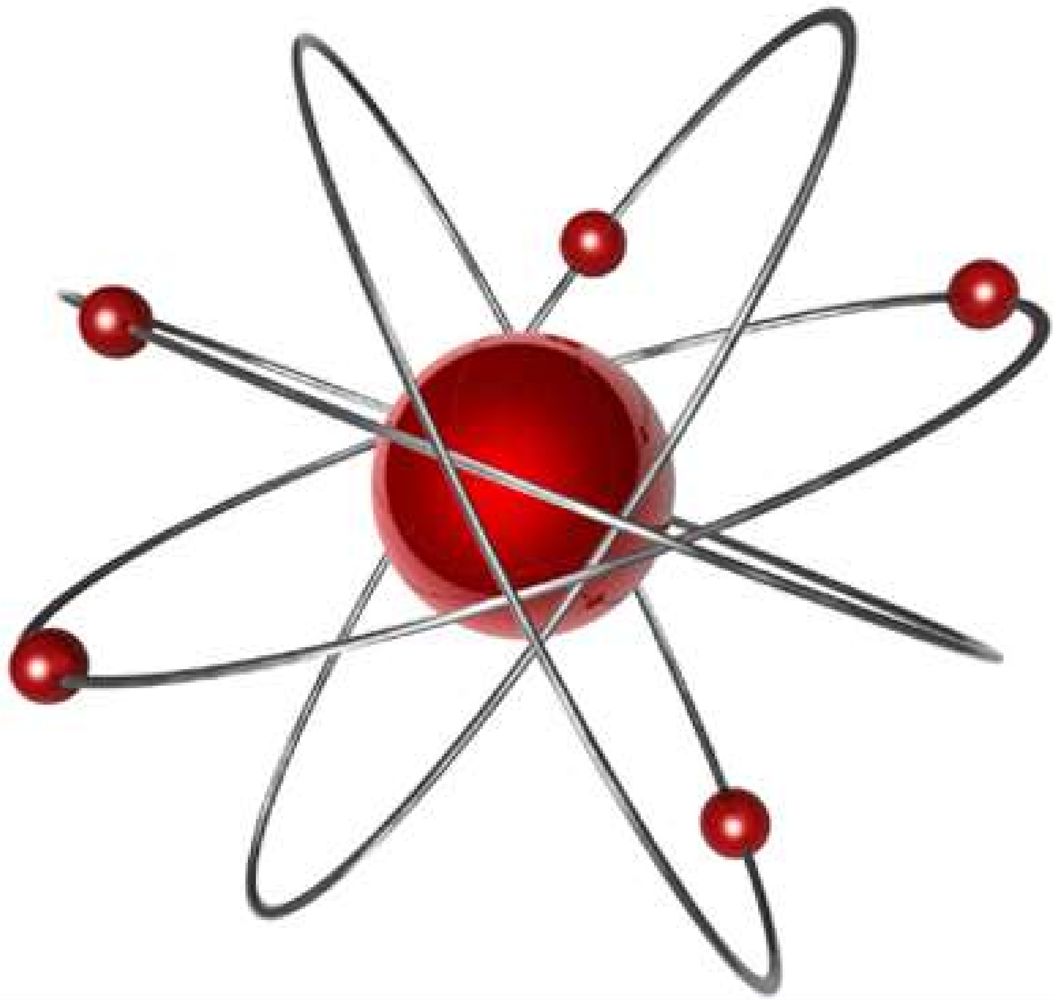
Quantum Electronics & Spintronics



Heath Salter

Kandice Buss

First Edition, 2012

ISBN 978-81-323-1359-5

Published by:
College Publishing House
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1

Quantum Mechanics

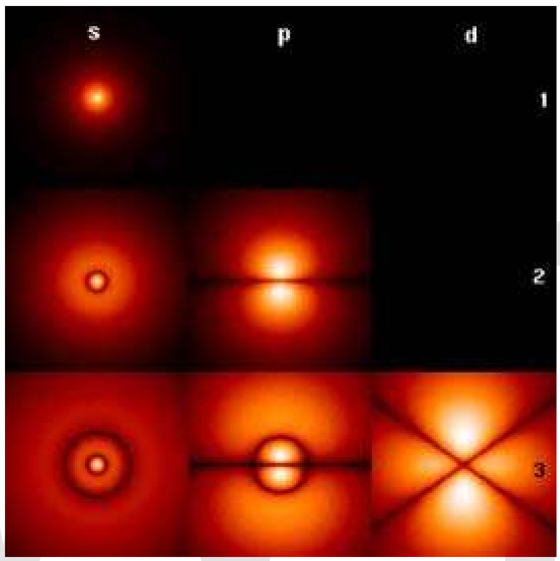


Fig. 1: Probability densities corresponding to the wavefunctions of an electron in a hydrogen atom possessing definite energy levels (increasing from the top of the image to the bottom: $n = 1, 2, 3, ...$) and angular momentum (increasing across from left to right: *s*, *p*, *d*, ...). Brighter areas correspond to higher probability density in a position measurement. Wavefunctions like these are directly comparable to Chladni's figures of acoustic modes of vibration in classical physics and are indeed modes of oscillation as well: they possess a sharp energy and thus a keen frequency. The angular momentum and energy are quantized, and only take on discrete values like those shown (as is the case for resonant frequencies in acoustics).

Quantum mechanics, also known as **quantum physics** or **quantum theory**, is a branch of physics providing a mathematical description of much of the dual particle-like and wave-like behavior and interactions of energy and matter. It departs from classical mechanics primarily at the atomic and subatomic scales, the so-called quantum realm. In advanced topics of quantum mechanics, some of these behaviors are macroscopic and only emerge at very low or very high energies or temperatures. The name, coined by Max Planck, derives from the observation that some physical quantities can be changed only by discrete amounts, or quanta, as multiples of the Planck constant, rather than being capable of varying continuously or by any arbitrary amount. For example, the angular momentum, or more generally the action, of an electron bound into an atom or molecule is quantized. While an unbound electron does not exhibit quantized energy levels, an electron bound in an atomic orbital has quantized values of angular momentum. In the context of quantum mechanics, the wave–particle duality of energy and matter and the uncertainty principle provide a unified view of the behavior of photons, electrons and other atomic-scale objects.

The mathematical formulations of quantum mechanics are abstract. Similarly, the implications are often non-intuitive in terms of classic physics. The centerpiece of the mathematical system is the wavefunction. The wavefunction is a mathematical function providing information about the probability amplitude of position and momentum of a particle. Mathematical manipulations of the wavefunction usually involve the bra-ket notation, which requires an understanding of complex numbers and linear functionals. The wavefunction treats the object as a quantum harmonic oscillator and the mathematics is akin to that of acoustic resonance. Many of the results of quantum mechanics do not have models that are easily visualized in terms of classical mechanics; for instance, the ground state in the quantum mechanical model is a non-zero energy state that is the lowest permitted energy state of a system, rather than a more traditional system that is thought of as simply being at rest with zero kinetic energy.

Historically, the earliest versions of quantum mechanics were formulated in the first decade of the 20th century at around the same time as the atomic theory and the corpuscular theory of light as updated by Einstein first came to be widely accepted as scientific fact; these latter theories can be viewed as quantum theories of matter and electromagnetic radiation. Quantum theory was significantly reformulated in the mid-1920s away from the old quantum theory towards the quantum mechanics formulated by Werner Heisenberg, Max Born, Wolfgang Pauli and their associates, accompanied by the acceptance of the Copenhagen interpretation of Niels Bohr. By 1930, quantum mechanics had been further unified and formalized by the work of Paul Dirac and John von Neumann, with a greater emphasis placed on measurement in quantum mechanics, the statistical nature of our knowledge of reality and philosophical speculation about the role of the observer. Quantum mechanics has since branched out into almost every aspect of 20th century physics and other disciplines such as quantum chemistry, quantum electronics, quantum optics and quantum information science. Much 19th century physics has been re-evaluated as the classical limit of quantum mechanics, and its more advanced developments in terms of quantum field theory, string theory, and speculative quantum gravity theories.

History

The history of quantum mechanics dates back to the 1838 discovery of cathode rays by Michael Faraday. This was followed by the 1859 statement of the black body radiation problem by Gustav Kirchhoff, the 1877 suggestion by Ludwig Boltzmann that the energy states of a physical system can be discrete, and the 1900 quantum hypothesis of Max Planck. Planck's hypothesis that energy is radiated and absorbed in discrete "quanta", or "energy elements", enabled the correct derivation of the observed patterns of black body radiation. According to Planck, each energy element *E* is proportional to its frequency *ν*:

$$E = h\nu$$

where *h* is Planck's action constant. Planck cautiously insisted that this was simply an aspect of the processes of absorption and emission of radiation and had nothing to do with the physical reality of the radiation itself. However, in 1905 Albert Einstein interpreted Planck's quantum hypothesis realistically and used it to explain the photoelectric effect, in which shining light on certain materials can eject electrons from the material. Einstein postulated that light itself consists of individual quanta of energy, later called photons.

The foundations of quantum mechanics were established during the first half of the twentieth century by Niels Bohr, Werner Heisenberg, Max Planck, Louis de Broglie, Albert Einstein, Erwin Schrödinger, Max Born, John von Neumann, Paul Dirac, Wolfgang Pauli, David Hilbert, and others. In the mid-1920s, developments in quantum mechanics quickly led to its becoming the standard formulation for atomic physics. In the summer of 1925, Bohr and Heisenberg published results that closed the "Old Quantum Theory". Out of deference to their dual state as particles, light quanta came to be called photons (1926). From Einstein's simple postulation was born a flurry of debating, theorizing and testing. Thus, the entire field of quantum physics emerged leading to its wider acceptance at the Fifth Solvay Conference in 1927.

The other exemplar that led to quantum mechanics was the study of electromagnetic waves such as light. When it was found in 1900 by Max Planck that the energy of waves could be described as consisting of small packets or quanta, Albert Einstein further developed this idea to show that an electromagnetic wave such as light could be described by a particle called the photon with a discrete energy dependent on its frequency. This led to a theory of unity between subatomic particles and electromagnetic waves called wave–particle duality in which particles and waves were neither one nor the other, but had certain properties of both. While quantum mechanics describes the world of the very small, it also is needed to explain certain macroscopic quantum systems such as superconductors and superfluids.

The word *quantum* derives from Latin meaning "how great" or "how much". In quantum mechanics, it refers to a discrete unit that quantum theory assigns to certain physical quantities, such as the energy of an atom at rest (see Figure 1). The discovery that particles are discrete packets of energy with wave-like properties led to the branch of physics that deals with atomic and subatomic systems which is today called quantum mechanics. It is the underlying mathematical framework of many fields of physics and chemistry, including condensed matter physics, solid-state physics, atomic physics, molecular physics, computational physics, computational chemistry, quantum chemistry, particle physics, nuclear chemistry, and nuclear physics. Some fundamental aspects of the theory are still actively studied. Quantum mechanics is essential to understand the behavior of systems at atomic length scales and smaller. For example, if classical mechanics governed the workings of an atom, electrons would rapidly travel towards and collide with the nucleus, making stable atoms impossible. However, in the natural world the electrons normally remain in an uncertain, non-deterministic "smeared" (wave–particle wave function) orbital path around or through the nucleus, defying classical electromagnetism. Quantum mechanics was initially developed to provide a better explanation of the atom, especially the spectra of light emitted by different atomic species. The quantum theory of the atom was developed as an explanation for the electron's staying in its orbital, which could not be explained by Newton's laws of motion and by Maxwell's laws of classical electromagnetism. Broadly speaking, quantum mechanics incorporates four classes of phenomena for which classical physics cannot account:

- The quantization (discretization) of certain physical quantities
- wave–particle duality
- uncertainty principle
- quantum entanglement

Mathematical formulations

In the mathematically rigorous formulation of quantum mechanics developed by Paul Dirac and John von Neumann, the possible states of a quantum mechanical system are represented by unit vectors (called "state vectors"). Formally, these reside in a complex separable Hilbert space (variously called the "state space" or the "associated Hilbert space" of the system) well defined up to a complex number of norm 1 (the phase factor). In other words, the possible states are points in the projectivization of a Hilbert space, usually called the complex projective space. The exact nature of this Hilbert space is dependent on the system; for example, the state space for position and momentum states is the space of square-integrable functions, while the state space for the spin of a single proton is just the product of two complex planes. Each observable is represented by a maximally Hermitian (precisely: by a self-adjoint) linear operator acting on the state space. Each eigenstate of an observable corresponds to an eigenvector of the operator, and the associated eigenvalue corresponds to the value of the observable in that eigenstate. If the operator's spectrum is discrete, the observable can only attain those discrete eigenvalues.

In the formalism of quantum mechanics, the state of a system at a given time is described by a complex wave function, also referred to as state vector in a complex vector space. This abstract mathematical object allows for the calculation of probabilities of outcomes of concrete experiments. For example, it allows one to compute the probability of finding an electron in a particular region around the nucleus at a particular time. Contrary to classical mechanics, one can never make simultaneous predictions of conjugate variables, such as position and momentum, with accuracy. For instance, electrons may be considered to be located somewhere within a region of space, but with their exact positions being unknown. Contours of constant probability, often referred to as "clouds", may be drawn around the nucleus of an atom to conceptualize where the electron might be located with the most probability. Heisenberg's uncertainty principle quantifies the inability to precisely locate the particle given its conjugate momentum.

As the result of a measurement, the wave function containing the probability information for a system collapses from a given initial state to a particular eigenstate of the observable. The possible results of a measurement are the eigenvalues of the operator representing the observable — which explains the choice of *Hermitian* operators, for which all the eigenvalues are real. We can find the probability distribution of an observable in a given state by computing the spectral decomposition of the corresponding operator. Heisenberg's uncertainty principle is represented by the statement that the operators corresponding to certain observables do not commute.

The probabilistic nature of quantum mechanics thus stems from the act of measurement. This is one of the most difficult aspects of quantum systems to understand. It was the central topic in the famous Bohr-Einstein debates, in which the two scientists attempted to clarify these fundamental principles by way of thought experiments. In the decades after the formulation of quantum mechanics, the question of what constitutes a "measurement" has been extensively studied. Interpretations of quantum mechanics have been formulated to do away with the concept of "wavefunction collapse"; see, for example, the relative state interpretation. The basic idea is that when a quantum system interacts with a measuring apparatus, their respective wavefunctions become entangled, so that the original quantum system ceases to exist as an independent entity. Generally, quantum mechanics does not assign definite values to observables. Instead, it makes predictions using probability distributions; that is, the probability of obtaining possible outcomes from measuring an observable. Often these results are skewed by many causes, such as dense probability clouds or quantum state nuclear attraction. Naturally, these probabilities will depend on the quantum state at the "instant" of the measurement. Hence, uncertainty is involved in the value. There are, however, certain states that are associated with a definite value of a particular observable. These are known as eigenstates of the observable ("eigen" can be translated from German as inherent or as a characteristic).

In the everyday world, it is natural and intuitive to think of everything (every observable) as being in an eigenstate. Everything appears to have a definite position, a definite momentum, a definite energy, and a definite time of occurrence. However, quantum mechanics does not pinpoint the exact values of a particle for its position and momentum (since they are conjugate pairs) or its energy and time (since they too are conjugate pairs); rather, it only provides a range of probabilities of where that particle might be given its momentum and momentum probability. Therefore, it is helpful to use different words to describe states having *uncertain* values and states having *definite* values (eigenstate). Usually, a system will not be in an eigenstate of the observable we are interested in. However, if one measures the observable, the wavefunction will instantaneously be an eigenstate (or generalized eigenstate) of that observable. This process is known as wavefunction collapse, a debatable process. It involves expanding the system under study to include the measurement device. If one knows the corresponding wave function at the instant before the measurement, one will be able to compute the probability of collapsing into each of the possible eigenstates. For example, the free particle in the previous example will usually have a wavefunction that is a wave packet centered around some mean position x_0, neither an eigenstate of position nor of momentum. When one measures the position of the particle, it is impossible to predict with certainty the result. It is probable, but not certain, that it will be near x_0, where the amplitude of the wave function is large. After the measurement is performed, having obtained some result x, the wave function collapses into a position eigenstate centered at x.

The time evolution of a quantum state is described by the Schrödinger equation, in which the Hamiltonian, the operator corresponding to the total energy of the system, generates time evolution. The time evolution of wave functions is deterministic in the sense that, given a wavefunction at an initial time, it makes a definite prediction of what the wavefunction will be at any later time.

During a measurement, on the other hand, the change of the wavefunction into another one is not deterministic, but rather unpredictable, i.e., random. A time-evolution simulation can be seen here. Wave functions can change as time progresses. An equation known as the Schrödinger equation describes how wave functions change in time, a role similar to Newton's second law in classical mechanics. The Schrödinger equation, applied to the aforementioned example of the free particle, predicts that the center of a wave packet will move through space at a constant velocity, like a classical particle with no forces acting on it. However, the wave packet will also spread out as time progresses, which means that the position becomes more uncertain. This also has the effect of turning position eigenstates (which can be thought of as infinitely sharp wave packets) into broadened wave packets that are no longer position eigenstates.

