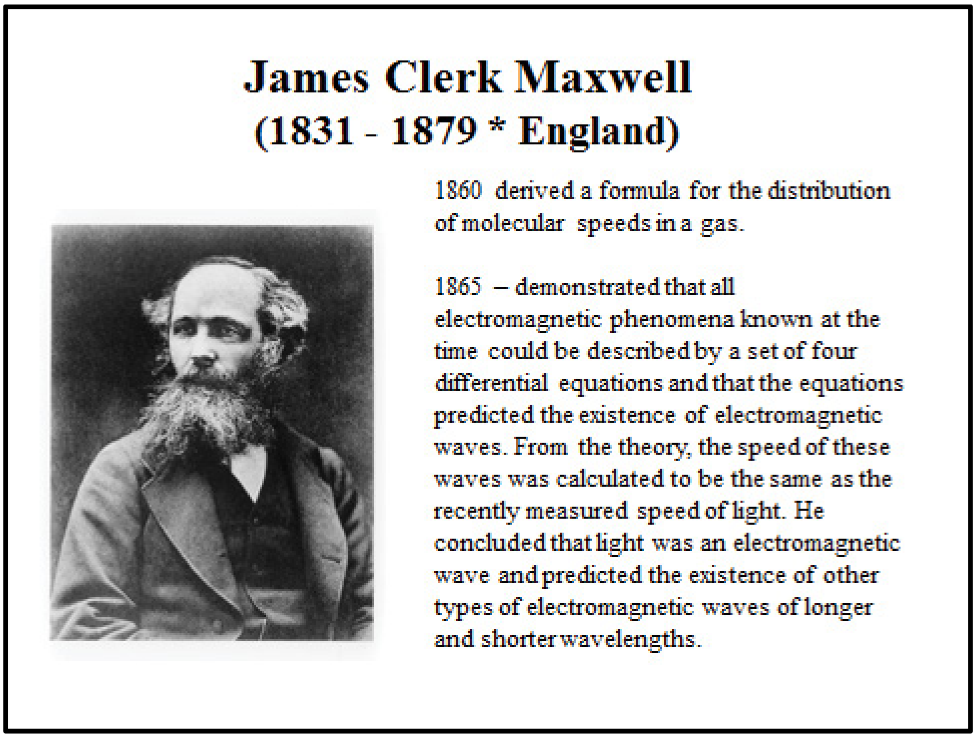


Figure 3.1 James Maxwell

In addition to accounting for all the known electromagnetic effects, the theory predicted several completely unsuspected oness. The most significant was the existence of electromagnetic waves – electromagnetic disturbances propagated through space. Maxwell calculated the speed of these waves from the basic equations of his theory and obtained the previously measured speed of light. To him this could not be a coincidence, and he concluded that light was an electromagnetic wave. In Maxwell's theory, light was just one small portion of the spectrum of electromagnetic waves. He realized that electromagnetic waves of much shorter and much longer wavelengths must also exist.

In 1888 the theoretical predictions of Maxwell's theory were experimentally confirmed in the laboratory by Heinrich Hertz when he produced and detected radio waves, a very long wavelength form of electromagnetic radiation.

In 1889, Hertz said,

> We know that light is a wave motion. We know the speed of the waves, we know their length. In a word, we know completely the geometric relationships of this motion. These things no longer permit of any doubt, and a refutation of this view is unthinkable to the physicist. In so far as human beings can know the truth, the wave theory is a certainty.

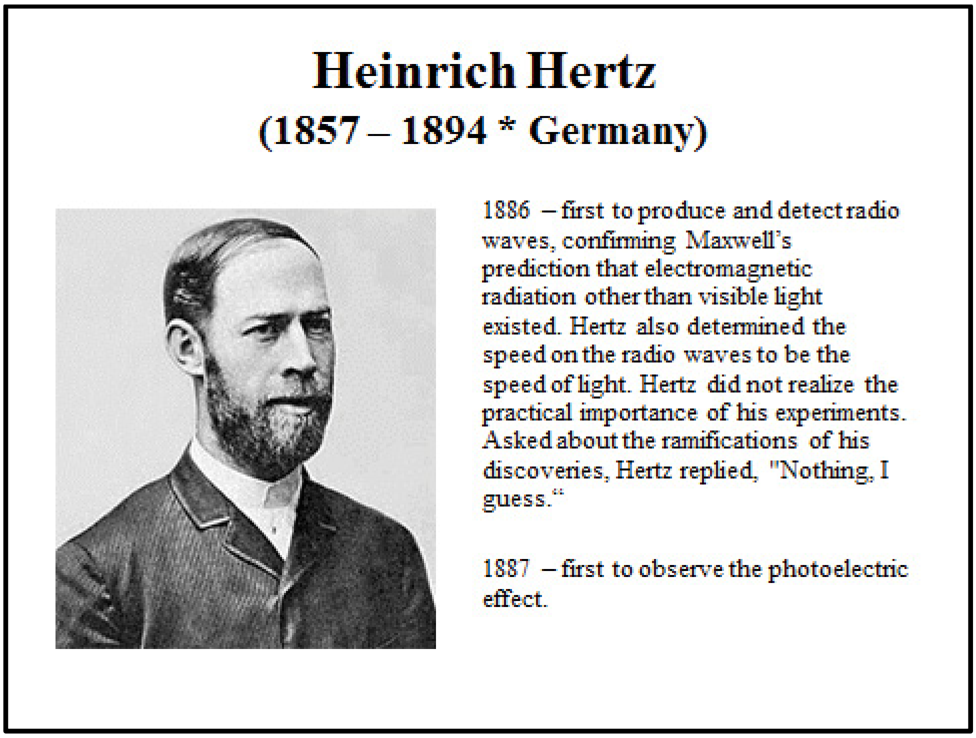


Figure 3.3 Heinrich Hertz

The two physical quantities most closely associated with the wave model are frequency and wavelength. The frequency of a wave is the number of complete cycles emitted per second, or, equivalently, the number of complete cycles passing a given point per second. This, together with the wavelength, the length of one complete cycle, determines the velocity with which the wave is propagated through the medium. For example, if the frequency of a wave is 15 cycles per second and the wavelength is 3 feet, the velocity of the wave is

$$3 \text{ feet/cycle} \times 15 \text{ cycles/second} = 45 \text{ feet/second}$$

This relationship is represented by the equation $\lambda\nu = v$, where the Greek letter lambda, λ, is the wavelength, the Greek letter nu, ν, is the frequency, and v is the speed of the wave. If a charged particle oscillates with a frequency of 4.9×10^{14} cycles per second, an electromagnetic wave will be produced whose frequency is 4.9×10^{14} cycles per second and whose wavelength is equal to the speed of light divided by the frequency.

$$\lambda = v/\nu = 3 \times 10^{8} \text{ meters/second}/4.9 \times 10^{14} \text{ cycles per second}$$

$$\lambda = 6.1 \times 10^{-7} \text{ meters.}$$

This wavelength is in the visible region of the electromagnetic spectrum and would appear to the human eye as red light.

In honor of Heinrich Hertz the unit for frequency, the cycle per second, has been renamed the Hertz and abbreviated Hz. Radio waves can be produced by an alternating current in an antenna. If the current oscillates with a frequency of 1×10^6 Hz or 1 megaHz, the resulting wavelength is 2 meters, in the AM radio region of the electromagnetic spectrum.

3.4 Energy

The events, processes, and interactions by which reality manifests itself consist of the transfer and/or transformation of energy. Perhaps the most fundamental law of nature is conservation of energy. That is, for an isolated system, energy can be neither created nor destroyed. It can be transformed from one type of energy into another and transferred from one object to another, but at the end of a process, the total amount of energy present in the system is exactly the same as it was at the start of the process. The application of this law and the understanding of its importance to science, along with Maxwell's development of electromagnetic theory, are perhaps the two most significant accomplishments of 19th century physics.

Electromagnetic radiation is a form of energy. The energy of motion, kinetic energy, is another. Depending on the nature of the processes involved, there are many ways to describe the energy of the various components of a system, but in all cases, energy is a conserved quantitative measure of the ability to create an effect – to do work.

Heat is also a form of energy, but not one that has an obvious description in terms of Newtonian physics. A cup of hot water appears pretty much the same as a cup of cold water. It is not until you put your finger in it that the difference becomes obvious. It is clear that the hotness or coldness of a body is determined by some internal property of the body. Because physicists were certain that all physical phenomena are ultimately explainable in terms of Newtonian physics, the development of a mechanical theory of heat became an important goal. The success of this effort was one of the great triumphs of physics in the 19^{th} century.

Perhaps the most obvious and fundamental property of heat is that when a hot object is in thermal contact with a cold object, the hot object will cool and the cold object will heat up. It is said that heat 'flows' from the hotter object to the cooler object. This phrase probably comes from the 18^{th} century concept of heat as literally a self-repellent fluid that flowed out of the hot object and into the cold object. This fluid was called 'caloric.' There were many problems with the theory and by the middle of the 19^{th} century it had been abandoned. By then, it was clear that heat was a form of energy, and thermodynamics, the study of the transformations between heat and mechanical energy, was being developed. This new science was stimulated by the Industrial Revolution, especially by the invention of the steam engine. By relating heat to energy, the laws of thermodynamics gave a firm indication of a connection between the theory of heat and the theory of motion, and hence, to Newton's laws.

By 1860, atomic theory was almost universally accepted. It provided a coherent, consistent model for chemical reactions and for the structure of matter. The basic particles, atoms or molecules, were considered to be in constant random motion, with the balance between the motions of the particles and the restraining forces of attraction between them determining whether the material was a solid, liquid, or a gas. Temperature was recognized as a measure of the average kinetic energy of the particles with the heat content determined by the temperature and the quantity of the material, that is, by the total kinetic energy of the particles. Heat transfer was accomplished when the faster moving particles of the hotter body collided with the slower moving particles of the cooler body. The collisions would slow down the particles of the hotter body and speed up the particles of the cooler body, thus cooling the hotter body and warming the cooler body.

Within a single body, the collisions will change the kinetic energies of the particles in an unpredictable way. As a result, some of the particles will be moving faster than others. The first person to develop a distribution function for the velocity of the particles as a function of temperature was James Clerk Maxwell. For any given substance at a specific temperature, there is a most probable speed. The most probable speed increases as the temperature increases. The probability that a particle will have a speed less than the most probable speed decreases as the speed approaches zero. Similarly, the probability that a particle will have a speed greater than the most probable speed decreases with the tail of the distribution extending to higher speeds as the temperature increases.

In 1885, Ludwig Boltzmann expanded on Maxwell's work, and the resulting expression is now known as the Maxwell-Boltzmann distribution function.

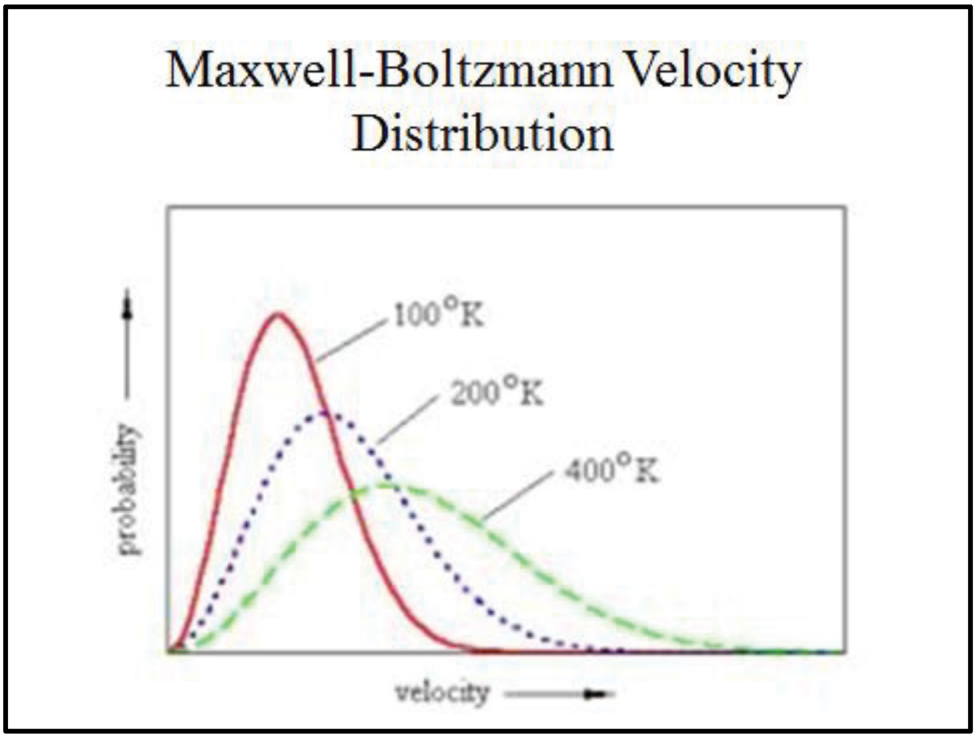


Figure 3.4 Velocity distribution

At the end of the 19th century the nature of light had been firmly established as an electromagnetic wave and other types of electromagnetic radiation had been discovered. The kinetic theory of heat provided a microscopic model for the concepts of temperature and heat transfer. Thermodynamics, with the introduction of the concept of entropy, was able to account for the macroscopic effects of work, heat and energy on a system. Additional applications of Newtonian mechanics had even further established its validity. It seemed as if every possible aspect of physical reality was accounted for. This body of knowledge is collectively known as classical physics. There were a few minor details, particularly some developments late in the century, that seemed to be puzzling, but there was no doubt in anyone's mind that, when they were eventually understood, the explanations would consist of some clever application of the physics at hand.

4 Unresolved Questions at the Beginning of the Twentieth Century

Your goals for this chapter are to know the following:

- What it was about the following phenomena that were in conflict with classical physics: Thermal radiation, the ether, atomic spectra, the orbit of Mercury, the electron, the photoelectric effect, radioactivity.
- How scientists at the end of the nineteenth century believed these problems would be resolved.

With Newtonian physics, electromagnetism, the kinetic theory of heat, and thermodynamics (what we now refer to as classical physics) well established, it appeared to physicists at the start of the 20th century that it was only a matter of time until all phenomena in the physical universe would be described in complete detail. True, certain perplexing experimental results had come to light in the last decade of the 19th century, but physicists had no doubt classical physics could eventually explain these phenomena. A leading physicist was advising promising students not to go into physics as he felt all of the fundamental work had been done and all that remained for 20th century physicists was the rather mundane task of applying the theory and determining additional decimal places for the physical constants.

4.1 Observations That Appeared To Defy Analysis Using Classical Physics

Thermal Radiation – It is a well-known fact that hot objects give off electromagnetic radiation. It is most obvious when the object is hot enough to emit radiation in the visible portion of the electromagnetic spectrum, as is the case with the sun or a light bulb filament. The fact is that all objects emit electromagnetic radiation, but cooler objects emit only in the radio and infrared regions.

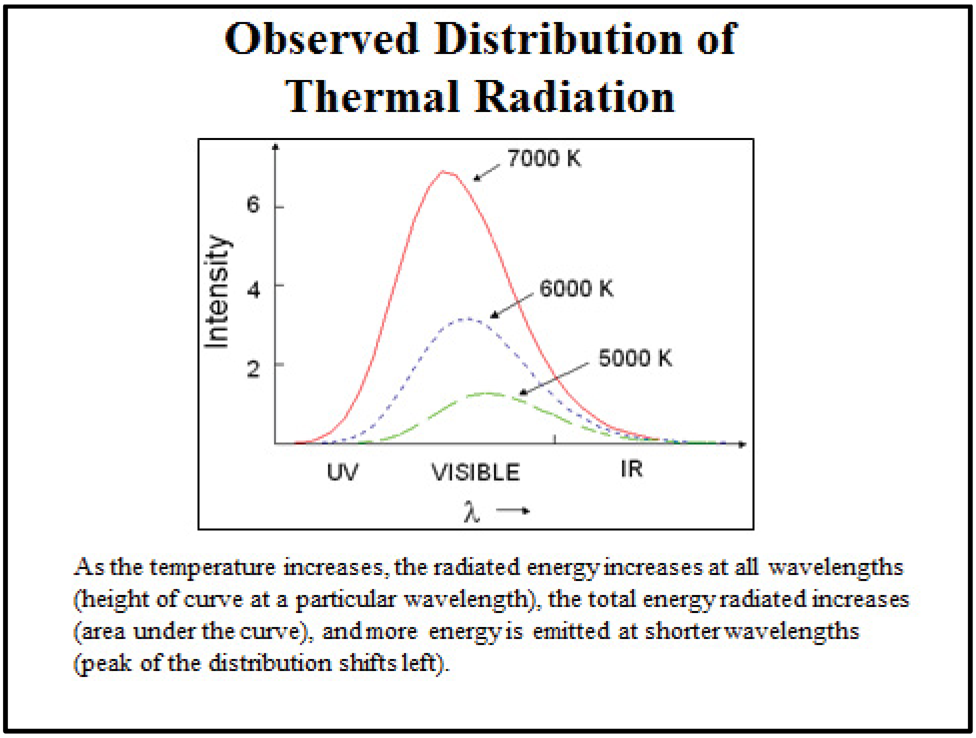


Figure 4.1 Thermal Radiation

In the late 1800s, Lord Rayleigh rigorously applied the principles of classical physics to the problem of thermal radiation. The results were blatantly wrong. They indicated any object, regardless of its temperature, will radiate more electromagnetic energy in the ultraviolet region (the shortest wavelength electromagnetic radiation known at that time) than it will in the visible or infrared. This classical treatment is aptly known as the 'ultraviolet catastrophe'. In fact, the actual distribution of thermal radiation had been accurately measured by that time and low temperature objects emit virtually no energy in the ultraviolet region.

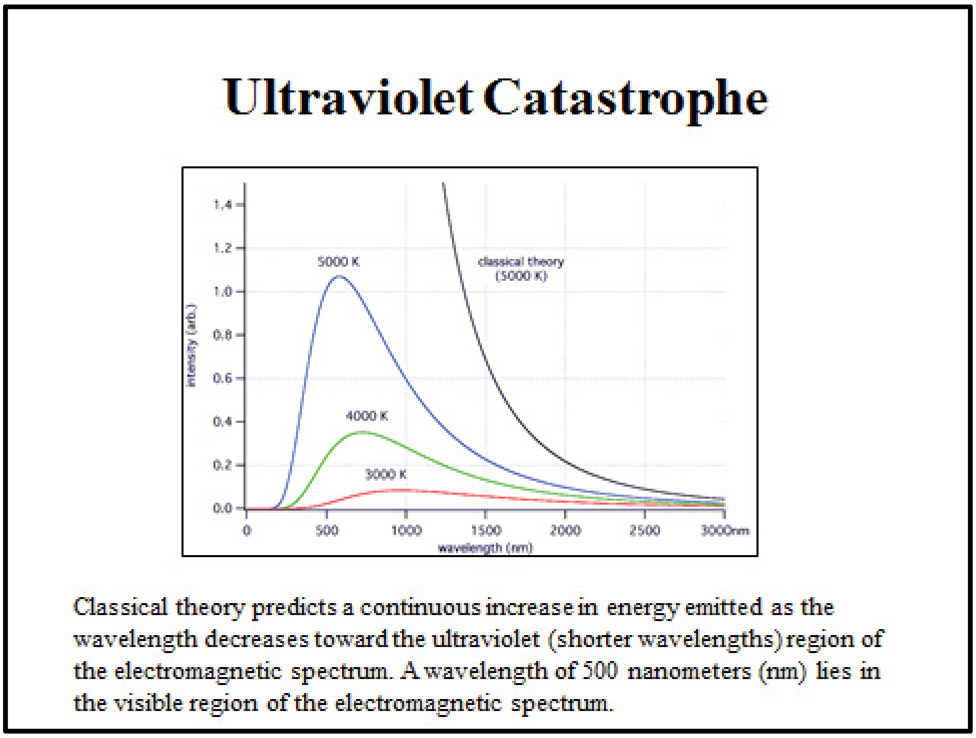


Figure 4.2 Ultraviolet Catastrophe

The Ether – An integral part of classical electromagnetic theory is the ether. The ether was supposed to be the invisible, weightless, substance which permeated the entire universe and served as the medium through which electromagnetic waves were propagated in much the same way that sound is propagated through the air. A wave by its very nature is an oscillating disturbance in some medium and ether was postulated as the medium that oscillated to propagate electromagnetic waves.

In 1887 Albert Michelson and Edward Morley performed experiments intended to demonstrate the existence of the ether. The speed of light as calculated by Maxwell's equations was taken to be the speed of light through the ether – that is the speed as measured by an observer at rest with respect to the ether. Thus in a frame of reference moving through the ether, an observer should obtain a different value than an observer at rest with respect to the ether, with the difference depending on the relative speed of the observer with respect to the ether. Michelson and Morley used the motion of the earth in its orbit around the sun to demonstrate this. They made simultaneous measurements of the speed of light in the direction of the earth's motion and the speed perpendicular to the motion. The classical analysis required that they differ. To everyone's great surprise, the speeds appeared to be exactly the same.

Atomic Spectra – In the last few decades of the 19th century, much experimental work was done on the electromagnetic radiation emitted by a low-density atomic gas. The distribution of the wavelengths emitted in this case differed considerably from that emitted in the case of thermal radiation. The radiation from an atomic gas did not form the continuous spectrum of thermal radiation, but rather a discrete spectrum. Only a limited number of wavelengths are emitted. Further, the discrete spectrum is characteristic of the emitting gas. An element could be identified simply by analyzing its emission spectrum. Classical physics could provide no explanation to account for this.

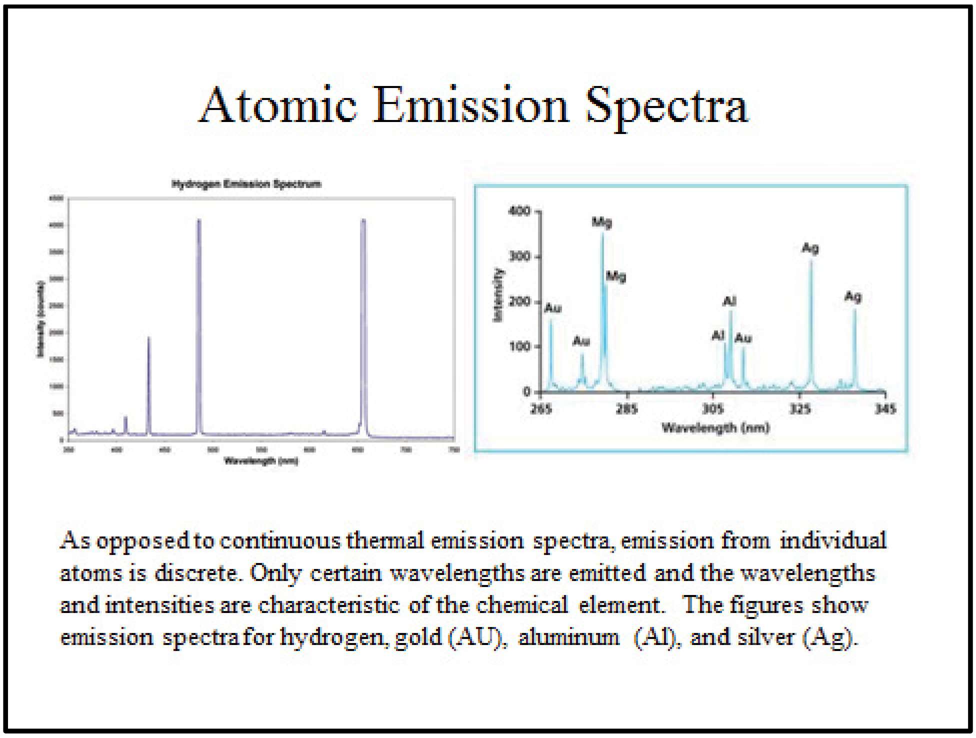


Figure 4.3 Atomic Spectra

Orbit of Mercury – Newton's law of gravity and his laws of motion could be used to calculate the orbits of the planets to an incredibly high degree of accuracy. The observed orbits of all the planets except Mercury correspond to the predictions based on classical Newtonian mechanics. For Mercury, the discrepancy was very, very slight, but undeniable. Newton's laws required that the axes of the elliptical orbit precess; that is, swing through a certain angle in a certain amount of time. However the amount of precession calculated was slightly different from that observed. Some astronomers attributed the discrepancy to the perturbing effects of an undiscovered planet located between Mercury and the sun. They even went so far as to name the planet Vulcan in anticipation of its discovery. This same strategy was used successfully to account for discrepancies in the orbit of Uranus, resulting in the discovery of Neptune. It soon became clear, however, that Vulcan did not exist.

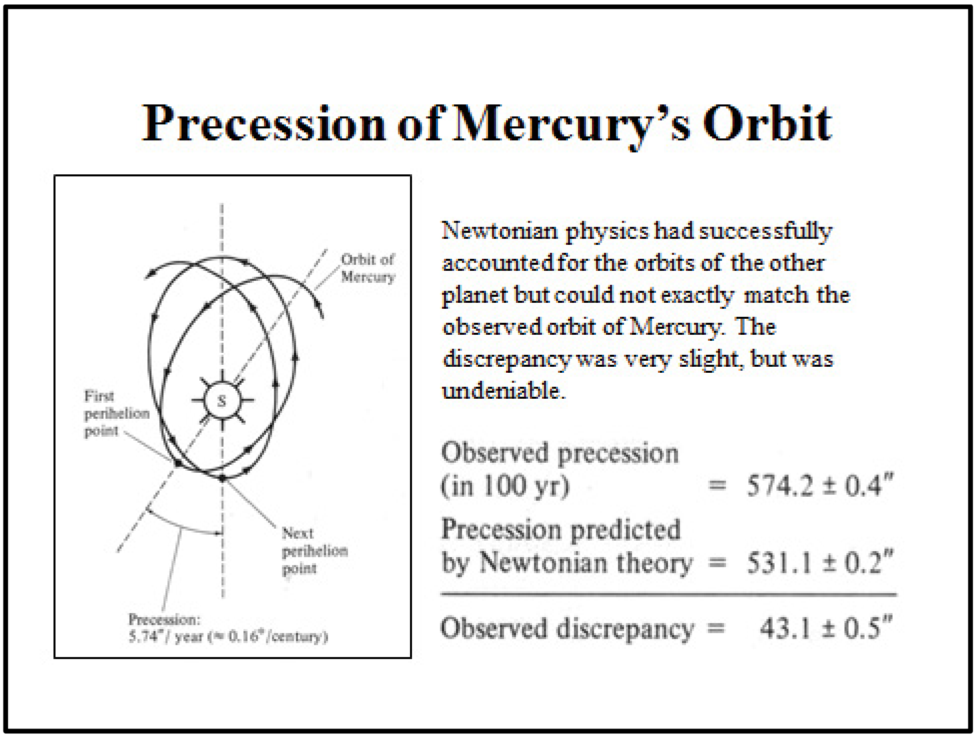


Figure 4.4 Mercury's Orbit

The Electron – If a high voltage is placed across the two ends of an evacuated glass tube, experiments demonstrate the existence of invisible rays emanating from the negative electrode or cathode. In 1869 these cathode rays were shown to travel in straight lines, and in 1870 they were shown to have both energy and momentum. In 1895 Jean Perrin discovered that the cathode rays carry negative charge by deflecting them in a magnetic field. J.J. Thomson was able to show, in 1897, that cathode rays are steams of a single type of negatively charged particle whose properties are independent of the material from which they are emitted. Thomson's experiments indicated that the mass of these particles is much, much less than the mass of even the lightest atom. These particles were recognized as being responsible for electricity and became known as electrons.

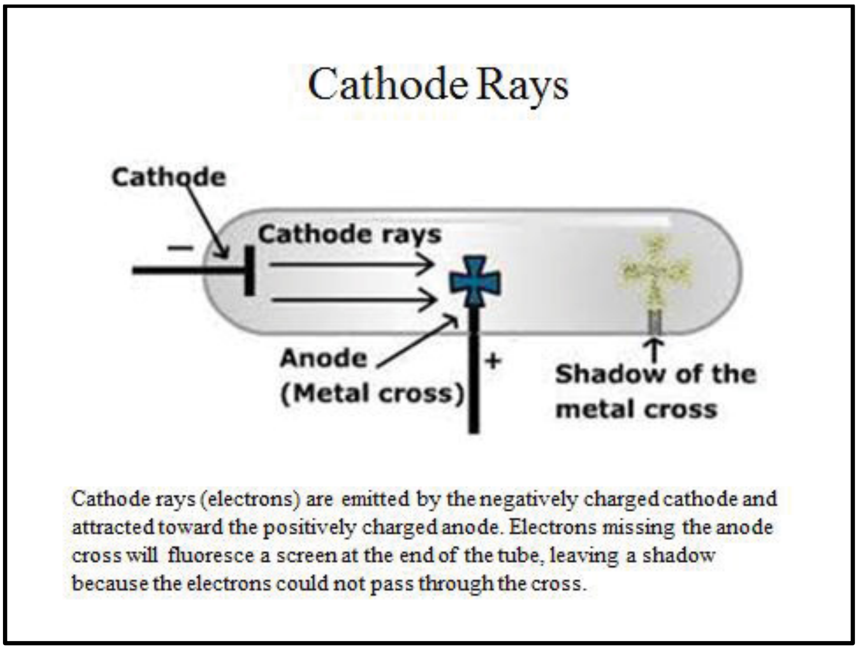


Figure 4.5 Cathode Rays

By 1900 it was well established that all atoms contain electrons as part of their internal structure and the race was on to develop an atomic model into which electrons could be incorporated in a logical way that would be consistent with the laws of classical physics.

Photoelectric Effect – In 1887 Heinrich Hertz discovered that it was possible for ultraviolet, or in some cases visible light, to free electrons from certain metallic surfaces. The photoelectric effect itself is not too surprising. Ultraviolet and visible light possess energy and therefore should have the ability to free electrons from the metal by transferring energy to the electrons. The difficulty comes when the photoelectric effect is studied in detail. For each type of metal, there is a minimum wavelength of radiation that will free electrons regardless of how much energy is transferred to the metal. On the other hand, below this minimum wavelength, electrons will be released, even for very low amounts of energy transferred. This behavior is incompatible with the classical wave model of electromagnetic radiation.

Radioactivity – The phenomenon of radioactivity was discovered in 1896 by the French physicist Henri Becequerel. In January of that year, Becquerel learned of an amazing discovery made by the German physicist Wilhelm Roentgen (1845–1923). When cathodes rays strike glass, they cause the glass to emit visible light. This phenomenon was well known and is called fluorescence. What Roentgen discovered is that in addition to the visible light, the fluorescent areas of the glass also emit an extremely penetrating radiation. Because the nature of this unexpected radiation was unknown, Roentgen simply called them X-rays. (Later X-rays were shown to be electromagnetic radiation with wavelengths shorter than ultraviolet light.) News of this mysterious radiation spread rapidly, and physicists all over the world began to study the properties of X-rays. Because X-rays could be used to produce dramatic photographs of bones inside a living body, the popular press splashed the story over the front pages.

When Becquerel learned about Roentgen's discovery, he immediately set out to try to discover whether the X-rays are simply a peculiar feature of cathode-ray tubes or whether they are associated with fluorescence in general. Becquerel knew that certain minerals will fluoresce when irradiated with ultraviolet light, so he set out to discover if X-rays are also associated with this fluorescence.

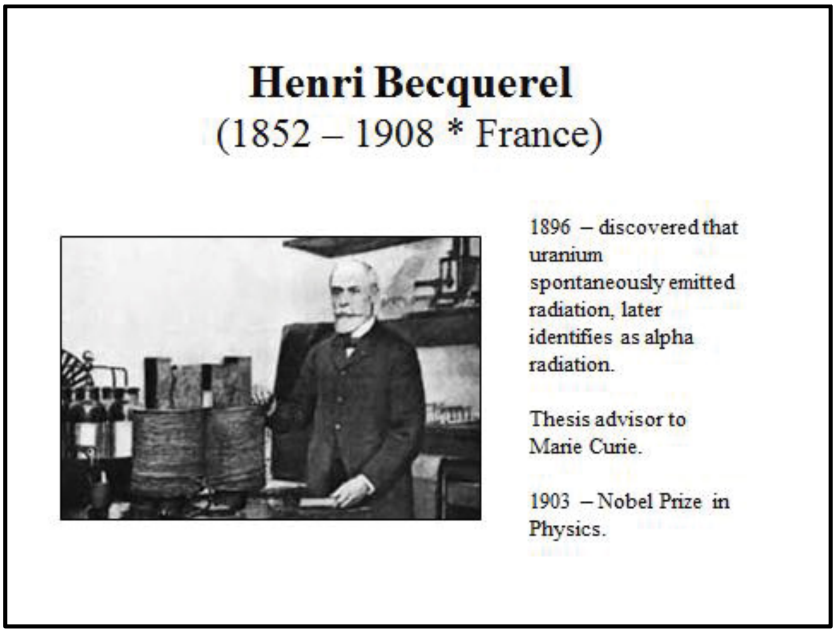


Figure 4.6 Henri Becquerel

Becquerel carefully wrapped a photographic plate with black paper to block visible and ultraviolet light, and placed the fluorescent mineral on top of the wrapped photographic plate. After irradiating the mineral with ultraviolet light to cause fluorescence, he developed the photographic plate to see if X-rays had penetrated the black paper to expose the film. His early experiments produced no exposure of the film. Then he used some fluorescent uranium minerals, which did cause exposure of the film, indicating that a very penetrating radiation was emitted by the fluorescing uranium minerals. Naturally Becquerel assumed that this radiation was X-rays.