Some wave functions produce probability distributions that are constant, or independent of time, such as when in a stationary state of constant energy, time drops out of the absolute square of the wave function. Many systems that are treated dynamically in classical mechanics are described by such "static" wave functions. For example, a single electron in an unexcited atom is pictured classically as a particle moving in a circular trajectory around the atomic nucleus, whereas in quantum mechanics it is described by a static, spherically symmetric wavefunction surrounding the nucleus (Fig. 1). (Note that only the lowest angular momentum states, labeled *s*, are spherically symmetric).

The Schrödinger equation acts on the entire probability amplitude, not merely its absolute value. Whereas the absolute value of the probability amplitude encodes information about probabilities, its phase encodes information about the interference between quantum states. This gives rise to the wave-like behavior of quantum states. It turns out that analytic solutions of Schrödinger's equation are only available for a small number of model Hamiltonians, of which the quantum harmonic oscillator, the particle in a box, the hydrogen molecular ion and the hydrogen atom are the most important representatives. Even the helium atom, which contains just one more electron than hydrogen, defies all attempts at a fully analytic treatment. There exist several techniques for generating approximate solutions. For instance, in the method known as perturbation theory one uses the analytic results for a simple quantum mechanical model to generate results for a more complicated model related to the simple model by, for example, the addition of a weak potential energy. Another method is the "semi-classical equation of motion" approach, which applies to systems for which quantum mechanics produces weak deviations from classical behavior. The deviations can be calculated based on the classical motion. This approach is important for the field of quantum chaos.

There are numerous mathematically equivalent formulations of quantum mechanics. One of the oldest and most commonly used formulations is the transformation theory proposed by Cambridge theoretical physicist Paul Dirac, which unifies and generalizes the two earliest formulations of quantum mechanics, matrix mechanics (invented by Werner Heisenberg) and wave mechanics (invented by Erwin Schrödinger). In this formulation, the instantaneous state of a quantum system encodes the probabilities of its measurable properties, or "observables". Examples of observables include energy, position, momentum, and angular momentum. Observables can be either continuous (e.g., the position of a particle) or discrete (e.g., the energy of an electron bound to a hydrogen atom). An alternative formulation of quantum mechanics is Feynman's path integral formulation, in which a quantum-mechanical amplitude is considered as a sum over histories between initial and final states; this is the quantum-mechanical counterpart of action principles in classical mechanics.

Interactions with other scientific theories

The fundamental rules of quantum mechanics are very deep. They assert that the state space of a system is a Hilbert space and the observables are Hermitian operators acting on that space, but do not tell us which Hilbert space or which operators, or if it even exists. These must be chosen appropriately in order to obtain a quantitative description of a quantum system. An important guide for making these choices is the correspondence principle, which states that the predictions of quantum mechanics reduce to those of classical physics when a system moves to higher energies or equivalently, larger quantum numbers. In other words, classical mechanics is simply a quantum mechanics of large systems. This "high energy" limit is known as the *classical* or *correspondence limit*. One can therefore start from an established classical model of a particular system, and attempt to guess the underlying quantum model that gives rise to the classical model in the correspondence limit.

When quantum mechanics was originally formulated, it was applied to models whose correspondence limit was non-relativistic classical mechanics. For instance, the well-known model of the quantum harmonic oscillator uses an explicitly non-relativistic expression for the kinetic energy of the oscillator, and is thus a quantum version of the classical harmonic oscillator.

Early attempts to merge quantum mechanics with special relativity involved the replacement of the Schrödinger equation with a covariant equation such as the Klein-Gordon equation or the Dirac equation. While these theories were successful in explaining many experimental results, they had certain unsatisfactory qualities stemming from their neglect of the relativistic creation and annihilation of particles. A fully relativistic quantum theory required the development of quantum field theory, which applies quantization to a field rather than a fixed set of particles. The first complete quantum field theory, quantum electrodynamics, provides a fully quantum description of the electromagnetic interaction. The full apparatus of quantum field theory is often unnecessary for describing electrodynamic systems. A simpler approach, one employed since the inception of quantum mechanics, is to treat charged particles as quantum mechanical objects being acted on by a classical electromagnetic field. For example, the elementary quantum model of the hydrogen atom describes the electric field of the hydrogen atom using a classical $-\frac{e^2}{4\pi \epsilon_0 r}$ Coulomb potential. This "semi-classical" approach fails if quantum fluctuations in the electromagnetic field play an important role, such as in the emission of photons by charged particles. Quantum field theories for the strong nuclear force and the weak nuclear force have been developed. The quantum field theory of the strong nuclear force is called quantum chromodynamics, and describes the interactions of the subnuclear particles: quarks and gluons. The weak nuclear force and the electromagnetic force were unified, in their quantized forms, into a single quantum field theory known as electroweak theory, by the physicists Abdus Salam, Sheldon Glashow and Steven Weinberg. These three men shared the Nobel Prize in Physics in 1979 for this work.

It has proven difficult to construct quantum models of gravity, the remaining fundamental force. Semi-classical approximations are workable, and have led to predictions such as Hawking radiation. However, the formulation of a complete theory of quantum gravity is hindered by apparent incompatibilities between general relativity, the most accurate theory of gravity currently known, and some of the fundamental assumptions of quantum theory. The resolution of these incompatibilities is an area of active research, and theories such as string theory are among the possible candidates for a future theory of quantum gravity. Classical mechanics has been extended into the complex domain and complex classical mechanics exhibits behaviours similar to quantum mechanics.

Quantum mechanics and classical physics

Predictions of quantum mechanics have been verified experimentally to a very high degree of accuracy. According to the correspondence principle between classical and quantum mechanics, all objects obey the laws of quantum mechanics, and classical mechanics is just an approximation for large systems (or a statistical quantum mechanics of a large collection of particles). The laws of classical mechanics thus follow from the laws of quantum mechanics as a statistical average at the limit of large systems or large quantum numbers. However, chaotic systems do not have good quantum numbers, and quantum chaos studies the relationship between classical and quantum descriptions in these systems.

Quantum coherence is an essential difference between classical and quantum theories, and is illustrated by the Einstein-Podolsky-Rosen paradox. Quantum interference involves the addition of *probability amplitudes*, whereas when classical waves interfere there is an addition of *intensities*. For microscopic bodies, the extension of the system is much smaller than the coherence length, which gives rise to long-range entanglement and other nonlocal phenomena characteristic of quantum systems. Quantum coherence is not typically evident at macroscopic scales, although an exception to this rule can occur at extremely low temperatures, when quantum behavior can manifest itself on more macroscopic scales. This is in accordance with the following observations:

- Many macroscopic properties of a classical system are a direct consequences of the quantum behavior of its parts. For example, the stability of bulk matter (which consists of atoms and molecules which would quickly collapse under electric forces alone), the rigidity of solids, and the mechanical, thermal, chemical, optical and magnetic properties of matter are all results of the interaction of electric charges under the rules of quantum mechanics.
- While the seemingly exotic behavior of matter posited by quantum mechanics and relativity theory become more apparent when dealing with extremely fast-moving or extremely tiny particles, the laws of classical Newtonian physics remain accurate in predicting the behavior of large objects—of the order of the size of large molecules and bigger—at velocities much smaller than the velocity of light.

Relativity and quantum mechanics

Even with the defining postulates of both Einstein's theory of general relativity and quantum theory being indisputably supported by rigorous and repeated empirical evidence and while they do not directly contradict each other theoretically (at least with regard to primary claims), they are resistant to being incorporated within one cohesive model.

Einstein himself is well known for rejecting some of the claims of quantum mechanics. While clearly contributing to the field, he did not accept the more philosophical consequences and interpretations of quantum mechanics, such as the lack of deterministic causality and the assertion that a single subatomic particle can occupy numerous areas of space at one time. He also was the first to notice some of the apparently exotic consequences of entanglement and used them to formulate the Einstein-Podolsky-Rosen paradox, in the hope of showing that quantum mechanics had unacceptable implications. This was 1935, but in 1964 it was shown by John Bell that, although Einstein was correct in identifying seemingly paradoxical implications of quantum mechanical nonlocality, these implications could be experimentally tested. Alain Aspect's initial experiments in 1982, and many subsequent experiements since, have verified quantum entanglement.

According to the paper of J. Bell and the Copenhagen interpretation (the common interpretation of quantum mechanics by physicists since 1927), and contrary to Einstein's ideas, quantum mechanics was not at the same time

- a "realistic" theory

- and a *local* theory.

The Einstein-Podolsky-Rosen paradox shows in any case that there exist experiments by which one can measure the state of one particle and instantaneously change the state of its entangled partner, although the two particles can be an arbitrary distance apart; however, this effect does not violate causality, since no transfer of information happens. Quantum entanglement is at the basis of quantum cryptography, with high-security commercial applications in banking and government.

Gravity is negligible in many areas of particle physics, so that unification between general relativity and quantum mechanics is not an urgent issue in those applications. However, the lack of a correct theory of quantum gravity is an important issue in cosmology and physicists' search for an elegant "theory of everything". Thus, resolving the inconsistencies between both theories has been a major goal of twentieth- and twenty-first-century physics. Many prominent physicists, including Stephen Hawking, have labored in the attempt to discover a theory underlying *everything*, combining not only different models of subatomic physics, but also deriving the universe's four forces —the strong force, electromagnetism, weak force, and gravity— from a single force or phenomenon. One of the leaders in this field is Edward Witten, a theoretical physicist who formulated the groundbreaking M-theory, which is an attempt at describing the supersymmetrical based string theory.

Attempts at a unified field theory

As of 2010 the quest for unifying the fundamental forces through quantum mechanics is still ongoing. Quantum electrodynamics (or "quantum electromagnetism"), which is currently (in the perturbative regime at least) the most accurately tested physical theory, has been successfully merged with the weak nuclear force into the electroweak force and work is currently being done to merge the electroweak and strong force into the electrostrong force. Current predictions state that at around 10^{14} GeV the three aforementioned forces are fused into a single unified field, Beyond this "grand unification," it is speculated that it may be possible to merge gravity with the other three gauge symmetries, expected to occur at roughly 10^{19} GeV. However — and while special relativity is parsimoniously incorporated into quantum electrodynamics — the expanded general relativity, currently the best theory describing the gravitation force, has not been fully incorporated into quantum theory.

Philosophical implications

Since its inception, the many counter-intuitive results of quantum mechanics have provoked strong philosophical debate and many interpretations. Even fundamental issues such as Max Born's basic rules concerning probability amplitudes and probability distributions took decades to be appreciated.

Richard Feynman said, "I think I can safely say that nobody understands quantum mechanics."

The Copenhagen interpretation, due largely to the Danish theoretical physicist Niels Bohr, is the interpretation of the quantum mechanical formalism most widely accepted amongst physicists. According to it, the probabilistic nature of quantum mechanics is not a temporary feature which will eventually be replaced by a deterministic theory, but instead must be considered to be a final renunciation of the classical ideal of causality. In this interpretation, it is believed that any well-defined application of the quantum mechanical formalism must always make reference to the experimental arrangement, due to the complementarity nature of evidence obtained under different experimental situations.

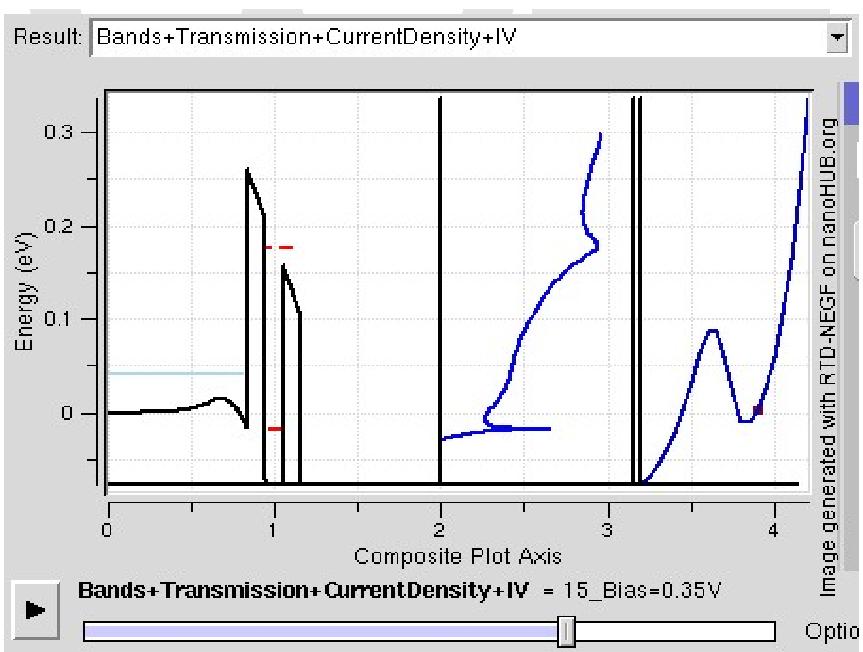
Albert Einstein, himself one of the founders of quantum theory, disliked this loss of determinism in measurement. (This dislike is the source of his famous quote, "God does not play dice with the universe.") Einstein held that there should be a local hidden variable theory underlying quantum mechanics and that, consequently, the present theory was incomplete. He produced a series of objections to the theory, the most famous of which has become known as the Einstein-Podolsky-Rosen paradox. John Bell showed that the EPR paradox led to experimentally testable differences between quantum mechanics and local realistic theories. Experiments have been performed confirming the accuracy of quantum mechanics, thus demonstrating that the physical world cannot be described by local realistic theories. The *Bohr-Einstein debates* provide a vibrant critique of the Copenhagen Interpretation from an epistemological point of view.

The Everett many-worlds interpretation, formulated in 1956, holds that all the possibilities described by quantum theory simultaneously occur in a multiverse composed of mostly independent parallel universes. This is not accomplished by introducing some new axiom to quantum mechanics, but on the contrary by *removing* the axiom of the collapse of the wave packet: All the possible consistent states of the measured system and the measuring apparatus (including the observer) are present in a *real* physical (not just formally mathematical, as in other interpretations) quantum superposition. Such a superposition of consistent state combinations of different systems is called an entangled state. While the multiverse is deterministic, we perceive non-deterministic behavior governed by probabilities, because we can observe only the universe, i.e. the consistent state contribution to the mentioned superposition, we inhabit. Everett's interpretation is perfectly consistent with John Bell's experiments and makes them intuitively understandable. However, according to the theory of quantum decoherence, the parallel universes will never be accessible to us. This inaccessibility can be understood as follows: Once a measurement is done, the measured system becomes entangled with both the physicist who measured it and a huge number of other particles, some of which are photons flying away towards the other end of the universe; in order to prove that the wave function did not collapse one would have to bring all these particles back and measure them again, together with the system that was measured originally. This is completely impractical, but even if one could theoretically do this, it would destroy any evidence that the original measurement took place (including the physicist's memory).

Applications

Quantum mechanics had enormous success in explaining many of the features of our world. The individual behaviour of the subatomic particles that make up all forms of matter—electrons, protons, neutrons, photons and others—can often only be satisfactorily described using quantum mechanics. Quantum mechanics has strongly influenced string theory, a candidate for a theory of everything and the multiverse hypothesis.

Quantum mechanics is important for understanding how individual atoms combine covalently to form chemicals or molecules. The application of quantum mechanics to chemistry is known as quantum chemistry. (Relativistic) quantum mechanics can in principle mathematically describe most of chemistry. Quantum mechanics can provide quantitative insight into ionic and covalent bonding processes by explicitly showing which molecules are energetically favorable to which others, and by approximately how much. Most of the calculations performed in computational chemistry rely on quantum mechanics.



A working mechanism of a resonant tunneling diode device, based on the phenomenon of quantum tunneling through the potential barriers.

Much of modern technology operates at a scale where quantum effects are significant. Examples include the laser, the transistor (and thus the microchip), the electron microscope, and magnetic resonance imaging. The study of semiconductors led to the invention of the diode and the transistor, which are indispensable for modern electronics.

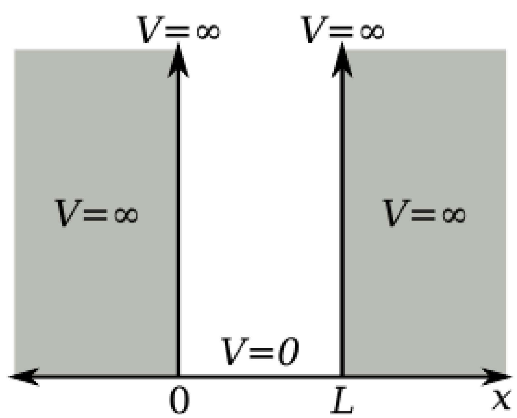
Researchers are currently seeking robust methods of directly manipulating quantum states. Efforts are being made to develop quantum cryptography, which will allow guaranteed secure transmission of information. A more distant goal is the development of quantum computers, which are expected to perform certain computational tasks exponentially faster than classical computers. Another active research topic is quantum teleportation, which deals with techniques to transmit quantum information over arbitrary distances.