One day Becquerel happened to develop some photographic plates that had been left in a drawer with samples of the uranium minerals. The minerals had not been exposed to ultraviolet light and had therefore not fluoresced. The plates had been wrapped in black paper, so there was no reason to expect any exposure of the plates, but they were exposed. Subsequent experiments showed that the uranium minerals spontaneously emit the penetrating radiation, even if they are not fluorescing. Becquerel was able to show that it was specifically the uranium atoms in the minerals that were emitting the radiation. Any sample of uranium spontaneously emits this radiation without any external energy supply. This phenomenon is quite different from that observed by Roentgen where X-rays are emitted only when glass is bombarded by cathode rays. Becquerel's new phenomenon was called radioactivity. A substance that emits this spontaneous radiation is said to be radioactive. Classical physics was at a loss to explain the nature of this radiation.

4.2 The Scientific Worldview at the Start of the 20-th Century

In spite of these perplexing experimental results, at the dawn of the 20th century, physicists believed that their theoretical understanding of the physical universe was complete. Newton had provided a description of gravity, the only force acting on a universal scale, as well as the laws governing the motion of objects acted upon by forces. Maxwell had provided a complete description of electromagnetic phenomena, including electromagnetic radiation. Heat was now understood as energy of motion and the laws of thermodynamics had been established. Atoms were understood as tiny, submicroscopic objects obeying the same laws of motion as ordinary-sized macroscopic objects. Space and time were understood as absolute and independent structures within which the phenomena of the physical universe unfolded.

This mechanistic-deterministic universe seemed so obvious, it must be true. There was little doubt that very soon some clever physicist would figure out how to account for all of the strange experimental results mentioned above, and that the explanations would lie completely within the theoretical framework of classical physics. Such hubris seldom goes unpunished.

5 Max Planck and the Concept of the Quantum

Your goals for this chapter are to know the following:

- How Max Planck solved the thermal radiation problem. How to explain in your own words why this is such a radical concept that even Planck did not believe it was true.
- What the symbols in the equation E = hν represent and how to solve problems involving this relationship.

Thermal radiation refers to the electromagnetic radiation emitted by a body as a consequence of its temperature. By the end of the 19th century, experimental physicists had completely determined the distribution of the radiation as a function of temperature.

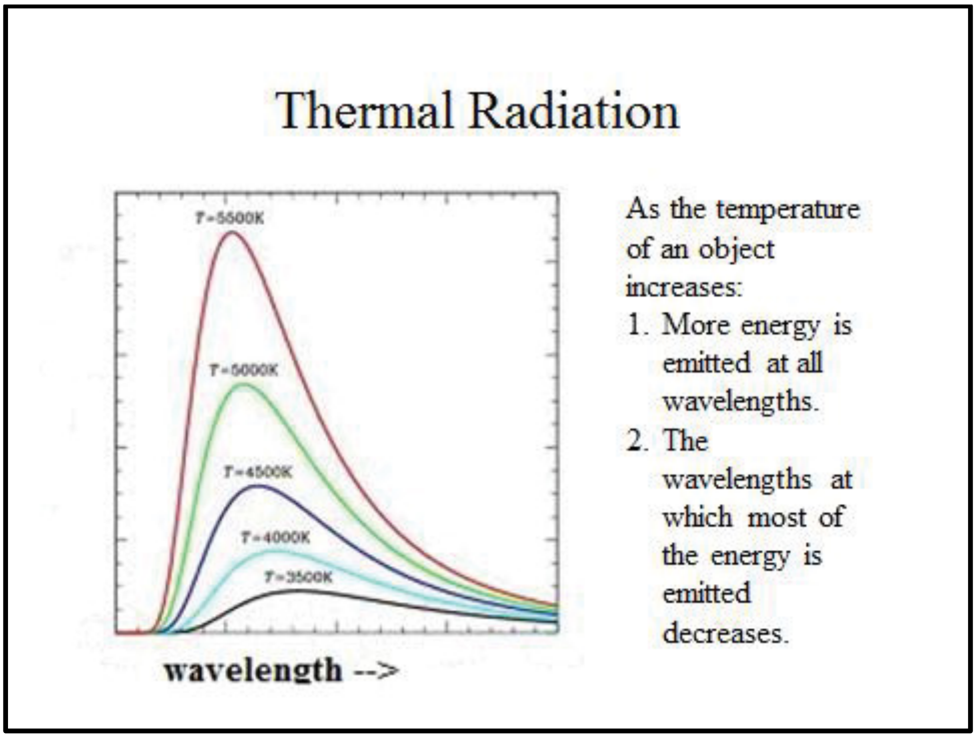


Figure 5.1 Thermal Radiation

Theoretical physicists, however, were at a complete loss to explain the fundamental physics behind this distribution. In fact, it was clear that using classical physics as a basis resulted in the nonsense outcome known as the ultraviolet catastrophe.

5.1 Max Planck

For years the German theoretical physicist, Max Planck, wrestled with this problem. In Maxwell theory of electromagnetism, electromagnetic radiation is emitted by oscillating charged particles with the frequency of the radiation being determined by the frequency of the oscillation. In Newtonian mechanics, energy of the oscillator is a continuous physical quantity; that is its value is not restricted in any way.

The energy of an oscillating object is determined only by the amplitude of the oscillation. In his study of thermal radiation, Planck assumed that both Maxwell's theory and the continuity of energy were correct. However, to simplify his calculations, he finally tried the mathematical trick of dividing the energy distribution of the oscillators into small units he called quanta, with the energy of the quanta being proportional to the frequency of the oscillator rather than the amplitude of the oscillation. Using these quanta, he was able to theoretically reproduce the observed wavelength distribution. That is, almost no energy emitted at either the very long or very short wavelengths and with most of the energy being radiated at intermediate values with the range depending on the temperature. This avoided the ultraviolet catastrophe of earlier calculations.

He had assumed from the beginning that once a suitable solution was found using these hypothetical quanta, he could use standard mathematical techniques to shrink the quanta down to infinitesimal size, yielding the expected continuous energy distribution for the oscillating charges. To his bitter disappointment, this was not the case. The calculation only produced the correct result if he kept the quanta of energy above some minimum size.

5.2 The Concept of the Quantum

The way the quanta avoided the ultraviolet catastrophe, was by making the size of the energy quanta proportional to the frequency. Thus the higher the frequency of the oscillator, the more energy it needed to oscillate at that frequency. Maxwell had earlier determined how the energy of a substance at a given temperature is distributed among the individual particles of the substance. The Maxwell distribution is bell-shaped with almost no particles having either very low or very high energies. Therefore, as the frequency increases, the number of charges oscillating with that frequency decreases and thus the amount of electromagnetic radiation emitted at the higher frequencies decreases to zero, in exactly the way that the data indicated it did. In this way, the ultraviolet catastrophe is prevented.

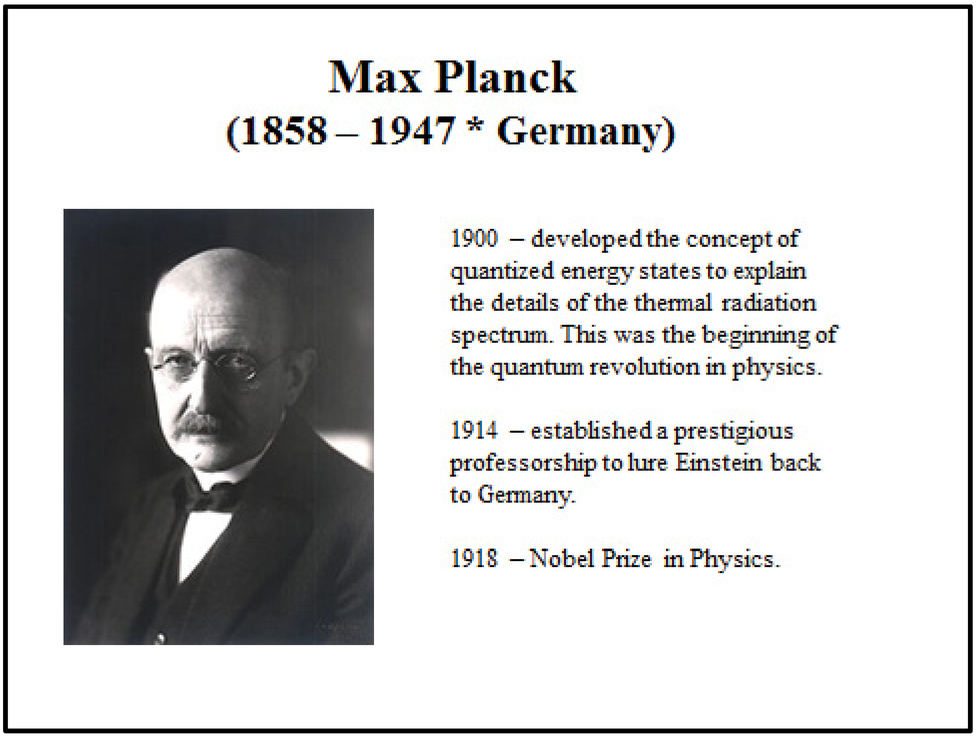


Figure 5.2 Max Planck

The proportionality between the energy of the quanta and the frequency of the oscillation is now called Planck's constant and is one of the most fundamental constants in nature.

$$E = h\nu$$

where E is the energy of the oscillator, h is Planck's constant, and nu, ν, is the frequency of the oscillation. Planck's constant, expressed in the standard set of units used in physics is

$$6.62 \times 10^{-35} \text{ joule-seconds.}$$

The concept of quantized physical quantities was not new to physics. Consider the mass of a sample of iron. Although on the macroscopic scale we can treat the mass of the iron as a continuous quantity, it is in fact quantized. The quantum unit in this case is the mass of a single iron atom. A mass of 3.2 iron atoms is not an allowed mass for iron. Because this quantum of mass is infinitesimal, we do not notice this on the macroscopic scale. Charge is another physical quantity that appears continuous on the macroscopic scale but is actually quantized with the quantum unit being the charge on the electron. All charges are just integer multiples of this basic unit.

In a similar way, the energy of an oscillator is quantized, but because of the incredibly small value of Planck's constant, the quantum unit, hν, is too small to be noticed on the macroscopic scale. The difference is that although there are obvious reasons why the mass of iron and the charge on an object are quantized, there is no explanation at all in classical physics why the energy of an oscillator should be quantized. Planck's explanation of thermal radiation is in direct conflict with classical physics, not to mention commonsense.

Imagine a pendulum, a macroscopic example of an oscillator. If a pendulum is set in motion, it will oscillate with a particular frequency depending on its length but independent of its amplitude. This fact was discovered by Galileo. The energy of the pendulum is determined by the amplitude. If you give it more energy it will oscillate with a larger amplitude and vice versa. To say, as Planck did, that the energy of an oscillator is quantized, is to say that the amplitude of the pendulum's oscillations is quantized; only certain amplitudes are allowed. On the surface of it, this is nonsense. For example, how does a pendulum get from one allowed amplitude to another without passing through an amplitude that is not allowed?

Planck was 42 years old at the time, and a thoroughgoing classical physicist. He refused to believe that the energy of oscillation really is quantized. He published his result only because it was a calculation that for the first time produced a correct result. He was absolute certain that the model was wrong, that the quanta did not exist, and that some clever physicist would soon discover where he had gone wrong. In retrospect, we know that it marked the beginning of a revolution that would shake the very foundations of our worldview. Planck's presentation of these ideas before a meeting of the German Physical Society on December 14, 1900, is considered the birthday of modern physics.

6 Albert Einstein and the Miracle Year

Your goals for this chapter are to know the following:

- Einstein's suggestion for the mechanism of the photoelectric effect.
- What is meant by wave-particle dualism.
- Einstein's explanation for Brownian motion. The significance of the confirmation of this explanation.
- Why nineteenth century scientists believed that Maxwell's equations did not satisfy the principle of relativity.
- The two postulates of the theory of relativity.
- What is meant by the equivalence of mass and energy. The role of the speed of light in this equivalence.

The year 1905 is often referred to as the Annus Mirabilis, the year Albert Einstein, a 23-year old with an undergraduate degree in physics, working as a Swiss patent clerk, published four papers in the Annalen der Physik, the leading German scientific journal, each of which contributed significantly to modern physics and one of which won him the Nobel Prize in Physics.

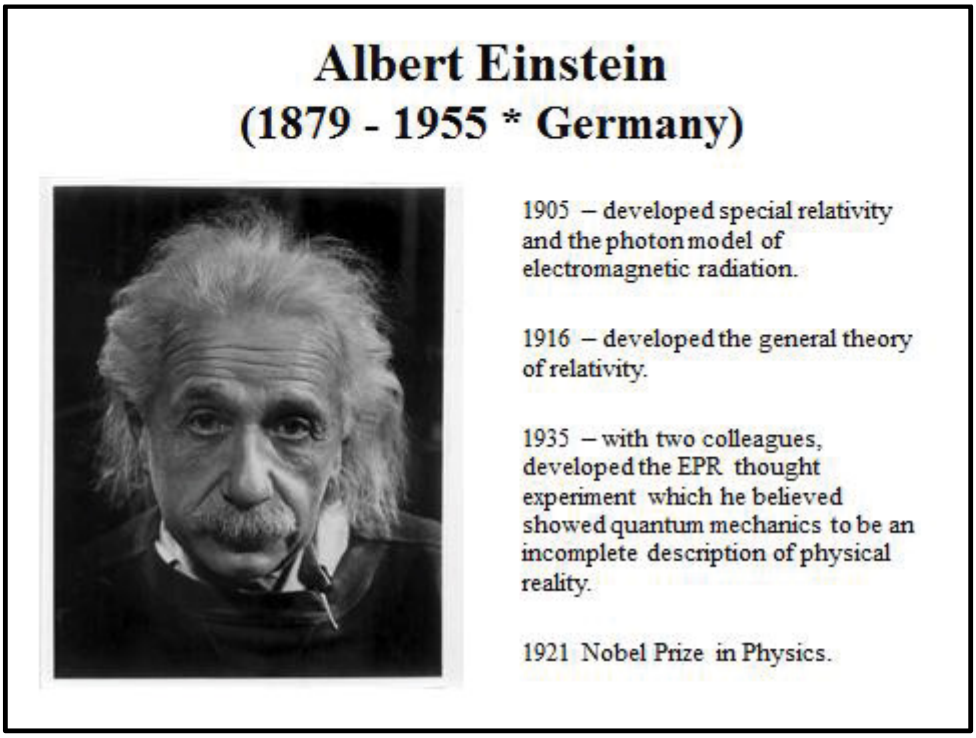


Figure 6.1 Albert Einstein

6.1 The Photoelectric Effect and the Quantum Nature of Electromagnetic Radiation

The first of these papers proposed that the classical wave model was not a complete description of electromagnetic radiation. While there could be little doubt that electromagnetic radiation sometimes behaved as if it were a wave, Einstein proposed that the energy of the radiation could be absorbed or emitted only in discrete amounts. Borrowing Planck's term, he called these discrete amounts quanta. In his paper, Einstein states,

> Energy, during the propagation of a ray of light, is not continuously distributed over steadily increasing spaces, but it consists of a finite number of energy quanta localized at points in space, moving without dividing and capable of being absorbed or generated only as entities.

The energy of these quanta was determined by the frequency of the radiation, being equal to Planck's constant times the frequency, again as in Planck's theory, but with a fundamental difference. In Planck's theory, hv represented the energy of the charged oscillators producing the electromagnetic radiation. Planck viewed the radiation itself according to the classical wave theory, while Einstein saw the radiation as consisting of the quanta; that is as a particle-like property of the radiation. In his paper, Einstein suggested that a thorough analysis of the photoelectric effect would support this model. At the time, only a few, very preliminary experimental results were available. The American experimental physicist, Robert Millikan, took up the challenge, undertaking a decade-long experimental program using the photoelectric effect to test Einstein's predictions. His motivation was to prove Einstein wrong, as he was convinced, as were all leading physicists at the time, that the classical view was the correct one.

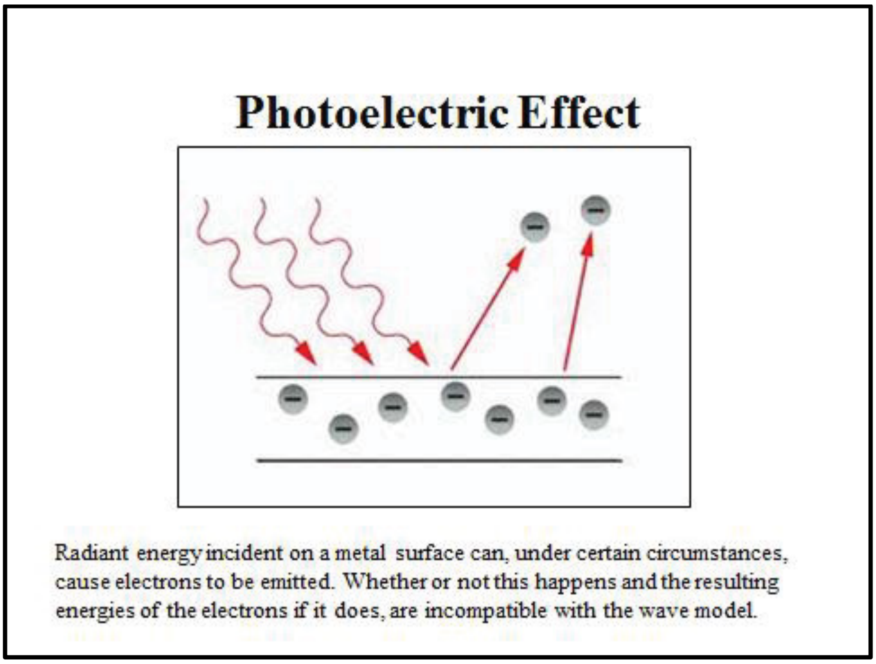


Figure 6.2 Photoelectric Effect

The photoelectric effect refers to the ability of electromagnetic radiation to free electrons from the surface of a metal. It was discovered in 1887 by Henrich Hertz during his experiments to produce and detect Maxwell's prediction of long wavelength electromagnetic radiation. Using Einstein's quanta model, the photoelectric effect occurs in the following way:

The electrons are bound to the surface of the metal by the force of attraction between their negative charge and the resulting positive charge of the surface losing the electrons. Therefore it takes a certain amount of work or energy to remove an electron, depending on the specific properties of the metal. This amount of energy is called the work function of the metal. If the quanta have sufficient energy, that is if the frequency of the radiation is sufficiently large, the radiation can free electrons. The frequency that will just do this is called the threshold frequency. The energy of the quanta above the work function goes to the kinetic energy of the freed electron.

Thus the model predicts that: 1) Below the threshold frequency, no electrons will be emitted even if the total energy of the irradiating beam is large. 2) Above the threshold frequency, electrons will be emitted. Their kinetic energy will be determined by the amount the frequency of the radiation is above the threshold frequency, that is, the amount of energy the quanta have above the work function. The number of emitted electrons will depend on the total number of quanta in the beam. These predictions are quite distinct from those of the classical wave model.

Millikan's results confirmed Einstein's predictions in every detail, but still Millikan was not able to accept Einstein's interpretation, and in 1916 wrote, "Einstein's photoelectric equation cannot in my judgment be looked upon at present as resting upon any sort of a satisfactory theoretical foundation," even though "it actually represents very accurately the behavior" of the photoelectric effect. Millikan was not alone in his skepticism of Einstein's theory. In 1913 in a letter to the Prussian Academy of Science recommending the establishment of a special position for Einstein, Max Planck wrote,

> There is hardly one among the great problems, in which modern physics is so rich, to which Einstein has not made an important contribution. That he may sometimes have missed the target in his speculations, as, for example, in his hypothesis of light quanta, cannot really be held too much against him, for it is not possible to introduce fundamentally new ideas without occasionally taking a risk.

As late as 1922, the great Danish physicist, Niels Bohr, in his 1922 Nobel address, stated, "The hypothesis of light-quanta is not able to throw light on the nature of radiation."

Einstein's quanta are now called photons, literally particles of light. The problem everyone was having with the concept of photons is that this particle model and the wave model are mutually exclusive. In commonsense logic, something cannot be both a wave and a particle, and the evidence for the wave behavior of electromagnetic radiation simply could not be dismissed. However, eventually the evidence for photons could not be dismissed either. In addition to the photoelectric effect, there was Compton scattering. In 1923, the American physicist Arthur Compton, published results of experiments on the scattering of X-rays by electrons. The wave model was completely unable to explain these results. Only by picturing the scattering of the radiation as one particle-like object colliding with another particle-like object could the data be accounted for.

The resolution of this dilemma lies partly in the realization that our descriptive abilities are limited. In classical physics, models are based on our experience with the everyday, macroscopic world. In this case something is either a wave or it is a particle. There are no ambiguities. However, on the microscopic scale this distinction seems no longer to exist, and we should properly say that electromagnetic radiation is wave-like in some situations and particle-like in others. Since there is no macroscopic analogy to this, we cannot visualize the true nature of electromagnetic radiation. Generally, when we think of electromagnetic radiation, we think in terms of a single model depending on which is more appropriate to the situation. In discussing interference we visualize electromagnetic radiation as a wave. In discussing the photoelectric effect or the Compton effect, we visualize it as a stream of particles. Although we often find it helpful to use models, we must be careful not to identify these pictures of reality with reality itself. In other words, don't confuse the map with the territory.

Wave-particle duality has actually been explicit in our treatment since introducing the quantum nature of electromagnetic radiation. The photon (particle description) has an energy that is proportional to the frequency (wave description) of the radiation. We will learn in a later chapter that wave-particle duality is a fundamental aspect of nature. It applies not just to electromagnetic radiation but to the entire micro-world. Electron, protons, atoms, etc., cannot be completely described in terms of particle properties. Rather, in some of their interactions, they exhibit wave properties. Einstein made a remark in a different context that is applicable here: "It is only with reluctance that one's desire for knowledge endures a dualism of this kind."

6.2 Brownian Motion and the Reality of Atoms

Although classical physicists generally accepted the reality of atoms, at the start of the 20th century there was no empirical evidence for their existence. Some influence physicists and chemists at the time saw the atom as simply a useful model rather than an actual real entity. In his second 1905 paper, Einstein provided a quantitative theory for the motion of suspended particles in a stationary fluid, a phenomenon described by Robert Brown in 1827 and known as Brownian motion. The theory required the existence of real atoms of definite, finite size. The experimental verification of Einstein's theory was made by the French physicists Jean Perrin. Since the publication of Perrin's results in 1908, no one has seriously doubted the atomic theory of matter. Perrin received the Nobel Prize in physics in 1926 for this work.

6.3 Special Relativity and Space-time

In his third 1905 paper, Einstein revolutionized our understanding of the nature of space and time and at the same time reconciled Maxwell's theory of electromagnetism with the laws of mechanics.

The classical laws of mechanics obey the principle of relativity. That is, they have the same mathematical form in all frames of reference that are in uniform motion with respect to one another. In physics, a frame of reference is simply a coordinate system used for the mathematical description of physical phenomena. The coordinate system is fixed to the state of motion of the observer of the phenomenon. How an object behaves when tossed in the air in a room on the surface of the earth (the earth frame of reference) is exactly the way it behaves on an airplane traveling with a constant velocity (the airplane frame of reference). The laws of physics are the same in both frames of reference. This principle of relativity for the laws of mechanics has been known since the time of Galileo. (Rather than an airplane, Galileo's example was a ship sailing on a perfectly smooth sea with a constant speed.)

A frame of reference in which the laws of mechanics have their simplest physical and mathematical form is called an inertial frame of reference. All other frames traveling at a constant velocity (a constant speed in a straight line) with respect to an inertial frame of reference are also inertial frames of reference. However, Maxwell's equations did not appear to satisfy the principle of relativity. They have their simplest physical and mathematical form only in a single inertial frame of reference and have more complex forms in all other inertial frames. At the time Maxwell proposed his theory, it was assumed that it was the inertial frame at rest with respect to the ether, the hypothetical medium through which electromagnetic radiation is propagated, that constituted this one special inertial frame.

For example, Maxwell's theory leads inescapably to the prediction that the speed of light in a vacuum is a constant, that is, it has a single unique value. This value is represented by the letter 'c' and its approximate value is 3×10^8 m/s or 186,000 miles per second.[1]

Footnote: In 1972 the speed of light in a vacuum was determined to be 299,792,456.2 ± 1.1 m/s. 3×10^8 m/s is within 0.067% of the correct value and is usually used. Light travels slightly slower in air than it does in a vacuum (about 0.37% under normal conditions) but 3×10^8 is also a very good approximation for the speed of light in air.

Philosophy played an important role in physics for Einstein. The principle of relativity had tremendous philosophical appeal to him "because it is so natural and simple." At the same time he had complete faith in Maxwell's theory of electromagnetism as it applied to the constancy of the speed of light in a vacuum. In his third paper he wrote, "As a result of an analysis of the physical concepts of time and space, it became evident (evident to Einstein maybe) that in reality there is not the least incompatibility between the principle of relativity and the law of propagation of light, and that by systematically holding fast to both these laws, a logical rigid theory could be arrived at." This is exactly what Einstein did in the third paper.

He developed his theory, later called the Special Theory of Relativity, based solely on two postulates.

Postulate I: The laws of physics are the same, that is, have the same mathematical form, in all inertial frames of reference.

Postulate II: The speed of light in a vacuum is an absolute constant.

He accepted these postulates as true and by "systematically holding fast" to them, logically deduced the physical consequences. When this is done, it is found that a host of very peculiar predictions result. For instance, this "logical rigid theory" of Einstein's will demand radical changes in our common sense, classical notions of time and space. Some of these very peculiar predictions will be treated in more detail in the following chapter.

The second postulate automatically provides an explanation for the negative results of the Michelson-Morley experiment where they failed to detect the predicted difference between the values of the speed of light in two frames of reference in motion with respect to one another. The speed of light is an absolute constant, the same in all inertial frames of reference. Although relativity explains the Michelson-Morley results, it seems fairly certain that Einstein's thinking was neither motivated nor influenced by the Michelson-Morley experiments. It is entirely possible that Einstein was even unaware of them.

In addition to explaining the Michelson-Morley results, relativity eliminated the troublesome concept of the ether. In the classical wave model some material substance must oscillate in order to produce a wave, much as air molecules oscillate to produce a sound wave and water molecules oscillate to produce a water wave. The ether was the proposed substance that oscillated to produce an electromagnetic wave.

In the classical theory, the speed with which a wave propagates is determined by the rigidity of the oscillating material. Thus sound travels much faster in a metal than it does in air. The extremely high speed of light would require the ether to be extremely rigid. However, in order for us to receive light from distant stars, the ether would have to fill all of the space of the universe, requiring the earth to move through the ether in its orbit around the sun. Because the earth does not lose energy as it does so, it would have to travel through this extremely rigid material experiencing no friction at all. The properties required of the ether were incompatible and physics was well rid of this hypothetical substance.

6.4 The Equivalence of Matter and Energy

In his fourth and final 1905 paper in the Annalen der Physik, Einstein developed an argument for mass as a form of energy. Prior to this paper the conservation of mass and the conservation of energy were considered two distinct and independent laws. That is energy could not be created or destroyed, and neither could mass. Conservation of mass was consistent with all observations, and, in the minds of physicists, mass was in no way related to energy. However, "systematically holding fast" to the postulates of the special theory of relativity, and its implications for time and space, required that mass could be created or destroyed as long as an equivalent amount of some other form of energy was destroyed or created.

Although at the time there was no evidence whatsoever for Einstein's ideas of time, space and the equivalence of mass and energy, Einstein knew, through physical intuition, that they must be true. When told shortly after the publication of his paper of an experiment that contradicted his theory, he said the theory is correct. Redo the experiment. And he was right.

This paper produced what is arguably the most famous equation in the field of physics: $E = mc^2$ where c is the speed of light. The speed of light has nothing to do with how mass is transformed to other forms of energy, c^2 is simply a proportionally constant that converts the traditional units for mass to those for energy.

In chemical reactions the amount of energy released or absorbed is not sufficient to produce a measurable change in the mass of the system, which is why the equivalence had not been noticed before. However, with the advent of nuclear physics in the early 20th century, the much greater energies involved relative to the mass of the particles, made the equivalence clear.

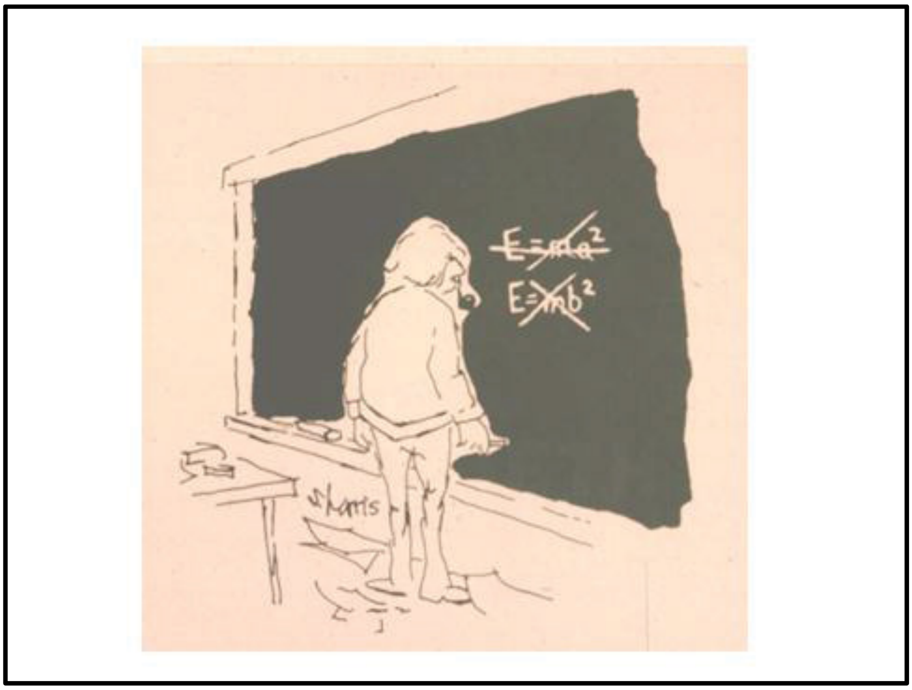


Figure 6.3 Einstein at blackboard

6.5 Einstein's Fifth Paper

The four 1905 papers described above were Einstein's most significant contributions in the miracle year. But there was also a fifth paper written that year, a paper entitled *A New Determination of Molecular Dimensions*. He submitted it to the University of Zurich as a doctoral thesis. Although the title of the dissertation focuses on the sizes of molecules, the technique Einstein describes also gives a measurement of the number of molecules (or atoms) present in a solution. This paper was selected for the dissertation primarily because it was the least revolutionary of the five and less likely to stir up opposition from a thesis committee. The only official comment Einstein received on the dissertation was that it was too short. In response, he added a single sentence, and it was accepted. Einstein was awarded his doctorate in 1906, meaning that his four revolutionary papers published the year before were published by someone with an undergraduate degree in physics.

7 Length Contraction and Time Dilation

Your goals for this chapter are to know the following:

- Which physical quantities are relative and which are absolute in classical physics.
- Which physical quantities are relative and which are absolute in Einstein's theory of relativity.
- How to solve problems involving time dilation and length contraction.
- What is meant by the twin paradox in relativity.

The laws of the Special Theory of Relativity derive from the logical consequences of Einstein's two postulates, first that the principle of relativity holds for all physical laws, both the laws of mechanics and the laws of electromagnetism, and that the speed of light is an absolute constant.

7.1 The Relativity of Speed

To say that something is an absolute constant means that its value is independent of the frame of reference in which it is measured. An example of an absolute constant is the charge on an electron. Speed, however, is obviously a relative, rather than an absolute, physical quantity. Consider a train moving at 60 mph with respect to an observer standing beside the tracks. Suppose a second observer is moving at 40 mph on a road parallel to the tracks in the same direction as the train. In the frame of reference of that observer, the train is moving forward with a speed of 20 mph. For an observer moving at 40 mph in the opposite direction, the train would move backwards through that frame of reference with a speed of 100 mph.