Quantum tunneling is vital in many devices, even in the simple light switch, as otherwise the electrons in the electric current could not penetrate the potential barrier made up of a layer of oxide. Flash memory chips found in USB drives use quantum tunneling to erase their memory cells.

Quantum mechanics primarily applies to the atomic regimes of matter and energy, but some systems exhibit quantum mechanical effects on a large scale; superfluidity (the frictionless flow of a liquid at temperatures near absolute zero) is one well-known example. Quantum theory also provides accurate descriptions for many previously unexplained phenomena such as black body radiation and the stability of electron orbitals. It has also given insight into the workings of many different biological systems, including smell receptors and protein structures. Recent work on photosynthesis has provided evidence that quantum correlations play an essential role in this most fundamental process of the plant kingdom. Even so, classical physics often can be a good approximation to results otherwise obtained by **quantum physics**, typically in circumstances with large numbers of particles or large quantum numbers. (However, some open questions remain in the field of quantum chaos.)

Examples

Particle in a box



1-dimensional potential energy box (or infinite potential well)

The particle in a 1-dimensional potential energy box is the most simple example where restraints lead to the quantization of energy levels. The box is defined as having zero potential energy inside a certain region and infinite potential energy everywhere outside that region. For the 1-dimensional case in the *x* direction, the time-independent Schrödinger equation can be written as:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi.$$

Writing the differential operator

$$\hat{p}_x = -i\hbar\frac{d}{dx}$$

the previous equation can be seen to be evocative of the classic analogue

$$\frac{1}{2m}\hat{p}_x^2 = E$$

with *E* as the energy for the state ψ, in this case coinciding with the kinetic energy of the particle.

The general solutions of the Schrödinger equation for the particle in a box are:

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \qquad\qquad E = \frac{\hbar^2 k^2}{2m}$$

or, from Euler's formula,

$$\psi(x) = C\sin kx + D\cos kx.$$

The presence of the walls of the box determines the values of *C*, *D*, and *k*. At each wall ($x = 0$ and $x = L$), $\psi = 0$. Thus when $x = 0$,

$$\psi(0) = 0 = C\sin 0 + D\cos 0 = D$$

and so $D = 0$. When $x = L$,

$$\psi(L) = 0 = C\sin kL.$$

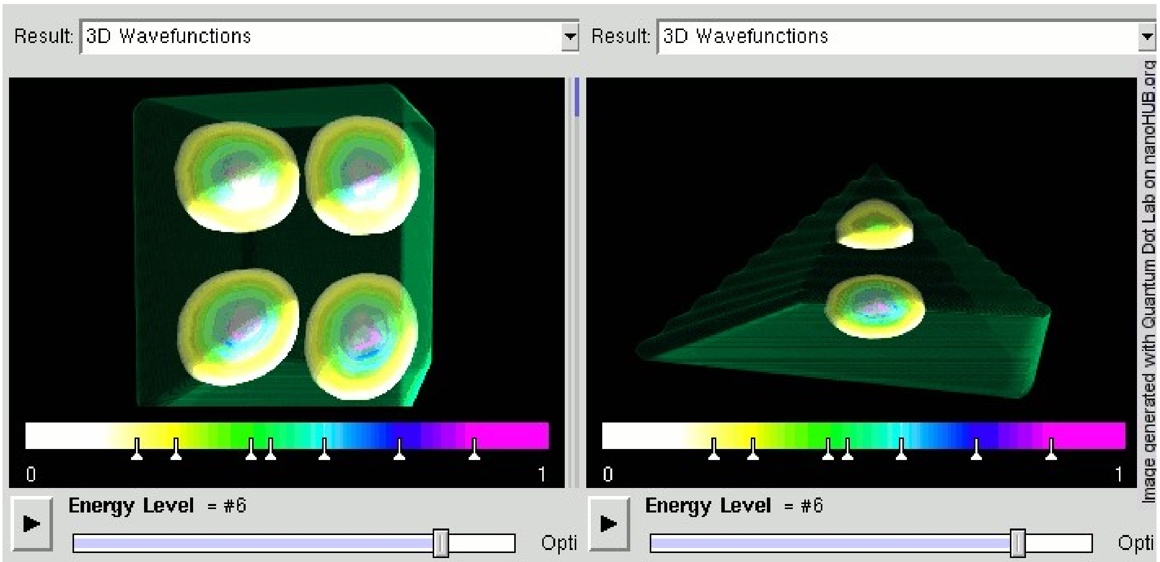
C cannot be zero, since this would conflict with the Born interpretation. Therefore $\sin kL = 0$, and so it must be that *kL* is an integer multiple of π. Therefore,

$$k = \frac{n\pi}{L} \qquad\qquad n = 1, 2, 3, \ldots.$$

The quantization of energy levels follows from this constraint on *k*, since

$$E = \frac{\hbar^2 \pi^2 n^2}{2mL^2} = \frac{n^2 h^2}{8mL^2}.$$

Free particle



3D confined electron wave functions for each eigenstate in a Quantum Dot. Here, rectangular and triangular-shaped quantum dots are shown. Energy states in rectangular dots are more ‘s-type’ and ‘p-type’. However, in a triangular dot the wave functions are mixed due to confinement symmetry.

For example, consider a free particle. In quantum mechanics, there is wave-particle duality so the properties of the particle can be described as the properties of a wave. Therefore, its quantum state can be represented as a wave of arbitrary shape and extending over space as a wave function. The position and momentum of the particle are observables. The Uncertainty Principle states that both the position and the momentum cannot simultaneously be measured with full precision at the same time. However, one can measure the position alone of a moving free particle creating an eigenstate of position with a wavefunction that is very large (a Dirac delta) at a particular position x and zero everywhere else. If one performs a position measurement on such a wavefunction, the result x will be obtained with 100% probability (full certainty). This is called an eigenstate of position (mathematically more precise: a *generalized position eigenstate (eigendistribution)*). If the particle is in an eigenstate of position then its momentum is completely unknown. On the other hand, if the particle is in an eigenstate of momentum then its position is completely unknown. In an eigenstate of momentum having a plane wave form, it can be shown that the wavelength is equal to h/p, where h is Planck's constant and p is the momentum of the eigenstate.

Chapter 2

Quantum Logic

In quantum mechanics, **quantum logic** is a set of rules for reasoning about propositions which takes the principles of quantum theory into account. This research area and its name originated in the 1936 paper by Garrett Birkhoff and John von Neumann, who were attempting to reconcile the apparent inconsistency of classical boolean logic with the facts concerning the measurement of complementary variables in quantum mechanics, such as position and momentum.

Quantum logic can be formulated either as a modified version of propositional logic or as a noncommutative and non-associative many-valued (MV) logic.

Quantum logic has some properties which clearly distinguish it from classical logic, most notably, the failure of the distributive law of propositional logic:

p and (*q* or *r*) = (*p* and *q*) or (*p* and *r*),

where the symbols *p*, *q* and *r* are propositional variables. To illustrate why the distributive law fails, consider a particle moving on a line and let

p = "the particle is moving to the right"
q = "the particle is in the interval [-1,1]"
r = "the particle is not in the interval [-1,1]"

then the proposition "*q* or *r*" is true, so

p and (*q* or *r*) = *p*

On the other hand, the propositions "*p* and *q*" and "*p* and *r*" are both false, since they assert tighter restrictions on simultaneous values of position and momentum than is allowed by the uncertainty principle. So,

(*p* and *q*) or (*p* and *r*) = false

Thus the distributive law fails.

Quantum logic has been proposed as the correct logic for propositional inference generally, most notably by the philosopher Hilary Putnam, at least at one point in his career. This thesis was an important ingredient in Putnam's paper *Is Logic Empirical?* in which he analysed the epistemological status of the rules of propositional logic. Putnam attributes the idea that anomalies associated to quantum measurements originate with anomalies in the logic of physics itself to the physicist David Finkelstein. However, this idea had been around for some time and had been revived several years earlier by George Mackey's work on group representations and symmetry.

The more common view regarding quantum logic, however, is that it provides a formalism for relating observables, system preparation filters and states. In this view, the quantum logic approach resembles more closely the C*-algebraic approach to quantum mechanics; in fact with some minor technical assumptions it can be subsumed by it. The similarities of the quantum logic formalism to a system of deductive logic may then be regarded more as a curiosity than as a fact of fundamental philosophical importance. A more modern approach to the structure of quantum logic is to assume that it is a diagram – in the sense of category theory – of classical logics.

Introduction

In his classic treatise *Mathematical Foundations of Quantum Mechanics*, John von Neumann noted that projections on a Hilbert space can be viewed as propositions about physical observables. The set of principles for manipulating these quantum propositions was called *quantum logic* by von Neumann and Birkhoff. In his book (also called *Mathematical Foundations of Quantum Mechanics*) G. Mackey attempted to provide a set of axioms for this propositional system as an orthocomplemented lattice. Mackey viewed elements of this set as potential *yes or no questions* an observer might ask about the state of a physical system, questions that would be settled by some measurement. Moreover Mackey defined a physical observable in terms of these basic questions. Mackey's axiom system is somewhat unsatisfactory though, since it assumes that the partially ordered set is actually given as the orthocomplemented closed subspace lattice of a separable Hilbert space. Piron, Ludwig and others have attempted to give axiomatizations which do not require such explicit relations to the lattice of subspaces.

Projections as propositions

The so-called *Hamiltonian* formulations of classical mechanics have three ingredients: *states*, *observables* and *dynamics*. In the simplest case of a single particle moving in $\mathbf{R}^3$, the state space is the position-momentum space $\mathbf{R}^6$. We will merely note here that an observable is some real-valued function f on the state space. Examples of observables are position, momentum or energy of a particle. For classical systems, the value $f(x)$, that is the value of f for some particular system state x, is obtained by a process of measurement of f. The propositions concerning a classical system are generated from basic statements of the form

- Measurement of *f* yields a value in the interval [*a*, *b*] for some real numbers *a*, *b*.

It follows easily from this characterization of propositions in classical systems that the corresponding logic is identical to that of some Boolean algebra of subsets of the state space. By logic in this context we mean the rules that relate set operations and ordering relations, such as de Morgan's laws. These are analogous to the rules relating boolean conjunctives and material implication in classical propositional logic. For technical reasons, we will also assume that the algebra of subsets of the state space is that of all Borel sets. The set of propositions is ordered by the natural ordering of sets and has a complementation operation. In terms of observables, the complement of the proposition $\{f \geq a\}$ is $\{f < a\}$.

We summarize these remarks as follows:

- The proposition system of a classical system is a lattice with a distinguished *orthocomplementation* operation: The lattice operations of *meet* and *join* are respectively set intersection and set union. The orthocomplementation operation is set complement. Moreover this lattice is *sequentially complete*, in the sense that any sequence $\{E_i\}_i$ of elements of the lattice has a least upper bound, specifically the set-theoretic union:

$$\operatorname{LUB}(\{E_i\}) = \bigcup_{i=1}^{\infty} E_i.$$

In the Hilbert space formulation of quantum mechanics as presented by von Neumann, a physical observable is represented by some (possibly unbounded) densely-defined self-adjoint operator *A* on a Hilbert space *H*. *A* has a spectral decomposition, which is a projection-valued measure E defined on the Borel subsets of **R**. In particular, for any bounded Borel function *f*, the following equation holds:

$$f(A) = \int_{\mathbb{R}} f(\lambda) \, d\mathrm{E}(\lambda).$$

In case *f* is the indicator function of an interval [*a*, *b*], the operator *f*(*A*) is a self-adjoint projection, and can be interpreted as the quantum analogue of the classical proposition

- Measurement of *A* yields a value in the interval [*a*, *b*].

The propositional lattice of a quantum mechanical system

This suggests the following quantum mechanical replacement for the orthocomplemented lattice of propositions in classical mechanics. This is essentially Mackey's *Axiom VII*:

- The orthocomplemented lattice Q of propositions of a quantum mechanical system is the lattice of closed subspaces of a complex Hilbert space H where orthocomplementation of V is the orthogonal complement $V^{\perp}$.

Q is also sequentially complete: any pairwise disjoint sequence $\{V_i\}_i$ of elements of Q has a least upper bound. Here disjointness of W_1 and W_2 means W_2 is a subspace of $W_1{}^{\perp}$. The least upper bound of $\{V_i\}_i$ is the closed internal direct sum.

Henceforth we identify elements of Q with self-adjoint projections on the Hilbert space H.

The structure of Q immediately points to a difference with the partial order structure of a classical proposition system. In the classical case, given a proposition p, the equations

$$I = p \vee q$$
$$0 = p \wedge q$$

have exactly one solution, namely the set-theoretic complement of p. In these equations I refers to the atomic proposition which is identically true and 0 the atomic proposition which is identically false. In the case of the lattice of projections there are infinitely many solutions to the above equations.

Having made these preliminary remarks, we turn everything around and attempt to define observables within the projection lattice framework and using this definition establish the correspondence between self-adjoint operators and observables: A *Mackey observable* is a countably additive homomorphism from the orthocomplemented lattice of the Borel subsets of **R** to Q. To say the mapping φ is a countably additive homomorphism means that for any sequence $\{S_i\}_i$ of pairwise disjoint Borel subsets of **R**, $\{\varphi(S_i)\}_i$ are pairwise orthogonal projections and

$$\varphi\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} \varphi(S_i).$$

Theorem. There is a bijective correspondence between Mackey observables and densely-defined self-adjoint operators on H.

This is the content of the spectral theorem as stated in terms of spectral measures.

Statistical structure

Imagine a forensics lab which has some apparatus to measure the speed of a bullet fired from a gun. Under carefully controlled conditions of temperature, humidity, pressure and so on the same gun is fired repeatedly and speed measurements taken. This produces some distribution of speeds. Though we will not get exactly the same value for each individual measurement, for each cluster of measurements, we would expect the experiment to lead to the same distribution of speeds. In particular, we can expect to assign probability distributions to propositions such as $\{a \leq \text{speed} \leq b\}$. This leads naturally to propose that under controlled conditions of preparation, the measurement of a classical system can be described by a probability measure on the state space. This same statistical structure is also present in quantum mechanics.

A *quantum probability measure* is a function P defined on *Q* with values in [0,1] such that P(0)=0, P(I)=1 and if $\{E_i\}_i$ is a sequence of pairwise orthogonal elements of *Q* then

$$\mathrm{P}\left(\sum_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(E_i).$$

The following highly non-trivial theorem is due to Andrew Gleason:

Theorem. Suppose *H* is a separable Hilbert space of complex dimension at least 3. Then for any quantum probability measure on *Q* there exists a unique trace class operator *S* such that

$$\mathrm{P}(E) = \mathrm{Tr}(SE)$$

for any self-adjoint projection *E*.

The operator *S* is necessarily non-negative (that is all eigenvalues are non-negative) and of trace 1. Such an operator is often called a *density operator*.

Physicists commonly regard a density operator as being represented by a (possibly infinite) density matrix relative to some orthonormal basis.

Automorphisms

An *automorphism* of *Q* is a bijective mapping $\alpha: Q \to Q$ which preserves the orthocomplemented structure of *Q*, that is

$$\alpha\left(\sum_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \alpha(E_i)$$

for any sequence $\{E_i\}_i$ of pairwise orthogonal self-adjoint projections. Note that this property implies monotonicity of α. If P is a quantum probability measure on *Q*, then $E \to \alpha(E)$ is also a quantum probability measure on *Q*. By the Gleason theorem characterizing quantum probability measures quoted above, any automorphism α induces a mapping α* on the density operators by the following formula:

$$\mathrm{Tr}(\alpha^*(S)E) = \mathrm{Tr}(S\alpha(E)).$$

The mapping α* is bijective and preserves convex combinations of density operators. This means

$$\alpha^*(r_1S_1 + r_2S_2) = r_1\alpha^*(S_1) + r_2\alpha^*(S_2)$$

whenever $1 = r_1 + r_2$ and r_1, r_2 are non-negative real numbers. Now we use a theorem of Richard V. Kadison:

Theorem. Suppose β is a bijective map from density operators to density operators which is convexity preserving. Then there is an operator *U* on the Hilbert space which is either linear or conjugate-linear, preserves the inner product and is such that

$$\beta(S) = USU^*$$

for every density operator *S*. In the first case we say *U* is unitary, in the second case *U* is anti-unitary.

Remark. This note is included for technical accuracy only, and should not concern most readers. The result quoted above is not directly stated in Kadison's paper, but can be reduced to it by noting first that β extends to a positive trace preserving map on the trace class operators, then applying duality and finally applying a result of Kadison's paper.

The operator *U* is not quite unique; if *r* is a complex scalar of modulus 1, then r *U* will be unitary or anti-unitary if *U* is and will implement the same automorphism. In fact, this is the only ambiguity possible.

It follows that automorphisms of *Q* are in bijective correspondence to unitary or anti-unitary operators modulo multiplication by scalars of modulus 1. Moreover, we can regard automorphisms in two equivalent ways: as operating on states (represented as density operators) or as operating on *Q*.