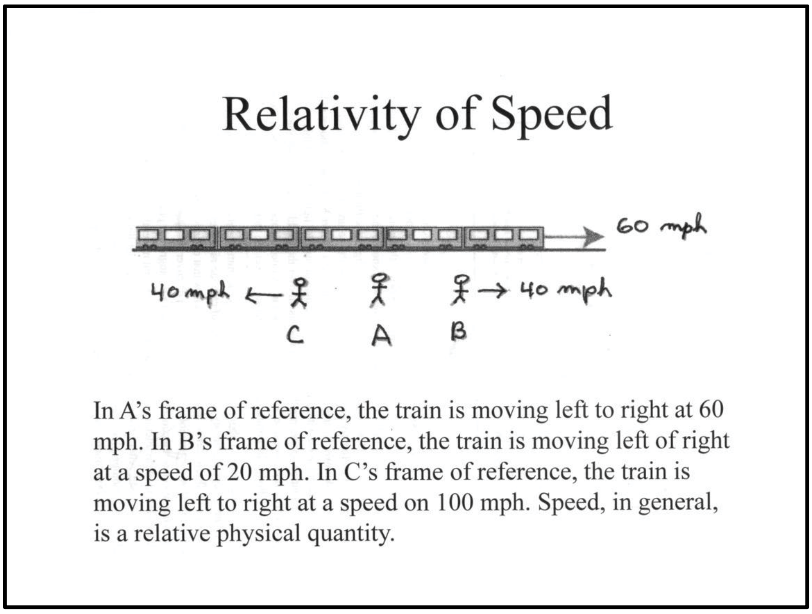


Figure 7.1 Relativity of Speed

The relative nature of speed in general is an obvious fact. However, Einstein is saying that there is something special about the speed of light; that it alone has a speed that is an absolute constant. In the diagram above, replace the train with a beam of light. Let observer B move at 80% of the speed of light, and observer C also move at 80% of the speed of light. According to Einstein, all three of the observers would measure the speed of the light beam moving through their frame of reference as the same absolute value, 3×10^8 meters per second. Only a complete revamping of our concepts of space and time could allow for this to be true.

7.2 The Relativity of Time

In classical physics space and time are absolute; space intervals and time intervals are completely independent of the frame of reference in which they are observed. However, a simple thought experiment can show that this is not true in relativity.

Imagine a train moving at close to the speed of light. An observer is positioned in the exact middle of the train. A second observer is at rest beside the tracks. At the exact instant the two observers pass one another, two flash bulb signals arrive simultaneously from the front and rear of the train.

Analyzing the situation from the frame of reference of the train, the observer on the train will conclude that the two flash bulbs went of simultaneously. Because each source of the light is equidistant from the train observer, and because the speed of light is an absolute constant, the two flashes of light must have traveled for the same amount of time. Thus, the two flash bulbs must have gone off at exactly the same time. That is they must have gone off simultaneously.

Now analyze the situation from the frame of reference of the observer standing beside the tracks. The two flash bulb signals arrive simultaneously from the front and rear of the train. Both the observer on the train and the observer beside the tracks agree on this. They are instantaneously face-to-face and at that instant each sees the two flashes. However, the speed of light is finite. Therefore it is clear that the light signals must have been emitted sometime in the past, sometime before the train reached its present position. The observer beside the tracks must have been closer to the front of the train and farther from the rear when the flashes were emitted. Thus, in the frame of reference of the observer beside the tracks, the flash from the rear had to travel a greater distance than the one from the front. Since the speed of light is an absolute constant, in order for the flashes to arrive simultaneously in the frame of reference of the observer beside the tracks, the flash from the rear must have been emitted before the flash from the front. In this frame, they did not go off simultaneously.

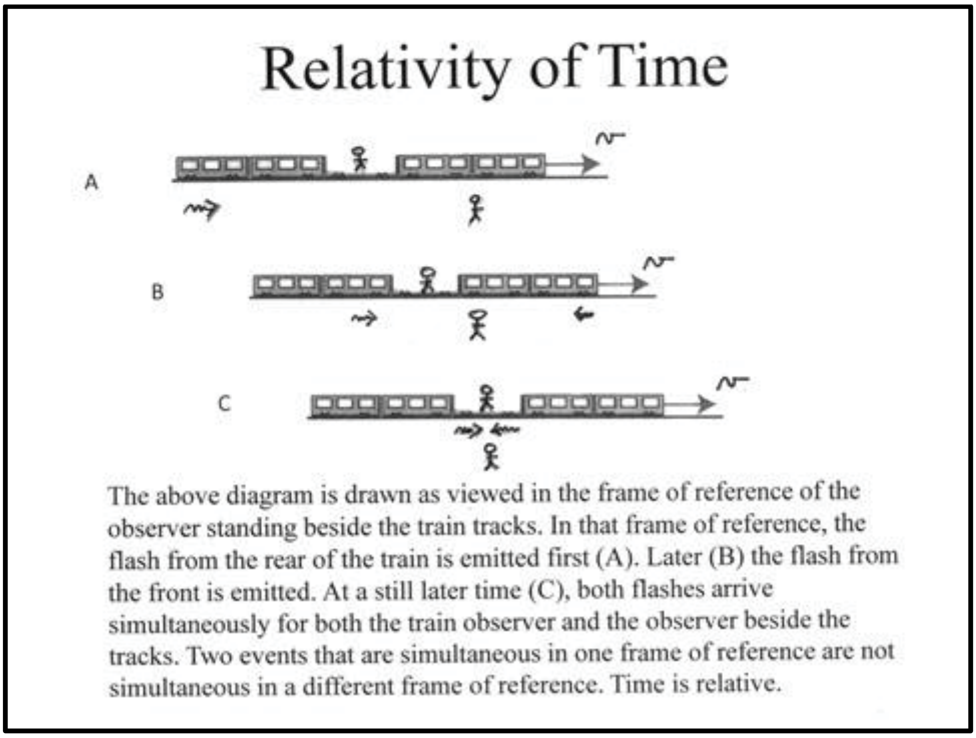


Figure 7.2 Relativity of Time

The train observer concludes that the light signals were emitted at exactly the same time. The observer beside the tracks concludes that the light signal from the rear of the train was emitted before the one from the front. Who is right? Both are right; time is a relative physical quantity. Events separated in space that are simultaneous for one observer are not simultaneous for a second observer in motion relative to the first one.

Not only is the time interval between events relative, the order of certain events can also be relative. To an observer in a frame of reference in which the train is moving right to left (i.e. an observer moving in the same direction as the train, but with a greater speed), the signal from the front of the train will have been emitted before the signal from the rear, the opposite of what the observer beside the tracks sees.

The relative nature of the order of certain events does not conflict with cause and effect. It two events are causally related, that is if one event produced the other, the order of the events is not relative. Otherwise a paradox could occur, such as the ball striking the wall before it is thrown. For all observers regardless of their relative motion, analysis using the theory of relativity will always yield the correct order for causally related events. However, the time interval between the events will be different. In fact, if two events that are not causally related but are separated in space by a distance small enough for light to travel between them, all observers will agree on the relative order of the events.

7.3 Time Dilation

As a further illustration of the relativity of time, consider the following situation. Observer A has a device for emitting and detecting light signals. It has a mirror located a certain distance above the device which will reflect the emitted light signal back to the device which will then detect it.

Suppose that when a light signal is received by the device another is immediately emitted. Suppose further that the distance is just right so that it takes the signal exactly one second for the round trip. That is, that a light signal is detected every second. (Actually this would require a device 93,000 miles high, but this is only a thought experiment.) The device is a clock – an instrument for measuring time, and the unit of time is one second, just as it is on ordinary clocks. The detection of the light signal is like the ticking of a clock.

An observer B has an identical clock. Imagine that observer B with his clock is moving left to right with respect to observer A and his clock. The situation, diagramed from the viewpoint of observer A, looks like the following.

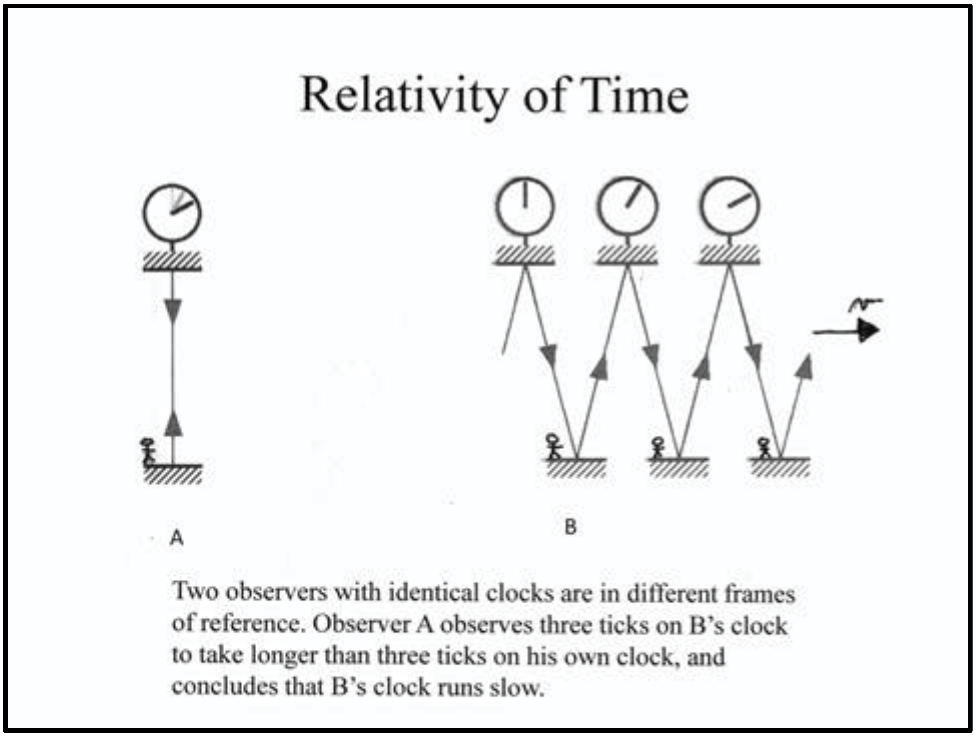


Figure 7.3 Clocks Relativity of Time

How does the time interval between the ticks of the two clocks compare? The second postulate of relativity tells us that A observers both light signals to travel at exactly the same speed. It is clear that the distance A observes B's light signal to travel between emission and detection is greater than for his own clock. Thus A will determine that the time interval between the emission and detection of the light signal by B's clock is greater than the time interval between these two events on his own clock. B's light signal must travel a greater distance at the same speed, and this requires more time. While A observes his clock to tick 60 times, or one minute on his clock, A will observe B's clock to tick less than 60 times, or less than a minute. Thus A concludes that B's clock runs slower than his own.

However, it should be clear that the situation just described from A's point of view, will have exactly the opposite reasoning when described from B's point of view. In B's frame of reference, it is A's light signal that will travel the greater distance and therefore take more time to complete one tick. Each observer claims that the other's clock runs slow. There appears to be a contradiction. However, stated differently there is no contradiction. Both thought experiments yield the same result: the clock in motion with respect to the observer will measure time more slowly than an identical clock at rest with respect to the observer. This phenomenon is known as time dilation.

This result was derived using light-signal clocks. Is it the same for all clocks? The answer is yes. In a frame of reference in motion with respect to an observer, all processes occurring in time will occur more slowly than in the observers frame. This includes biological as well as mechanical processes. Living beings age more slowly in moving frames of reference than will be the case in a frame of reference at rest with respect to the observer.

Going back to our thought experiment, it is clear that the added distance the light signal must travel depends on the relative speed of the moving frame of reference. The faster the clock is traveling with respect to the observer, the greater the distance the signal will have to travel between emission and detection. Thus time dilation must be a function of the relative speed.

The time-dilation equation is written

$$T = \frac{To}{\sqrt{1 - v^2/c^2}}$$

where T is the time interval indicated on a clock at rest with respect to the observer, T_o is the time interval indicated on a clock moving with a speed v relative to the observer, and c is the speed of light. Of course, if v = 0, the two times are the same.

In this equation, the relative speed must be on the order of the speed of light in order for the difference in the time intervals to be significant. This explains why this effect was not noticed empirically before 1905. Our ordinary experiences do not involve speeds anywhere near the speed of light and thus the consequence of time dilation produces no measurable effect. However, early in the 20th century, physicists began studying the motion of subatomic particles and time dilation was soon confirmed.

The speed at which the consequences of relativity become significant is usually taken to be about one-tenth the speed of light. The speed of light is 186,000 miles per second, so one-tenth would be 18,600 miles per second or 1,146,000 mph. This is a speed that will probably never be reached by ordinary macroscopic objects, but one easily obtained by submicroscopic particles such as electrons and protons. Even at this high speed, the different in the rate at which time passes in the two frames of reference is only about 2%.

7.4 Length Contraction

The Special Theory of Relativity united space and time, physical quantities that in classical physics were each absolute and independent. What is absolute in relativistic physics are space-time intervals rather than time intervals and space intervals separately. If space-time intervals are absolute and time intervals are relative, then space intervals must be relative as well.

For the space-time interval to be absolute, ΔT ΔL for the time and space intervals between two events must be a constant, that is it must have the same value for all observers. Let To be the time measured on a clock moving as a speed v with respect to the observer and Lo be the space interval (distance) measured for the moving frame. Let T and L be the time and distance intervals measured in the at rest frame. Thus, T L = To Lo. Substituting the time dilation equation, into this equation, we get

$$L = Lo\sqrt{1 - v^2/c^2}.$$

This result is known as length contraction. It is important to realize that the phenomenon of length contraction involves only the space interval in the direction of the relative motion. Thus, the shape of an object in motion relative to the observer will be altered. For example, if a square is moving relative to an observer, it will be a rectangle in the observer's frame of reference. The dimension in the direction of the motion will be contracted, while the perpendicular dimension will be unaffected. To say that the moving square will be a rectangle in the observer's frame of reference is not the same as saying the square will appear to be a rectangle. It will actually be a rectangle with all the properties of a rectangle in that frame of reference. The space and time intervals do not just appear to be different in different inertial frames, they must actually be different in order for the speed of light to be an absolute constant. This real difference has physical consequences that have been verified time and time again by experiments and observations.

Time dilation and length contraction can be summarizes as:

1) Every clock goes at its fastest rate when it is at rest relative to the observer. If it moves relative to the observer with a speed v, its rate is slowed by a factor of $\sqrt{1-\ v^2/c^2}$.
2) Every object is largest when it is at rest relative to the observer. If it moves relative to the observer with a speed v, it is contracted in the direction of motion by a factor of $\sqrt{1-\ v^2/c^2}$.

As mentioned earlier, it is subatomic phenomena that are likely to show significant relativistic effects. One such phenomenon is the production of muons in the upper atmosphere of the earth. Muons are essentially overweight electrons. They are unstable, and after a very short time, will decay to electrons. Muons have been studied extensively in the laboratory and muons created with speeds small compared to the speed of light have a lifetime of about 2.2×10^{-6} seconds, that is, after about 2.2 microseconds, they will spontaneously decay.

In addition to being created in the laboratory, they are also created when high speed cosmic rays strike the upper atmosphere of the earth. This occurs about 2 miles or 3000 meters above the earth's surface. Because of the high energies involved when cosmic rays strike the molecules of the upper atmosphere, the muons are created with extremely high speeds, speeds approaching the speed of light. Most of the muons created in this way pass through the atmosphere and bombard the surface of the earth.

This would be inexplicable to a pre-1905 physicist. Even if they could travel at the speed of light, 3×10^8 m/sec, with a lifetime of 2.2×10^{-6} seconds, they could only travel a distance of 660 meters before decaying, well short of the 3000 m needed to reach the surface of the earth.

Because the muons are created with speeds approaching the speed of light, a relativistic analysis is required. In the earth frame of reference, the muon represents a moving clock and thus its clocks run slower and time dilation needs to be taken into account. If the muon is traveling at a speed of 98% of the speed of light, (typical of the speeds involved), the factor $\sqrt{1-\ v^2/c^2}$ will have the value

$$\sqrt{1-(0.98c)^2/c^2} =.\sqrt{1-\ 0.96} = \sqrt{0.04} = 0.20.$$

Thus, time on the muon clock will pass at only 20% of the rate of earth clocks. This means that the muon will not decay until earth clocks register 5 times the laboratory life time of the muon,

$5 \times 2.2 \times 10^{-6}$ seconds or 11×10^{-6} seconds.

Traveling at 98% of the speed of light, or 2.94×10^8 m/sec, the average distance traveled by the muon in 11×10^{-6} seconds is 3230 meters. This analysis explains why most muons created in the upper atmosphere can make it to the earth's surface, 3000 meters below, before decaying, which, in fact, they are observed to do.

In the previous example, we analyzed the problem from the point of view of an earth observer. In that frame of reference the muon is a moving clock and time dilation accounts for the fact that muons are able to strike the surface. Now let's analyze the situation from the point of view of the muon. In the muon's frame of reference it is not a moving clock and therefore time dilation does not apply. The lifetime of the muon is not 11×10^{-6} seconds but 2.2×10^{-6} seconds. However, space intervals are relative, and the 3000 meters the muon must travel to reach the earth will be contracted in the muon's frame of reference.

In the frame of reference of the muon, the muon is at rest while the upper atmosphere is receding at 98% of the speed of light and the surface of the earth is approaching at 98% of the speed of light. The atmosphere is the moving frame of reference and the space interval or distance from the top of the atmosphere to the earth's surface is contracted. Using our value of $\sqrt{1 - v^2/c^2} = 0.20$ from the previous calculation, and Lo, the distance in the moving frame of reference (3000 meters) gives L, the separation between the upper atmosphere and the surface in the muon's frame of reference, as 600 meters. In the muon's frame of reference, its 2.2×10^{-6} second lifetime is enough time to travel the 600 meters of atmosphere, resulting in a collision with the surface.

7.5 The Twin Paradox

Because time passes at different rates in different frames of reference the theoretical possibility exists for twins in different frames of reference being reunited with one twin older than the other. Suppose that one twin stays home and the other is in a rocket ship traveling at, say 98% of the speed of light. The time dilation factor is 0.20 as in the earlier calculations. From the point of view of the stay-at-home twin, the traveling twin will only age 6 years in the same time interval he ages 30 years. If the traveling twin turns around and immediately returns home, again at 98% of the speed of light, the return trip will take 30 years on the stay-at-home twin's clock, while again only 6 years pass on the rocket ship clocks. The twins will be reunited with the stay-at-home 60 years older than when the trip started while the traveling twin will be only 12 years older.

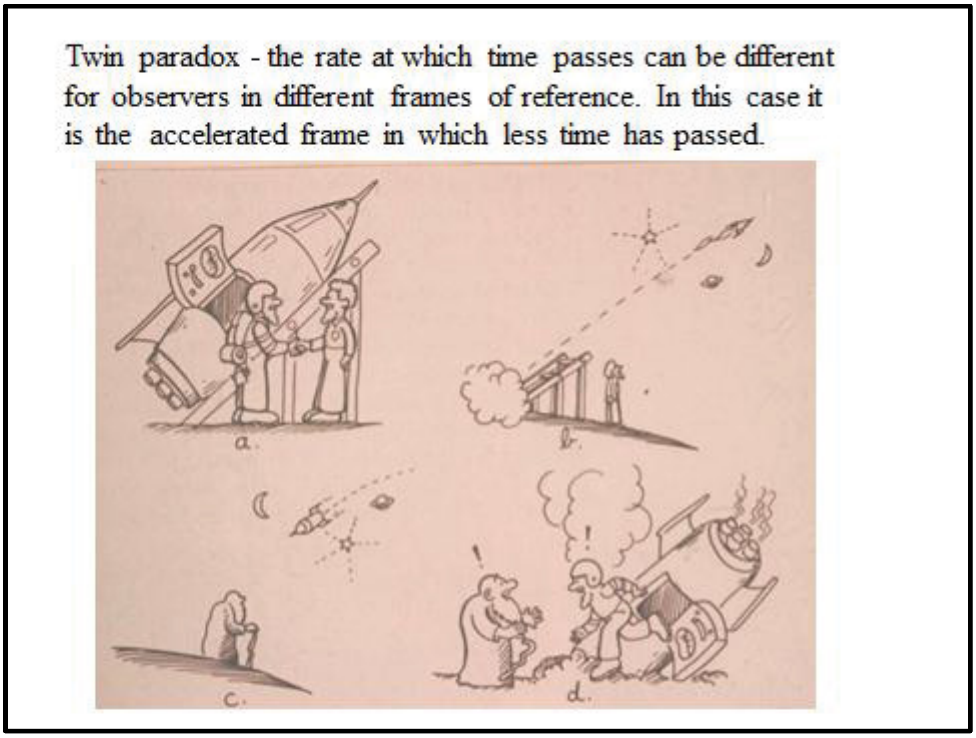
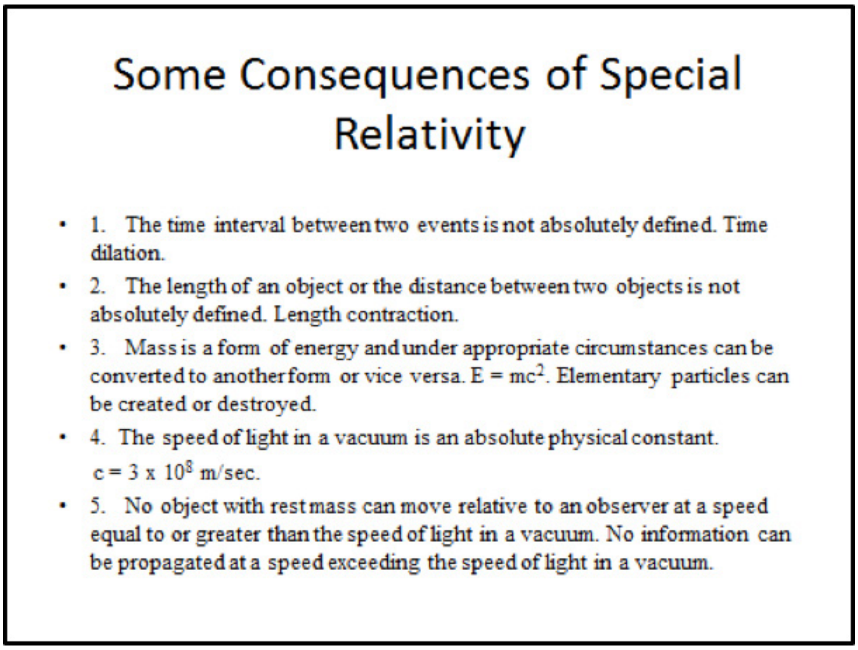


Figure 7.4 Twin paradox

This scenario is known as the Twin Paradox, although it is not a paradox at all. There is nothing paradoxical about the above calculation. The above situation was described from the point of view of the stay-at-home twin. This is because of the two, his frame of reference is the inertial frame while that of the rocket ship is not. It must turn around in order for the two to be reunited. That is, it must decelerate to a stop, then accelerate back up to 98% of the speed of light. During the time it takes to do this, the rocket ship frame is not an inertial frame, and the laws of the special theory do not apply.

However, the later General Theory of Relativity, as discussed in the next chapter, can be used in a non-inertial frame of reference. For the first half of the trip, both frames are inertial (neglecting the initial acceleration). The traveling twin agrees that he has aged 6 years, but because the stay-at-home twin is the moving frame, the rocket observer see his twin age only 1.2 years (6 years times 0.20). On the return half of the trip, both are again inertial observers and the rocket travelers again ages 6 years and sees his twin age another 1.2 years. However, during the turn around, the equations of special relativity do not apply, and the equations of the general theory must be used. According to the general theory, during the turn around, the rocket observer will see the stay-at-home clocks suddenly speed up dramatically. According to precise calculations, the stay-at-home clock will speed forward and tick off 57.6 years during the turn-around time, even if it takes only a few minutes on the rocket ship clocks. Thus each twin will be able to explain why the stay-at-home twin has aged 60 years while the traveling twin has only aged 6 years.



Although the twin paradox is strictly a thought experiment, the equivalent experiment has been done many times with atomic clocks. Atomic clocks can keep time with extreme precision. Two identical atomic clocks are prepared and set to exactly the same time. One is put on an earth-orbiting satellite and the other remains at rest in the laboratory. When they are reunited, the clocks show different times in exact accord with Einstein's theories. The speeds involved are very small compared to the speed of light, but the precision of the clocks is such that effects on the order of tiny fractions of a second are observable.

8 The General Theory of Relativity

Your goals for this chapter are to know the following:

- What is meant by the principle of equivalence.
- The experiment Einstein suggested as a test of the general theory and what this suggests about the nature of space-time.
- Why the general theory is important to our current understanding of the universe.

The Special Theory of Relativity discussed in the two previous chapters is restricted in its application to inertial frames of reference. The first postulate says that the laws of physics have the same mathematical form in all inertial frames of reference, an inertial frame being one in which the laws have their simplest mathematical form. Once an inertial frame is found, all other frames moving in a straight line with a constant speed are also inertial frames. This creates a special class of observers and excludes from the theory any frame of reference accelerating (changing either its speed or direction of motion) with respect to an inertial observer. A frame of reference in which there is a gravitational field is also a non-inertial frame and is excluded from the special theory. The laws of physics have a more complicated mathematical form when gravity is present. Einstein's physical intuition told him that the principle of relativity could be extended to include non-inertial frames of reference and that by doing this gravity would emerge as a natural property of space-time.

Einstein set out almost immediately to develop this more general theory that would include non-inertial as well as inertial frames. The task was formidable, requiring a degree of mathematical sophistication previously unknown in physics. By late 1915 the task was complete.

It is significant to note that Einstein was led to the more general theory not by the need to explain some experimental data, but rather the 'need' to conform to certain philosophical principles. In his own words,

> Our experience hitherto justifies us in believing that nature is the realization of the simplest conceivable mathematical ideas. I am convinced that we can discover, by means of purely mathematical constructions, those concepts and those lawful connections between them which furnish the key to the understanding of natural phenomena. Experience may suggest the appropriate mathematical concepts, but they certainly cannot be deduced from it. Experience remains, of course, the sole criterion of the physical utility of a mathematical construction. But the creative principle resides in mathematics. In a certain sense, therefore, I hold it true that pure thought can grasp reality, as the ancients dreamed.

This approach may be criticized as being unscientific (and it is in the case of most ordinary people), but in the case of Einstein it produced a physical theory of incredible beauty and power.

8.1 The Principle of Equivalence

The concept of mass is familiar. In most people's mind it is associated with weight. More massive objects weigh more. This property of mass is called gravitational mass because it determines the force that gravity will exert on the object. There is another, seemingly unrelated, property of mass. The more massive an object is, the harder is it to change either the direction or the speed of its motion: that is, the more difficult it is to accelerate the object. This property of mass is called inertial mass. In classical physics, inertial effects are unrelated to gravitational effects. In developing his theories of motion and gravity, Newton assumed that they were the same property, although there was no theoretical justification for this assumption. It just seemed to work. By the end of the 19th century it had been empirically shown that, to a high degree of accuracy, inertial and gravitational mass can be considered numerically equal, although there was still no theoretical explanation for this.

Having two completely independent properties that are accidentally equal to each other is, philosophically, a very unsatisfactory situation. Einstein reasoned that there must be some underlying physical significance to this equality which would constitute a single interpretation for mass.

As with the special theory, Einstein utilized thought experiments to guide his thinking. Imagine a person in a small, widowless compartment with two objects of different mass. Suppose this compartment is located on the surface of the earth. If the two objects are dropped simultaneously from the same height, they will hit the floor at the same time. Galileo was the first to show this and to measure the acceleration due to gravity as 9.8 m/sec^2. That is, near the surface of the earth a free-falling object will change its speed at a rate of 9.8 m/sec each second.

Suppose now that the compartment is in space far removed from all massive objects. There will be no gravitational forces acting on the compartment and its contents. If there are no forces acting on the compartment, the person and the two objects will float around in the compartment much as you see the astronauts doing when in orbit around the earth. This is often referred to as weightlessness, but that is not strictly correct. The astronauts are not weightless; the earth is still pulling them and the space craft downward. That weight is what is keeping them in orbit rather than from flying off into space. However, with the rockets turned off, the space craft is free falling around the earth and this is, as we will see, equivalent to being weightless.

Now imagine that there are rockets attached to the bottom of the compartment causing it to accelerate, just as the astronauts are accelerated on takeoff. The inertial mass of the person, with its tendency to maintain its state of motion, will resist this change in velocity. The person will be pressed against the floor of the compartment just as the astronauts are pressed against their seats on takeoff. With effort, the person will be able to stand. If the two objects are dropped as before, their inertial mass will tend to keep them in a constant state of motion and the floor of the compartment will accelerate up and overtake them at an ever increasing speed.

In the frame of reference of the observer, the objects will 'fall' to the floor. In fact, if the acceleration of the compartment is 9.8 m/sec per second, they will 'fall' to the floor exactly as they would if the compartment was at rest on the surface of the earth. As Einstein thought about this, he realized that there is no possible measurement the person in the compartment could make that would distinguish between the two situations. In 1911, based only on his thought experiments, Einstein made the following bold statement. "It is impossible to distinguish, by any experiment whatsoever, between the effects of acceleration and the effects of gravity." This statement is known as the principle of equivalence.

Footnote. The principle of equivalence only applies in a small region of space. Otherwise non-uniformities in the gravitational field can be distinguished from the uniform effects produced by acceleration. The principle of equivalence should not be confused with the equivalence of mass and energy, which is an entirely different matter.

The thought experiment above is a mechanical one. What about experiments involving electromagnetic phenomena? Again, as was the case in 1905, Einstein turned to the propagation of light. Suppose that there is a source of light in an inertial frame of reference and that a beam of light from the source enters the compartment horizontally from the side. If the compartment is at rest with respect to the light source, it is also an inertial frame. The light beam will enter the compartment, travel horizontally across the compartment, and strike the wall the same distance above the floor as the place it entered. By considering this situation in the frame of reference of the accelerating compartment in outer space (a non-inertial frame of reference) Einstein was able to make an important prediction.

In an inertial frame the speed of light is an absolute constant. Thus the beam of light in its inertial frame travels at a constant speed as it passes through the compartment. However, the vertical speed of the compartment is constantly increasing. Thus, in fixed time intervals the beam of light will always travel the same horizontal distance while the compartment will travel increasingly greater vertical distances.

In this case, the path of the beam of light is not a straight line in the frame of reference of the compartment. It will be bent toward the floor. If an accelerating frame of reference is in no way distinguishable from a frame of reference in a gravitational field, this thought experiments predicts that a beam of light will not be a straight line in a gravitational field. Einstein concluded that the path of light is altered by gravity such that a light beam passing near a massive object ought to travel as if attracted toward the object.

Calculations showed that the amount of bending produced even by an object as massive as the sun is extremely small. However, in a 1911 paper, Einstein proposed an experiment that should in principle allow a measurement to be made. For reasons discussed later, this experiment was, fortunately, not carried out until 1919.

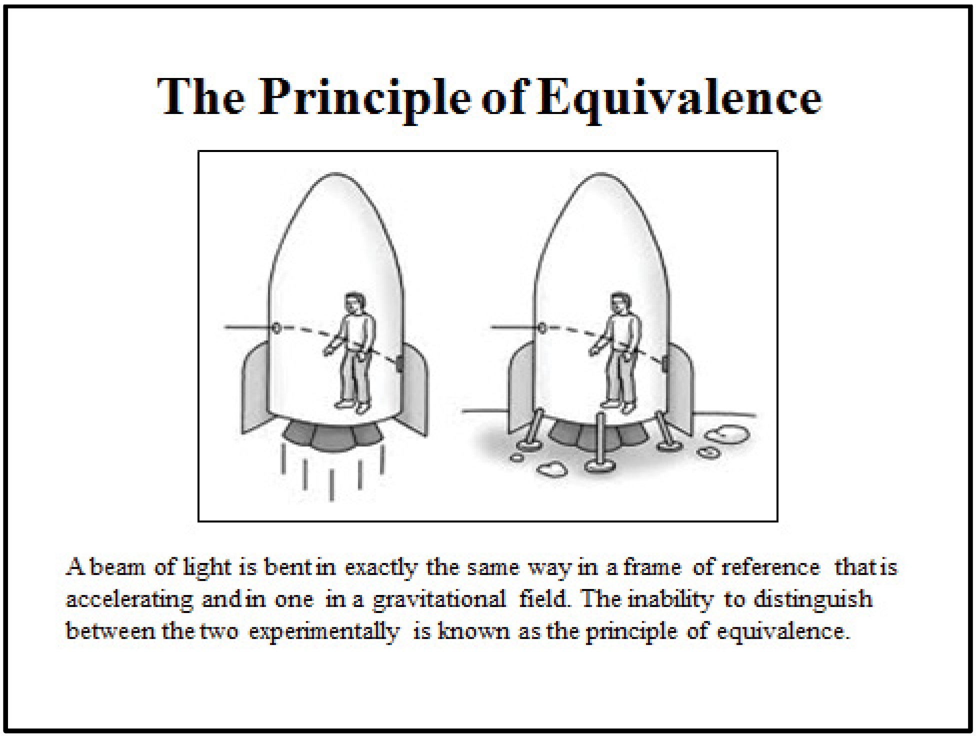


Figure 8.1 The Principle of Equivalence

8.2 The General Theory of Relativity

Einstein's development of the General Theory of Relativity was strikingly similar to his development of the Special Theory of Relativity. In the special theory, Einstein took the law of propagation of light and the principle of relativity (for inertial frames of reference), which appeared to be incompatible and showed that they were both true. In the case of the general theory, he took the principle of equivalence and the general principle of relativity (that the laws of physics are the same in all frames of reference, both inertial and non-inertial) which appeared to be incompatible and showed that they were both true.