Non-relativistic dynamics

In non-relativistic physical systems, there is no ambiguity in referring to time evolution since there is a global time parameter. Moreover an isolated quantum system evolves in a deterministic way: if the system is in a state *S* at time *t* then at time $s > t$, the system is in a state $F_{s,t}(S)$. Moreover, we assume

- The dependence is reversible: The operators $F_{s,t}$ are bijective.
- The dependence is homogeneous: $F_{s,t} = F_{s-t,0}$.
- The dependence is convexity preserving: That is, each $F_{s,t}(S)$ is convexity preserving.

- The dependence is weakly continuous: The mapping **R**→ **R** given by $t \to \mathrm{Tr}(\mathrm{F}_{s,t}(S)\, E)$ is continuous for every E in Q.

By Kadison's theorem, there is a 1-parameter family of unitary or anti-unitary operators $\{U_t\}_t$ such that

$$\mathrm{F}_{s,t}(S) = U_{s-t} S U^*_{s-t}$$

In fact,

Theorem. Under the above assumptions, there is a strongly continuous 1-parameter group of unitary operators $\{U_t\}_t$ such that the above equation holds.

Note that it easily from uniqueness from Kadison's theorem that

$$U_{t+s} = \sigma(t,s) U_t U_s$$

where σ(t,s) has modulus 1. Now the square of an anti-unitary is a unitary, so that all the U_t are unitary. The remainder of the argument shows that σ(t,s) can be chosen to be 1 (by modifying each U_t by a scalar of modulus 1.)

Pure states

A convex combination of statistical states S_1 and S_2 is a state of the form $S = p_1 S_1 + p_2 S_2$ where p_1, p_2 are non-negative and $p_1 + p_2 = 1$. Considering the statistical state of system as specified by lab conditions used for its preparation, the convex combination S can be regarded as the state formed in the following way: toss a biased coin with outcome probabilities p_1, p_2 and depending on outcome choose system prepared to S_1 or S_2

Density operators form a convex set. The convex set of density operators has extreme points; these are the density operators given by a projection onto a one-dimensional space. To see that any extreme point is such a projection, note that by the spectral theorem S can be represented by a diagonal matrix; since S is non-negative all the entries are non-negative and since S has trace 1, the diagonal entries must add up to 1. Now if it happens that the diagonal matrix has more than one non-zero entry it is clear that we can express it as a convex combination of other density operators.

The extreme points of the set of density operators are called pure states. If S is the projection on the 1-dimensional space generated by a vector ψ of norm 1 then

$$\mathrm{Tr}(SE) = \langle E\psi | \psi \rangle$$

for any E in Q. In physics jargon, if

$$S = |\psi\rangle\langle\psi|,$$

where ψ has norm 1, then

$$\operatorname{Tr}(SE) = \langle \psi | E | \psi \rangle.$$

Thus pure states can be identified with *rays* in the Hilbert space *H*.

The measurement process

Consider a quantum mechanical system with lattice *Q* which is in some statistical state given by a density operator *S*. This essentially means an ensemble of systems specified by a repeatable lab preparation process. The result of a cluster of measurements intended to determine the truth value of proposition *E*, is just as in the classical case, a probability distribution of truth values **T** and **F**. Say the probabilities are *p* for **T** and $q = 1 - p$ for **F**. By the previous section $p = \operatorname{Tr}(S\,E)$ and $q = \operatorname{Tr}(S\,(I - E))$.

Perhaps the most fundamental difference between classical and quantum systems is the following: regardless of what process is used to determine *E* immediately after the measurement the system will be in one of two statistical states:

- If the result of the measurement is **T**

$$\frac{1}{\operatorname{Tr}(ES)} ESE.$$

- If the result of the measurement is **F**

$$\frac{1}{\operatorname{Tr}((I-E)S)} (I-E)S(I-E).$$

(We leave to the reader the handling of the degenerate cases in which the denominators may be 0.) We now form the convex combination of these two ensembles using the relative frequencies *p* and *q*. We thus obtain the result that the measurement process applied to a statistical ensemble in state *S* yields another ensemble in statistical state:

$$\mathrm{M}_E(S) = ESE + (I-E)S(I-E).$$

We see that a pure ensemble becomes a mixed ensemble after measurement. Measurement, as described above, is a special case of quantum operations.

Limitations

Quantum logic derived from propositional logic provides a satisfactory foundation for a theory of reversible quantum processes. Examples of such processes are the covariance transformations relating two frames of reference, such as change of time parameter or the transformations of special relativity. Quantum logic also provides a satisfactory understanding of density matrices. Quantum logic can be stretched to account for some kinds of measurement processes corresponding to answering yes-no questions about the state of a quantum system. However, for more general kinds of measurement operations (that is quantum operations), a more complete theory of filtering processes is necessary. Such an approach is provided by the consistent histories formalism. On the other hand, quantum logics derived from MV-logic extend its range of applicability to irreversible quantum processes and/or 'open' quantum systems.

In any case, these quantum logic formalisms must be generalized in order to deal with super-geometry (which is needed to handle Fermi-fields) and non-commutative geometry (which is needed in string theory and quantum gravity theory). Both of these theories use a partial algebra with an "integral" or "trace". The elements of the partial algebra are not observables; instead the "trace" yields "greens functions" which generate scattering amplitudes. One thus obtains a local S-matrix theory.

Since around 1978 the Flato school has been developing an alternative to the quantum logics approach called deformation quantization.

In 2004, Prakash Panangaden described how to capture the kinematics of quantum causal evolution using System BV, a deep inference logic originally developed for use in structural proof theory. Alessio Guglielmi, Lutz Straßburger, and Richard Blute have also done work in this area.

Chapter 3

Quantum Electrodynamics

Quantum electrodynamics (**QED**) is the relativistic quantum field theory of electrodynamics. In essence, it describes how light and matter interact and is the first theory where full agreement between quantum mechanics and special relativity is achieved. QED mathematically describes all phenomena involving electrically charged particles interacting by means of exchange of photons and represents the quantum counterpart of classical electrodynamics giving a complete account of matter and light interaction. One of the founding fathers of QED, Richard Feynman, has called it "the jewel of physics" for its extremely accurate predictions of quantities like the anomalous magnetic moment of the electron, and the Lamb shift of the energy levels of hydrogen.

In technical terms, QED can be described as a perturbation theory of the electromagnetic quantum vacuum.

History

The first formulation of a quantum theory describing radiation and matter interaction is due to Paul Adrien Maurice Dirac, who, during 1920, was first able to compute the coefficient of spontaneous emission of an atom.

Paul Dirac

Dirac described the quantization of the electromagnetic field as an ensemble of harmonic oscillators with the introduction of the concept of creation and annihilation operators of particles. In the following years, with contributions from Wolfgang Pauli, Eugene Wigner, Pascual Jordan, Werner Heisenberg and an elegant formulation of quantum electrodynamics due to Enrico Fermi, physicists came to believe that, in principle, it would be possible to perform any computation for any physical process involving photons and charged particles. However, further studies by Felix Bloch with Arnold Nordsieck, and Victor Weisskopf, in 1937 and 1939, revealed that such computations were reliable only at a first order of perturbation theory, a problem already pointed out by Robert Oppenheimer. At higher orders in the series infinities emerged, making such computations meaningless and casting serious doubts on the internal consistency of the theory itself. With no solution for this problem known at the time, it appeared that a fundamental incompatibility existed between special relativity and quantum mechanics .

Hans Bethe

Difficulties with the theory increased through the end of 1940. Improvements in microwave technology made it possible to take more precise measurements of the shift of the levels of a hydrogen atom, now known as the Lamb shift and magnetic moment of the electron. These experiments unequivocally exposed discrepancies which the theory was unable to explain. A first indication of a possible way out was given by Hans Bethe. In 1947, while he was traveling by train to reach Schenectady from New York, after giving a talk at the conference at Shelter Island on the subject, Bethe completed the first non-relativistic computation of the shift of the lines of the hydrogen atom as measured by Lamb and Retherford. Despite the limitations of the computation, agreement was excellent. The idea was simply to attach infinities to corrections at mass and charge that were actually fixed to a finite value by experiments. In this way, the infinities get absorbed in those constants and yield a finite result in good agreement with experiments. This procedure was named renormalization.

Shelter Island Conference group photo (Courtesy of Archives, National Academy of Sciences)

Feynman (center) and Oppenheimer (left) at Los Alamos

Based on Bethe's intuition and fundamental papers on the subject by Sin-Itiro Tomonaga, Julian Schwinger, Richard Feynman and Freeman Dyson, it was finally possible to get fully covariant formulations that were finite at any order in a perturbation series of quantum electrodynamics. Sin-Itiro Tomonaga, Julian Schwinger and Richard Feynman were jointly awarded with a Nobel prize in physics in 1965 for their work in this area. Their contributions, and those of Freeman Dyson, were about covariant and gauge invariant formulations of quantum electrodynamics that allow computations of observables at any order of perturbation theory. Feynman's mathematical technique, based on his diagrams, initially seemed very different from the field-theoretic, operator-based approach of Schwinger and Tomonaga, but Freeman Dyson later showed that the two approaches were equivalent. Renormalization, the need to attach a physical meaning at certain divergences appearing in the theory through integrals, has subsequently become one of the fundamental aspects of quantum field theory and has come to be seen as a criterion for a theory's general acceptability. Even though renormalization works very well in practice, Feynman was never entirely comfortable with its mathematical validity, even referring to renormalization as a "shell game" and "hocus pocus".

QED has served as the model and template for all subsequent quantum field theories. One such subsequent theory is quantum chromodynamics, which began in the early 1960s and attained its present form in the 1975 work by H. David Politzer, Sidney Coleman, David Gross and Frank Wilczek. Building on the pioneering work of Schwinger, Gerald Guralnik, Dick Hagen, and Tom Kibble, Peter Higgs, Jeffrey Goldstone, and others, Sheldon Glashow, Steven Weinberg and Abdus Salam independently showed how the weak nuclear force and quantum electrodynamics could be merged into a single electroweak force.

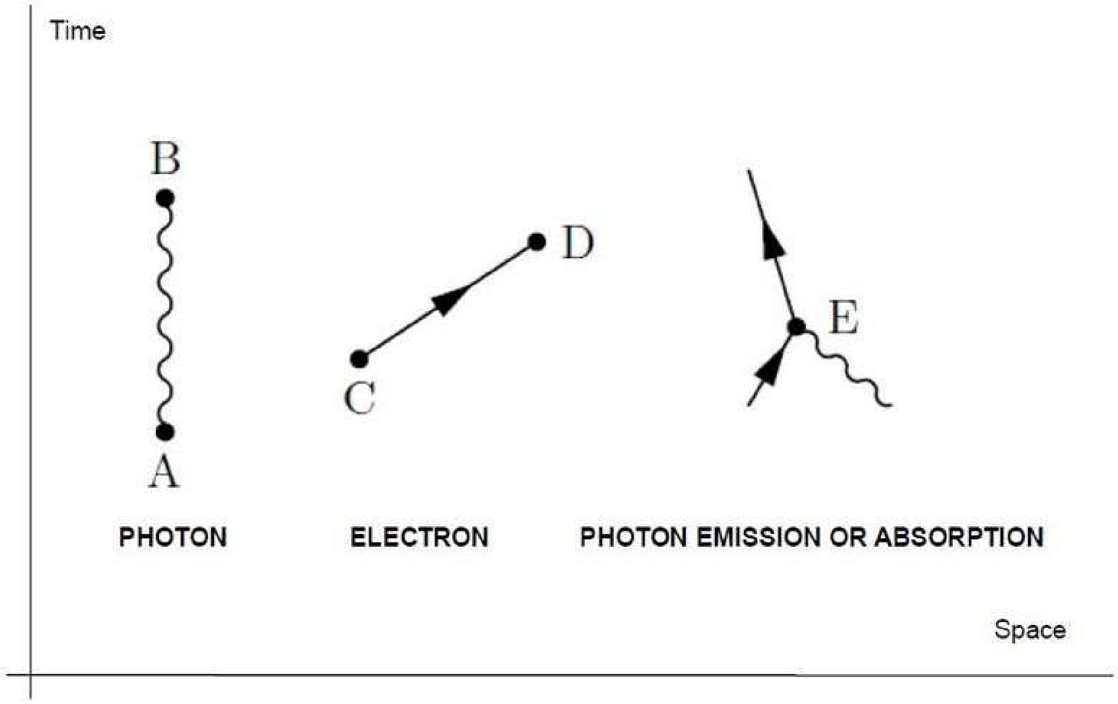
Feynman's view of quantum electrodynamics

Introduction

Near the end of his life, Richard P. Feynman gave a series of lectures on QED intended for the lay public. These lectures were transcribed and published as Feynman (1985), *QED: The strange theory of light and matter*, a classic non-mathematical exposition of QED from the point of view articulated below.

The key components of Feynman's presentation of QED are three basic actions.

- A photon goes from one place and time to another place and time.
- An electron goes from one place and time to another place and time.
- An electron emits or absorbs a photon at a certain place and time.



Feynman diagram elements

These actions are represented in a form of visual shorthand by the three basic elements of Feynman diagrams: a wavy line for the photon, a straight line for the electron and a junction of two straight lines and a wavy one for a vertex representing emission or absorption of a photon by an electron. These may all be seen in the adjacent diagram.

It is important not to over-interpret these diagrams. Nothing is implied about *how* a particle gets from one point to another. The diagrams do *not* imply that the particles are moving in straight or curved lines. They do *not* imply that the particles are moving with fixed speeds. The fact that the photon is often represented, by convention, by a wavy line and not a straight one does *not* imply that it is thought that it is more wavelike than is an electron. The images are just symbols to represent the actions above: photons and electrons do, somehow, move from point to point and electrons, somehow, emit and absorb photons. We do not know how these things happen, but the theory tells us about the probabilities of these things happening. Trajectory is a meaningless concept in quantum mechanics.

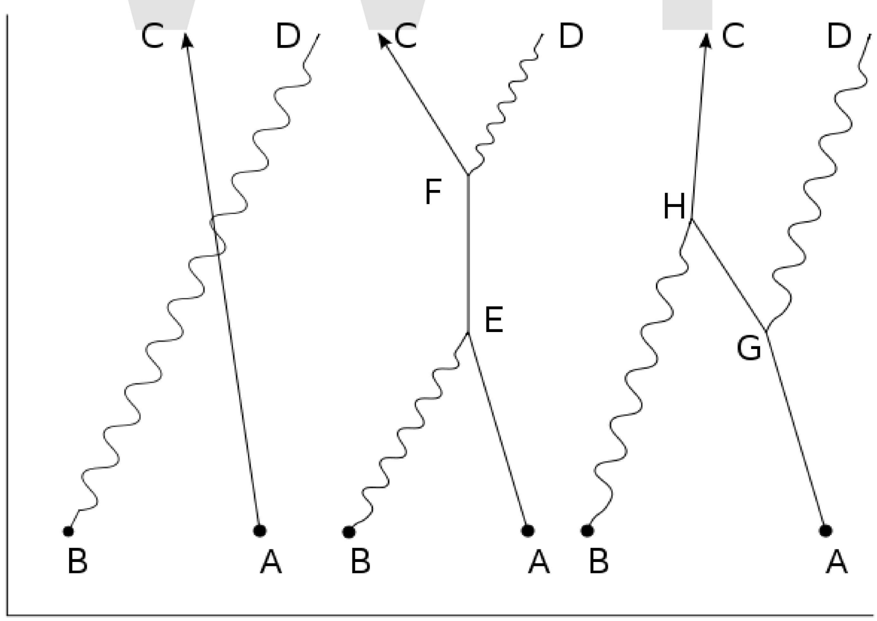
As well as the visual shorthand for the actions Feynman introduces another kind of shorthand for the numerical quantities which tell us about the probabilities. If a photon moves from one place and time – in shorthand, A – to another place and time – shorthand, B – the associated quantity is written in Feynman's shorthand as P(A to B). The similar quantity for an electron moving from C to D is written E(C to D). The quantity which tells us about the probability for the emission or absorption of a photon he calls 'j'. This is related to, but not the same as, the measured electron charge 'e'.

QED is based on the assumption that complex interactions of many electrons and photons can be represented by fitting together a suitable collection of the above three building blocks, and then using the probability-quantities to calculate the probability of any such complex interaction. It turns out that the basic idea of QED can be communicated while making the assumption that the quantities mentioned above are just our everyday probabilities. (A simplification of Feynman's book.) Later on this will be corrected to include specifically quantum mathematics, following Feynman.

The basic rules of probabilities that will be used are that a) if an event can happen in a variety of different ways then its probability is the **sum** of the probabilities of the possible ways and b) if a process involves a number of independent subprocesses then its probability is the **product** of the component probabilities.

Basic constructions

Suppose we start with one electron at a certain place and time (this place and time being given the arbitrary label A) and a photon at another place and time (given the label B). A typical question from a physical standpoint is: 'What is the probability of finding an electron at C (another place and a later time) and a photon at D (yet another place and time)?'. The simplest process to achieve this end is for the electron to move from A to C (an elementary action) and that the photon moves from B to D (another elementary action). From a knowledge of the probabilities of each of these subprocesses – E(A to C) and P(B to D) – then we would expect to calculate the probability of both happening by multiplying them, using rule b) above. This gives a simple estimated answer to our question.



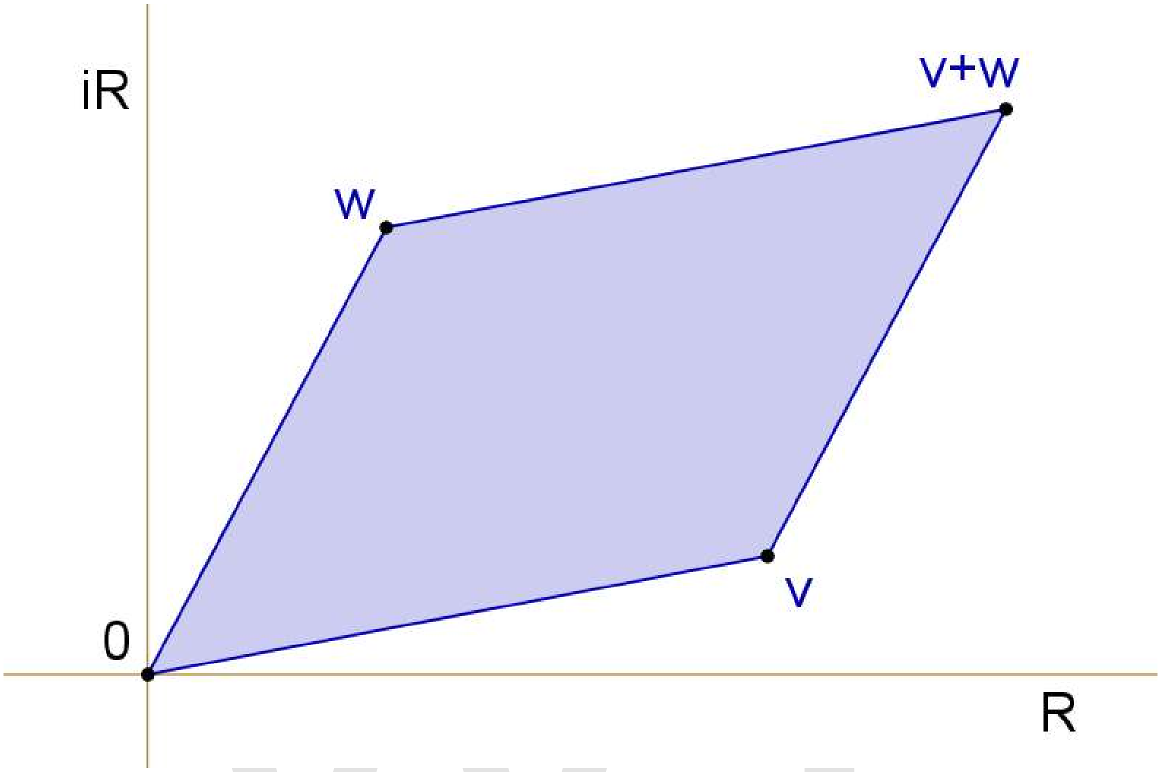
Compton scattering

But there are other ways in which the end result could come about. The electron might move to a place and time E where it absorbs the photon; then move on before emitting another photon at F; then move on to C where it is detected, while the new photon moves on to D. The probability of this complex process can again be calculated by knowing the probabilities of each of the individual actions: three electron actions, two photon actions and two vertexes – one emission and one absorption. We would expect to find the total probability by multiplying the probabilities of each of the actions, for any chosen positions of E and F. We then, using rule a) above, have to add up all these probabilities for all the alternatives for E and F. (This is not elementary in practice, and involves integration.) But there is another possibility: that is that the electron first moves to G where it emits a photon which goes on to D, while the electron moves on to H, where it absorbs the first photon, before moving on to C. Again we can calculate the probability of these possibilities (for all points G and H). We then have a better estimation for the total probability by adding the probabilities of these two possibilities to our original simple estimate. Incidentally the name given to this process of a photon interacting with an electron in this way is Compton Scattering.

There are an *infinite number* of other intermediate processes in which more and more photons are absorbed and/or emitted. For each of these possibilities there is a Feynman diagram describing it. This implies a complex computation for the resulting probabilities, but provided it is the case that the more complicated the diagram the less it contributes to the result, it is only a matter of time and effort to find as accurate an answer as one wants to the original question. This is the basic approach of QED. To calculate the probability of **any** interactive process between electrons and photons it is a matter of first noting, with Feynman diagrams, all the possible ways in which the process can be constructed from the three basic elements. Each diagram involves some calculation involving definite rules to find the associated probability.

That basic scaffolding remains when one moves to a quantum description but some conceptual changes are requested. One is that whereas we might expect in our everyday life that there would be some constraints on the points to which a particle can move, that is **not** true in full quantum electrodynamics. There is a certain possibility of an electron or photon at A moving as a basic action to *any other place and time in the universe*. That includes places that could only be reached at speeds greater than that of light and also *earlier times*. (An electron moving backwards in time can be viewed as a positron moving forward in time.)

Probability amplitudes



Addition of probability amplitudes as complex numbers

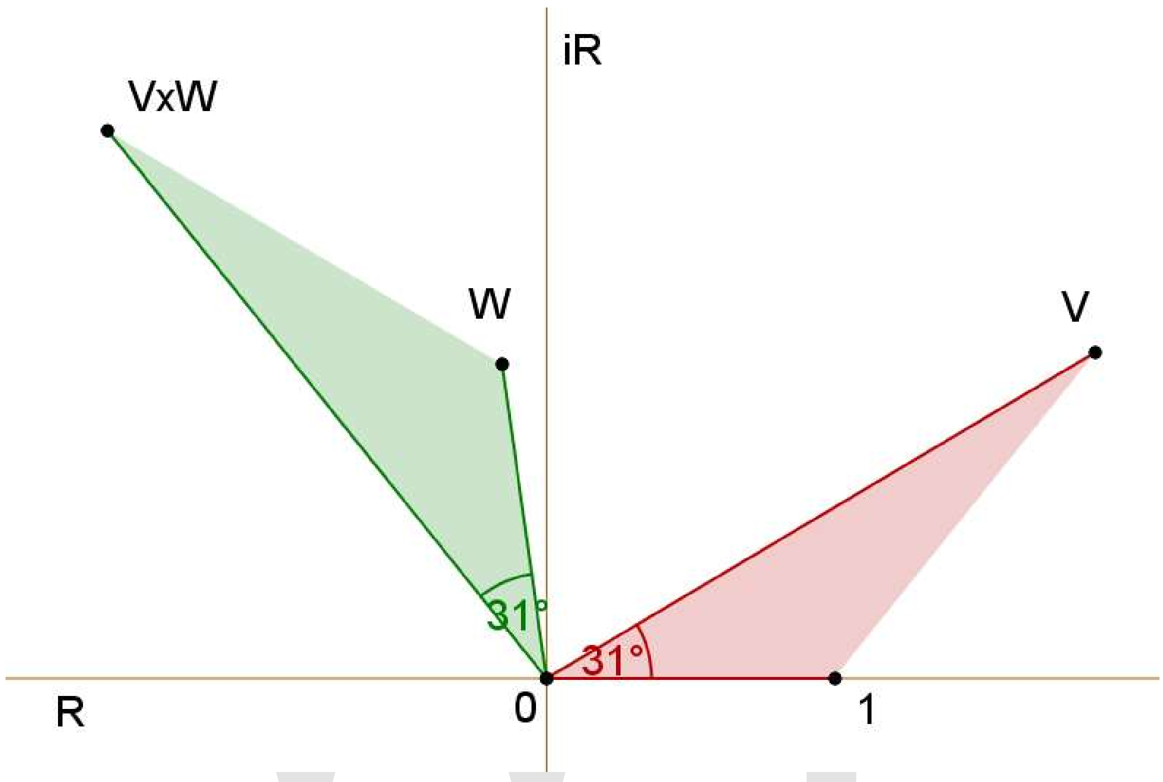
Quantum mechanics introduces an important change on the way probabilities are computed. It has been found that the quantities which we have to use to represent the probabilities are not the usual real numbers we use for probabilities in our everyday world, but complex numbers which are called probability amplitudes. Feynman avoids exposing the reader to the mathematics of complex numbers by using a simple but accurate representation of them as arrows on a piece of paper or screen. (These must not be confused with the arrows of Feynman diagrams which are actually simplified representations in two dimensions of a relationship between points in three dimensions of space and one of time.) The amplitude-arrows are fundamental to the description of the world given by quantum theory. No satisfactory reason has been given for *why* they are needed. But pragmatically we have to accept that they are an essential part of our description of all quantum phenomena. They are related to our everyday ideas of probability by the simple rule that the probability of an event is the **square** of the length of the corresponding amplitude-arrow. So, for a given process, if two probability amplitudes, **v** and **w**, are involved, the probability of the process will be given either by

$$P = |\mathbf{v} + \mathbf{w}|^2$$

or

$$P = |\mathbf{v} \times \mathbf{w}|^2.$$

The rules as regards adding or multiplying, however, are the same as above. But where you would expect to add or multiply probabilities, instead you add or multiply probability amplitudes that now are complex numbers.



Multiplication of probability amplitudes as complex numbers

Addition and multiplication are familiar operations in the theory of complex numbers and are given in the figures. The sum is found as follows. Let the start of the second arrow be at the end of the first. The sum is then a third arrow that goes directly from the start of the first to the end of the second. The product of two arrows is an arrow whose length is the product of the two lengths. The direction of the product is found by adding the angles that each of the two have been turned through relative to a reference direction: that gives the angle that the product is turned relative to the reference direction.

That change, from probabilities to probability amplitudes, complicates the mathematics without changing the basic approach. But that change is still not quite enough because it fails to take into account the fact that both photons and electrons can be polarized, which is to say that their orientation in space and time have to be taken into account. Therefore P(A to B) actually consists of 16 complex numbers, or probability amplitude arrows. There are also some minor changes to do with the quantity "j", which may have to be rotated by a multiple of 90º for some polarizations, which is only of interest for the detailed bookkeeping.

Associated with the fact that the electron can be polarized is another small necessary detail which is connected with the fact that an electron is a Fermion and obeys Fermi-Dirac statistics. The basic rule is that if we have the probability amplitude for a given complex process involving more than one electron, then when we include (as we always must) the complementary Feynman diagram in which we just exchange two electron events, the resulting amplitude is the reverse – the negative – of the first. The simplest case would be two electrons starting at A and B ending at C and D. The amplitude would be calculated as the "difference", E(A to B)xE(C to D) – E(A to C)xE(B to D), where we would expect, from our everyday idea of probabilities, that it would be a sum.

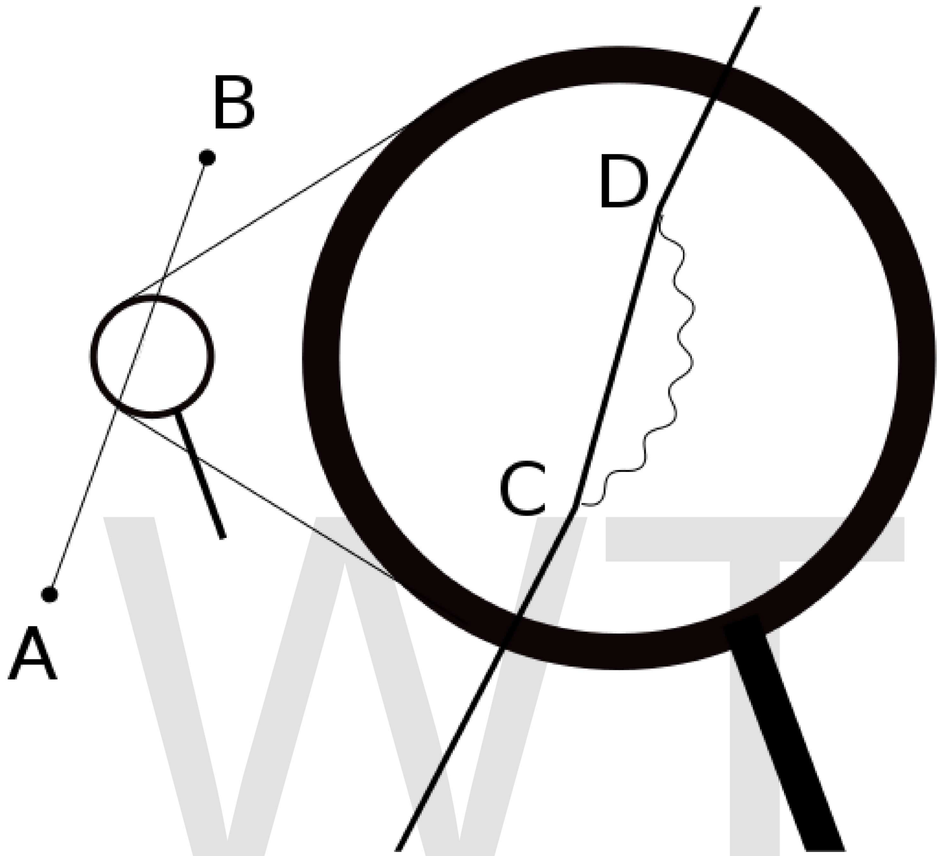
Propagators

Finally, one has to compute P(A to B) and E (C to D) corresponding to the probability amplitudes for the photon and the electron respectively. These are essentially the solutions of the Dirac Equation which describes the behavior of the electron's probability amplitude and the Klein-Gordon equation which describes the behavior of the photon's probability amplitude. These are called Feynman propagators. The translation to a notation commonly used in the standard literature is as follows:

$$P(\text{A to B}) \rightarrow D_F(x_B - x_A), \quad E(\text{C to D}) \rightarrow S_F(x_D - x_C)$$

where a shorthand symbol such as x_A stands for the four real numbers which give the time and position in three dimensions of the point labeled A.

Mass renormalization



Electron self-energy loop

A problem arose historically which held up progress for twenty years: although we start with the assumption of three basic "simple" actions, the rules of the game say that if we want to calculate the probability amplitude for an electron to get from A to B we must take into account **all** the possible ways: all possible Feynman diagrams with those end points. Thus there will be a way in which the electron travels to C, emits a photon there and then absorbs it again at D before moving on to B. Or it could do this kind of thing twice, or more. In short we have a fractal-like situation in which if we look closely at a line it breaks up into a collection of "simple" lines, each of which, if looked at closely, are in turn composed of "simple" lines, and so on *ad infinitum*. This is a very difficult situation to handle. If adding that detail only altered things slightly then it would not have been too bad, but disaster struck when it was found that the simple correction mentioned above led to *infinite* probability amplitudes. In time this problem was "fixed" by the technique of renormalization. However, Feynman himself remained unhappy about it, calling it a "dippy process".

Conclusions

Within the above framework physicists were then able to calculate to a high degree of accuracy some of the properties of electrons, such as the anomalous magnetic dipole moment. However, as Feynman points out, it fails totally to explain why particles such as the electron have the masses they do. "There is no theory that adequately explains these numbers. We use the numbers in all our theories, but we don't understand them – what they are, or where they come from. I believe that from a fundamental point of view, this is a very interesting and serious problem."

Mathematics

Mathematically, QED is an abelian gauge theory with the symmetry group U(1). The gauge field, which mediates the interaction between the charged spin-1/2 fields, is the electromagnetic field. The QED Lagrangian for a spin-1/2 field interacting with the electromagnetic field is given by the real part of

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \ ,$$

where

γ^μ are Dirac matrices;

ψ a bispinor field of spin-1/2 particles (e.g. electron-positron field);

$\bar{\psi} \equiv \psi^\dagger\gamma_0$, called "psi-bar", is sometimes referred to as Dirac adjoint;

$D_\mu = \partial_\mu + ieA_\mu + ieB_\mu$ is the gauge covariant derivative;

e is the coupling constant, equal to the electric charge of the bispinor field;

A_μ is the covariant four-potential of the electromagnetic field generated by the electron itself;

B_μ is the external field imposed by external source;

$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor.

Equations of motion

To begin, substituting the definition of *D* into the Lagrangian gives us:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - e\bar{\psi}\gamma_\mu(A^\mu + B^\mu)\psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

Next, we can substitute this Lagrangian into the Euler-Lagrange equation of motion for a field:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) - \frac{\partial \mathcal{L}}{\partial \psi} = 0 \qquad (2)$$

to find the field equations for QED.