In the previous section it was shown that the principle of equivalence led to the fact that while light in an inertial frame of reference travels in a straight line, light in an accelerating frame (or in a gravitational field) travels in a curved path. It appears that the laws governing the motion of light in an inertial and in a non-inertial frame are different, violating the general principle of relativity. That Einstein was able to resolve this paradox with the General Theory of Relativity, in spite of unimaginable mathematical and conceptual difficulties, stands as a monument to the human intellect. As was the case with the special theory, the solution lies in our concept of space-time.

The main problem Einstein was having with developing his new theory was mathematical. Although by ordinary standards, Einstein might be considered a mathematical genius, by the standards of theoretical physicists, Einstein was not exceptional. It was his unprecedented physical intuition (with the possible exception of Isaac Newton) rather that his mathematical ability that made him the greatest physicist of his time. It was only in the process of work on his theory that Einstein gradually acquired the mathematical techniques with which to express the theory. In late 1912, he wrote to a friend,

> I occupy myself exclusively with the problem of gravitation and now believe that I will overcome all difficulties with the help of a friendly mathematician here. But one thing is certain: that in all my life I have never before labored at all as hard, and that I have become imbued with a great respect for mathematics, the subtle parts of which, in my innocence, I had till now regarded as pure luxury. Compared with this problem, the original theory of relativity is child's play.

The 'here' was his alma mater, Zurich Polytechnic Institute, where he had just returned as professor of physics, and the 'friendly mathematician' was Marcel Gossmann, an old school friend. The necessary mathematical technique was tensor calculus, Grossmann's specialty.

In addition to the principle of equivalence and the general principle of relativity, Einstein put an additional constraint on the theory, one dictated by aesthetic values which he held to be of paramount importance in physics. Out of the literally thousands of formalisms provided by the tensor calculus consistent with the principle of equivalence and the general principle of relativity, Einstein insisted that only the mathematically simplest formalism would provide a correct description of nature.

In 1914, Einstein left Zurich for Berlin. There, late in 1915, after years of almost constant effort, Einstein arrived at a formalism that seemed to satisfy all his requirements. As a test, Einstein used the theory to calculate the orbit of Mercury. The classical Newtonian theory of gravity was very, very slightly, though undeniably, inconsistent with the observed orbit. The difference was an incredibly small 43 seconds of arc per century in the precession rate of the orbit.

The observed and the calculated orbit matched perfectly; the additional 43 seconds of arc per century came naturally and of necessity from the theory. It was immediately clear that Einstein's theory is a more accurate description of gravity than Newton's, a theory that for centuries had been assumed to be absolutely correct.

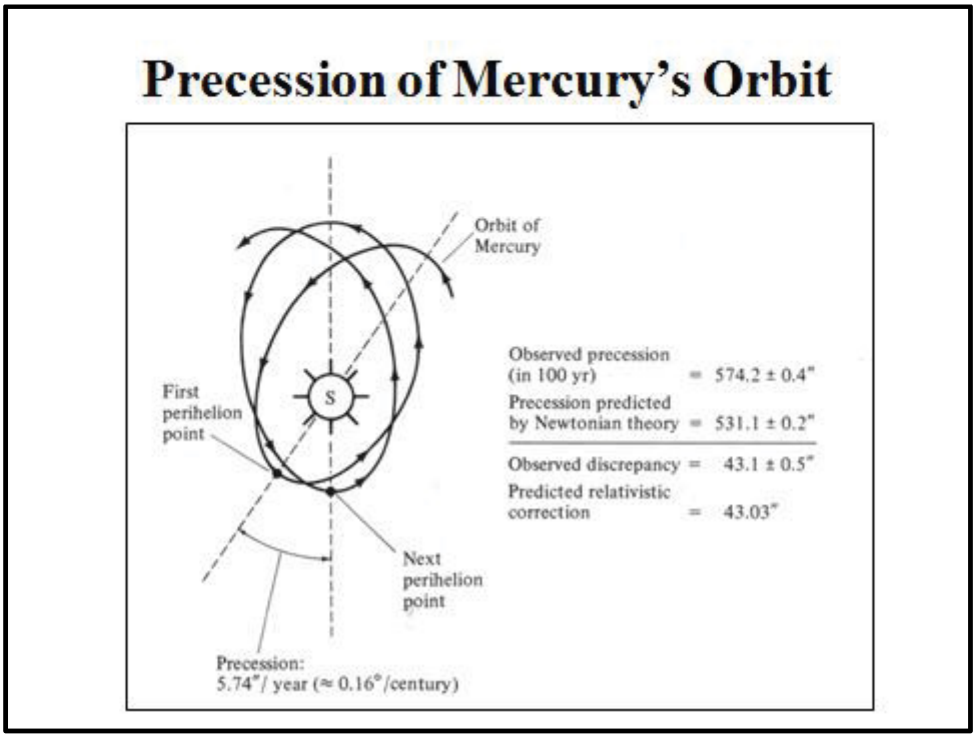


Figure 8.2 Precession of Mercury's Orbit

The General Theory of Relativity is, as Einstein had intended from the beginning, a theory of gravity. However, it did not simply produce a new, more general force law. It changed in a very fundamental way our concept of gravity. According to the theory, the effects of gravity are not the result of a force being exerted on an object, but rather are the result of the natural, inertial motion of the object through space-time, the properties of which are determined by the presence of other massive objects.

In the special theory, space and time are interwoven, and separate models of each cannot be constructed. In the general theory, space-time and matter lose their independent meaning. They are different aspects of a single unity; each is meaningless in the absence of the other.

In addition to providing a more accurate description of the effects of gravity, the general theory also eliminated one of the most disturbing aspects of the classical theory – action-at-a-distance. The sun does not pull on the earth, rather it changes the properties of the space-time in its vicinity. In effect, it curves space-time in a way that inertial motion through it is no longer a straight line.

This is much like an object moving on a curved two-dimensional surface. Imagine a horizontal, flat rubber sheet. A bowling ball placed on it will depress the sheet creating a curved surface surrounding the ball. If a marble is rolled across the rubber surface, it will not travel in a straight line. It will be deflected by the curved surface with the amount of the deflection being greatest where the curvature is greatest. In fact, the speed of the marble can be such that the marble will orbit the bowling ball, much as the earth orbits the sun. The bowling ball is not exerting a force on the marble to produce the orbit. Rather it is the curvature of the surface that is producing the orbit.

The amount of the curvature of the rubber sheet depends on the distance from the ball. In this analogy, the amount of the curvature of the rubber sheet models the strength of the gravitational field. Just as is the case with the strength of the gravitational field, the amount of curvature of the rubber sheet surrounding the bowling ball decreases with distance.

This analogy is not exact, as it explains the motion in terms of a curved two-dimensional surface whereas gravity is a consequence of curved four-dimensional space-time. However, it does provide a visualizable way to represent the nature of gravity in the General Theory of Relativity.

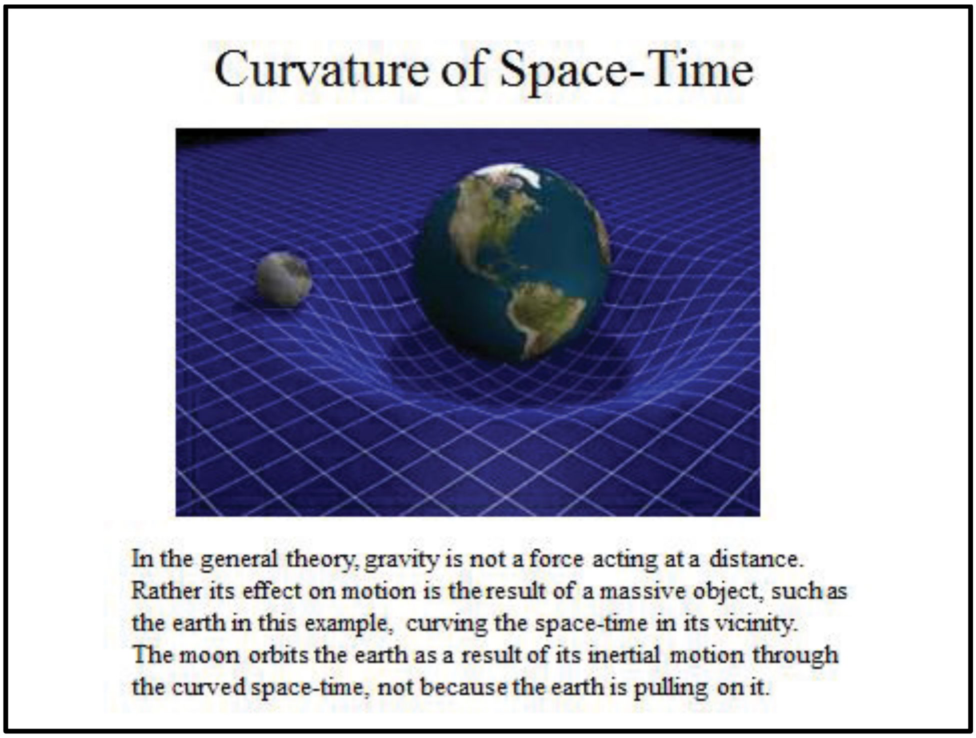


Figure 8.3 Curvature of Space-Time

8.3 Evidence for the General Theory of Relativity

The calculation of the correct motion for the planet Mercury was a tremendous success for the general theory. However, physical theories are judged primarily on predictions of new, unsuspected physical phenomenon that are subject to experimental or observational verification. The Mercury result, though completely unforced, was not a prediction but an explanation of a previously known fact.

The first crucial test of Einstein's theory was made during the total eclipse of the sun in May, 1919. In 1911 Einstein had proposed an experiment to measure the amount of bending of light as it passed near the sun. During a total eclipse of the sun, the sky is as dark as at night, and the stars can be seen. A photograph of the stars taken during a total eclipse can be compared to an earlier photograph of the same region of sky taken at night. Light rays passing near the eclipsed sun are bent. For these stars, their relative positions with respect to the other stars will be slightly different on the two photographs.

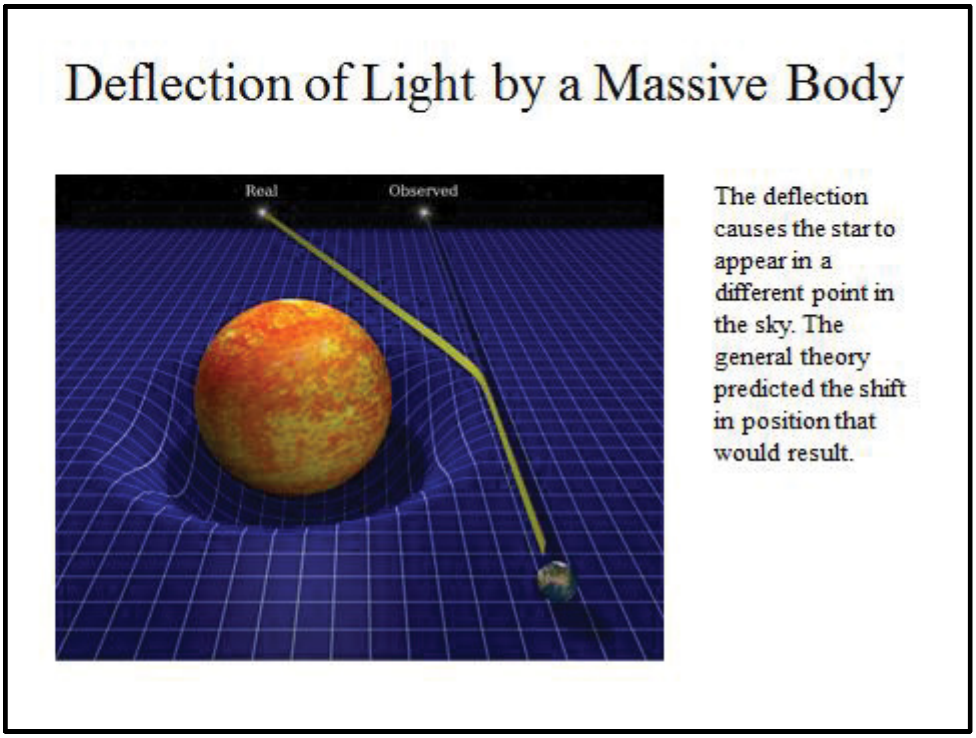


Figure 8.4 Deflection of Light

For the star in the figure, the photograph taken during the eclipse would show a different location in the sky compared to a photograph taken at some other time. The difference between the apparent and the true positions of the star is a direct measure of the amount of bending produced by the gravitational effect of the sun.

Einstein predicted the bending of light in a gravitational field in 1911, well before his theory was complete. He made this prediction based on the principle of equivalence and Newton's theory of gravity. The calculation indicated that a light beam passing very near the surface of the sun would be deflected through an angle of 0.87 seconds of arc or 0.00024 degrees. Although this angle is extremely small, there was a possibility that it could be measured. In 1914, the German astronomer Erwin Finlay-Freundlich, set off for Russia to observe the total eclipse of the sun and try to verify Einstein's prediction. However, he was prevented from doing so by the outbreak of the First World War. Einstein wrote the following in a letter to a friend:

> Europe, in her insanity, has started something unbelievable. In such times one realizes to what a sad species of animal one belongs. I quietly pursue my peaceful studies and contemplations and feel only pity and disgust. My dear astronomer Freundlich will become a prisoner of war in Russia instead of being able there to observe the eclipse of the sun. I am worried about him.

In fact, it was fortunate that Freundlich was not able to make his measurements. By the end of 1915, it was clear that his 1911 prediction was in error. In 1911 he had assumed that Newton's gravitational theory was appropriate for the calculation, but his general theory indicated that this was not the case. A new calculation based on the general theory produced a value of 1.75 seconds of arc, about twice the earlier prediction.

There was, of course, no chance to test the eclipse prediction until the war ended. However, as early as 1917, Sir Arthur Eddington, a British physicist, began preparing for two 1919 expeditions. Eddington was one of the first to appreciate the significance of the general theory, but his efforts to organize the test expeditions were motivated by more than just scientific curiosity. Eddington was a Quaker, and like Einstein, was profoundly disturbed by the war. He felt that if a British expedition verified the work of a German theoretical physicist, this would help heal the wounds of war, and, in particular, would reestablish scientific relations between the warring nations.

The results confirmed Einstein's prediction. 1919 was the last year of Einstein's private life. The announcement of the verification of his theory to a war-weary and heartsick people made him a world-wide hero. He became the personification of intelligence, an identification that has survived in spite of the fact that Einstein has been dead for more than a half-century. This fame, which he neither sought nor enjoyed, would later cause him to refer to his years of obscurity in the Bern Patent Office as the happiest of his life.

8.4 The General Theory of Relativity Today

In the decades following its publication, the number of observable effects that distinguished Einstein's General Theory of Relativity from the much simpler Newtonian theory was small, and the magnitude of the differences between the predictions of the two theories was almost negligible. For these and other reasons, interest in the general theory soon all but disappeared – almost, but not quite. During the twenties and thirties, a handful of theoretical physicists and cosmologists were applying the general theory and arriving at results which staggered the imagination, results so inconceivable and incomprehensible that they were generally ignored.

This attitude toward the General Theory of Relativity changed rather dramatically in the 1960s due primarily to developments in the field of astronomy. The discoveries of such exotic objects as quasars and pulsars suddenly made the bizarre predictions of Einstein's theory seem less unreasonable. Today the General Theory of Relativity is again at the forefront of physics. The two most important applications of the theory are the physics of gravitationally collapsed objects, black holes being the most extreme example, and cosmology, the science of the Universe as a whole.

A black hole is a region of space-time where gravity is so strong that anything, even light, that enters the region will be trapped there. The possibility of the existence of a black hole was recognized as a direct prediction of the general theory almost immediately after the formulation of the theory. However, it was not until 1939 that Robert Oppenheimer and one of his students, Hartland Snyder, suggested a possible mechanism whereby one might actually form.

When all possible fuels of a star are exhausted, it will begin to collapse gravitationally under its own weight. Using the general theory, Oppenheimer and Snyder were able to show that if the mass of the collapsing star exceeded a certain value, today believed to be about three times the mass of our sun, no known force could prevent complete collapse. The star will collapse to smaller and smaller volumes, collapse to the point where electrons, protons, and neutrons are crushed out of existence, crushed to an object of infinite density surrounded by a volume of space where gravity will prevent the emission of light. It is little wonder that at first physicists were reluctant to accept this. Arthur Eddington characterized the idea as absurd.

Today the existence of black holes is a virtual certainty. Black holes from 3 to 10 or so solar masses have been identified in orbit around ordinary stars. A supermassive black hole containing 2.2 million solar masses has been found in the center of our galaxy. Other supermassive black holes have been detected in the centers of other galaxies.

Modern cosmology, one of the greatest intellectual adventures ever undertaken, owes its existence to Einstein's General Theory of Relativity. In 1916, soon after the calculation of Mercury's orbit, Einstein applied his theory to the universe as a whole and got an unwelcome result. The theory indicated that the universe could not be static. The theory clearly predicted that the amount of space in the universe must be either increasing or decreasing. Einstein did not believe that this was true, and the other physicists and astronomers he consulted assured him that it was not. Everyone was sure that the universe was static and unchanging on the large scale.

For once in his life, Einstein lost faith in the basic simplicity of nature, and added an ad hoc term, which he called the cosmological constant, to his theory. Though it marred the beauty and simplicity of the theory, Einstein believed that it would eliminated the offending prediction of expanding or contracting space. The effect of the cosmological constant is to provide a universal repulsive force that Einstein thought could balance the attractive force of gravity and allow for a static universe.

In the 1920s two theoretical physicists independently showed that Einstein's 'fix' of his theory did not work. Not only did the original theory require that space be expanding or contracting, but the new version with the cosmological constant made exactly the same prediction. Hardly anyone at the time studied the general theory and no one paid much attention to these papers, including Einstein himself. One of the involved physicists, the Belgium priest, George Lemaitre, was so discouraged by the lack of interest in his paper that he switched his research to another field.

In 1929, the American astronomer, Edwin Hubble, published a law stating that the radiation received from distant galaxies is stretched out, that is, has longer wavelengths, than radiation received from local sources. Furthermore the amount of the stretching, called redshift by astronomers, is directly proportional to the distance to the galaxy. Hubble knew that this was a very significant result, but had no idea what was the cause of it.

When Arthur Eddington heard of this result, he immediately knew the explanation. Eddington was among the few physicists who actually understood the details of the general theory. He was also one of the few who were aware of the papers showing that, regardless of the version of the theory used, the general theory required either an expanding or a contracting universe. Either was consistent with the general theory. As to which described our actual universe, it was an empirical question. Eddington immediately recognized Hubble's Law as empirical evidence that space is expanding. When Einstein heard of this, he called his inclusion of the cosmological constant the "greatest blunder of my scientific career." We will later see that this may not be the case.

Eddington concluded space was expanding because expanding space would stretch out the wavelengths of light from distant galaxies just has Hubble had observed. Further, the light from the more distant galaxies would spend more time traveling through expanding space and thus be stretched more, making the amount of redshift directly proportional to the distance. Had space been contracting, the wavelengths would be compressed, or blueshifted rather than redshifted.

When George Lemaitre learned of these developments, he returned his interest to cosmology. He was the first to think scientifically about what the universe must have been like in the past if space is expanding. He concluded that if time could be run backwards, the universe would become increasingly dense. Using the general theory, it was clear that this increasing density would become infinite at some finite time in the past, that is; that the universe must have had a beginning in time.

The prevailing belief among scientists at the time was that the universe was infinitely old. Thus the prediction by a Catholic priest that the universe had a finite age, (he had probably believed this from childhood) was not taken very seriously. Lemaitre's theory had an additional strike against it. Almost nothing was known about nuclear physics at the time, so Lemaitre was unable to use his theory to make testable predictions regarding the present universe.

The idea was taken up again in the late 1940s and early 1950s by the Russian-American physicist, George Gamow. By that time considerable progress had been made in the area of nuclear physics and Gamow was able to make two significant, testable predictions. The first of these was that the very early universe could not have produced any appreciable amount of chemical elements more massive than helium. Thus, the early universe must have nearly pure hydrogen and helium. The second was that the entire universe would be filled with thermal radiation at a temperature of a few degrees above absolute zero; the cosmic background radiation. Gamow's theory soon became known as the Big Bang theory.

By the late 1950s, it had become clear that the elements heavier than helium are actually created in the interiors of massive stars, and that the explosive deaths of these stars distributed the heavier elements throughout the galaxy, making them available to later generations of stars, including our sun. The evidence was also accumulating that in the past, the galaxy was more nearly pure hydrogen and helium.

The clinching piece of evidence in favor of the Big Bang was the discovery in 1964 of the predicted cosmic background radiation. Arno Penzias and Robert Wilson, two radio astronomers working for Bell Laboratories, detected the radiation using a radio telescope built for trans-Atlantic telephone conversations. The pendulum of scientific opinion immediately swung from the Steady State theory, with an infinitely old universe, to the Big Bang and a finite age for the universe. The latest studies of this cosmic background radiation indicates that the universe began 13.7 billion years ago.

It is now universally accepted that space is expanding. It was also universally accepted that gravity, as understood by the General Theory of Relativity, required that the rate of expansion must be decreasing, In the mid-1990s, two independent research teams set out to measure the rate of slow down. By 1998, their results were in. Much to their (and everyone else's) amazement, both determined that the expansion rate was actually increasing. The only possible explanation for this is that, in addition to the attraction of gravity, the universe contains a global repulsive force, and, at the present time, the effect of this repulsive force is greater than that of gravity. The nature of this force is unknown and it is just referred to as dark energy, but Einstein's cosmological constant of 1916 seems to be in accord with what has been observed so far of the changing expansion rate of the universe. Perhaps its inclusion was not a 'blunder' after all.

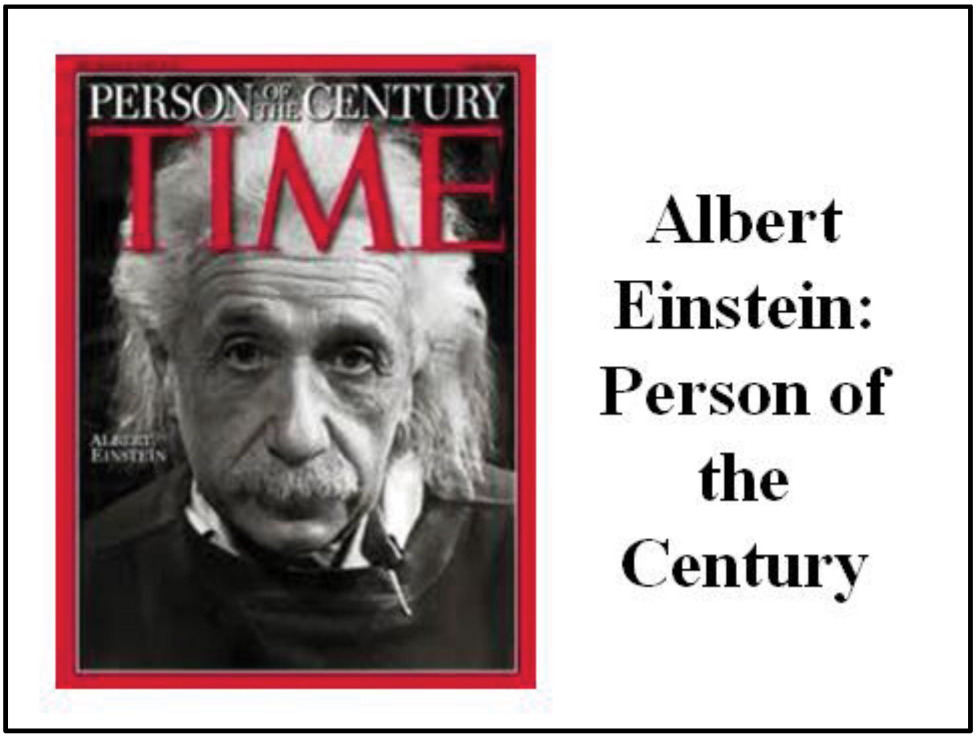


Figure 8.5 Person of the Century

9 The Bohr Model of the Atom

Your goals for this chapter are to know the following:

- What is meant by the nuclear model of the atom.
- What distinguishes the Bohr model of the atom from the classical models.
- How to calculate the allowed energies of the Bohr atom.
- How the Bohr model is used to explain atomic spectra.
- The major problem with the Bohr model.

The Greeks are credited with introducing the concept of the atom. Leucippus is thought to have first introduced the idea, and his student, Democritus, developed the concept in more detail. They viewed matter, which appears to the senses as continuous, as actually constructed by discrete building blocks too small to be apparent. Atoms were considered to have no internal structure; they were indivisible and indestructible. The word 'atom' comes from the Greek word meaning indivisible. Democritus taught, "Nothing exists except atoms and the void, all else is mere opinion."

The atomic theory never gained wide acceptance in Greek thought and was rejected by Aristotle. During the Middle Ages, atomic theory was considered subversive by the Church as it was seen to be incompatible with the transubstantiation of the host into the body of Christ. It was not until the 17th century that the atomic theory gained any semblance of respectability. Galileo, Newton, and most of their contemporaries were atomists, though more for philosophical reasons than scientific ones. There certainly was no empirical evidence at that time to even suggest the existence of atoms.

9.1 Classical Models of the Atom

Atomic theory made its first significant step toward scientific respectability in the first part of the 19th century, when John Dalton, an English chemist and schoolmaster, put forth a quantitative atomic theory. Dalton's ideas provided a theoretical basis for the science of chemistry and led to predictions concerning chemical reactions that were later verified. Because of the usefulness of the concept in forming a rational picture of many natural phenomena, it became easier to believe that atoms exist than to deny it.

Belief in the reality of atoms, however, was far from universal. Because of the lack of direct experimental evidence, some influential physicists and chemists were not convinced. They recognized the value of the concept in forming a picture of chemical reactions, but were skeptical of the real, physical existence of atoms. As discussed earlier, the issue was finally settled in the early 20th century when the Brownian motion experiments suggested by Einstein proved the reality of atoms.

Dalton's atomic model might well be called the marble model. Dalton viewed atoms as hard spheres with no internal structure – indestructible and indivisible, as the name implies. However, toward the end of the 19th century, it became clear that this could not possibly be correct.

If a high voltage is placed across two ends of an evacuated tube, invisible rays are produced that emanate from the negative (cathode) electrode. These negatively charged particles were obviously coming from the atoms of the electrode, and with their very small mass relative to atoms, could only be constituents of the atoms. As they were also responsible for the phenomenon of electricity, they were called electrons.

In 1909 the most widely accepted model of the atom was the one proposed by the discoverer of the electron, J.J. Thomson. Thomson visualized all of the positive charge of the atom as being spread uniformly throughout a sphere of about 10^{-10} meters in diameter, with the electrons as smaller particles distributed in shells throughout the atom. For a neutral atom, the total positive charge would be exactly balanced by the negative charge of the electrons. The model was known as the plum pudding model, with the positive charged matter being the pudding and the electrons being the plums.

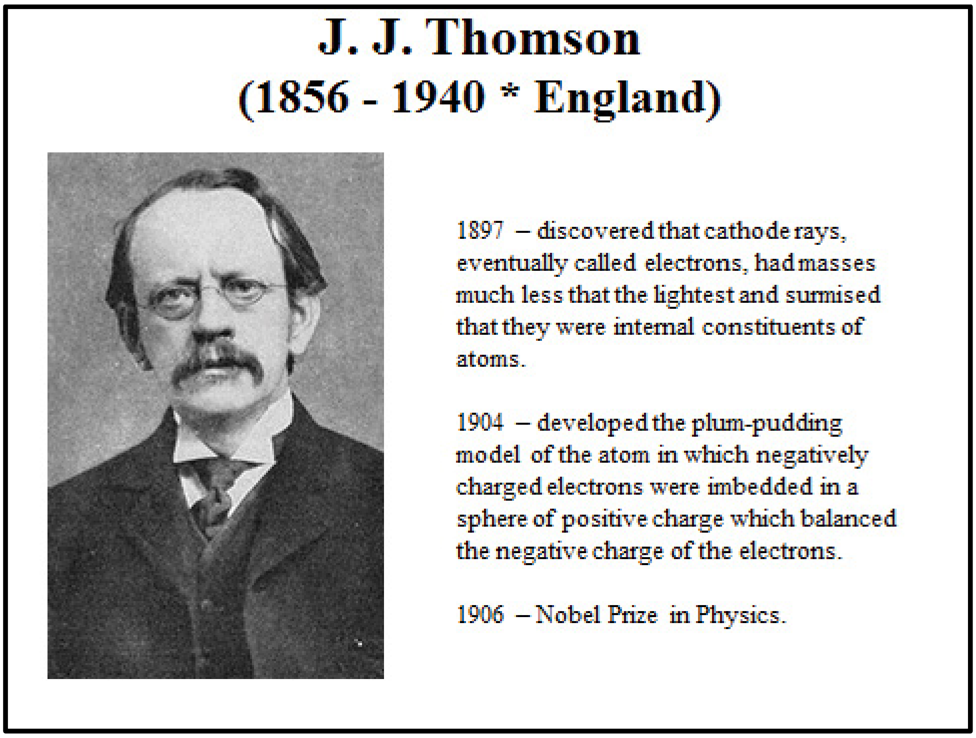


Figure 9.1 J.J. Thomson

In order for a model to be useful, it must account for a substantial body of experimental data and also should generate predictions that can then be tested. Thomson's model was consistent with the existence of electrons as discrete components of the atom, but that was just about it.

In 1911, Ernest Rutherford, a physicist from New Zealand working in England at the time, proposed an alternative model for the atom, one more solidly based in experimental physics. In order to understand how Rutherford came up with this model, we must go back to the end of the 19th century and the discovery of radioactivity.

In 1896, the French physicist Henri Becquerel accidently discovered unusual radiation coming from uranium. Rutherford was quickly able to show that this radiation was complex and consisted of at least two distinct types. Having no idea of the nature of these, he simply called them alpha and beta, the first two letters of the Greek alphabet (the origin of the word 'alphabet'). The alpha radiation could be stopped by several layers of foil. The beta radiation had much greater penetrating ability. Later, a third type was added called gamma (the third letter in the Greek alphabet). It has the greatest penetrating ability of the three.

In 1903, Rutherford determined that the alpha particles are positively charged, and later concluded that they were helium atoms with two electrons removed. He realized that if alpha particles were shot through thin sheets of material, they would be deflected from their original straight line paths and the resulting scattering would provide evidence concerning the internal structure of the atom.

Rough calculations based on the Thomson model predicted that, for metallic foil a few atoms thick, the scattering angle would be small. Rutherford himself was not willing to embark on this line of research as he, like most other physicist at the time, felt Thomson's model was correct. Thus, the alpha scattering experiments would in all likelihood confirm the principle details of the Thomson model rather than produce new insights into the nature of atoms. Fortunately, Rutherford did consider alpha scattering experiments appropriate for his young research assistants.

On the suggestion of Hans Geiger (of later Geiger counter fame) Ernest Marsden, who was still an undergraduate student at the time, was assigned the project. Much to Rutherford's amazement, a few days later, Geiger rushed into his office in great excitement and said "We have been able to get some of the alpha particles coming backwards!"

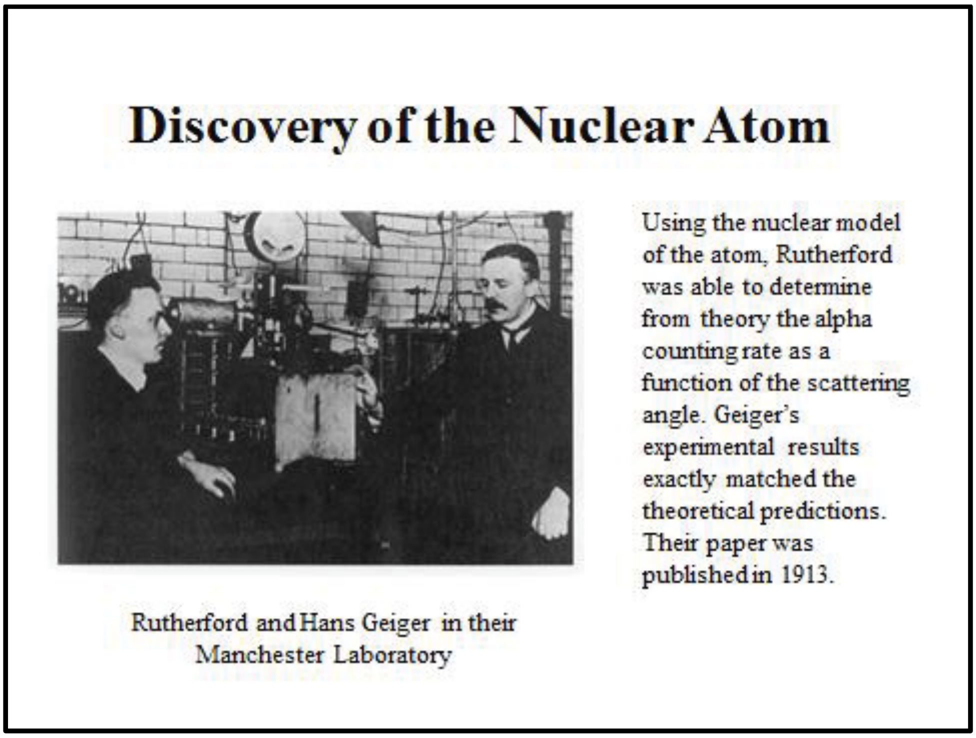


Figure 9.2 Nuclear Atom

Rutherford realized that Marsden's experiment bode ill for the Thomson model, and set his mind to explaining these strange results. Finally, in 1911, he came upon the answer. What Rutherford proposed was the nuclear model of the atom with the positive charge and almost all of the mass of the atom concentrated in a tiny nucleus at the center of the atom, the diameter of the nucleus being just one ten-thousandth of the diameter of the atom. Rutherford's dimensionally correct analogy was a fly in a cathedral. The electrons were viewed as situated outside the nucleus with a distribution that corresponded to the volume of the atom. In this model, atoms are almost entirely empty space.