The two terms from this Lagrangian are then:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) = \partial_\mu \left(i\bar{\psi}\gamma^\mu \right)$$
$$\frac{\partial \mathcal{L}}{\partial \psi} = -e\bar{\psi}\gamma_\mu (A^\mu + B^\mu) - m\bar{\psi}.$$

Substituting these two back into the Euler-Lagrange equation (2) results in:

$$i\partial_\mu \bar{\psi}\gamma^\mu + e\bar{\psi}\gamma_\mu (A^\mu + B^\mu) + m\bar{\psi} = 0$$

with complex conjugate:

$$i\gamma^\mu \partial_\mu \psi - e\gamma_\mu (A^\mu + B^\mu)\psi - m\psi = 0.$$

Bringing the middle term to the right-hand side transforms this second equation into:

$$\boxed{i\gamma^\mu \partial_\mu \psi - m\psi = e\gamma_\mu (A^\mu + B^\mu)\psi.}$$

The left-hand side is like the original Dirac equation and the right-hand side is the interaction with the electromagnetic field.

One further important equation can be found by substituting the Lagrangian into another Euler-Lagrange equation, this time for the field, A^μ:

$$\partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right) - \frac{\partial \mathcal{L}}{\partial A_\mu} = 0. \qquad (3)$$

The two terms this time are:

$$\partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right) = \partial_\nu \left(\partial^\mu A^\nu - \partial^\nu A^\mu \right)$$
$$\frac{\partial \mathcal{L}}{\partial A_\mu} = -e\bar{\psi}\gamma^\mu \psi$$

and these two terms, when substituted back into (3) give us:

$$\boxed{\partial_\nu F^{\nu\mu} = e\bar{\psi}\gamma^\mu \psi.}$$

Interaction picture

This theory can be straightforwardly quantized treating bosonic and fermionic sectors as free. This permits to build a set of asymptotic states to start a computation of the probability amplitudes for different processes. In order to be able to do so, we have to compute an evolution operator that, for a given initial state, will give a final state in such a way to have

$$M_{fi} = \langle f|U|i\rangle.$$

This technique is also known as the S-Matrix. Evolution operator is obtained in the interaction picture where time evolution is given by the interaction Hamiltonian. So, from equations above is

$$V = e\int d^3x\bar{\psi}\gamma^\mu\psi A_\mu$$

and so, one has

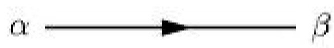
$$U = T\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} dt' V(t')\right]$$

being T the time ordering operator. This evolution operator has only a meaning as a series and what we get here is a perturbation series with a development parameter being fine structure constant. This series is named Dyson series.

Feynman diagrams

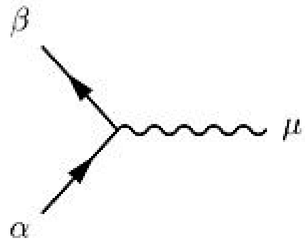
Despite the conceptual clarity of this Feynman approach to QED, almost no textbooks follow him in their presentation. When performing calculations it is much easier to work with the Fourier transforms of the propagators. Quantum physics considers particle's momenta rather than their positions, and it is convenient to think of particles as being created or annihilated when they interact. Feynman diagrams then *look* the same, but the lines have different interpretations. The electron line represents an electron with a given energy and momentum, with a similar interpretation of the photon line. A vertex diagram represents the annihilation of one electron and the creation of another together with the absorption or creation of a photon, each having specified energies and momenta.

Using Wick theorem on the terms of the Dyson series, all the terms of the S-matrix for quantum electrodynamics can be computed through the technique of Feynman diagrams. In this case rules for drawing are the following

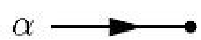


$$\rightarrow \quad \left(\frac{i}{\not{p} - m + i\varepsilon}\right)_{\beta\alpha}$$

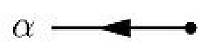
$$\rightarrow \quad \frac{-i\eta_{\mu\nu}}{p^2 + i\varepsilon}$$



$$\rightarrow \quad -ie\gamma^{\mu}_{\beta\alpha}(2\pi)^4\delta^{(4)}(p_1 + p_2 + p_3).$$

Incoming fermion: 

$\rightarrow \quad u_\alpha(\vec{p}, s)$

Incoming antifermion: 

$\rightarrow \quad \overline{v}_\alpha(\vec{p}, s)$

Outgoing fermion: 

$\rightarrow \quad \overline{u}_\alpha(\vec{p}, s)$

Outgoing antifermion: 

$\rightarrow \quad v_\alpha(p, s)$

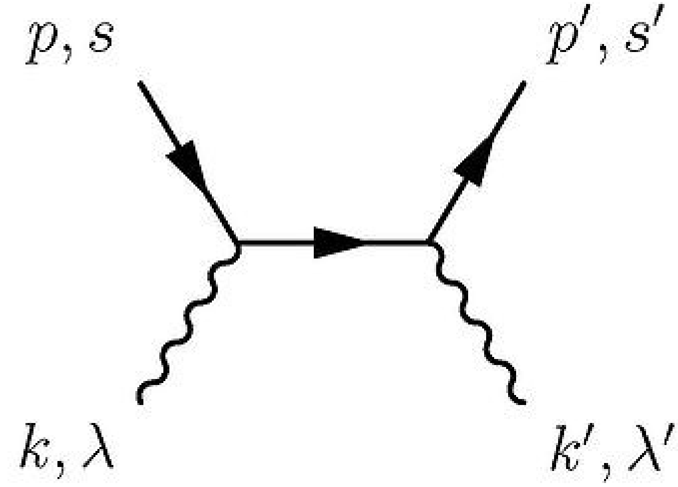
Incoming photon: 

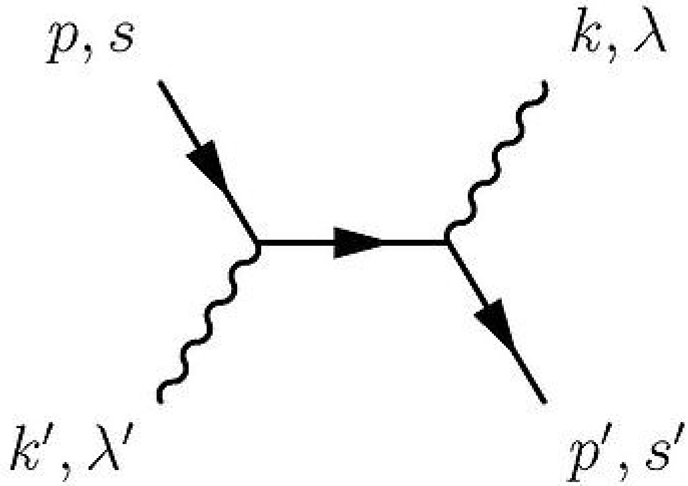
$\rightarrow \quad \epsilon_\mu(\vec{k}, \lambda)$

Outgoing photon: 

$\rightarrow \quad \epsilon_\mu(\vec{k}, \lambda)^*$

To these rules we must add a further one for closed loops that implies an integration on momenta $\int d^4p/(2\pi)^4$. From them, computations of probability amplitudes are straightforwardly given. An example is Compton scattering, with an electron and a photon undergoing elastic scattering. Feynman diagrams are in this case





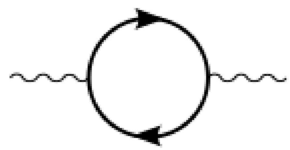
and so we are able to get the corresponding amplitude at the first order of a perturbation series for S-matrix:

$$M_{fi} = (ie)^2 \overline{u}(\vec{p}\,', s') \not{\epsilon}\,'(\vec{k}\,', \lambda')^* \frac{\not{p} + \not{k} + m_e}{(p+k)^2 - m_e^2} \not{\epsilon}(\vec{k}, \lambda) u(\vec{p}, s) + (ie)^2 \overline{u}(\vec{p}\,', s') \not{\epsilon}(\vec{k}, \lambda) \frac{\not{p} - \not{k}' + m_e}{(p-k')^2 - m_e^2} \not{\epsilon}\,'(\vec{k}\,', \lambda')^* u(\vec{p}, s)$$

from which we are able to compute the cross section for this scattering.

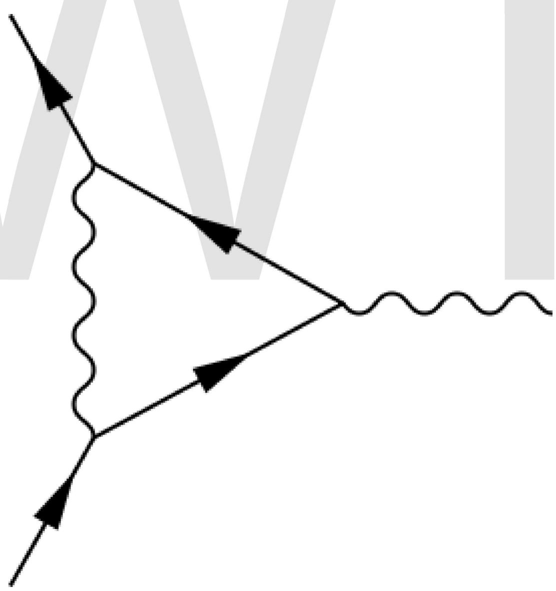
Renormalizability

Higher order terms can be straightforwardly computed for the evolution operator but these terms display diagrams containing the following simpler ones



One-loop contribution to the vacuum polarization function Π

One-loop contribution to the electron self-energy function Σ



One-loop contribution to the vertex function Γ

that, being closed loops, imply the presence of diverging integrals having no mathematical meaning. To overcome this difficulty, a technique like renormalization has been devised, producing finite results in very close agreement with experiments. It is important to note that a criterion for theory being meaningful after renormalization is that the number of diverging diagrams is finite. In this case the theory is said **renormalizable**. The reason for this is that to get observables renormalized one needs a finite number of constants to maintain the predictive value of the theory untouched. This is exactly the case of quantum electrodynamics displaying just three diverging diagrams. This procedure gives observables in very close agreement with experiment as seen e.g. for electron gyromagnetic ratio.

Renormalizability has become an essential criterion for a quantum field theory to be considered as a viable one. All the theories describing fundamental interactions, except gravitation whose quantum counterpart is presently under very active research, are renormalizable theories.

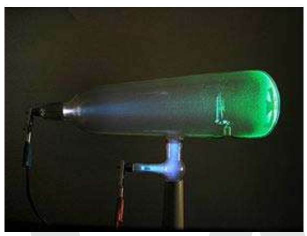
Nonconvergence of series

An argument by Freeman Dyson shows that the radius of convergence of the perturbation series in QED is zero. The basic argument goes as follows: if the coupling constant were negative, this would be equivalent to the Coulomb force constant being negative. This would "reverse" the electromagnetic interaction so that *like* charges would *attract* and *unlike* charges would *repel*. This would render the vacuum unstable against decay into a cluster of electrons on one side of the universe and a cluster of positrons on the other side of the universe. Because the theory is sick for any negative value of the coupling constant, the series do not converge, but are an asymptotic series. This can be taken as a need for a new theory, a problem with perturbation theory, or ignored by taking a "shut-up-and-calculate" approach.

Chapter 4

Electron

Electron



Experiments with a Crookes tube first demonstrated the particle nature of electrons. In this illustration, the profile of the cross-shaped target is projected against the tube face at right by a beam of electrons.

Composition:	Elementary particle
Particle statistics:	Fermionic
Group:	Lepton
Generation:	First
Interaction:	Gravity, Electromagnetic, Weak
Symbol(s):	e^-, β^-
Antiparticle:	Positron (also called antielectron)
Theorized:	Richard Laming (1838–1851), G. Johnstone Stoney (1874) and others.
Discovered:	J. J. Thomson (1897)
Mass:	$9.10938215(45)\times10^{-31}$ kg $5.4857990943(23)\times10^{-4}$ u $[1,822.88850204(77)]^{-1}$ u $0.510998910(13)$ MeV/c^2
Electric charge:	−1 e $-1.602176487(40)\times10^{-19}$ C -4.803×10^{-10} esu
Magnetic moment:	$-1.00115965218111\ \mu_B$

Spin: ½

The **electron** is a subatomic particle carrying a negative electric charge. It has no known components or substructure. Therefore, the electron is generally believed to be an elementary particle. An electron has a mass that is approximately 1/1836 that of the proton. The intrinsic angular momentum (spin) of the electron is a half-integer value in units of $\hbar$, which means that it is a fermion. The antiparticle of the electron is called the positron. The positron is identical to the electron except that it carries electrical and other charges of the opposite sign. When an electron collides with a positron, both particles may either scatter off each other or be totally annihilated, producing a pair (or more) of gamma ray photons. Electrons, which belong to the first generation of the lepton particle family, participate in gravitational, electromagnetic and weak interactions. Electrons, like all matter, have quantum mechanical properties of both particles and waves, so they can collide with other particles and be diffracted like light. However, this duality is best demonstrated in experiments with electrons, due to their tiny mass. Since an electron is a fermion, no two electrons can occupy the same quantum state, in accordance with the Pauli exclusion principle.

The concept of an indivisible amount of electric charge was theorized to explain the chemical properties of atoms, beginning in 1838 by British natural philosopher Richard Laming; the name *electron* was introduced for this charge in 1894 by Irish physicist George Johnstone Stoney. The electron was identified as a particle in 1897 by J. J. Thomson and his team of British physicists.

In many physical phenomena, such as electricity, magnetism, and thermal conductivity, electrons play an essential role. An electron in motion relative to an observer generates a magnetic field, and will be deflected by external magnetic fields. When an electron is accelerated, it can absorb or radiate energy in the form of photons. Electrons, together with atomic nuclei made of protons and neutrons, make up atoms. However, electrons contribute less than 0.06% to an atom's total mass. The attractive Coulomb force between an electron and a proton causes electrons to be bound into atoms. The exchange or sharing of the electrons between two or more atoms is the main cause of chemical bonding.

According to theory, most electrons in the universe were created in the big bang, but they may also be created through beta decay of radioactive isotopes and in high-energy collisions, for instance when cosmic rays enter the atmosphere. Electrons may be destroyed through annihilation with positrons, and may be absorbed during nucleosynthesis in stars. Laboratory instruments are capable of containing and observing individual electrons as well as electron plasma, whereas dedicated telescopes can detect electron plasma in outer space. Electrons have many applications, including welding, cathode ray tubes, electron microscopes, radiation therapy, lasers and particle accelerators.

History

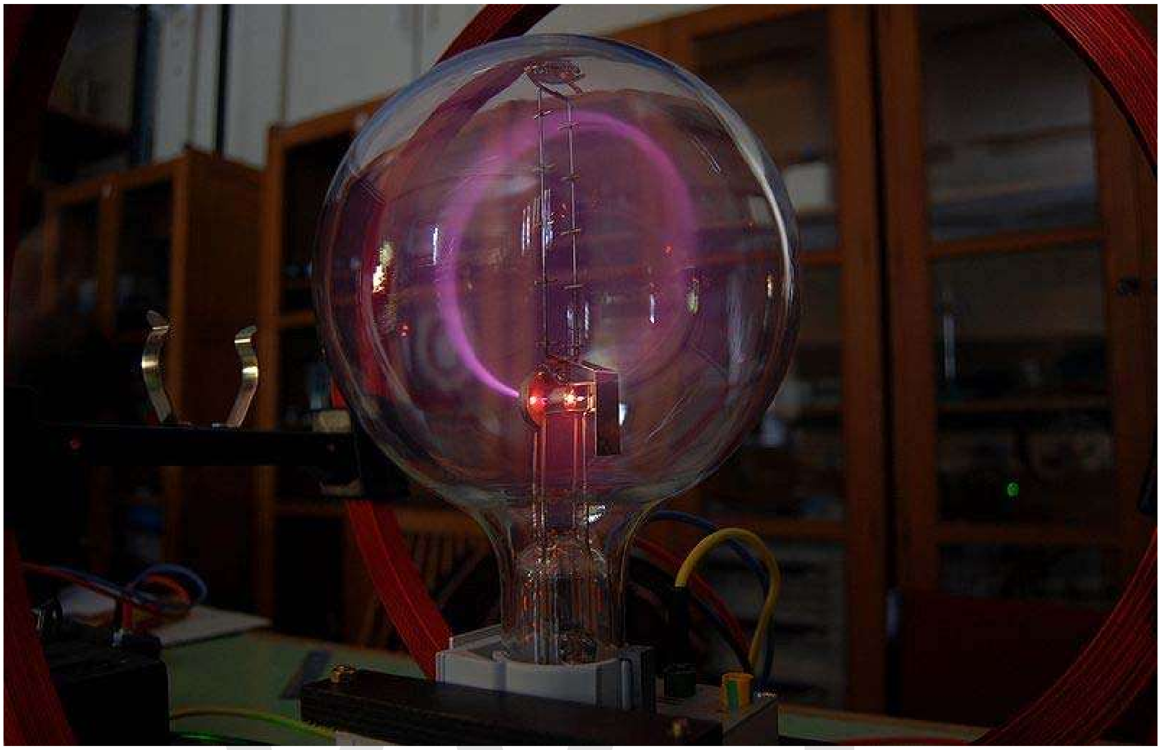
The ancient Greeks noticed that amber attracted small objects when rubbed with fur. Apart from lightning, this phenomenon is humanity's earliest recorded experience with electricity. In his 1600 treatise *De Magnete*, the English scientist William Gilbert coined the New Latin term *electricus*, to refer to this property of attracting small objects after being rubbed. Both *electric* and *electricity* are derived from the Latin *ēlectrum* (also the root of the alloy of the same name), which came from the Greek word ήλεκτρον (*ēlektron*) for amber.