The statistics of the alpha-scattering experiments were completely explained by the nuclear model. Because the atom is mostly empty space, most of the alpha particles passed through the thin gold foil with little or no deflection. If the alpha did interact with the much less massive electrons (the alpha is 8000 times more massive than an electron), it would be scattered through a small angle, much as a high speed marble would be slightly deflected by a collision with a B-B. An extremely small fraction of the alphas would interact with the much more massive gold nuclei. This could result in a large scattering angle. A head on collision with the more massive gold nucleus would result in the alphas being scattered back in the direction from which they had come, much as a high speed marble could bounce back in a collision with a billiard ball.

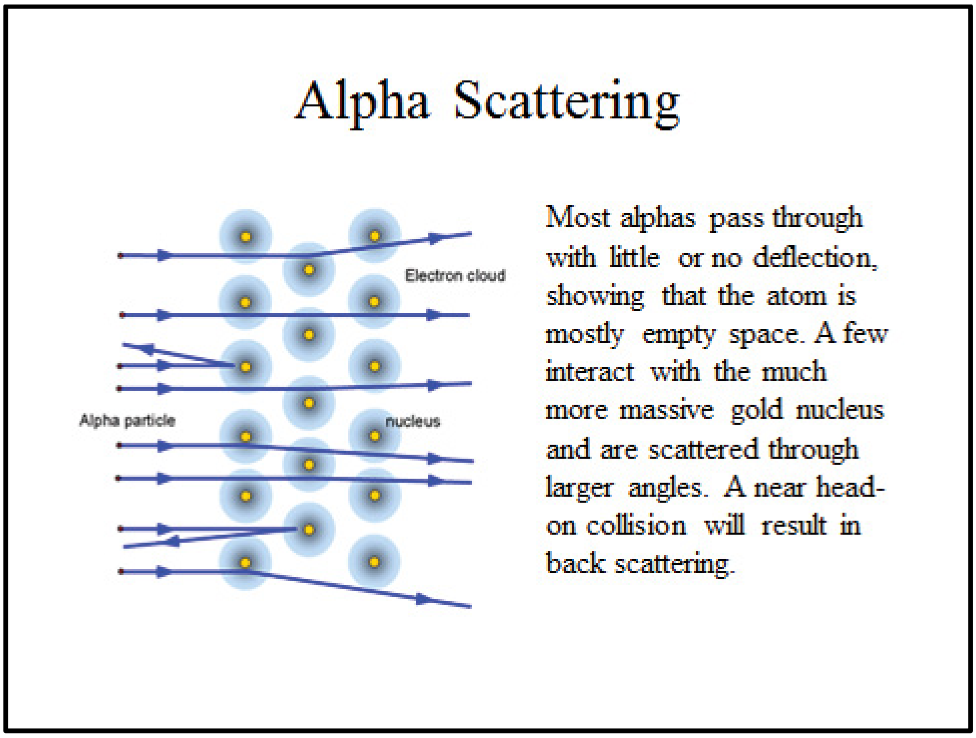


Figure 9.3 Alpha Scattering

It is often the case in physics that the solution to one problem results in another problem of more fundamental significance. Such was the case with Rutherford's nuclear model. The experimental work of Geiger and Marsden established the unambiguous existence of the nucleus. However, if all of the positive charge and almost all of the mass are concentrated in a tiny sphere at the center of the atom, what are the electrons doing? They could not be at rest outside the nucleus. The negatively charged electrons would be attracted to the positively charged nucleus (unlike charges attract each other) and the atom would collapse down to nuclear size.

Attractive forces do not always lead to collapse however. The earth and all the other planets have a tendency to fall toward the sun because of gravity. It is the motion of the planets that prevents the collapse of the solar system. Perhaps the electrons are in stable orbits around the nucleus in much the same way that the planets are in stable orbits around the sun, the difference being that it is the attractive force of unlike charges, rather than gravity, which holds the atom together.

This speculation is particularly attractive because of the similarity between the force the nucleus exerts on the electrons and Newton's equation for he gravitational force. In Newton's theory, the force of gravity is proportional to the product of the masses and inversely proportional to the square of the distance between them. In the law for the force between charges, the force is proportional to product of the charges and inversely proportional to the square of the distance between them. These two inverse-square forces produce identical calculations using Newton's laws of motion. If planets can orbit the sun, it seemed obvious that electrons should be able to orbit the nucleus.

However, there is another factor that needs to be taken into account. Maxwell's theory requires that any charged particle moving in a closed path must emit electromagnetic radiation. Thus, if the electron is orbiting the nucleus, the atom must be continuously emitting electromagnetic radiation. Electromagnetic radiation is a form of energy. The continuous emission of energy would result in the electron spiraling into the nucleus. Calculations showed that the atom would collapse in a tiny fraction of a second. This obviously does not happen. It appears that orbital motion of the electrons cannot provide the stability original hoped for.

The stability of the atom was not the only difficulty with Rutherford's nuclear model of the atom. Over the last few decades of the 19th century, a tremendous body of data had been collected on the phenomenon of atomic spectra. Individual atoms of a gas could, under certain circumstances, emit electromagnetic radiation. The nature of the radiation must be determined by the internal structure of the atom. Any satisfactory model of the atom must be able to account for the radiation both qualitatively and quantitatively. The planetary nuclear model could do neither. It required that the atom emit electromagnetic radiation, but the predicted a continuous spectrum for the radiation and almost immediate collapse of the atom were in sharp contrast with actual experimental observations.

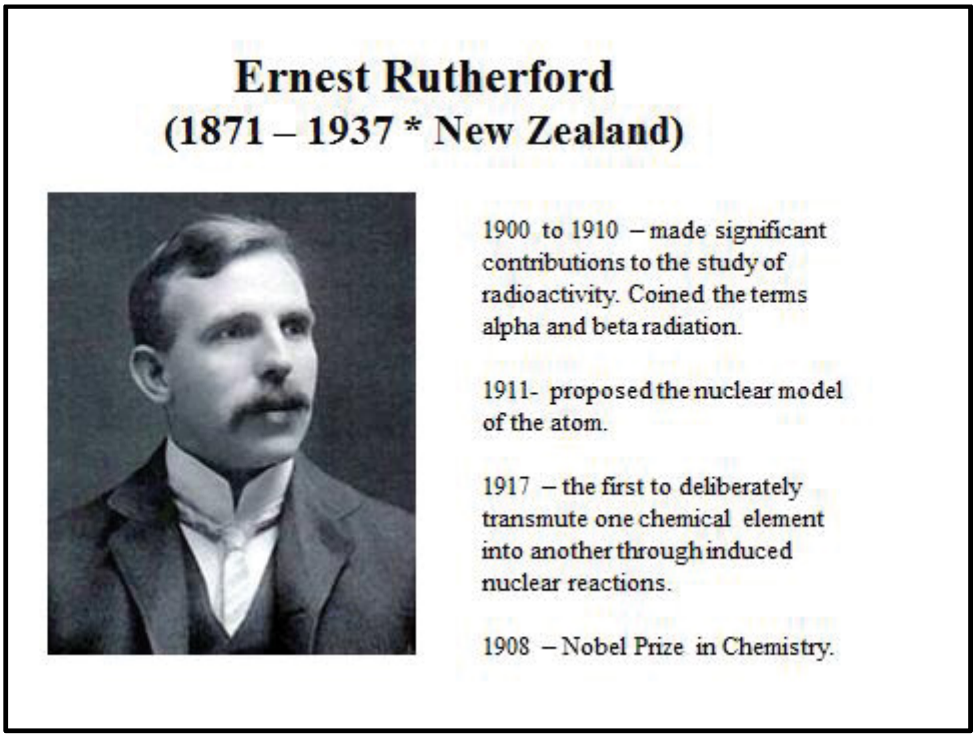


Figure 9.4 Ernest Rutherford

9.2 The Bohr Model of the Atom

An electric discharge tube is an evacuated tube containing a small quantity of gas. When an electric current is established across the tube, the atoms of the gas will emit electromagnetic radiation at certain distinct wavelengths, the spectra being characteristic of the chemical element present in the tube. A common example of a discharge tube is a neon sign. Related to these emission spectra, are absorption spectra produced when a continuous spectrum is passed through a low-density gas. In this case, instead of only certain wavelengths being emitted by the low-density gas, only certain wavelengths are absorbed from the continuous spectrum by the low-density gas. Like emission spectra, absorption spectra are characteristic of the chemical nature of the low-density gas.

The spectra of the various gasses were known quite well by the end of the 19^{th} century. In 1885, Johann Balmer was able to develop a mathematical relationship which correctly represented the wavelengths of the visible portion of the hydrogen spectrum. The Balmer equation was determined simply by manipulating the experimentally determined wavelengths. It was not based on any theory or physical picture of the atoms or the emission process, and, therefore, was not of much immediate interest to physicists. However, any relationship as accurate as Balmer's is very likely to contain hidden physical significance, and it remained for later physicists to discover what this significance was and thereby gain insight into the structure of atoms.

In 1911, the year Rutherford proposed the nuclear model of the atom, a young Danish physicist, Niels Bohr, obtained his Ph.D. and started a one-year fellowship at Cambridge under J.J. Thomson. The young Dane's training had been strongly theoretical, and he brought with him a thorough knowledge of the quantum physics of Planck and Einstein (though, like everyone else, he did not accept Einstein's particle model for electromagnetic radiation). In England, quantum theory was not as popular as it was in continental Europe. This was particularly true in Thomson's laboratory where much time was spent trying to explain atomic spectra using Thomson's plum-pudding model of the atom and the classical physics of Newton and Maxwell.

In 1912, Rutherford visited Cambridge where he met Bohr. They took an immediate liking to one another, while Bohr's relationship with Thomson was becoming increasingly strained. The strain with Thomson was caused, at least in part, because Bohr felt the nuclear model was closer to being correct than was the plum-pudding model. Soon it was decided that it that it might be in the best interest of all concerned if Bohr completed his fellowship at Manchester with Rutherford. At Manchester, Bohr's ideas on the quantum nature of the atom began to coalesce.

Bohr's theory of the atom consisted of four postulates. It addressed only the hydrogen atom, the simplest in nature, with one positive charge on the nucleus and one electron in orbit. He used Rutherford's planetary model as a basis. He then postulated that Maxwell's theory did not apply to electrons orbiting the nucleus of an atom. This was necessary to prevent the collapse of the atom.

Postulate I: An electron orbiting the nucleus does not emit electromagnetic radiation.

In order to explain the existence of discrete emission and absorption spectra, Bohr made two additional assumptions.

Postulate II: The energy of the atom is quantized, that is, only certain values of energy are allowed.

Postulate III: Energy is emitted or absorbed by an atom only when it makes a transition from one allowed energy state to another.

Bohr called the allowed energy states of Postulate II, stationary states to stress their non-radiating nature. Using Newtonian mechanics, the energy of an orbiting system is determined by the radius of the orbit. Thus the allowed stationary states are equivalent to restricted values for the radius of the orbit of the electron around the nucleus. Certain values of the radius are allowed and all others are forbidden. In order for the electron to make a transition from one of the allowed energy states to another, it must emit or absorb energy depending on whether the transition is to a higher or lower energy state.

If only certain energies are allowed for the atom, only energies corresponding to the difference between allowed energies can be emitted or absorbed. Bohr's quantum concepts were based on Planck's quantum hypothesis. Bohr visualized the electron as oscillating with a frequency determined by Planck's quantum relationship $E = h\nu$ with E equal to the difference in energy between allowed energy states. Thus, the discrete energies of the atom result in discrete frequencies or wavelengths.

Bohr had qualitatively accounted for discrete emission and absorption spectra. His next step was to determine the rule that would account for them quantitatively; that is, the rule that would restrict the energy states in just the right way to account for the experimentally observed wavelengths of emission and absorption. Before Bohr had time to do this, his fellowship in England was up and he returned to Copenhagen to marry and assume a teaching position.

In January 1913, a lucky event occurred. An old classmate asked Bohr if the theory he was working on might explain the Balmer equation. Bohr was not familiar with the equation and the friend suggested that he look it up. Bohr later recalled, "As soon as I saw Balmer's formula, the whole thing was immediately clear to me." Bohr now had the necessary orbit rule.

Postulate IV: The angular momentum of the electron in its orbit, mvr, must be an integer multiple of Planck's constant divided by 2π.

In the fourth postulate, m is the mass of the electron, v is its speed (the product, mv, is the linear momentum of the electron), and r is the radius of the orbit. Quantizing the angular momentum is equivalent to quantizing the energy. Now that he had an equation for angular momentum, he was able to use the classical relationship between the two physical quantities to determine the allowed energy states. Because the energy involved in atomic physics are so small, the energy unit used for events on the macroscopic scale is not appropriate. A new unit called the electron-volt, eV, (one eV is the change in the kinetic energy of an electron accelerated through a potential difference of one volt) is used. If the energy of a hydrogen nucleus (later determined to be a proton) at rest relative to an electron is taken to be zero, then the energy of the hydrogen atom is negative. That is, energy must be added to a hydrogen atom to remove the electron producing the zero energy state.

Based on the fourth postulate, and expressed in eVs, the allowed energies for the hydrogen atom are:

$$E_n = -\frac{1}{n^2}(13.6 \text{ eV})$$

where n is an integer, 1, 2, 3, etc.

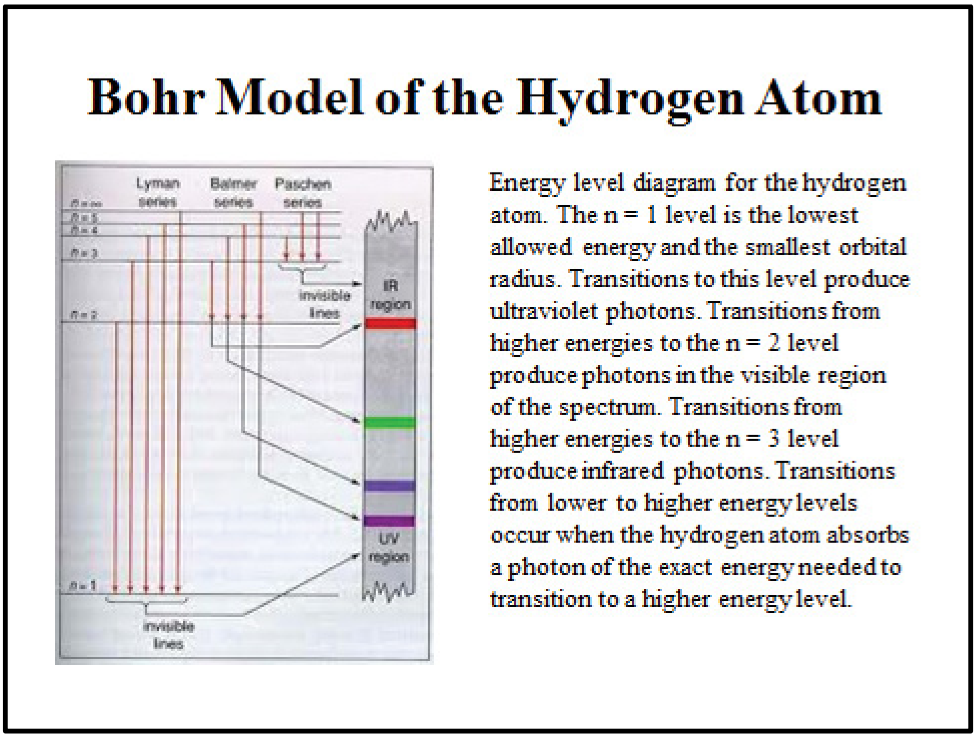


Figure 9.5 Hydrogen Atom

Each value of 'n' corresponds to a particular allowed energy state. The lowest possible energy for the hydrogen atom is called the ground state and all states above are excited states. They are called excited because the atom will not remain in these states for long, usually less than 10^{-8} seconds. The atom will spontaneously transition to a lower energy state by emitting its excitation energy. It is not necessary for the transition to be directly to the ground state, just to a lower state. Several transitions may occur before the ground state is reached. It is these spontaneous transitions from higher energy states to lower ones that constitute discrete emission spectra.

When an atom is in its ground state, it has no excitation energy and cannot emit energy. It can however, absorb energy and transition to a higher allowed stationary state. It is these upward transitions that are responsible for absorption spectra. Because the allowed energy states are unique for the atoms of a particular element, so are the differences between them, explaining why atomic spectra are characteristic of the chemical element of the low-density gas that produces them.

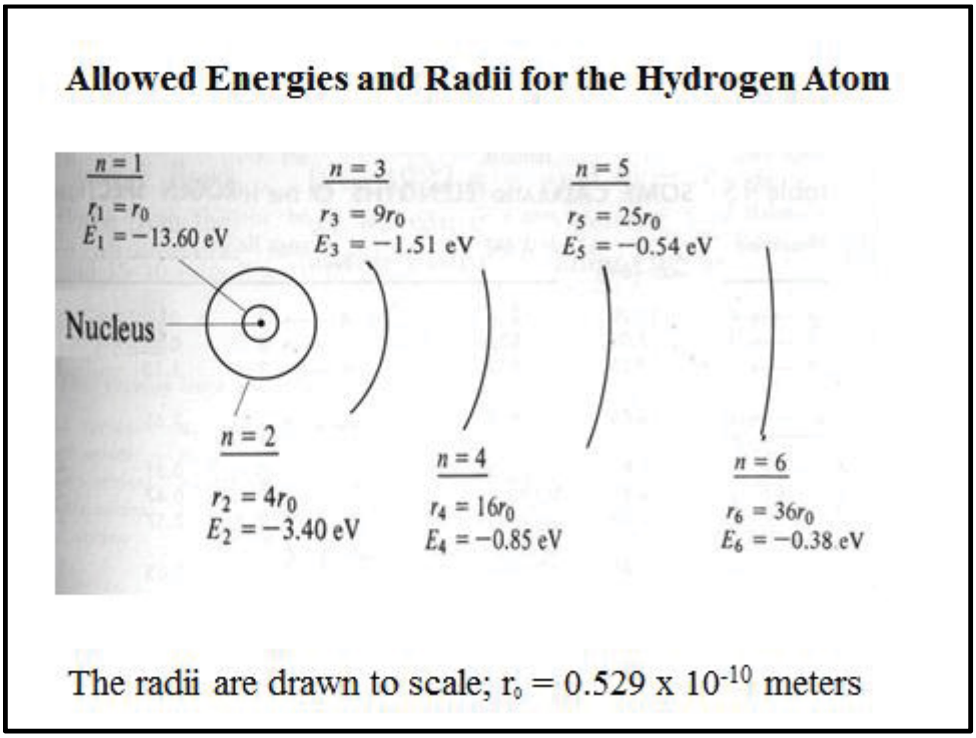


Figure 9.6 Allowed energies

Bohr's theory of the hydrogen atom explains the Balmer equation. The wavelengths of the emitted radiation in transitions between higher energy states and the n = 2 state (the first excited state of hydrogen) exactly correspond to the wavelengths generated by the Balmer equation.

Example calculation: What are the energies and wavelengths emitted in the transitions to the n = 2 state of hydrogen?

From Bohr's equations, the energies of the allowed states are: n = 2, -3.40 eV; n = 3, -1.51 eV; n = 4, – 0.85 eV; n = 5, – 0.54 eV, and n = 6, – 0.38 eV.

A transition from the n = 3 to the n = 2 state reduces the energy of the atom from -1.51 eV to – 3.4 eV, a drop of 1.89 eV. From the equation $E = h\nu = hc/\lambda$ we get

$$\lambda = hc/E = (4.14 \times 10^{-15} \text{ eV sec})(3 \times 10^{8} \text{ m/sec})/1.89 \text{ eV} = 6.57 \times 10^{-7} \text{ meters.}$$

This wavelength is in the red region of the visible spectrum and exactly corresponds to the first term in the Balmer equation.

The transition from n = 4 to n = 2, corresponds to a drop in energy of 2.55 eV and a wavelength of 4.87×10^{-7} meters. This wavelength is in the blue-green region of the visible spectrum and exactly corresponds to the second term in the Balmer equation.

The transitions from the n = 5 and 6 states yield wavelengths in the violet region of the visible spectrum that exactly correspond to the third and forth terms in the Balmer equation.

The energy differences between the n = 1 state and the higher levels are larger and the corresponding wavelengths shorter than is the case for the Balmer series. These wavelengths lie in the ultraviolet region of the spectrum. Transitions down to the n = 3 state have less energy and longer wavelengths than the Balmer series and lie in the infrared region. In hydrogen, only transition down to the n = 2 state result in the emission of visible light.

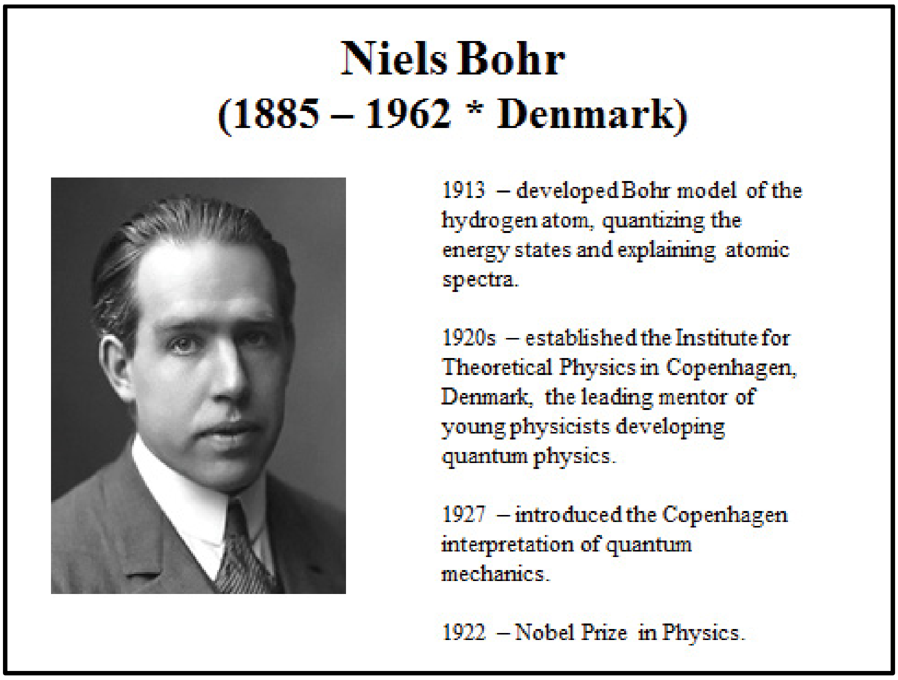


Figure 9.7 Niels Bohr

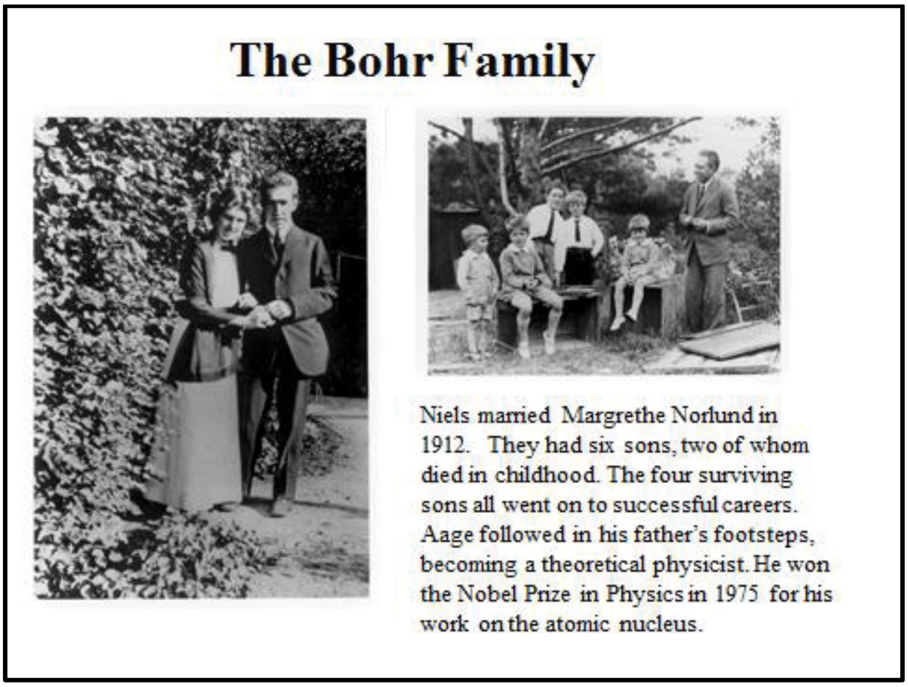


Figure 9.8 Bohr family

9.3 Problems with the Bohr Model

The Bohr theory of the atom represents an important theoretical breakthrough in our understanding of the atom. It was the first atomic model to account for atomic emission and absorption spectra. However, there were several very troubling aspects of the theory. First and foremost, Bohr simply stated that well accepted laws of physics did not apply in the case of atoms. He offered no justification for this (except, perhaps, that it seemed to work) and did not provide a comprehensive theoretical explaanation as to why classical physics did not work in this case.

Bohr's theory was neither a classical nor a quantum theory, but a combination of the two. He used the classical theory of Newton to discuss the electron in its allowed orbit and to calculate the energy associated with the state. On the other hand, the existence of stationary states, the mechanism for emission and absorption, and the quantization of angular momentum are strictly non-classical. Conceptually, this awkward hybrid of old and new made little sense. Even Rutherford, whose case for the nuclear model received strong support from the Bohr theory, had some concern on this point. In a letter to Bohr, he wrote,

> Your ideas as to the mode of origin of the spectra in hydrogen are very ingenious and seem to work out very well; but the mixture of Planck's ideas and the old mechanics makes it very difficult to form a physical idea of what is the basis of it all.

In addition, and perhaps most disturbingly, the transition between energy states is a discontinuous event. That is, it cannot be represented as an electron moving continuously in space and time from one orbit to another. At one instant of time the atom is in one state; at the next instant it is in another with energy being either emitted or absorbed. To emphasize the discontinuous nature of the transitions, Bohr called them 'quantum jumps.'

Physicists at the time had difficulty with this point. Rutherford's initial response to the theory was:

> There appears to me one grave difficulty with your hypothesis, which I have no doubt you fully realize, namely, how does an electron decide what frequency it is going to vibrate at when it passes from one stationary state to the other? It seems to me that you would have to assume that the electron knows beforehand where it is going to stop.

It was clear that classical physics could never offer an acceptable theory of the atom. What was needed was a completely quantum description of the atom with no carry-over from classical physics. No one realized this more than Bohr himself, and he immediately set to work on such a theory.

It is worth noting that, although the simplest way to visualize the emission and absorption of discrete energies is to use Einstein's photon model, Bohr continued to use the wave model. At the time, he, like Rutherford, Planck, and virtually all other physicists, believed that the photon model could not possibly be correct. Bohr thought of emission of radiation in terms of Maxwell's classical theory. In this theory, radiation of a certain frequency is produced by a charge vibrating at that frequency. This is not the case with the photon model. Photon creation and absorption is, by its very nature, discontinuous. In retrospect, it is clear that the photon model is much more appropriate than the wave model for use with the Bohr model. Bohr did not openly accept the idea of particles of light until 1925. In his Noble Prize acceptance speech in 1922, he said, "In spite of its heuristic value, the hypothesis of light quanta, which is quite irreconcilable with the so-called interference phenomena, is not able to throw light on the nature of radiation."

10 Werner Heisenberg and Matrix Mechanics

Your goals for this chapter are to know the following:

- Why Heisenberg felt it was necessary to develop a new theory of the atom.
- What initially troubled Heisenberg about the mathematics of his new theory and how this was resolved by Max Born.
- The contribution Wolfgang Pauli made to Heisenberg's theory.

The start of the 20th century witnessed a revolution in physics – the quantum revolution. It began with Max Planck's theory of thermal radiation in 1900 and continued through Einstein's explanation of the photoelectric effect and Bohr's model of the atom. This quantum theory was clearly a mishmash of the old classical ideas and ad hoc quantum postulates. Physics cannot be based on such a flimsy foundation. A physical theory with the same sort of generality as Newton's mechanics and Maxwell's electromagnetic theory is required. A theory based not on a large number of ad hoc postulates, each applying to a restricted range of phenomena, but one based on a few simple, elegant, powerful postulates. Neils Bohr took the lead in the effort to develop his theory.

10.1 Werner Heisenberg

In 1922, Bohr visited Germany to deliver a series of lectures setting out the view on quantum theory developed in Copenhagen. A young German graduate student, Werner Heisenberg was in the audience. Heisenberg had read the recent papers coming out of Copenhagen and did not fully agree with them. He had the nerve to speak up from the back of the room with his objections. After the lecture was over, Bohr invited Heisenberg on a long walk to continue the discussion on the issues Heisenberg raised. At the end of the conversation, Bohr invited Heisenberg to spend some time in Copenhagen. Years later Heisenberg would say, "My real scientific career only began that afternoon."

In 1924, Heisenberg completed his doctorate at the University of Munich and joined his friend Wolfgang Pauli at the University of Gottingen. There he studied under the mathematical physicists, Max Born. His thesis at Munich had nothing to do with quantum theory, but ever since his meeting with Bohr, the problem of the atom had become his main interest. Although Heisenberg had remained in contact with Bohr, it was not until September of 1924 that he was able to spend several months in Copenhagen. By this time, his focus was shifting away from what atoms are and more toward what atoms do. He was no longer much concerned about the position and velocity of the electron, but with the frequencies and intensities of the emitted radiation. He had completely abandoned the old idea of electrons following well-defined orbits governed by classical mechanics, but had nothing yet to put in the place of the orbits.

After Copenhagen, Heisenberg returned to Gottingen to continue trying to force his new approach into a coherent theory. The mathematical difficulties seemed insurmountable. In June of 1925, a severe bout of hay fever, sent him to a small barren island, Helgoland, off the northern coast of Germany where he felt the fresh air and isolation would allow his mind to concentrate on the problem. Slowly his health returned and his mind was able to better focus on his new mechanics. He was able to devise a method of multiplication that suited his needs. He was confident he now had the tool he needed.

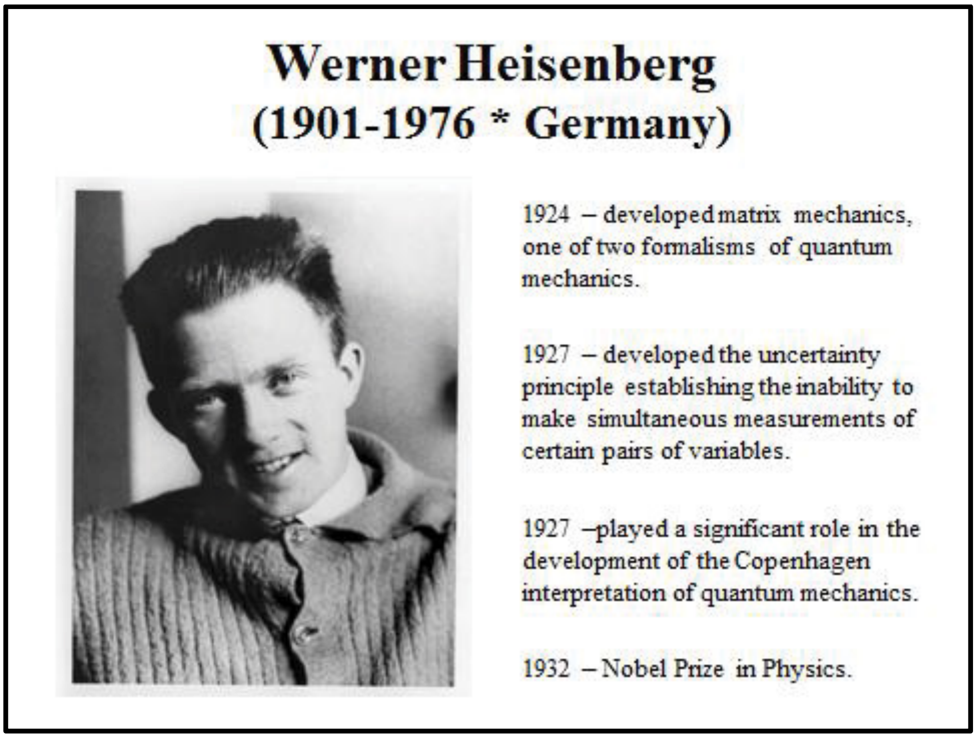


Figure 10.1 Werner Heisenberg

10.2 Matrix Mechanics

Lying sleeplessly in his bed one night, he suddenly saw clearly how to proceed. He jumped up and began feverishly to lay out the theory. Finally he reached the end and got an answer that was more than he could have dreamed of. He discovered that his strange mathematics did indeed yielded a consistent result for the energy of a mechanical system, but only so long as the energy was one of a restricted set of values. His new form of mechanics was, as he had hoped, a quantized form of mechanics. Unlike all previous attempts, the quantization came not from ad hoc postulates restricting the physics but came naturally from the mathematics itself.

One thing, however, disturbed him. His multiplication rule was not reversible. That is, × times y was not necessarily the same as y time x. This confused Heisenberg, but it was exactly this feature that quantized his theory. Upon his return to Gottingen from Helgoland, Heisenberg showed his theory to Max Born. Something in the theory seemed familiar to Born, but he could not place it at the time. It was not until a month later that Born, suddenly realized that the strange algebra Heisenberg was using was actually an arcane branch of mathematics known as matrix algebra, one that had never before been used in physics. Without knowing it, Heisenberg had reinvented matrix algebra.

After showing his new ideas to Born, Heisenberg left for a short vacation. It was while he was away that Born recognized Heisenberg's mathematics as matrix algebra. Born and his student Pascual Jordan immediately rewrote Heisenberg's theory in the formal language of matrix algebra. When Heisenberg returned, he joined Born and Jordan in what became known as the three-man paper. This further refined and extended Heisenberg's work. From this point forward the theory was called matrix mechanics.

Matrix mechanics was not immediately accepted by most physicists. It required learning a new branch of mathematics for one thing. Quantum mechanics in matrix form was extremely complicated, and it was not easy to understand what the matrices represented physically. This was not at all what physicists were used to.

Pauli was, at first, very skeptical of Heisenberg's theory, sometimes scathingly so. Heisenberg eventually lost his temper and challenged Pauli to either put up or shut up about the theory. Stung, Pauli set to work, and in less than a month had used matrix mechanics to derive the Balmer series of spectral lines of the hydrogen atom. It was Bohr's ability to do this in 1913 that established the credibility of his, much simpler, model of the atom. But while Bohr's model was an awkward hybrid of classical mechanics and ad hoc quantum postulates, Heisenberg's was a fully quantum theory with applicability to any mechanical system, not just atoms.

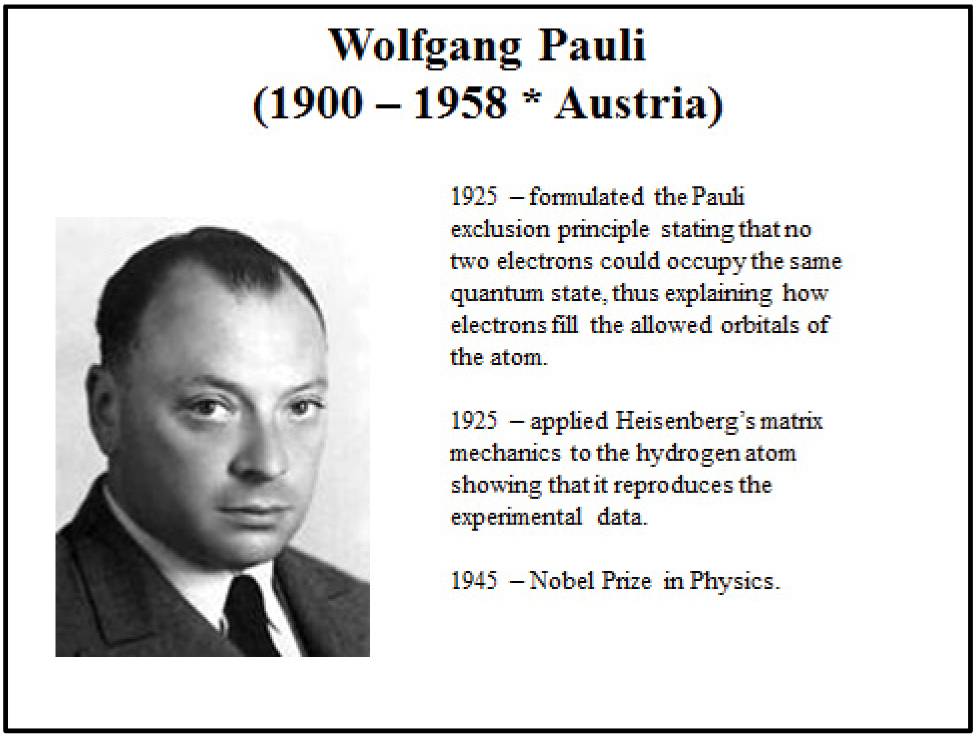


Figure 10.2 Wolfgang Pauli

Pauli's calculation was a powerful and convincing demonstration that matrix mechanics was more than a mathematical formalism. A mollified Heisenberg wrote Pauli saying "I hardly need to tell you how thrilled I am about the new hydrogen theory, and how amazed that you worked it out so quickly."

Physicists had waited 25 years for a comprehensive quantum theory. Now that they had one, they didn't like it. It was formulated in a weird form of mathematics that they were not familiar with. Worst of all, it retained the quantum jumps first introduced by the Bohr model – transitions from one state to another that were discontinuous in space and time. To the immense relief of most, within a year a new version, much more compatible with their ideas of what a physical theory should look like, was developed.

11 Erwin Schrodinger and Wave Mechanics

Your goals for this chapter are to know the following:

- De Broglie's contribution to wave mechanics.
- How de Brogile's concept is related to the Bohr model of the atom.
- What differentiated Schrodinger's theory from Heisenberg.
- Why most physicists preferred Schrodinger's theory to Heisneberg's.

While matrix mechanics was being developed in Germany, events were occurring in Switzerland that would lead to a second version of quantum theory. This version had its beginning in the somewhat metaphysical musings of a young French aristocrat.

11.1 De Broglie Waves

Louis Victor de Broglie was a member of one of the most prestigious noble families of Europe. As was traditional with a man of his position, he received a humanist education. In 1910, he earned a degree in history from the Sorbonne with a view of becoming a diplomat. Through the influence of his older brother Maurice, who, against his family's wishes, had become an experimental physicist, de Broglie acquired an interest in physics and particularly in the work of Einstein. In 1913 he received a degree in science. After service in the First World War, de Broglie resumed his study of science and began work toward his doctorate in physics.

In 1924, de Broglie presented his thesis to the Faculty of Science at Paris University. It was inspired by Einstein's theory of light which ascribed a particle-like nature to electromagnetic radiation which classically had been considered to be a wave. By 1924 the dual nature of electromagnetic radiation was beginning to be accepted. Only a particle model could explain the photoelectric and Compton effects.

De Broglie reasoned that if nature consisted of electromagnetic radiation and material particles, and electromagnetic radiation had a dual nature, then symmetry arguments suggested that material particles should also possess a dual nature. That is, material particles should, under certain circumstanced, behave as if they were waves. What an absurd idea. Everyone knew that electrons and atoms are like tiny baseballs and are not in the least wave-like.

De Broglie's thesis presented quite a problem to the Faculty of Science. On the one hand, de Broglie was obviously intelligent, from one of the most influential families in France, and the brother of a distinguished colleague. On the other hand, his thesis was truly bizarre. This convert from the humanities had put forward the most outrageous hypothesis, offered no experimental evidence in its favor, and offered no real physical interpretation of his matter waves. De Broglie himself characterized his theory as "a formal scheme whose physical content is not yet determined."

The problem was resolved by sending a copy of the thesis to Einstein for his opinion. Einstein replied, "It may look crazy, but it is completely sound." The faculty was off the hook. With no less an authority than Einstein standing behind it, the thesis was accepted and the doctorate awarded. In retrospect this was a very wise decision, for in 1929 de Broglie became the first person to receive the Nobel Prize in Physics for a doctoral thesis.

De Broglie based his theory of matter waves on an analogy with electromagnetic radiation. For electromagnetic radiation, the wave properties (wavelength, frequency) are related to the particles properties (energy, momentum) in a precise, well-defined way.

$$E = h\nu \text{ and } p = h/\lambda$$

where E and p are the energy and momentum and ν and λ are the frequency and wavelength. De Broglie assumed that these relationships, which had been shown to apply to electromagnetic radiation, also applied to particles such as electrons and atoms. Therefore the wavelength associated with a material particle is

$$\lambda = h/p = h/mv$$

where m is the mass of the particle and v is its speed.

When this equation is applied to electrons traveling at speeds on the order of 1% of the speed of light, the resulting wavelengths are on the order of 10^{-10} meters, approximately the wavelengths of ultraviolet radiation.

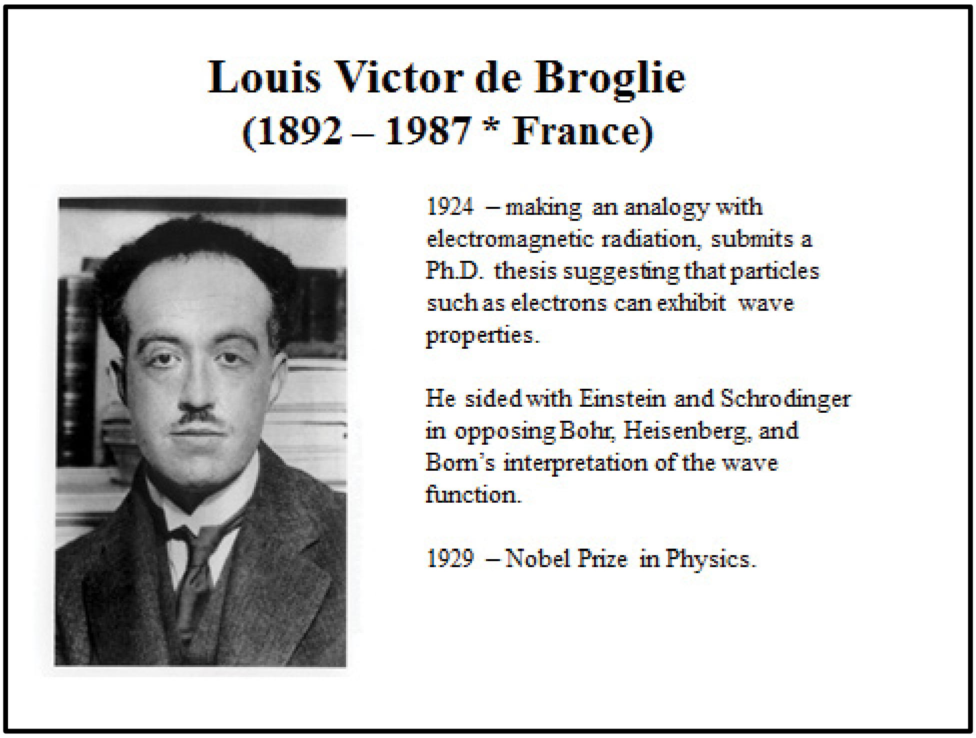


Figure 11.1 Louis Victor de Broglie

When de Broglie's relationship is applied to a macroscopic object, such as a baseball, the extremely small value of h, 6.63×10^{-34} joule-seconds, insures that the resulting wavelengths are infinitesimal and immeasurable. This is important because baseballs and other macroscopic objects exhibit only the particle-like behavior consistent with our everyday experience. If de Broglie's relationships had resulted in measurable wave properties for such objects, his theory could not possibly be correct.

In addition to symmetry, de Broglie's thesis had one other argument in favor of ascribing a wave nature to material particles. According to the Bohr model of the atom, the electron orbits are quantized – only certain orbits are allowed. It is difficult to imagine why this is so if we think of the electron as a small sphere circling the nucleus. However when the electron is viewed as a wave, the quantized orbits are a natural feature of the model. For instance, a guitar string can vibrate only with certain wavelengths, all others dying out almost immediately. If the guitar string analogy is used for the atom, only certain wavelengths are allowed for the orbital electron and perhaps Bohr's ad hoc quantum rule can be reproduced from a more fundamental postulate – the wave nature of the electron.

Imagine the electron orbits as circular guitar strings. Only vibrations containing a whole number of wavelengths can be established. With the guitar string analogy, all others would quickly die out. Thus the circumference of the orbit must be equal to a whole number times the wavelength of the electron.

$$2\pi r = n\lambda \text{ (where } n = 1, 2, 3, \text{)}$$

where r is the radius of the orbit and λ is the wavelength of the electron. Using de Broglie's relationship, we get

$$2\pi r = nh/p = nh/mv$$

$$\text{or, } mvr = nh/2\pi$$

which is exactly Bohr's fourth postulate – his quantum rule.

Though the derivation of Bohr's quantum rule lent some plausibility to de Broglie's ideas, experimental verification was required before they could be accepted. This came in 1927.

For many years, two American physicists, Clinton Davisson and Lester Germer, had been experimenting with the scattering of electrons by a solid. A beam of electrons was directed perpendicular to the face of a crystal, and the rate of electrons scattering as a function of scattering angle was measured. The results were inexplicable in terms of a particle model for the electrons. However, if one viewed the beam of electrons as a wave, the scattering results could be successfully interpreted as an interference effect associated with the scattered waves. (Exactly the same method had been used in 1913 to demonstrate the wave nature of X-rays.) Moreover, when the results were used to determine the wavelength associated with the electron beam, a value in complete agreement with the de Broglie relationship was obtained. Since 1927 the wave nature of other elementary particles, atoms, and even molecules has been successfully demonstrated.

11.2 Schrodinger's Wave Equation

De Broglie suggested that material particles have associated with them wave properties, with the wavelength being determined by their momentum. However there is much more that must be known about these mysterious matter waves if they are to be of use in physics. Wave equations exist for such macroscopic phenomena as sound and water waves, and Maxwell's electromagnetic theory provides a wave equation for electromagnetic waves. What was needed is a general wave equation whose solution is the de Broglie wave to be associated with a particular microscopic system. De Broglie's thesis did not provide this.

As noted earlier, Einstein was among the first to become aware of de Broglie's ideas. Early in 1925, in a paper on ideal gases, he referred to de Broglie's work and stated, "I believe it involves more than merely an analogy." The remark started Erwin Schrodinger, a physicist at the University of Zurich, thinking about matter waves and wave equations. At this time, however, he was not able to come up with a formalism whose results were in agreement with observations. Disappointed, Schrodinger temporarily abandoned his efforts.

Several months later, the eminent physical chemist Peter Debye was instrumental in rekindling Schrodinger's interest. Debye was in charge of the physics colloquium at the Swiss Federal Institute in Zurich (Einstein's alma mater). He suggested to Schrodinger that since he was not working on anything important at the time, that he prepare a colloquium on the recent thesis of de Broglie. After the presentation, Debye casually remarked that it seemed childish to discuss waves if there was no corresponding wave equation from which the properties of the waves could be calculated. Just a few weeks later, Schrodinger started a colloquium by saying, "My colleague Debye suggested that one should have a wave equation; well I have one."

By making an analogy between de Broglie waves and mechanical waves, Schrodinger had produced a general wave equation whose solution corresponded at all points in space and time to the de Broglie wave describing the particular microscopic system. The form of the equation is known mathematically as a partial differential equation, a type of equation much more familiar to physicists than Heisenberg's matrix algebra. Physics laws had, since the time of Newton, been expressed in the form of differential equations. A good deal of experience with calculus is required to solve this type of equation, but the equation itself has some interesting features:

$$(- \frac{h^2}{2m} \nabla^2 \ E_p) \ \Psi = i \, \frac{h}{2\pi} \frac{\partial}{\partial t} \Psi$$

In this equation, ∇^2 and $\frac{\partial}{\partial t}$ represent mathematical operations in differential calculus to be performed on the wave function, Ψ; ∇^2differentiation with respect to the space coordinates, and $\frac{\partial}{\partial t}$ differentiation with respect to time. E_p represents the potential energy of the system and 'i' is the symbol for the imaginary number, the square root of negative one. Schrodinger's equation represents the first time an imaginary number, familiar in mathematical equations, had appeared in a fundamental physics equation. If the potential energy is known, the equation can, at least in principle, be solved for the wave function. The symbol for the wave function is a Greek letter spelled 'psi' and pronounced 'sigh.' It is a mathematical function of space and time that contains all of the information represented by the de Broglie wave.

As a test of his theory, Schrodinger used the potential energy of the hydrogen atom and solved for the wave functions. The result, published in 1926, was astounding. For an electron bound to the hydrogen nucleus, the Schrodinger theory yields a series of wave functions, each corresponding to a definite energy state. These quantized energies were identical to those of the Bohr theory. The associated wave functions are known as standing waves or as stationary states.

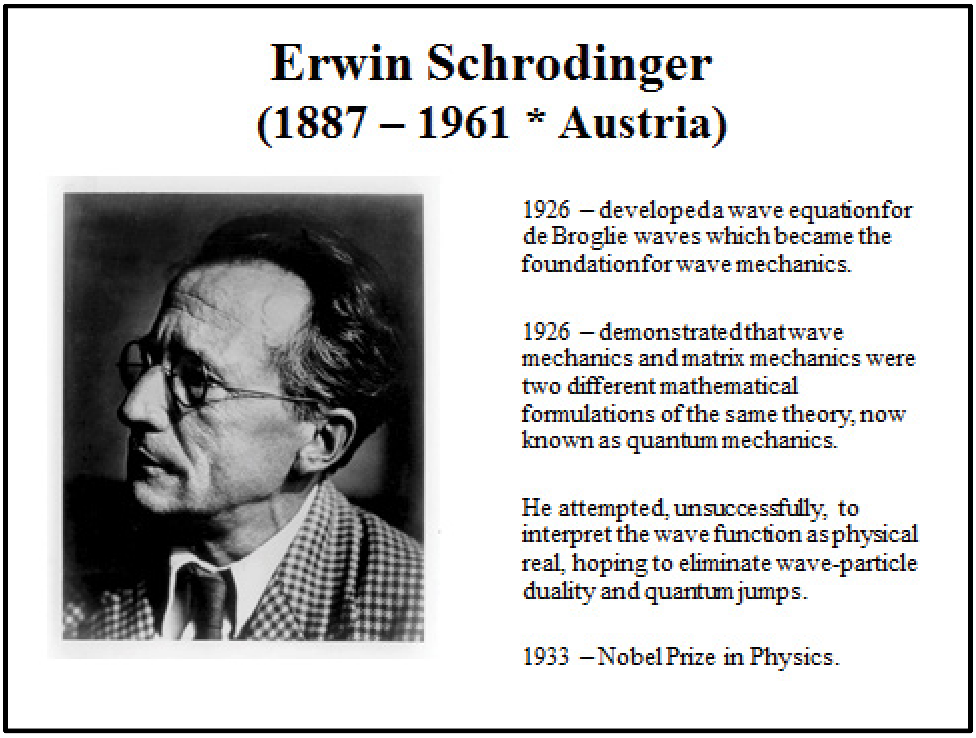


Figure 11.2 Erwin Schrodinger

For years physicists had realized the inadequacy of the half classical-half quantum Bohr theory. What was clearly needed was a new theory, one that could stand on its own without using classical mechanics as a crutch. Now suddenly, in the space of less than a year, there was not one but two such theories. An embarrassment of riches, to be sure. Almost immediately, however, Schrodinger, much to the relief of the physics community, was able to show that the two theories, Heisenberg's matrix mechanics and his wave mechanics, were mathematically equivalent. Matrix mechanics and wave mechanics are now considered just two different formalisms for a single physical theory – quantum mechanics.

Quantum mechanics completely supplants the ad hoc quantum theories of Planck, Einstein, and Bohr. Not only does it duplicate all their successes, it extends well beyond them to encompass phenomena inexplicable using the old quantum theory. It is every bit as elegant and comprehensive as the theories of classical mechanics and electromagnetism.

Schrodinger's formulation of quantum mechanics seemed to many of the old guard physicists as a less radical break with the old classical physics. There was even some hope that, as Schrodinger himself hinted in several papers, wave mechanics could replace what Bohr and Heisenberg referred to as quantum jumps, with smooth transitions from one state to another. That is, as fluid transformations of one standing wave to another, an event that would occur rapidly but still continuously in space and time. This as opposed to the discontinuous events of the original Bohr theory and of matrix mechanics. The physics community quickly became divided as to which was the preferred theory. One side, including most of the older more conservative physicists like Einstein, Planck, de Broglie, and Schrodinger, strongly supported Schrodinger's point of view. On the other side, the young Turks, Heisenberg, Pasqual Jordan, Pauli, and even a few of the older physics such a Bohr and Born, believed that there was no turning back from the discontinuous implications of the new physics.

Physicists working in other areas of physics were looking for a version of quantum mechanics they could understand. That is why wave mechanics received a warm welcome relative to the almost complete disregard that greeted matrix mechanics. Wave mechanics contained no weird algebra, just old fashion differential equations, not unlike those of classical physics.

Both Heisenberg and Schrodinger's versions of quantum mechanics of are non-relativistic theories. Not surprisingly, some difficulties arise when dealing with electrons, which due to of their small mass, are easily accelerated to relativistic speeds. In 1928, an eccentric, but brilliant, young English physicist, P. A. M. Dirac, succeeded in formulating a relativistic wave equation for electrons, thus completing the development of the theory. This work predicted the existence of a strange particle called an anti-electron. This particle, now called the positron, has the same mass and several other properties as an electron, except that is has a positive charge. Positrons were discovered in 1932 by the American physicist, Carl Anderson. Since then anti-protons and anti-neutrons have been discovered.

In his 1933 Nobel acceptance speech, Schrodinger said his intention in developing wave mechanics was to save "the soul of the old system" of mechanics. However, wave mechanics was not able to save the soul of the old system. In the late 20s and early 30s, it became increasingly clear that fundamental classical ideas simply did not apply on the microscopic scale. The only reason they appeared to work is that, in the limit as the microscopic approaches the macroscopic, the laws of quantum mechanics approach those of classical physics. Quantum mechanics is the more general theory and classical physics is just an approximation applicable to the macroworld. A few, Einstein among them, held out to the bitter end, but for the next generation of physicists, there was little doubt that the physics of quantum mechanics was the correct physics. Some of its philosophical implications are still debated today, but the physics itself is now beyond question. Quantum mechanics has no competition.

12 Max Born and the Interpretation of the Wave Function

Your goals for this chapter are to know the following:

- Schrodinger's interpretation of his wave function.
- Born' interpretation of Schrodinger's wave function.
- Why most physicists were initially reluctant to accept Born's interpretation.
- The significance of the double-slit experiment.

In developing his wave equation, Schrodinger used analogy with mechanical waves such as water waves or sound waves. For mechanical waves, the wave function that is the solution to the wave equation has a simple physical interpretation; for example, the amplitude of the water wave or the pressure of the sound wave. But what is the interpretation of Ψ in wave mechanics? Initially it was strictly a computational device, and un-interpreted mathematical function from which certain physically meaningful quantities, such as energy or wavelength, could be deduced. Is there more to it than that? Just how does Ψ correspond to 'reality' in the traditional sense of the word? A satisfactory interpretation is necessary before quantum mechanics can be considered a completed theory. One attempt at this was made by Schrodinger soon after completing the mathematical formalism of his theory.

12.1 Schrodinger's Interpretation

Schrodinger viewed Ψ as representing a continuous distribution of mass and charge in space. That is, as an alternative to the non-visualizible wave-particle duality, Schrodinger considered the electron to be only a wave with its mass and charge smeared out over the region of space where Ψ was not equal to zero. This seemed to work well for the atom, for if the electron could be thought of as a standing wave surrounding the nucleus rather than as a particle orbiting it, the model of the atom would no longer be in conflict with Maxwell's theory of electromagnetism. Maxwell's theory requires that a charged particle traveling in a closed path emit electromagnetic radiation, and this loss of energy would quickly lead to the collapse of the atom. However, if the electrons are standing waves, the charge distribution does not change with time, and hence no electromagnetic radiation would be emitted. This problem had previously been handled by an ad hoc postulate of the Bohr theory, which, of course, satisfied no one, especially Bohr.

Perhaps the most appealing aspect of Schrodinger's interpretation is that it appeared to restore continuity to physics. The idea of quantum jumps, processes not amenable to continuous space-time descriptions, had been introduced by Bohr in his theory of the hydrogen atom. It was also an integral part of matrix mechanics. This notion profoundly disturbed most physicists for obvious reasons. The mechanistic-deterministic worldview was still fundamental to most physicists' thinking. Atomic transitions, as viewed by Schrodinger, consisted of the gradual dissolving of one wave pattern into another and were continuous in space and time. For these reasons Schrodinger's interpretation received enthusiastic support from the more traditionally minded physicists such as Einstein, Planck, and de Broglie.

Schrodinger's interpretation viewed electrons as waves. But what about the many experiments that seemed to indicate a particle nature for an electron? That is, experiments in which the mass and charge associated with an electron seemed not to be smeared out over a region of space, but to be localized in a very small region of space. How could Schrodinger's electron waves, which in the hydrogen atom are spread out over a volume a million, million times greater than the nucleus, be used to represent free electrons which were known to be localized to a volume at least as small as the nucleus?

Schrodinger initially thought he had the solution to this problem. It is a well known property of waves that two or more waves occupying the same region of space at the same time will interfere with each other. This interference can be destructive; that is, the waves can cancel each other out. Or the interference can be constructive and the waves will reinforce each other. If a large number of waves of slightly differing wavelengths interfere with each other, the resulting wave pattern will be confined to a very small region of space with complete destructive interference everywhere else. This localized wave is called a wave packet and has the properties of a particle. With the appropriate combination of waves, wave packets of the correct size and velocity can be constructed to represent free electrons.

The problem is, and this eventually led to the rejection of Schrodinger's interpretation, that these wave packets refuse to stay small. Even the smallest wave packet would, according to Schrodinger's own wave equation, very quickly spread out to occupy large volumes of space. Electrons clearly do not behave in this manner, and therefore, Ψ must be interpreted differently.

12.2 Max Born's Interpretation

In Schrodinger's interpretation, it was not the wave function itself but the square of the wave function, Ψ^2, that represented the electron wave. Specifically, Ψ^2 gave the density of the smeared out electron at all points in space and time. In general, Ψ is a complex function, consisting of both a real and an imaginary part. It can also have negative values. Only a real, non-negative function can represent a density. Multiplying Ψ by its complex conjugate, Ψ^*, always produces a real, non-negative function. This is technically represented as $\Psi^*\Psi$, but frequently abbreviated Ψ^2.

At about the same time Schrodinger was formulating his interpretation, Max Born at the University of Gottingen, one of the principle developers of matrix mechanics, was also working on an interpretation of the wave function. Born was strongly influenced by Einstein's interpretation of wave-particle duality as it applied to electromagnetic theory. For Einstein, the electromagnetic wave was a kind of 'phantom wave' that served to guide the photon. The square of the wave amplitude at any point in space determined the probability of finding a photon at that point. Carrying this idea over directly to matter waves, Born proposed that the square of a wave function representing an electron at a particular point in space gave the probability of finding the electron at that point. In other words, the electron is most likely to be found where Ψ^2 is large, less likely to be found where Ψ^2 is small, and will never be found where Ψ^2 is zero.

While Schrodinger tried to treat electrons only as waves, Born appears to be taking the exact opposite point of view. It appears that Born sees only the particles as being real while the wave associated with them are only 'phantom waves,' mathematical functions that determine the probability of finding the particle at a particular point in space and time. Although the Born interpretation is extremely useful in accounting for a wide range of experimental results in atomic physics, it cannot be the complete story.

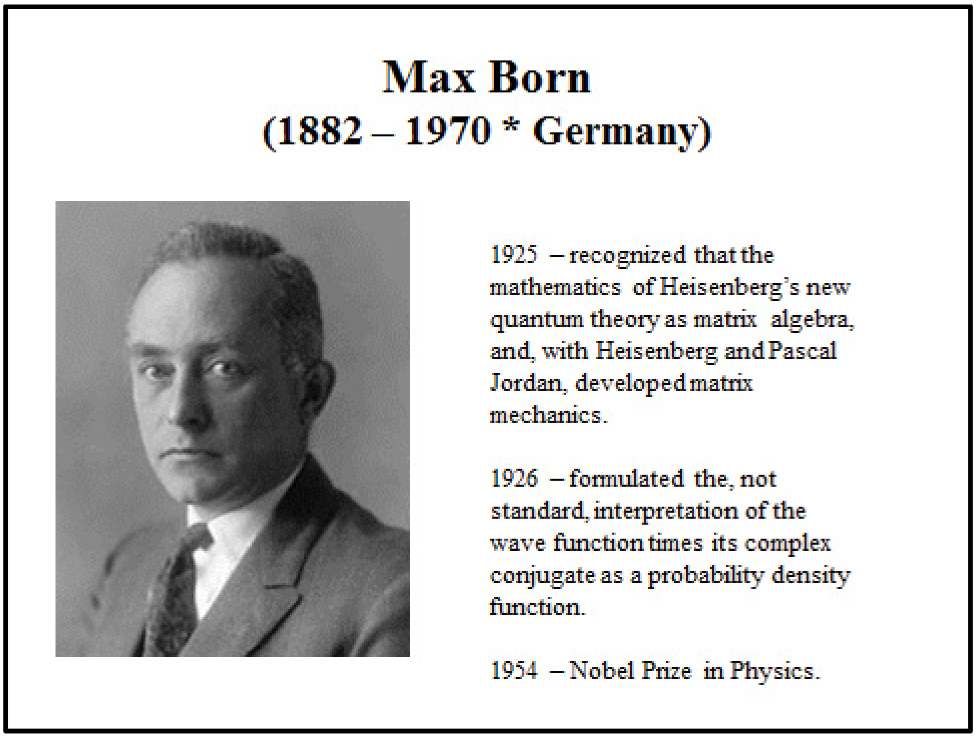


Figure 12.1 Max Born

12.3 The Double-Slit Experiment

In 1801, Thomas Young was able to offer very strong evidence that light was a wave using a double-slit experiment. An opaque object with two small, close-together slits is placed between a source of monochromatic (single wavelength) light and a screen. The pattern of light that reaches the screen is determined by whether light is a wave or a stream of particles. The results were just as predicted by the wave model. The two light beams, one from each slit, interfered with each other producing an alternating pattern of bright and dark lines. When the source of light is replaced with a source of electrons, the results are exactly the same.

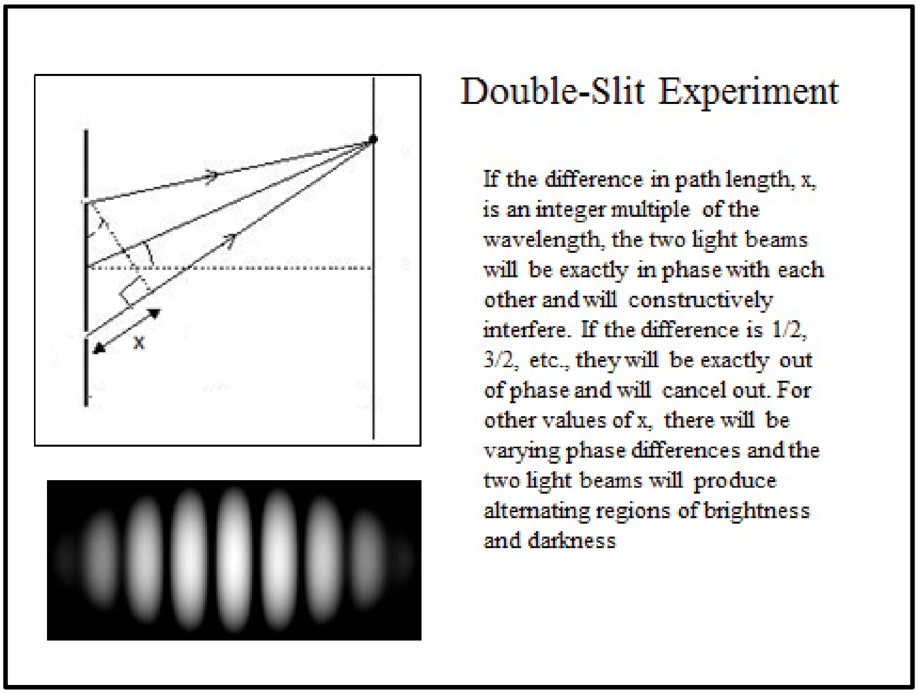


Figure 12.2 Double-Slit Experiment

Now consider a variation of the experiment. The double slits are in an object that is opaque to electrons, so the electrons must pass through one or the other of the slits. The intensity of the electron source is reduced to the point that only one electron reaches the screen at a time. In this way the distribution of electrons hitting the screen can be determined. First, the experiment is done with one of the slits closed, a single-slit experiment. When an individual electron hits the screen, a flash of light is emitted at a single point on the screen, thus showing the particle-like property of the electron. However, the point is not necessarily on a straight line from the source through the slit as the particle model would predict. Later, another electron strikes the screen after passing through the same slit, but strikes it at a different point. Finally, after an extremely large number of electrons have had time to pass through the slits, it is clear that they are striking the screen in a definite pattern, and that the pattern is exactly the one predicted by the wave model. The electron interacts with the screen as a particle would; it produces a single localized flash of light. However, the path it takes to get to the screen seems to be determined by the wave function.