In 1737 C. F. du Fay and Hawksbee independently discovered what they believed to be two kinds of frictional electricity; one generated from rubbing glass, the other from rubbing resin. From this, Du Fay theorized that electricity consists of two electrical fluids, "vitreous" and "resinous", that are separated by friction and that neutralize each other when combined. A decade later Benjamin Franklin proposed that electricity was not from different types of electrical fluid, but the same electrical fluid under different pressures. He gave them the modern charge nomenclature of positive and negative respectively. Franklin thought that the charge carrier was positive.

Between 1838 and 1851, British natural philosopher Richard Laming developed the idea that an atom is composed of a core of matter surrounded by subatomic particles that had unit electric charges. Beginning in 1846, German physicist William Weber theorized that electricity was composed of positively and negatively charged fluids, and their interaction was governed by the inverse square law. After studying the phenomenon of electrolysis in 1874, Irish physicist George Johnstone Stoney suggested that there existed a "single definite quantity of electricity", the charge of a monovalent ion. He was able to estimate the value of this elementary charge *e* by means of Faraday's laws of electrolysis. However, Stoney believed these charges were permanently attached to atoms and could not be removed. In 1881, German physicist Hermann von Helmholtz argued that both positive and negative charges were divided into elementary parts, each of which "behaves like atoms of electricity".

In 1894, Stoney coined the term *electron* to describe these elementary charges, saying, "... an estimate was made of the actual amount of this most remarkable fundamental unit of electricity, for which I have since ventured to suggest the name *electron*". The word *electron* is a combination of the word *electric* and the suffix -*on*, with the latter now used to designate a subatomic particle, such as a proton or neutron.

Discovery



A beam of electrons deflected in a circle by a magnetic field

The German physicist Johann Wilhelm Hittorf undertook the study of electrical conductivity in rarefied gases. In 1869, he discovered a glow emitted from the cathode that increased in size with decrease in gas pressure. In 1876, the German physicist Eugen Goldstein showed that the rays from this glow cast a shadow, and he dubbed them cathode rays. During the 1870s, the English chemist and physicist Sir William Crookes developed the first cathode ray tube to have a high vacuum inside. He then showed that the luminescence rays appearing within the tube carried energy and moved from the cathode to the anode. Furthermore, by applying a magnetic field, he was able to deflect the rays, thereby demonstrating that the beam behaved as though it were negatively charged. In 1879, he proposed that these properties could be explained by what he termed 'radiant matter'. He suggested that this was a fourth state of matter, consisting of negatively charged molecules that were being projected with high velocity from the cathode.

The German-born British physicist Arthur Schuster expanded upon Crookes' experiments by placing metal plates in parallel to the cathode rays and applying an electric potential between the plates. The field deflected the rays toward the positively charged plate, providing further evidence that the rays carried negative charge. By measuring the amount of deflection for a given level of current, in 1890 Schuster was able to estimate the charge-to-mass ratio of the ray components. However, this produced a value that was more than a thousand times greater than what was expected, so little credence was given to his calculations at the time.

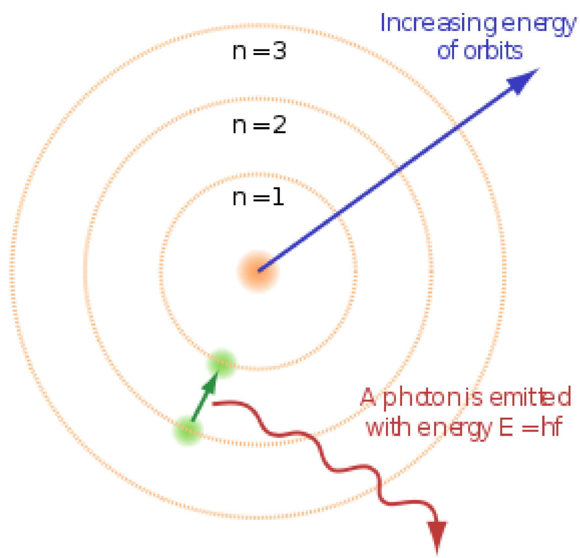
In 1896, the British physicist J. J. Thomson, with his colleagues John S. Townsend and H. A. Wilson, performed experiments indicating that cathode rays really were unique particles, rather than waves, atoms or molecules as was believed earlier. Thomson made good estimates of both the charge e and the mass m, finding that cathode ray particles, which he called "corpuscles," had perhaps one thousandth of the mass of the least massive ion known: hydrogen. He showed that their charge to mass ratio, e/m, was independent of cathode material. He further showed that the negatively charged particles produced by radioactive materials, by heated materials and by illuminated materials were universal. The name electron was again proposed for these particles by the Irish physicist George F. Fitzgerald, and the name has since gained universal acceptance.

While studying naturally fluorescing minerals in 1896, the French physicist Henri Becquerel discovered that they emitted radiation without any exposure to an external energy source. These radioactive materials became the subject of much interest by scientists, including the New Zealand physicist Ernest Rutherford who discovered they emitted particles. He designated these particles alpha and beta, on the basis of their ability to penetrate matter. In 1900, Becquerel showed that the beta rays emitted by radium could be deflected by an electric field, and that their mass-to-charge ratio was the same as for cathode rays. This evidence strengthened the view that electrons existed as components of atoms.

The electron's charge was more carefully measured by the American physicist Robert Millikan in his oil-drop experiment of 1909, the results of which he published in 1911. This experiment used an electric field to prevent a charged droplet of oil from falling as a result of gravity. This device could measure the electric charge from as few as 1–150 ions with an error margin of less than 0.3%. Comparable experiments had been done earlier by Thomson's team, using clouds of charged water droplets generated by electrolysis, and in 1911 by Abram Ioffe, who independently obtained the same result as Millikan using charged microparticles of metals, then published his results in 1913. However, oil drops were more stable than water drops because of their slower evaporation rate, and thus more suited to precise experimentation over longer periods of time.

Around the beginning of the twentieth century, it was found that under certain conditions a fast moving charged particle caused a condensation of supersaturated water vapor along its path. In 1911, Charles Wilson used this principle to devise his cloud chamber, allowing the tracks of charged particles, such as fast-moving electrons, to be photographed.

Atomic theory



The Bohr model of the atom, showing states of electron with energy quantized by the number n. An electron dropping to a lower orbit emits a photon equal to the energy difference between the orbits.

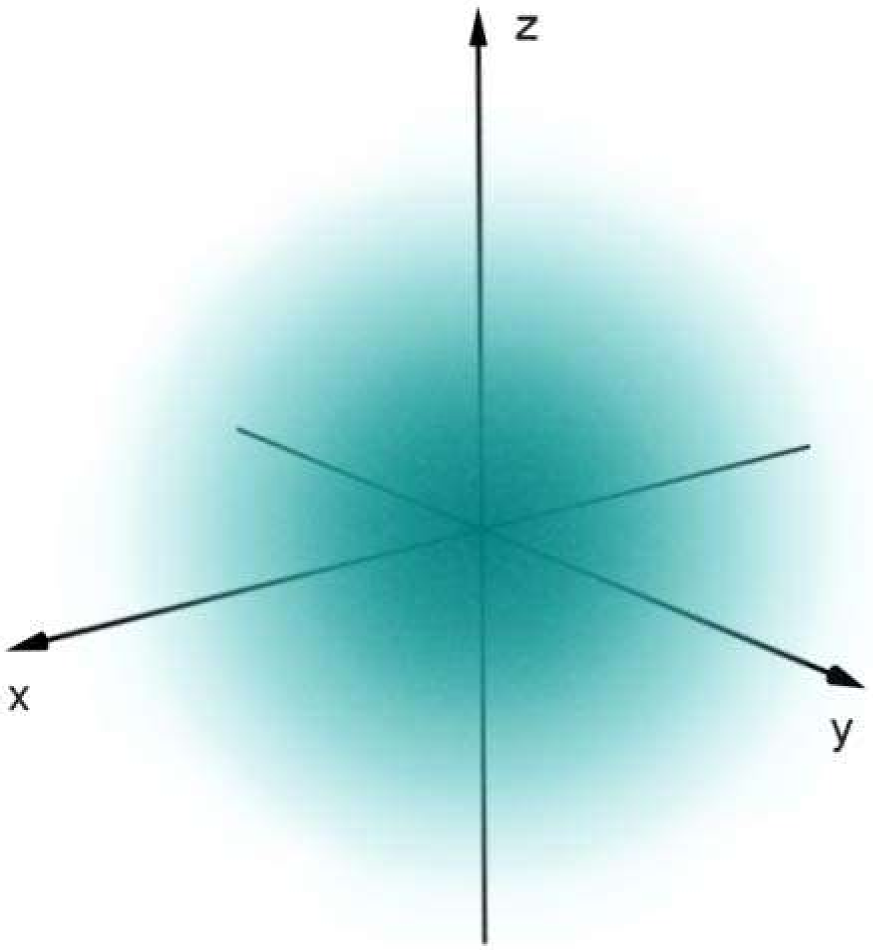
By 1914, experiments by physicists Ernest Rutherford, Henry Moseley, James Franck and Gustav Hertz had largely established the structure of an atom as a dense nucleus of positive charge surrounded by lower-mass electrons. In 1913, Danish physicist Niels Bohr postulated that electrons resided in quantized energy states, with the energy determined by the angular momentum of the electron's orbits about the nucleus. The electrons could move between these states, or orbits, by the emission or absorption of photons at specific frequencies. By means of these quantized orbits, he accurately explained the spectral lines of the hydrogen atom. However, Bohr's model failed to account for the relative intensities of the spectral lines and it was unsuccessful in explaining the spectra of more complex atoms.

Chemical bonds between atoms were explained by Gilbert Newton Lewis, who in 1916 proposed that a covalent bond between two atoms is maintained by a pair of electrons shared between them. Later, in 1923, Walter Heitler and Fritz London gave the full explanation of the electron-pair formation and chemical bonding in terms of quantum mechanics. In 1919, the American chemist Irving Langmuir elaborated on the Lewis' static model of the atom and suggested that all electrons were distributed in successive "concentric (nearly) spherical shells, all of equal thickness". The shells were, in turn, divided by him in a number of cells each containing one pair of electrons. With this model Langmuir was able to qualitatively explain the chemical properties of all elements in the periodic table, which were known to largely repeat themselves according to the periodic law.

In 1924, Austrian physicist Wolfgang Pauli observed that the shell-like structure of the atom could be explained by a set of four parameters that defined every quantum energy state, as long as each state was inhabited by no more than a single electron. (This prohibition against more than one electron occupying the same quantum energy state became known as the Pauli exclusion principle.) The physical mechanism to explain the fourth parameter, which had two distinct possible values, was provided by the Dutch physicists Abraham Goudsmith and George Uhlenbeck when they suggested that an electron, in addition to the angular momentum of its orbit, could possess an intrinsic angular momentum. This property became known as spin, and explained the previously mysterious splitting of spectral lines observed with a high-resolution spectrograph; this phenomenon is known as fine structure splitting.

Quantum mechanics

In his 1924 dissertation *Recherches sur la théorie des quanta* (Research on Quantum Theory), French physicist Louis de Broglie hypothesized that all matter possesses a De Broglie wave similar to light. That is, under the appropriate conditions, electrons and other matter would show properties of either particles or waves. The corpuscular properties of a particle are demonstrated when it is shown to have a localized position in space along its trajectory at any given moment. Wave-like nature is observed, for example, when a beam of light is passed through parallel slits and creates interference patterns. In 1927, the interference effect was demonstrated with a beam of electrons by English physicist George Paget Thomson with a thin metal film and by American physicists Clinton Davisson and Lester Germer using a crystal of nickel.



Orbital s ($\ell = 0$, $m_\ell = 0$)

In quantum mechanics, the behavior of an electron in an atom is described by an orbital, which is a probability distribution rather than an orbit. In the figure, the shading indicates the relative probability to "find" the electron, having the energy corresponding to the given quantum numbers, at that point.

The success of de Broglie's prediction led to the publication, by Erwin Schrödinger in 1926, of the Schrödinger equation that successfully describes how electron waves propagated. Rather than yielding a solution that determines the location of an electron over time, this wave equation can be used to predict the probability of finding an electron near a position. This approach was later called quantum mechanics, which provided an extremely close derivation to the energy states of an electron in a hydrogen atom. Once spin and the interaction between multiple electrons were considered, quantum mechanics allowed the configuration of electrons in atoms with higher atomic numbers than hydrogen to be successfully predicted.

In 1928, building on Wolfgang Pauli's work, Paul Dirac produced a model of the electron - the Dirac equation, consistent with relativity theory, by applying relativistic and symmetry considerations to the hamiltonian formulation of the quantum mechanics of the electro-magnetic field. In order to resolve some problems within his relativistic equation, in 1930 Dirac developed a model of the vacuum as an infinite sea of particles having negative energy, which was dubbed the Dirac sea. This led him to predict the existence of a positron, the antimatter counterpart of the electron. This particle was discovered in 1932 by Carl D. Anderson, who proposed calling standard electrons *negatrons*, and using *electron* as a generic term to describe both the positively and negatively charged variants. This usage of the term 'negatron' is still occasionally encountered today, and it may be shortened to 'negaton'.

In 1947 Willis Lamb, working in collaboration with graduate student Robert Rutherford, found that certain quantum states of hydrogen atom, which should have the same energy, were shifted in relation to each other, the difference being the Lamb shift. About the same time, Polykarp Kusch, working with Henry M. Foley, discovered the magnetic moment of the electron is slightly larger than predicted by Dirac's theory. This small difference was later called anomalous magnetic dipole moment of the electron. To resolve these issues, a refined theory called quantum electrodynamics was developed by Sin-Itiro Tomonaga, Julian Schwinger and Richard P. Feynman in the late 1940s.

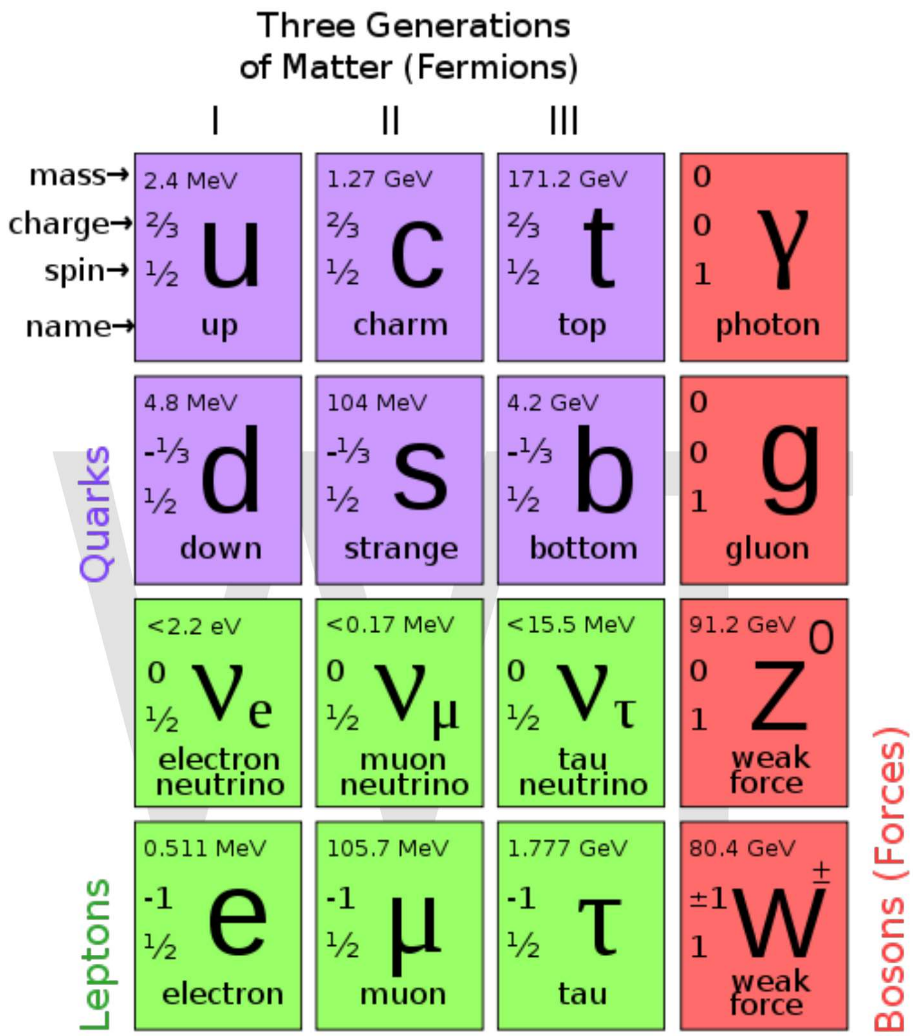
Particle accelerators

With the development of the particle accelerator during the first half of the twentieth century, physicists began to delve deeper into the properties of subatomic particles. The first successful attempt to accelerate electrons using Electromagnetic induction was made in 1942 by Donald Kerst. His initial betatron reached energies of 2.3 MeV, while subsequent betatrons achieved 300 MeV. In 1947, synchrotron radiation was discovered with a 70 MeV electron synchrotron at General Electric. This radiation was caused by the acceleration of electrons, moving near the speed of light, through a magnetic field.