Stranger yet are the results of experiments using both slits. One might suspect that, since each electron has to pass through one slit or the other, the pattern after a large number of electrons have had time to strike the screen, would be a combination of two individual one-slit patterns. If the slits are opened one at a time, even if they are alternated back and forth very quickly, this expected pattern is exactly the result. However, if both slits are open at the same time, a completely different pattern results – the interference pattern predicted by the wave model.

Even though the electrons are shot through one at a time, the accumulated pattern of electrons striking the screen depends on whether the two slits are open simultaneously or alternately. It is as if the electron passing through one of the slits 'knows' whether the other slit is open or closed, and this influences the direction in which it leaves the slit. The presence or absence of the other slit, the one the electron did not pass through, can influence its motion. Since the electrons are shot out one at a time, it is not possible that this is a collective effect with electrons passing through different slits interfering with or otherwise influencing each other. It is as if the single electron is passing through both slits simultaneously! Thus, the wave function is more than just an abstract mathematical function. It must be something physically real. The wave function is not quite a thing, but it is more than just an idea.

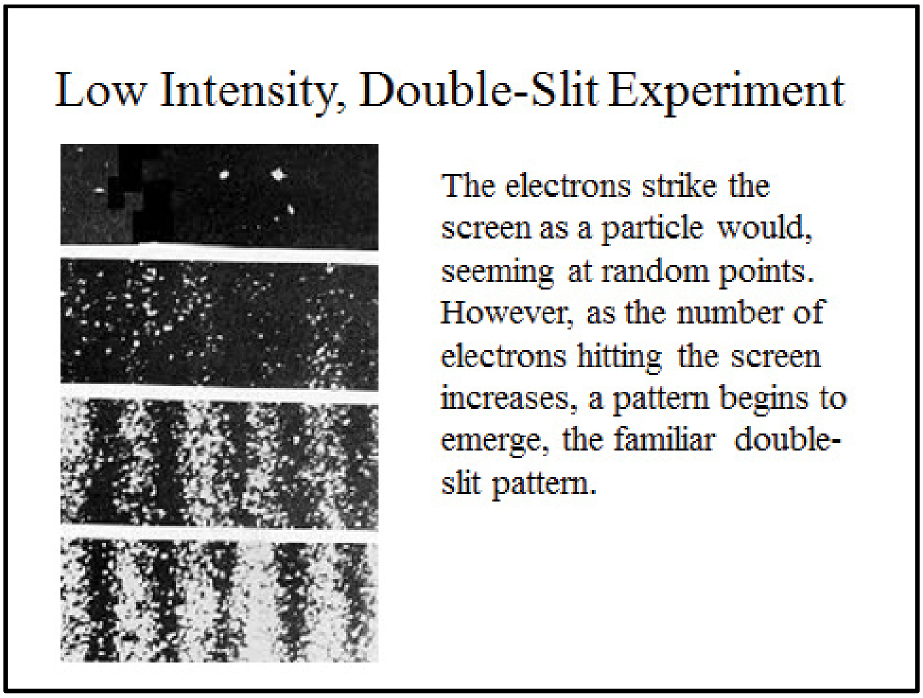


Figure 12.3 Low Intensity

These experiments can also be carried out using electromagnetic radiation with the intensity of the beam reduced so that only a single photon strikes the screen at a time. The results using photons are absolutely identical to those using electrons. The concept of wave-particle duality on the microscopic scale appears to be a necessary component of our understanding to the physical universe in spite of our inherent dislike of it.

12.4 Indeterminacy

Newtonian mechanics is a deterministic physical theory. Given a set of initial conditions, the subsequent development of the system can be determined with certainty. In its extreme form, this deterministic philosophy led the mathematical physicist Pierre Laplace to assert that if some super intelligence could know the exact state of the universe at some particular time, the entire future and past could be calculated. Of course, he knew that this was not possible in practice, but was confident that it was possible in principle.

With Born's interpretation of Ψ, probability was introduced into physics. The initial conditions of the system completely determine the wave function, and the wave function's behavior in time is exactly determined by the Schrodinger's wave equation. This part of the physics is deterministic. However, the result of observations made on the system is not. For example, suppose the wave function is that of an electron confined to a certain region of space. The wave function has nonzero amplitude within the region of confinement and zero amplitude everywhere else. The probability of detecting the electron at a particular point in space at a particular time is given by Ψ^2 at that time. Where the wave function has its largest amplitude is where the electron is most likely to be found, but it may be found somewhere else, just with a smaller probability. It is the detection of the electron, that constituted the observably reality, and this has only a probabilistic connection to the wave function.

As this consequence of quantum mechanics became increasing clear, the opponents, in an attempt to salvage classical realism and space-time continuity, concluded that quantum mechanics was not a complete theory. It was not wrong (in fact every prediction made turned out to be correct) it was just incomplete. That is, there are variables within the system that completely determined where the electron would be at any particular time, but that these variables do not show up in the wave function. They are simply not included in the formalism of quantum mechanics. The dissenters claimed the system is deterministic even though the wave function itself does not provide deterministic information. Thus, quantum mechanics is incomplete. These variables became known as 'hidden variables.' In the 1930s Einstein and two colleagues proposed a clever thought experiment that appeared to show that hidden variables must exist. This experiment is known as the Einstein-Podolsky-Rosen experiment, and is discussed in some detail in a later chapter. In the 1980s, a version of this experiment was performed. The results were very clear – hidden variables do not exist.

It is now established beyond any reasonable doubt that quantum mechanics is a complete theory, and that the information resulting from its application constitute the most we can know about a microscopic system. On the microscopic scale, it is a nondeterministic theory. However, when applied to a macroscopic situation, with it innumerable individual microscopic components, quantum mechanics reduces to a deterministic theory in essentially the classical sense. The behavior of a baseball is completely deterministic, and follows classical laws.

13 Werner Heisenberg and the Uncertainty Principle

Your goals for this chapter are to know the following

- The principle debaters and their contributions to the development of quantum mechanics.
- How the necessary uncertainty in in momentum and position cam still be compatible with the concept of an orbit.
- How to solve problems involving the uncertainty principle.
- The positions of Bohr and Einstein during their famous debates.

Niels Bohr established the Institute for Theoretical Physics in Copenhagen in 1921. There he gathered around him some of the most brilliant young physicists in the world. The atmosphere at the Institute was intense and exciting with constant, sometimes around the clock, debates on atomic physics and the meaning of the new mechanics. This unrelenting probing and questioning was quick to eliminate untenable ideas and served as a catalyst for some of the most significant contributions to 20th century physics.

13.1 Werner Heisenberg

Among the most intense debaters at Copenhagen were Werner Heisenberg and Niels Bohr himself. Toward the end of 1926 the physical interpretation of quantum mechanics dominated their discussions. (Discussions is perhaps too weak a work. Heisenberg was often reduced to tears by these 'discussions.') Bohr and Heisenberg were approaching the problem from different points of view, and as a result, a satisfactory conclusion was seldom reached. The effect was to completely exhaust each of them both mentally and physically. In order to recuperate, Bohr, in February 1927, went skiing in Norway, with Heisenberg glad to be left to his own thoughts.

Heisenberg had come to realize that just as the theory of relativity required a fundamental change in our concepts of space and time, quantum mechanics would require similar changes in our views of position and momentum (the mass of an object times its velocity). With Bohr in Norway, Heisenberg could now concentrate all his efforts on this problem.

In quantum mechanics, contrary to classical physics, assignments of arbitrarily precise values of both position and momentum are incompatible with each other. The momentum of an object is represented in quantum mechanics by the wavelength of the wave function, $p = h/\lambda$. Thus, in order to have a precisely defined momentum, the wave function must have a single wavelength. But a wave with a single wavelength is infinitely long and therefore contains no information on the position of the object. On the other hand, the wave function of an object with a well defined position is represented by a wave packet which extends over a very small region of space. Wave packets are produced by the superposition of a very large range of wavelengths. In fact, a wave packet confined to a single location in space requires an infinite range of wavelengths, and therefore contains no information on the momentum of the particle. Thus, the concept of an orbit or a path for an elementary particle makes no sense in quantum mechanics because a path through space contains information about both the position and momentum of the particle.

This presented grave difficulties for both Bohr and Heisenberg because it was an experimental fact that the path of an electron could be seen in a cloud chamber. (When a charged particle such as an electron, passes through air, it ionizes some of the molecules along its path. In a cloud chamber, water vapor condenses along the ionized trail to make the path of the electron visible. This is similar to a vapor trail left by a jet plane.) During Bohr's absence it suddenly occurred to Heisenberg that if he determined quantitatively the theoretically allowed precision with which both position and momentum could be measured, he could see if this was consistent with the path of an electron through a cloud chamber. That is, perhaps the uncertainties required by quantum mechanics were not of sufficient scale to significantly affect the macroscopic vapor trail representing the electron's path. A brief calculation showed Heisenberg that this was indeed the case. The relationship he derived, now known as Heisenberg's uncertainty principle, is

$$\Delta p \, \Delta x \geq h/4\pi$$

where Δp represents the uncertainty in momentum and Δx the uncertainty in position. The equation expresses the theoretical result that the product of the uncertainties in momentum and position is at least as great as Planck's constant divided by 4π. Thus, the greater the precision of, say, the momentum measurement, the greater the uncertainty in the location of the particle and vice versa. As in the example discussed above, if the uncertainty in either is zero, the uncertainty in the other is infinite.

Because of the very small numerical value of Planck's constant, it is possible for the relationship to be satisfied and still have a well-defined macroscopic path, such as the trail of water vapor in a cloud chamber or the trajectory of a baseball.

Example calculation: According to the uncertainty principle, what would be the minimum uncertainty in the speed to a 0.2 kilogram baseball if its position was measured to an accuracy of 0.1 millimeter?

$\Delta p \; \Delta x \geq h/4\pi$

$m\Delta v \; \Delta x \geq h/4\pi$

(0.2 kilogram) Δv (1×10^{-3} meters) $\geq$ (6.62×10^{-34} joule-seconds)/ 4π

A joule-second is equivalent to a kilogram-meter^2 per second.

Δv e 2.63×10^{-31} meters per second, clearly insignificant.

Suppose on the other hand, an electron is measured to be in a region the size of the nucleus, not an unusual measurement. What would then be the minimum uncertainty in its velocity?

$\Delta p \; \Delta x \geq h/4\pi$

(9.11×10^{-31} kilogram) Δv (1×10^{-14} meters) $\geq$ (6.62×10^{-34} joule-seconds)/ 4π

$\Delta v \geq 5.78 \times 10^{8}$ meters per second, or almost twice the speed of light.

If however the uncertainty in the electron's location is only about 1 millimeter, as is the case with the cloud chamber, the uncertainty in its velocity would be reduced to 5.78×10^{-3} meters per second, a negligible uncertainty.

At this point the uncertainty principle was a completely theoretical result, and Heisenberg was concerned as to whether or not it would be borne out by experiment. But the more he thought about how the actual measurements might be made, the more he was convinced of the validity of the relationship.

Any measurement is an interaction with the system and will necessarily disturb the system being measured. In macroscopic physics this disturbance is either completely negligible (as when photons bounce off a baseball to allow us to determine its position) or can be taken into account in a precise way (as when a thermometer is used to measure the temperature of a small quantity of liquid). Classically, it is possible to measure all physical quantities simultaneously with a precision limited only by our ingenuity. But on the microscopic scale it is different. A quantum mechanical treatment of measurement experiments involving elementary particles and photons implies that the observed object will be disturbed in an unpredictable way. The unpredictable nature of the disturbance is such that certain pairs of physical quantities cannot be simultaneously measured (measured by a single interaction) to an arbitrarily high degree of precision. Momentum and position are one such pair as indicated earlier. Energy and time are another such pair, as expressed in the following relationship:

$$\Delta E\ \Delta t \geq h/4\pi$$

This limit is not one that can be exceeded by devising better experimental equipment or more careful experimental techniques; it is a limit inherent in nature. It is not just impossible to exceed in practice, it is impossible to exceed in principle.

Heisenberg developed his uncertainty relationship while Bohr was away skiing in Norway. He was well aware of Bohr's intense scrutiny of new developments in quantum theory and was apprehensive about the criticism his ideas would face. Perhaps because of this, he submitted his paper before Bohr returned. When Bohr did return and read the paper, he interpreted Heisenberg's relationship differently, seeing it not in the simple physical terms of the paper, but as an example of a deeper philosophical concept. He asked that Heisenberg contact the journal and delay publication until they worked out the differences in their interpretations and how best to present this important result to the physics community. Heisenberg at first angrily refused. He had developed an elegant result in a straightforward way and now Bohr wanted him to present it in what Heisenberg considered philosophical gobbly-gook. Eventually Heisenberg agreed to add a footnote to the paper thanking Bohr for clarifications and suggesting that the source of the uncertainty was perhaps more complex than the author had implied.

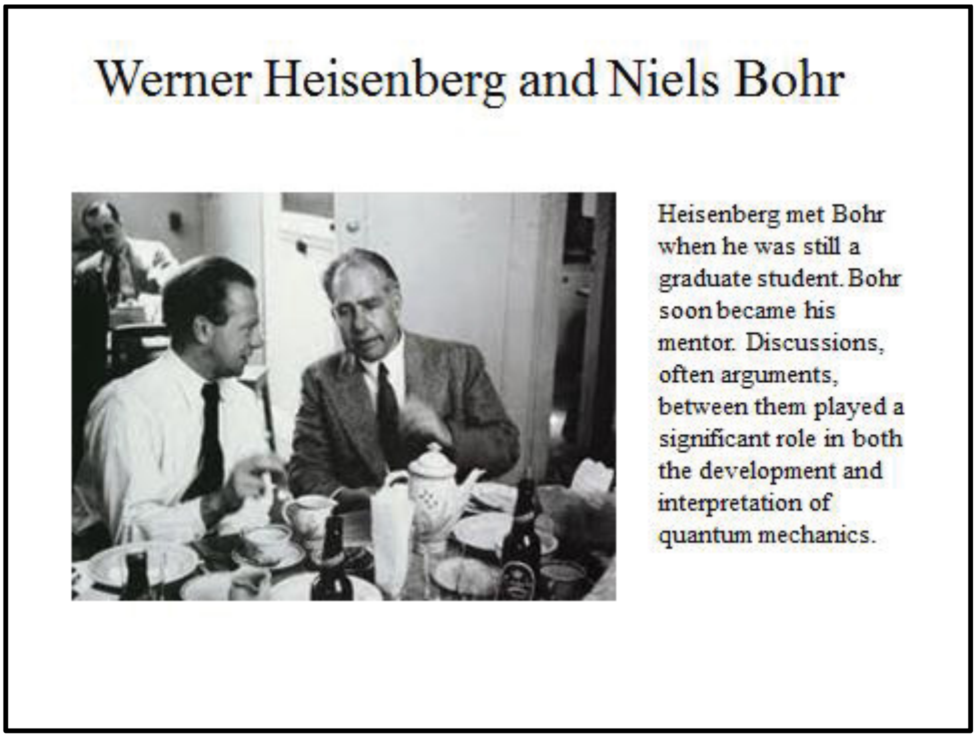


Figure 13.1 Heisenberg and Bohr

13.2 The 1927 Solvay Conference

Ernest Solvay was a Belgian amateur scientist who made a fortune from an industrial chemical process he developed. He became intrigued by the latest developments in physics, and in 1911, sponsored an invitation-only conference where twenty leading physicists came together to discuss and debate the new discoveries. This first conference included Einstein, Planck, Rutherford, and Mme Curie.

Because the conference was so well received, Solvay decided to repeat the conference every three years. The schedule was interrupted by the First World War, but resumed afterwards. The fifth conference was held in 1927. It was the first opportunity for leading physicists to weigh-in on quantum mechanics and the uncertainty principle in particular. At the conference, a striking division became apparent. On the one hand, there was the young Turks such as Heisenberg, Pauli and Dirac who were mostly concerned with applying the new physics and not much concerned with what it meant philosophically. On the other hand, there was the older, more conservative physicists such as Planck, de Broglie, and Schrodinger who were reluctant to throw out all of classical physics. Max Born, though somewhat older, sided with the Heisenberg group and was sharply critical of Schrodinger's presentation. Then there was Bohr and Einstein, perhaps the two deepest thinkers at the conference. They were passionately committed to understanding the meaning of this new way of looking at the world.

Einstein did not make a formal presentation at the meeting, saying he needed to study the matter in more detail. However, informally, over meals and late into the night, he questioned the advocates of quantum mechanics and expressed his own, mostly intuitive and philosophical, reservations. Bohr, unlike the Heisenberg group, took Einstein's concerns seriously and tried to answer his objections.

Einstein was particularly concerned about the probabilistic, non-realistic interpretation that seemed to be at the heart of quantum mechanics. To counter this idea, he proposed a thought experiment involving electrons passing through a small hole and hitting a screen. Their wave properties would cause a characteristic distribution on the screen, but each electron would hit the screen as an individual particle. Once it hit the screen, the wave function at other places on the screen would immediately disappear, implying instantaneous action-at-a-distance. Didn't a realistic interpretation, that there was in fact no probability that that particular electron would have hit somewhere else, make more sense? Bohr responded, but, as would be the case throughout their lives, Einstein remained unconvinced.

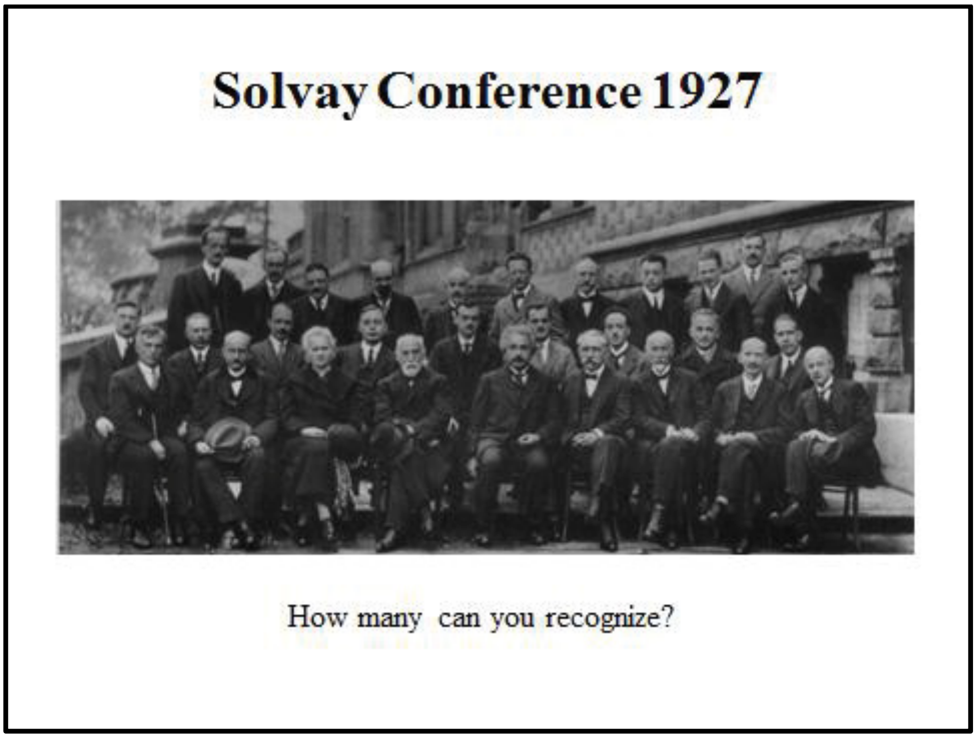


Figure 13.2 Solvay Conference 1927

13.3 The 1930 Solvay Conference

In 1930, the sixth conference was held. The official topic was magnetism, but outside the scheduled lectures the only significant topic was quantum mechanics. Although almost everyone had been won over by this time, there were some important hold-outs, Einstein, chief among them. In the 1927 Solvay conference, Einstein raised several questions which the younger physicists mostly ignored. Bohr, both out of personal respect for Einstein and to qualm some of his own philosophical worries, spent hours in discussion trying to persuade Einstein, but the discussions took place on a metaphysical level and neither could make their point of view clear to the other.

In 1930, Einstein came prepared with a specific quantitative demonstration that the uncertainty principle was wrong, and therefore quantum mechanics was as well. The experiment, of course, was a thought experiment, Einstein's traditional approach to getting to the heart of the physics. One version of the uncertainty principle stated that the more accurately you try to measure the energy of some quantum event, the less well you could know the time at which it occurred. Einstein proposed what he was convinced was a situation that violated this restriction. The experiment was very clever and Bohr was completely stumped by it.

The next morning Bohr was radiant. Overnight he had figured it out. Ironically Einstein had neglected one of the consequences of his own General Theory of Relativity, the effect that gravity has on the rate of time. Bohr showed, to his obvious delight, that a proper analysis based on relativistic physics produced exactly the Heisenberg uncertainty relationship. Einstein was forced to concede that Bohr's analysis was correct in this case, but remained unconvinced that the way Bohr and Heisenberg were describing physical reality was correct.

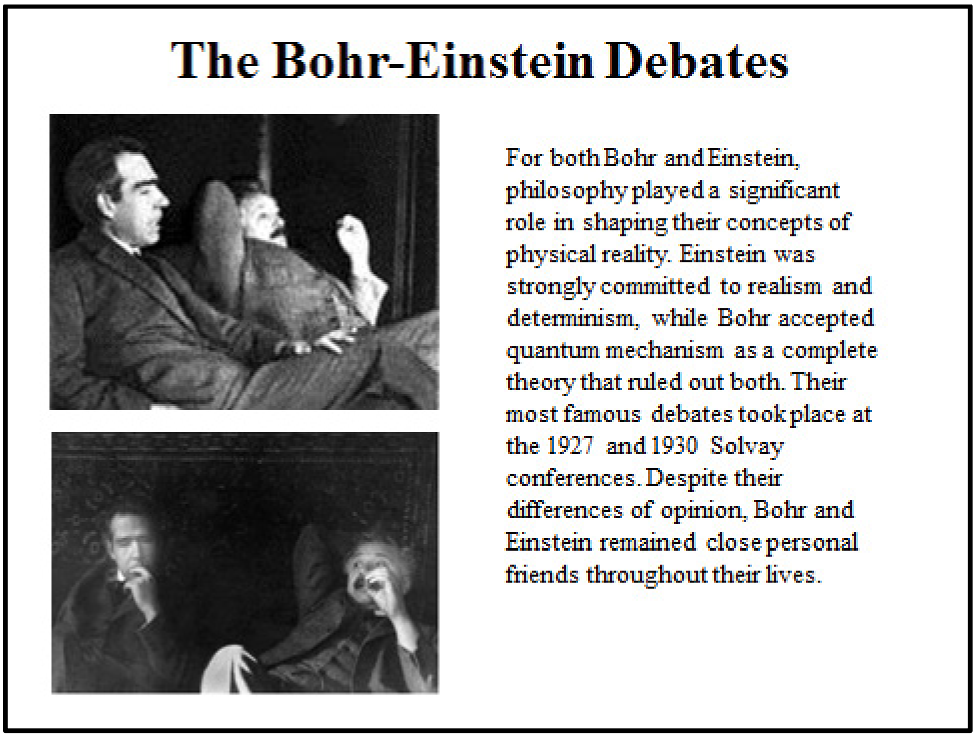


Figure 13.3 Bohr Einstein Debates

[People you might have recognized in the photograph: Back row, left to right: 6th Edwin Schrodinger, 8th Wolfgang Pauli, 9th Werner Heisenberg. Middle row, left to right: 5th P.A.M. Dirac, 7th Louis de Broglie, 8th Max Born, 9th Niels Bohr. Front row, left to right: 2nd Max Planck, 3rd Marie Curie, 5th Albert Einstein.]

14 Niels Bohr and the Copenhagen Interpretation

Your goals for this chapter are to know the following:

- What is meant by a quantum jump and why the majority of physicist, Schrodinger, for example, were opposed to it.
- The basic point of the correspondence principle.
- The basic point of complementarity.
- The basic principles of the Copenhagen interpretation of quantum mechanics and how they contrast with classical mechanics.

By 1927, physicists had given up hope of using the classical physics of Newton and Maxwell to account for empirical data on the atomic scale. By then it was also clear that calculations based on quantum mechanics corresponded, with an extremely high degree of accuracy, to the actual experimental measurements. What was very unclear was what quantum mechanics is telling us about the nature of physical reality. The person most obsessed with this question was Niels Bohr. From the time of his 1913 publication of the Bohr model of the hydrogen atom, through the developments of matrix mechanics and wave mechanics, it motivated intense discussions with the select few engaged in quantum physics.

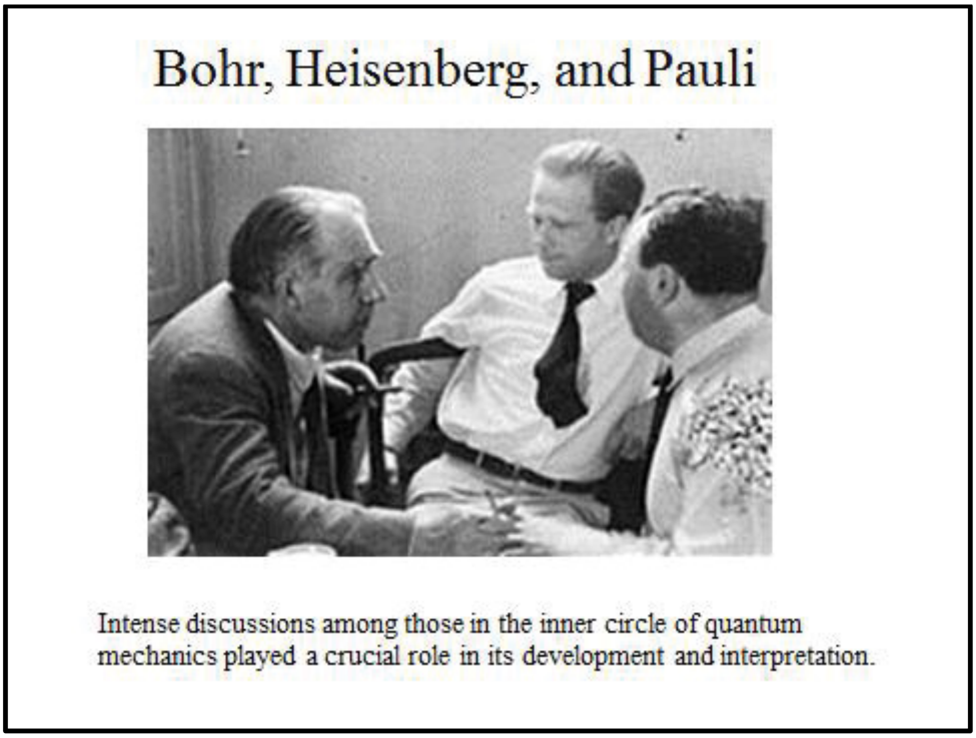


Figure 14.1 Bohr, Heisenberg, and Pauli

The famous Bohr-Schrodinger debate is another example of the intensity of the discussions in Copenhagen. Toward the end of the 1926 summer term, Schrodinger addressed the Munich physics seminar on wave mechanics. Heisenberg left Copenhagen primarily to hear this lecture and to represent the Copenhagen point of view in subsequent discussions. Schrodinger's talk was well received, particularly his physical interpretation of the wave function. His interpretation seemed to eliminate many of the features of matrix mechanics and the old quantum theory that the more traditional minded physicist found so objectionable. During the discussion Heisenberg attempted to point out several difficulties with Schrodinger's interpretation, but was not very convincing. In fact, he was taken to task by some of the more mature physicists present. Discouraged, Heisenberg wrote Bohr of the unhappy outcome. Probably as a result, Bohr invited Schrodinger to spend some time in September in Copenhagen.

The discussions began immediately upon Schrodinger's arrival and lasted each day from early morning until late at night. After a few days Schrodinger fell ill, possibly as a direct result of the intensity of the confrontation. However, the debate did not end. Bohr would sit on the edge of Schrodinger's bed constantly questioning, probing, and expounding his own point of view. Heisenberg was sometimes present during these discussions. Later he recalled hearing Schrodinger yell at Bohr, "If we are going to have to put up with these damn quantum jumps, I am sorry that I ever had anything to do with quantum theory." The visit ended with Schrodinger having given ground and Bohr more certain than ever that the Copenhagen group was on the right track. This debate played an important role in the development of the Copenhagen interpretation.

14.1 The Correspondence Principle and Complementarity

The correspondence principle was introduced by Bohr in 1920. In its simplest (though very incomplete) interpretation, it says that quantum calculations must reproduce the results of classical physics in those situations in which classical physics is known to apply. That is, there cannot be one theory, one set of rules, for the atomic realm and another for the macro-world. The laws of quantum physics must naturally reduce to those of classical physics as the scale of the application increases.

Bohr also understood that any description of a quantum object's properties or behavior must be expressed in the language of classical physics. The outcome of any experiment is a concrete datum not a statistical list of possibilities. Measurements are made using macroscopic objects obeying classical laws and the results of measurements on quantum objects are expressed in terms exactly like those of measurements on macroscopic objects; that is in terms of location, momentum, energy, and the like.

Although Bohr was never fully successful in explaining to others, even to close associates such as Heisenberg, exactly what he meant by the correspondence principle, it is clear that Bohr relied heavily on it in fleshing out his ideas about the atom. A textbook published in the early 1920s said that the correspondence principle "cannot be expressed in exact quantitative laws [but] in Bohr's hands it has been extraordinarily fruitful." The physicist Emilio Segre later said that the correspondence principle amounted to saying, "Bohr would have proceeded in this way."

Another philosophical concept that Bohr relied on heavily, and likewise could never explain exactly what he meant by it, is complementarity. In a letter to Schrodinger, Einstein refers to the concept and confesses, "Bohr's principle of complementarity, the sharp formulation of which, however, I have been unable to achieve despite much effort which I have expended on it." To again greatly simplify a very profound concept, Bohr saw the wave behavior and the particle behavior not as mutually exclusive but as complementary to one another. They are, of course, contradictory, but each is equally necessary to the understanding of atomic phenomena. Bohr's first reaction to Heisenberg's uncertainty principle was to view it as an example of complementary, provoking an intense argument between them.

Complementarity became somewhat of an obsession with Bohr. In later years, he often spoke to scientists other than physicists – biologists, sociologists, anthropologists, psychologists – stressing the importance of complementarity in their particular fields and giving examples where it might be applicable. In psychology, for example, he pointed out that we are creatures of both reason and emotion, a dualism he saw as not unlike that of waves and particles.

14.2 The Copenhagen Interpretation

Finally, in September of 1927 at a conference in Como, Italy, Bohr formally presented what is now known as the Copenhagen Interpretation. Bohr was a disaster as a public speaker. He generally spoke in a low tone, made worse by a strong accent, and would often switch from language to language. The main problem, however, was the fact that his sentence structure was convoluted. Some claimed that listening to Bohr was a lot like reading James Joyce's *Ulysses*, more a stream of consciousness than a coherent, structured presentation. The wife of a visiting professor at Bohr's institute, after listening to a welcoming lecture that ended to enthusiastic applause, turned to her neighbor and said she looked forward to hearing the English translation. He looked at her puzzled and said, "That was the English translation."

Como was no exception. His tortured and tortuous remarks left most of the audience utterly baffled. Others thought that Bohr had taken some well-known physics and clothed it in mysterious philosophical language. It is not clear that anyone outside Bohr's inner circle realized the significance of what Bohr was trying to say. Both Max Born and Heisenberg stood to say that they agreed with Bohr, this, only months after Heisenberg's fierce, tense standoff with Bohr over many of these same ideas. Despite much internal disagreement, the Bohr camp presented a united front at the meeting. Some later saw this as some sort of conspiracy to stifle criticism.

The Copenhagen Interpretation, like the correspondence principle and complementarity before it, is in many ways vague and has sometimes been interpreted differently by different proponents. However, there are several basic principles that are generally accepted as being part of the interpretation.

1. A system is completely described by its wave function. There are no hidden variables. Any uncertainty that exists within the wave function is true uncertainty, not just a representation of our lack of knowledge about the system.
2. If the wave function is a mixed state with respect to a particular variable (that is if several different values of the variable are possible results of a measurement of the variable), the square of the wave function gives the probabilities of obtaining each of the possible results.
3. The result of a measurement on a mixed state is completely random. There is no factor within either the wave function or the measurement device that will determine which of the allowed values will result, only the probability with which they might result. Quantum mechanics is a non-deterministic theory.
4. When a variable is measured, one of the possibilities is actualized. The wave function immediately collapses to one that is a pure state for the value that results from the measurement. All other possibilities cease to exist.
5. Measurements are not passive determinations of an objective world but active interactions with the thing measured. The way in which one chooses to measure the system becomes a part of the system and inseparably influences the outcome.

The Copenhagen Interpretation is based on the assumption that nothing is real until it is measured. For example, an electron's position becomes real only when it is measured; its momentum is real only when it is measured. Heisenberg's uncertainty principle does not permit the simultaneous measurement of the location and momentum of a particle, and therefore it makes no sense to think about a particle having simultaneous values for these variables. The orbit or trajectory of a microscopic particle is a meaningless concept. To ask, "Where is the particle now or how fast is it moving?" does not make any sense in the context of quantum mechanics.

Consider the wave function associated with the hydrogen atom. Each of the allowed energy states has its own wave function, which represents, among other things, information on the location of the electron with respect to the nucleus. When a measurement is made, the location is determined. Before the measurement, it may have been equally likely to be found on the opposite side of the atom. Does that mean that since it was found on one side that it could not have been on the opposite side? In the Copenhagen Interpretation the answer is no. The reason is that the measurement is not a passive determination of the location of the electron, as it would be in classical physics. In classical physics, the electron would have been at the location it was determined to be at whether or not a measurement is made. In the Copenhagen Interpretation, the measurement created the location of the electron and could just as likely have created it on the opposite side. The electron has no location until it is measured.

Although the Copenhagen Interpretation is still the most widely accepted interpretation of quantum mechanics, there are several alternative interpretations. For example, in 1957, Hugh Everett first proposed a many-worlds interpretation which asserts the objective reality of the wave function, but denies the collapse of the wave function as a result of measurement. It contends that all of the possibilities inherent in the wave function actually do occur, but in the sense of a many-branched tree rather than in a single unfolding universal history. The claim is that this reconciles the fully deterministic equations of quantum physics with the apparent non-deterministic observations. Each possibility actually does occur, just in branching universes.

More recently, another alternative to the Copenhagen Interpretation has arisen. The Copenhagen Interpretation describes what happens when an observer makes a measurement, but the observer and the act of measurement are themselves treated classically. But actually, physicists and their apparatus must be governed by the same quantum mechanical rules that govern everything else in the universe. This problem becomes particularly acute in the field of quantum cosmology, where the quantum system is the universe as a whole. Modern developments show, however, that the emergence of classical properties can be understood within the framework of quantum theory itself, eliminating the half-quantum half-classical character of the Copenhagen Interpretation.

Most physicists today consider the question of interpretation as a philosophical rather than a scientific one and simply use quantum mechanics as a tool in their scientific endeavors without worrying about it one way or another. As David Mermin expressed it, "Shut up and calculate."

15 The Einstein-Podolsky-Rosen Challenge and Bell's Inequality

Your goals for this chapter are to know the following:

- Einstein's fundamental objections to quantum mechanics.
- Why he thought the EPR thought experiment could produce results contrary to the predictions of quantum mechanics.
- The experimental set-up for testing Bell's inequality.
- The results of the testing of Bell's inequality.

Twentieth century physics destroyed many of our commonsense concepts of physical reality. Albert Einstein was one of the prime executioners in this assault, shattering classical notions of space, time, and gravity. However, as we have seen, when it came to the small-scale universe of the atom, he was unwilling to accept the loss of objective, visualizible reality that quantum mechanics seemed to require. He opposed the theory from its inception, frequently by way of intimate discussions with Bohr. At the 1927 and 1930 Solvay conferences, Einstein presented thought experiments to Bohr that he believed contradicted key aspects of the theory. However, Bohr was able in both cases to satisfactorily repute Einstein's objections. Einstein was defeated but not convinced.

15.1 The Einstein-Podolsky-Rosen Experiment

Einstein was sure that the properties of an object exist independently of whether or not they are observed. This point of view is called realism. It requires that if an electron is detected at a particular point in space at a particular time, it was very near that point an instant earlier and would have existed at that point at that particular time even if there had been no measurement. He also was certain that events which are isolated from one another in space cannot affect one another. This is known as separability. Two events are truly isolated if the time interval between measurements on the two is less than the time required for a signal traveling at the speed of light to pass between them. This is a consequence of relativity and is known as Einstein separability.

Einstein's last and most famous attempt to disprove quantum mechanics came in 1935. He and two younger colleagues, Boris Podolsky and Nathan Rosen, proposed a thought experiment that they believed clearly demonstrated the existence of objective reality at the subatomic level. The title of their paper was, Can Quantum-Mechanical Description of Reality Be Considered Complete? Their proof hinged on the assumption that measurements made on one particle could not in any way influence the properties of another particle widely separated in space. That is, their proof depended on the separability assumption. The thought experiment they proposed became known as the EPR experiment, named for the three authors of the paper.

According to quantum mechanics, each individual isolated system has its own wave function. If such a system, say a single isolated particle, can be transformed into two individual particles, doing so does not create two wave functions. Rather each particle shares a single wave function that evolves continuously and deterministically from the original single-particle wave function in accord with Schrodinger's wave equation. Remember there is no collapse of the wave function unless a measurement is made. Otherwise it evolves according to deterministic laws. In the language of quantum mechanics, the two separate particles sharing a single wave function are said to be entangled.

If we call the two particles A and B, measuring the location of A will collapse the wave function in such a way as to destroy all information about both A and B's momentum. However, knowing the location of A, the exact location of B can be determined. Conversely, measuring the momentum of A will collapse all information about location. But again, knowing the exact momentum of A, the exact momentum of B can be determined. Thus, without actually measuring anything about B, depending of what we arbitrarily choose to measure for A, either the exact location or the exactly momentum of B has been determined. The argument now concludes that since we can choose freely to measure either position or momentum, it "follows" that both must simultaneously be elements of reality. The EPR paper seems to have presented a way to establish the exact values of **either** the momentum **or** the position of B due to measurements made on particle A, without disturbing particle B in any way. (The either-or conclusion will be important in Bohr's criticism of the paper.)

This can be explained in one of two ways. Either the measurement made on A instantly determined the corresponding property of B, as quantum mechanics would imply, or, contrary to quantum mechanics, the information about the outcome of all possible measurements was already present in both particles at the time they separated. EPR used separability to conclude that the first possibility could not be true. If the two particles are separated by a very large distance at the time of the measurement, the measurement on A could not possibly affect the properties of B. If the second possibility were true, then the information on the outcome of either a location or a momentum measurement must have been encoded in some hidden variables not included in the quantum mechanical treatment. Einstein's criterion for a complete theory is that "every element of physical reality must have a counterpart in the physical theory." Because, according to the second explanation, this in not true for quantum mechanics, Einstein concluded that quantum mechanics is incomplete.

The EPR paradox challenges the prediction of quantum mechanics that it is impossible to know both the position and the momentum of a quantum particle. This challenge can be extended to other pairs of physical properties covered by the Heisenberg uncertainty principle such as energy and time. The EPR paper ends by saying: "While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible."

Bohr's response was published, five months later in the same journal and with exactly the same title. Bohr did not specifically attack separability. If he had, as in retrospect he should have, the EPR argument would have collapsed immediately. Instead, he focused on the "either-or' argument. In his paper he said, "the extent to which an unambiguous meaning can be attributed to such an expression as 'physical reality'… must be founded on a direct appeal to experiments and measurements." This is clearly a philosophical approach, typical of Bohr and typically somewhat obscure. Most physicists, having little patience for philosophy, did not pay much attention to the EPR debate.

Part of the reason they did not is that, unfortunately, the argument could not be settled empirically. The EPR experiment was not one that could actually be done; it involved a strictly hypothetical situation. In 1964, however, John Bell, an Irish physicists working at CERN, a high energy facility in Switzerland, was able to apply a version of the EPR argument to an experiment that, although very difficult to do, was at least in principle possible.

15.2 Bell's Inequality

Each atom and subatomic particle has a measurable property associated with it called spin angular momentum or usually just spin. This property is very similar, but not identical, to the spin of an ordinary object such as a top or the earth. If an object is rotating counterclockwise when viewed along a particular direction, it is said to have spin up, If it is rotating clockwise, it is said to have spin down. If you wrap the fingers of your right hand around the object in the direction of the spinning motion, your thumb will point in the direction of spin. Atoms and subatomic particles can also be said to have spin up or spin down.

Atoms and subatomic particles, unlike ordinary-sized objects, cannot have arbitrary quantities of spin nor can the angle of their axis of rotation with respect of a particular direction take on arbitrary values. On the microscopic scale, both the amount of spin and the orientation of the spin are quantized.

Particle spins are restricted to half-integer multiples of Planck's constant divided by 2π: 0 h/ 2π, ½ h/ 2π, 1 h/ 2π, 3/2 h/ 2π, etc. Atoms and subatomic particles with spin ½ h/ 2π, (usually called spin-½ particles) when measured can have only two orientations. These two orientations ard called spin up or spin down. Nothing in between is ever observed. For these particles, a measurement of the spin in any arbitrarily chosen direction is always two valued – either one-half up or one-half down. The 'in any arbitrarily chosen direction' is important. Even if the two directions are perpendicular, the measurement can only yield spin up or spin down. It is this type of particle that will be used in our example of Bell's inequality.

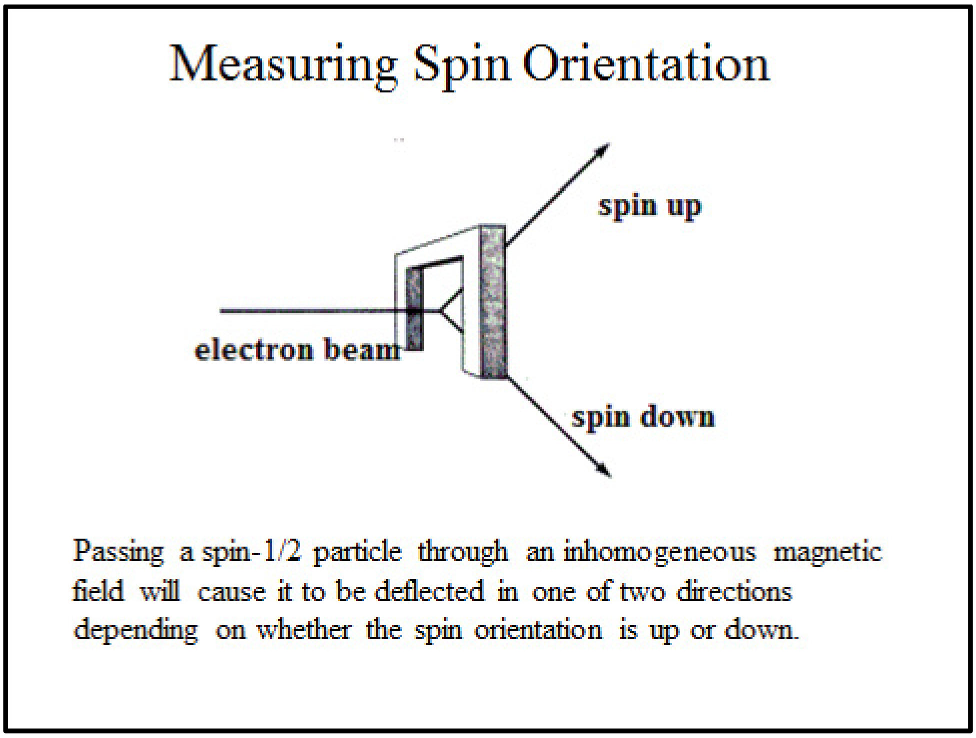


Figure 15.1 Measuring Spin Orientation

Measuring devices can measure the spin of a particle in any arbitrarily chosen direction, but no measuring device exists that can simultaneously measure the values of the spin in two different directions. Nor is a measurement of the spin in two different directions, one immediately after the other meaningful. The first measurement in a particular direction collapses the wave function in such a way that destroys information about spin in any other direction. This is completely analogous to a measurement of the location destroying information about the possibilities of other locations.

As with the EPR experiment, a single particle system splits into an entangled pair sharing a single wave function. If the spin of the original one-particle system is zero, the total spin of the two-particle system must also be zero. Spin, like energy, is a conserved physical quantity. If the two particles are of the two-valued, spin one-half variety, in order for the total spin to be zero, one must have spin up and the other spin down. In this way the two spins cancel out leaving a total spin of zero as required.

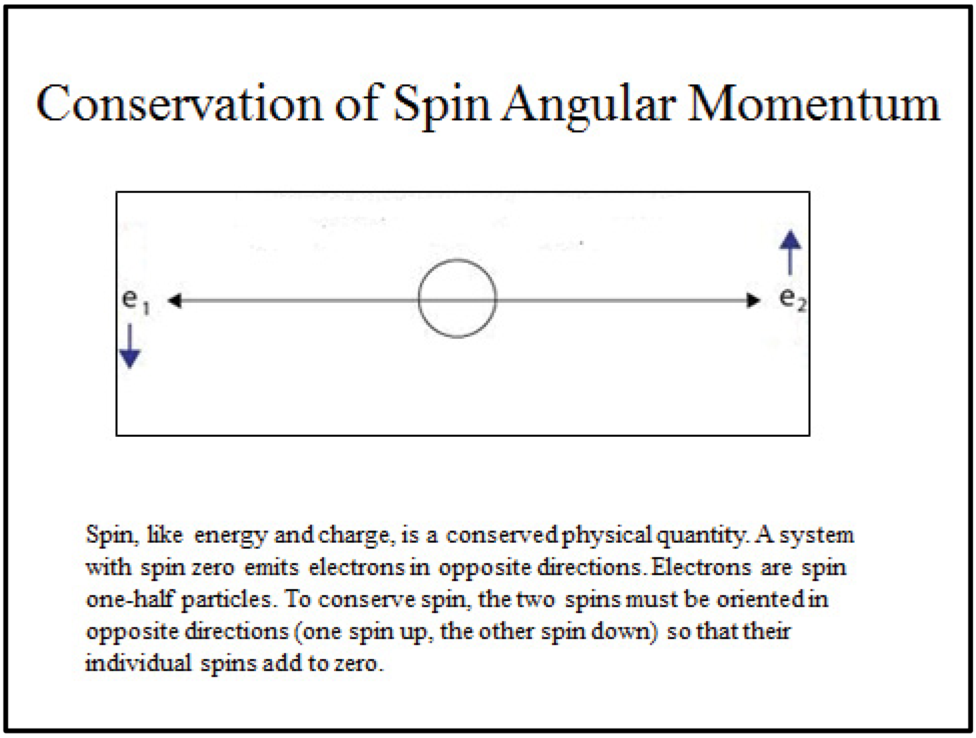
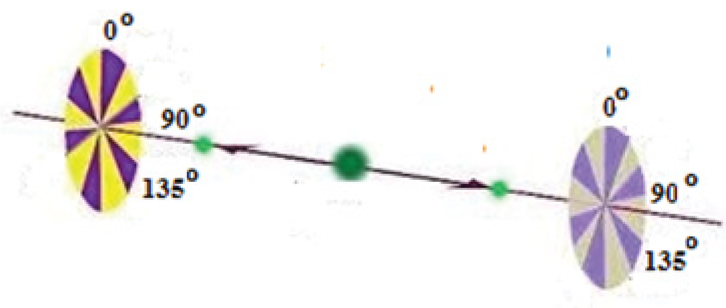


Figure 15.2 Conservation of Spin

As in the EPR example, one particle is A and the other is B. In our example, devices are set up to measure spin for each of the particles. Each device randomly and independently selects between three directions in which to measure the spin, at an angle of either 0°, 90°, or 135° with respect to some well specified direction. The result of each measurement is either spin up or spin down. The data recorded is the angle selected and the result of the measurement: 90° up, 135° down, 135° up, 0° down, etc. A tacit assumption of this experiment is that the choice of which orientation to measure, either by some random quantum event or by the free will of an observer, is not determined in any way by past events.

After a very large number of events have been measured, the two sets of data are brought together and compared. As expected, if for a single event, the angle randomly selected turned out to be the same, each device recorded opposite values for spin orientation, spin up for one of the particles and spin down for the other one. The crux of the experiment is to determine the probability with which certain pairs of measurements occur. That is, for example, how often is A measured to have spin up in the 90° direction while B is measured to have spin up in the 135° direction?

Bell's Inequality



A single particle system with spin zero emits two spin one-half particles in opposite directions. The two particles are described by a single wave function requiring their spins to be in opposite directions. Observers randomly choose which of the three orientations to make a measurement in. Local realism requires Bell's inequality must hold. Quantum mechanics predicts that it will be violated.

Figure 15.3 Bell's Inequality

Applying the assumptions of realism (each of the emitted particles has a well-defined spin from the moment of emission) and separability (the measurement of one of the particles can have no effect on a measurement of the other particle), the relative probabilities of certain combinations is easily determined. If, for example, we represent the number of events for which A is measured to have spin up in the 0° direction and B is measured to have spin down in the 135° direction as N_{AB} (0° up, 135° down), it is necessary that

$$N_{AB}\,(0^\circ \text{ up}, 135^\circ \text{ down}) + N_{AB}\,(135^\circ \text{ up}, 90^\circ \text{ up}) \geq N_{AB}\,(0^\circ \text{ up}, 90^\circ \text{ up}).$$

An equation based on the assumptions of realism and separability for this type of experiment is known as Bell's inequality. It is a relatively straightforward statement. If all possible combinations of measurements are equally likely, the left-hand side will be exactly twice the right-hand side. If the EPR argument is correct, the inequality must be reflected in the data. Because the inequality is based on probabilities, statistical fluctuations may cause it to be violated. However, the chance of this happening can be made arbitrarily small by using a large number of events.

In 1964 John Bell realized that there was an experiment involving atoms or subatomic particles that could be done to test this relationship. He also realized that by selecting certain orientations at which to measure the spin, 0°, 90°, or 135° being one such example, quantum mechanics predicts the opposite inequality.

$$N_{AB}\,(0^\circ \text{ up}, 135^\circ \text{ down}) + N_{AB}\,(135^\circ \text{ up}, 90^\circ \text{ up}) < N_{AB}\,(0^\circ \text{ up}, 90^\circ \text{ up}).$$

Thus, the EPR argument could be tested. If the first inequality corresponds to the data, EPR was correct and quantum mechanics is incomplete. It the second equation holds, then either one or both of EPR assumptions of realism or separability is wrong.[2]

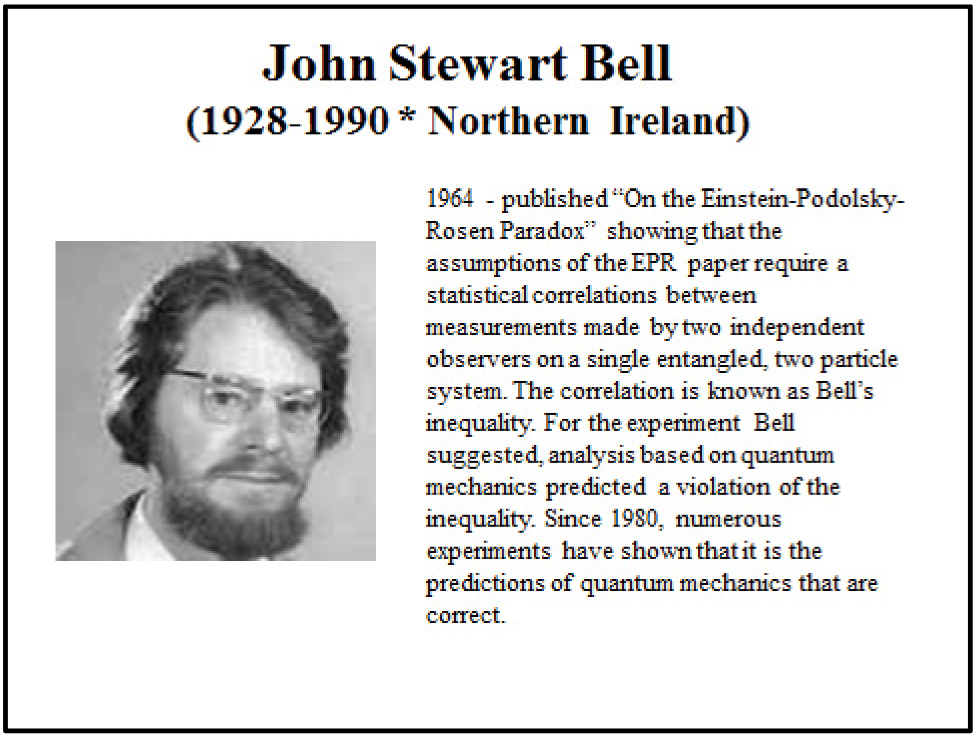


Figure 15.4 John Stewart Bell

The experiment, though possible in principle, is extremely difficult to perform and it wasn't until the 1980s that reliable results were obtained. But the results were clear. The predictions of quantum mechanics are confirmed. Bell's inequality is violated. Something about the assumption of realism and separability that went into its derivation is incorrect.

Trying to decide which of the two basic assumptions is the culprit, however, may not be the right approach. That might be the equivalent of trying to decide whether the assumption that an electron is a particle or the assumption that it is a wave is at fault, since it clearly cannot be both. In fact, there are certain situations where, regardless of what is assumed about realism, it appears that results obtained in one region depend of variables measured in a remote region. On the other hand, certain aspects of quantum behavior seem to demand renunciation of realism regardless of what is assumed about separability. Perhaps realism and separability are just extrapolations of our everyday experiences, which like the wave and particle models are useful, but not totally appropriate on the atomic level.

16 The Delayed-Choice Experiment and Schrodinger's Cat

Your goals for this chapter are to know the following:

- What is meant by a delayed-choice experiment.
- Schrodinger argued that until the box is opened, the wave function representing the cat is a superposition of "dead cat" and "alive cat" states and that wave function collapses to one or the other when an observer opens the box. The key to the paradox is, what counts as an observation. Describe present-day thinking on this.

Quantum mechanics has completely destroyed most of our commonsense concepts regarding the micro world. We simply cannot extrapolate our experiences in the macro world to the behavior of elementary particles and atoms. We know this not just from the theory of quantum mechanics, but from empirical data, the only reliable path to knowledge. Experiments such as the low-intensity, double-slit experiment and the experiments to test Bell's inequality are unequivocal. Perhaps the strangest of all consequences of quantum mechanics, however, is one proposed by John Wheeler in 1978. His delayed-choice experiment suggested that, on the micro scale, the past only comes into existence when it is recorded in the present. In his experiment, something that commonsense tells us has already happened can be determined by the free choice of an observer as to what will be observed. The act of observation creates the past.

16.1 The Delayed-Choice Experiment

Wheeler's experiment consists of a low-intensity, double-slit apparatus except that the screen can be removed at the last moment, allowing the photon to reach two more remote telescopes, each one focused on one of the slits. If the screen is removed, only one of the telescopes will record the photon, determining which of the two slits the photon passed through. In this case, the path of the photon will be determined, and the photon is behaving as if it were a particle. If the photon is allowed to hit the screen, however, there is no information on which slit the photon passed through, and the photon will behave as if it were a wave, eventually producing the interference pattern predicted by the wave model.

The key to the experiment is to delay the choice of whether or not to remove the screen until after the photon has already passed the location of the slits and is in the region between the slits and the screen. If the screen is left in place, an interference pattern will eventually develop showing that the photon must have passed through both slits. If the screen is removed, only one of the telescopes will detect it determining which one of the two slits the photon passed through. Because the choice is made after the photon is past the slits, the experiment appears to determine what happened (the photon passed either though one slit or both slits) after it has already happened. The past is being created by the free choice of which measuring device to use. Wheeler says, "The past only exists as it is recorded in the present – it is senseless to talk of what the particle was doing before it was observed."

In 2007, Jean-François Roch of the Ecole Normale Supérieure de Cachan in France and colleagues performed a version of Wheeler's experiment, and the results were exactly as he predicted; the past is not the past until it is the recorded past.

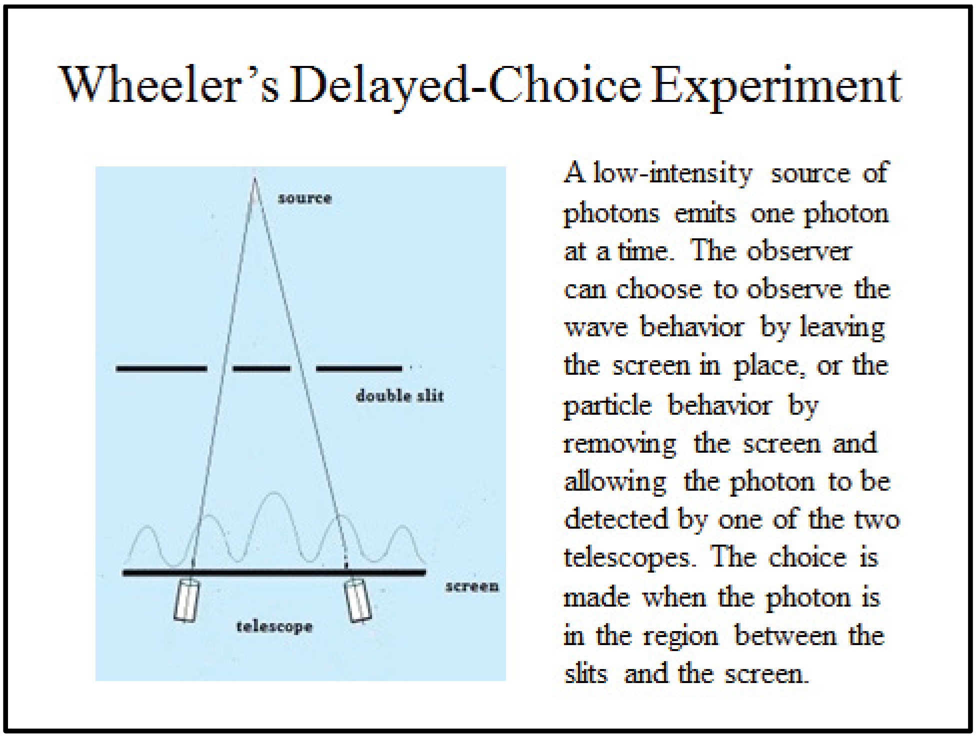


Figure 16.1 Delayed Choice

16.2 Schrodinger's Cat

Perhaps the most widely known thought experiment in quantum mechanics is Schrodinger's cat. It was designed by Schrodinger in 1935 to illustrate what he saw as the problem with the Copenhagen interpretation of quantum mechanics. It was intended as further discussion of the Einstein-Podolsky-Rosen paper of the same year.

The thought experiment imagines a closed box containing a cat, a small sample of radioactive material, a Geiger counter to register the decay, and a device that, when triggered by the Geiger counter, will release a poison gas and kill the cat. If the probability that one of the atoms will decay in, say, one hour is 50-50, the wave function of the unobserved sample after one hour will be a superposition of states, one with an atom decayed and one with no atoms decayed. In the Copenhagen interpretation, a measurement is required to collapse the wave function to one or the other of the two states. According to Schrödinger, this implies that the superposition of the states of the sample can be extended to two states of the cat, half dead and half alive and that the cat's wave function does not collapse until the box is opened and an observation made. To Schrodinger (and to most everyone else), this seemed absurd, and he intended the thought experiment to demonstrate that the Copenhagen interpretation is wrong.

This thought experiment has been continuously discussed since it was first proposed. It has even generated an extension known as 'Wigner's Friend,' proposed by the theoretical physicist Eugene Wigner. It posits a friend of Wigner who performs the Schrödinger's cat experiment after Wigner leaves the laboratory. Only when he returns does Wigner learn the result of the experiment from his friend, that is, whether the cat is alive or dead. The question is raised: was the state of the system a superposition of "dead cat/sad friend" and "live cat/happy friend," only determined when Wigner learned the result of the experiment? Is it only consciousness of an outcome that can cause the collapse?

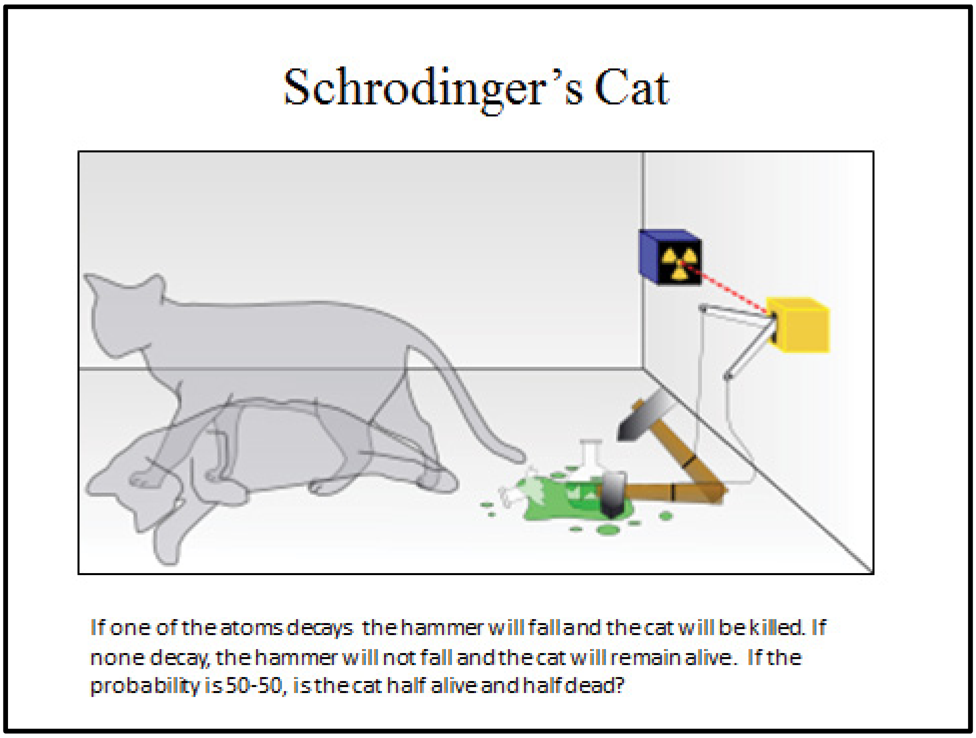


Figure 16.2 Schrodinger's Cat

These kinds of questions are not taken very seriously by most physicists for many reasons, one of which is that there is a somewhat commonsense interpretation of the Schrodinger's cat experiment. The explanation is detailed nicely and understandably in an article by Martin Gardner in the October 1982 issue of Discover magazine entitled 'Quantum Weirdness.'. He begins the article with a quote attributed to various major-league umpires:

> Some is balls and some is strikes, but until I calls 'em, they ain't nothin.'

The question discussed in the article is "What constitutes an observation?" For Wigner, the observation is completed only when a conscious mind becomes aware of the result. For Gardner and most physicists, it is when the micro scale event results in a macro scale event that is not time-reversible. For example, when a radioactive nucleus at rest emits an alpha particle, the emitted particle has a certain kinetic energy and the nucleus recoils with a certain kinetic energy, the process conserving both energy and momentum. This, however, is a time-reversible process. If a 'film' of the decay is run backwards, the reaction represented is a possible reaction. The process in either direction is allowed. Neither violates any law of physics. Micro scale events are in general time-reversible.

On the other hand, macro scale events are in general not time-reversible. A film of a Geiger counter registering the decay is not time-reversible. It is not possible for the counter to register the decay before it occurs. The absurd interpretation of Schrodinger's cat experiment is eliminated if any non-time-reversible process resulting from a time-reversible process constitutes a measurement and will collapse the wave function. The superposed wave function representing the radioactive sample collapses when the Geiger counter registers the event. Before this the cat is 100% alive and after the event it is 100% dead, independent of the box being opened and an observation of the cat being made. The observation has already been made by the Geiger counter.

This line of reasoning also applies to Wheeler's delayed choice experiment. The photon hitting the screen or being seen by one of the telescopes is a macro scale event. Until then the process is time-reversible. The non-time-reversible observation does not alter the past but rather brings the past into existence. The photon has no precise past until it is measured.

No physicist doubts that the micro scale swarms with quantum weirdness, but most physicists agree with physicist and author Heinz Pagels that only the micro world is weird. "Once information about the quantum world is irreversibly in the macroscopic world, we can safely attribute objective significance to it – it can't slip back into the quantum never-never land."

Endnotes

1. This is a law of physics in Maxwell's theory, but was assumed to be true only in the inertial frame at rest with respect to the ether. In any other inertial frame, it would be expected to have a different value; two observers moving with respect to one another should obviously measure different values for the speed of a light beam. Because this law is not the same in all inertial frames of reference, Maxwell's theory does not satisfy the principle of relativity. It was this predicted difference of the speed of light in different frames of reference that Michelson and Morley attempted, unsuccessfully, to measure in 1887.
2. The quantum mechanical probability that A will be measured to have spin up and B will be measured to have spin down is given by $0.25[1 + \cos(\Theta_A - \Theta_B)]$ and the probability that both will be measured up is $0.25[1 - \cos(\Theta_A - \Theta_B)]$. These equations give 0.073 for each term on the left and 0.250 for the term on the right, violating Bell's inequality. Notice that if $\Theta_A = \Theta_B$, the probability that both spins will be up is zero, consistent with the requirement that the total spin of the system is zero. The probability that A will be up and B down is 0.50, and the probability that A will be down and B up is also 0.50 giving a probability of one that the two spins are in opposite directions.