With a beam energy of 1.5 GeV, the first high-energy particle collider was ADONE, which began operations in 1968. This device accelerated electrons and positrons in opposite directions, effectively doubling the energy of their collision when compared to striking a static target with an electron. The Large Electron-Positron Collider (LEP) at CERN, which was operational from 1989 to 2000, achieved collision energies of 209 GeV and made important measurements for the Standard Model of particle physics.

Characteristics

Classification



Standard Model of elementary particles. The electron is at lower left.

In the Standard Model of particle physics, electrons belong to the group of subatomic particles called leptons, which are believed to be fundamental or elementary particles. Electrons have the lowest mass of any charged lepton (or electrically charged particle of any type) and belong to the first-generation of fundamental particles. The second and third generation contain charged leptons, the muon and the tau, which are identical to the electron in charge, spin and interactions, but are more massive. Leptons differ from the other basic constituent of matter, the quarks, by their lack of strong interaction. All members of the lepton group are fermions, because they all have half-odd integer spin; the electron has spin $\frac{1}{2}$.

Fundamental properties

The invariant mass of an electron is approximately 9.109×10^{-31} kilogram, or 5.489×10^{-4} atomic mass unit. On the basis of Einstein's principle of mass–energy equivalence, this mass corresponds to a rest energy of 0.511 MeV. The ratio between the mass of a proton and that of an electron is about 1836. Astronomical measurements show that the proton-to-electron mass ratio has held the same value for at least half the age of the universe, as is predicted by the Standard Model.

Electrons have an electric charge of -1.602×10^{-19} coulomb, which is used as a standard unit of charge for subatomic particles. Within the limits of experimental accuracy, the electron charge is identical to the charge of a proton, but with the opposite sign. As the symbol *e* is used for the elementary charge, the electron is commonly symbolized by e^-, where the minus sign indicates the negative charge. The positron is symbolized by e^+ because it has the same properties as the electron but with a positive rather than negative charge.

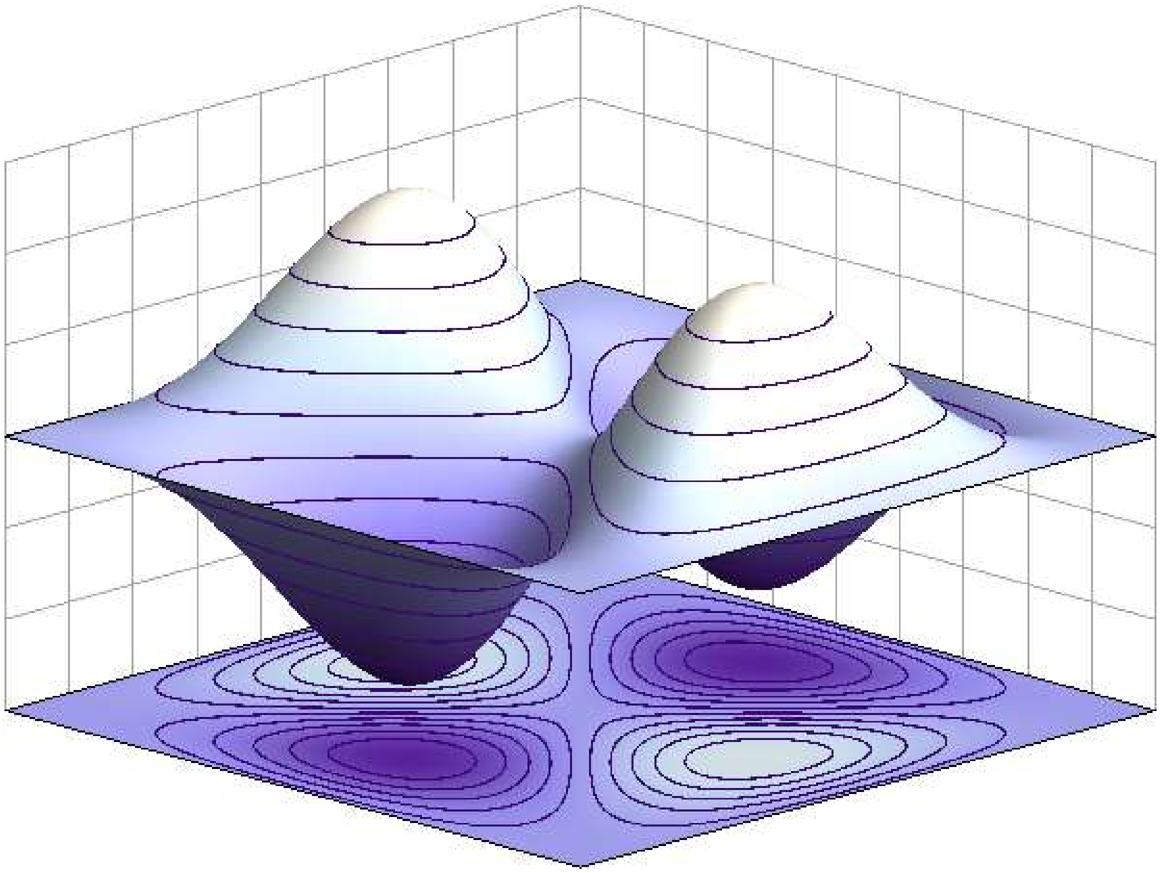
The electron has an intrinsic angular momentum or spin of $\frac{1}{2}$. This property is usually stated by referring to the electron as a spin-$\frac{1}{2}$ particle. For such particles the spin magnitude is $\frac{\sqrt{3}}{2}\hbar$. while the result of the measurement of a projection of the spin on any axis can only be $\pm\frac{\hbar}{2}$. In addition to spin, the electron has an intrinsic magnetic moment along its spin axis. It is approximately equal to one Bohr magneton, which is a physical constant equal to $9.274\ 009\ 15(23)\times10^{-24}$ joules per tesla. The orientation of the spin with respect to the momentum of the electron defines the property of elementary particles known as helicity.

The electron has no known substructure. Hence, it is defined or assumed to be a point particle with a point charge and no spatial extent. Observation of a single electron in a Penning trap shows the upper limit of the particle's radius is 10^{-22} meters. There *is* a physical constant called the "classical electron radius", with the much larger value of 2.8179×10^{-15} m. However, the terminology comes from a simplistic calculation that ignores the effects of quantum mechanics; in reality, the so-called classical electron radius has little to do with the true fundamental structure of the electron.

There are elementary particles that spontaneously decay into less massive particles. An example is the muon, which decays into an electron, a neutrino and an antineutrino, with a mean lifetime of 2.2×10^{-6} seconds. However, the electron is thought to be stable on theoretical grounds: the electron is the least massive particle with non-zero electric charge, so its decay would violate charge conservation. The experimental lower bound for the electron's mean lifetime is 4.6×10^{26} years, at a 90% confidence level.

Quantum properties

As with all particles, electrons can act as waves. This is called the wave–particle duality and can be demonstrated using the double-slit experiment. The wave-like nature of the electron allows it to pass through two parallel slits simultaneously, rather than just one slit as would be the case for a classical particle. In quantum mechanics, the wave-like property of one particle can be described mathematically as a complex-valued function, the wave function, commonly denoted by the Greek letter psi (ψ). When the absolute value of this function is squared, it gives the probability that a particle will be observed near a location—a probability density.



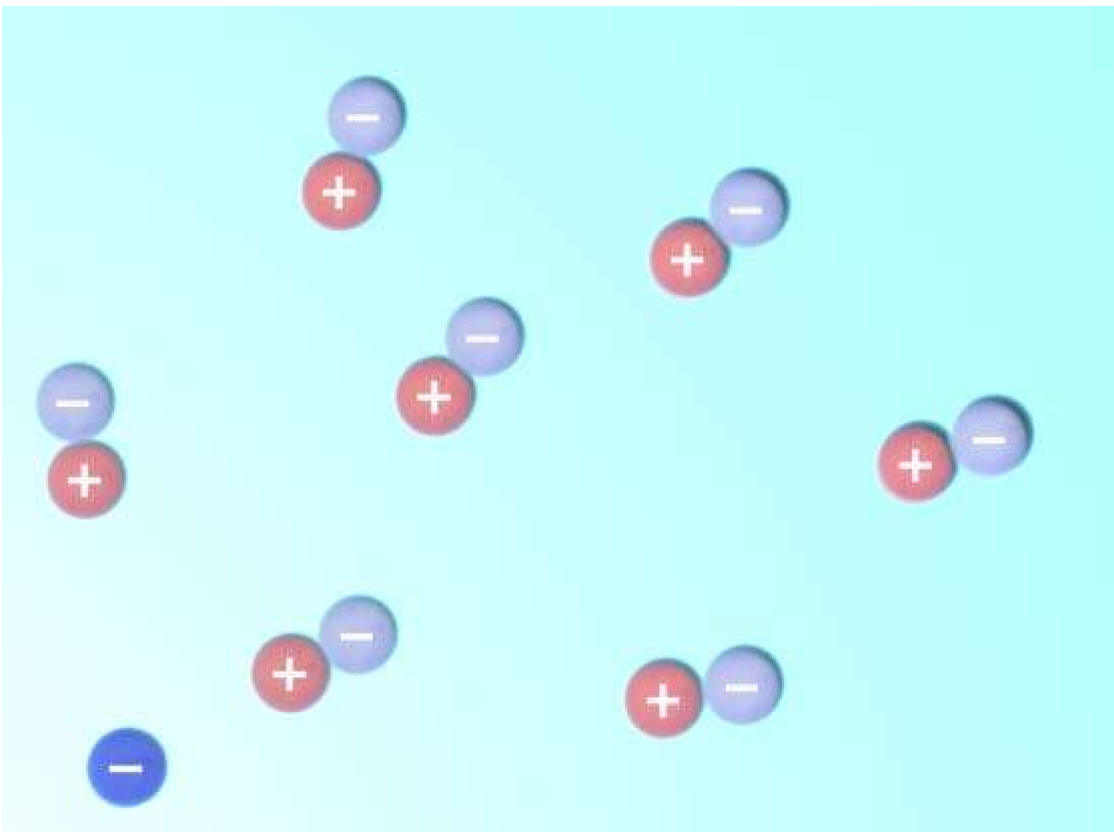
Example of an antisymmetric wave function for a quantum state of two identical fermions in a 2-dimensional box. If the particles swap position, the wave function inverts its sign.

Electrons are identical particles because they cannot be distinguished from each other by their intrinsic physical properties. In quantum mechanics, this means that a pair of interacting electrons must be able to swap positions without an observable change to the state of the system. The wave function of fermions, including electrons, is antisymmetric, meaning that it changes sign when two electrons are swapped; that is, $\psi(r_1, r_2) = -\psi(r_2, r_1)$, where the variables r_1 and r_2 correspond to the first and second electrons, respectively. Since the absolute value is not changed by a sign swap, this corresponds to equal probabilities. Bosons, such as the photon, have symmetric wave functions instead.

In the case of antisymmetry, solutions of the wave equation for interacting electrons result in a zero probability that each pair will occupy the same location or state. This is responsible for the Pauli exclusion principle, which precludes any two electrons from occupying the same quantum state. This principle explains many of the properties of electrons. For example, it causes groups of bound electrons to occupy different orbitals in an atom, rather than all overlapping each other in the same orbit.

Virtual particles

Physicists believe that empty space may be continually creating pairs of virtual particles, such as a positron and electron, which rapidly annihilate each other shortly thereafter. The combination of the energy variation needed to create these particles, and the time during which they exist, fall under the threshold of detectability expressed by the Heisenberg uncertainty relation, $\Delta E \bullet \Delta t \geq \hbar$. In effect, the energy needed to create these virtual particles, ΔE, can be "borrowed" from the vacuum for a period of time, Δt, so that their product is no more than the reduced Planck constant, $\hbar \approx 6.6\times10^{-16}$ eV•s. Thus, for a virtual electron, Δt is at most 1.3×10^{-21} s.



A schematic depiction of virtual electron–positron pairs appearing at random near an electron (at lower left)

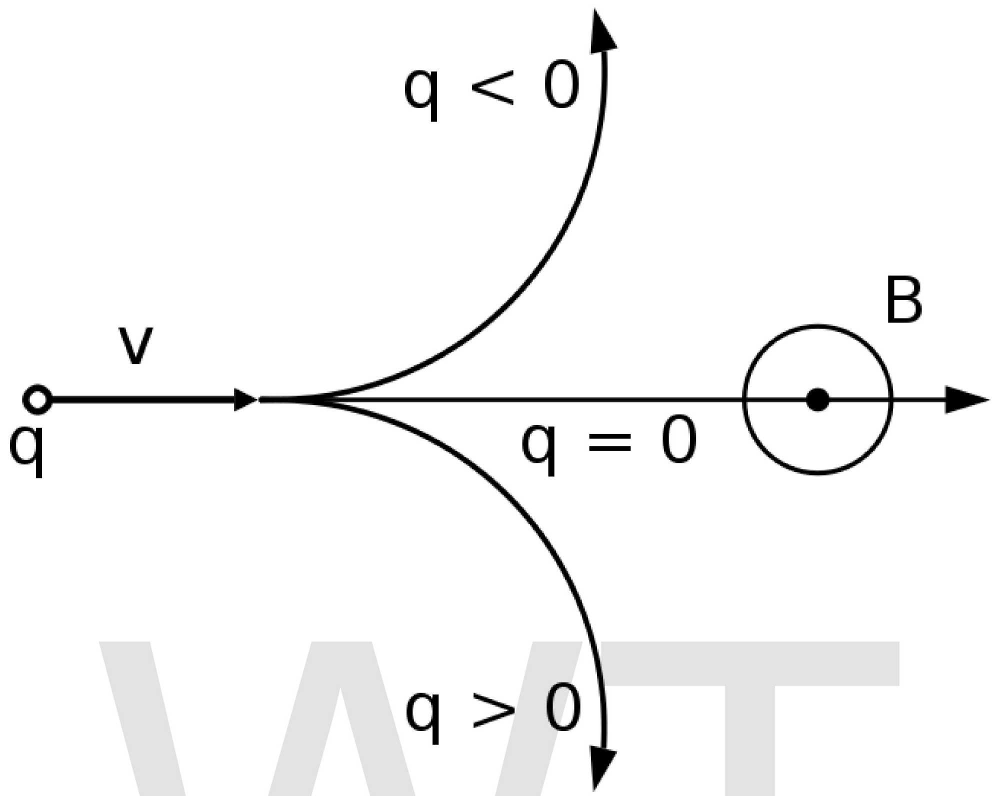
While an electron–positron virtual pair is in existence, the coulomb force from the ambient electric field surrounding an electron causes a created positron to be attracted to the original electron, while a created electron experiences a repulsion. This causes what is called vacuum polarization. In effect, the vacuum behaves like a medium having a dielectric permittivity more than unity. Thus the effective charge of an electron is actually smaller than its true value, and the charge decreases with increasing distance from the electron. This polarization was confirmed experimentally in 1997 using the Japanese TRISTAN particle accelerator. Virtual particles cause a comparable shielding effect for the mass of the electron.

The interaction with virtual particles also explains the small (about 0.1%) deviation of the intrinsic magnetic moment of the electron from the Bohr magneton (the anomalous magnetic moment). The extraordinarily precise agreement of this predicted difference with the experimentally determined value is viewed as one of the great achievements of quantum electrodynamics.

In classical physics, the angular momentum and magnetic moment of an object depend upon its physical dimensions. Hence, the concept of a dimensionless electron possessing these properties might seem inconsistent. The apparent paradox can be explained by the formation of virtual photons in the electric field generated by the electron. These photons cause the electron to shift about in a jittery fashion (known as zitterbewegung), which results in a net circular motion with precession. This motion produces both the spin and the magnetic moment of the electron. In atoms, this creation of virtual photons explains the Lamb shift observed in spectral lines.

Interaction

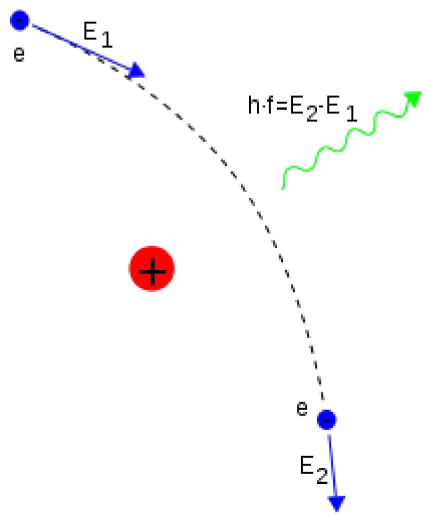
An electron generates an electric field that exerts an attractive force on a particle with a positive charge, such as the proton, and a repulsive force on a particle with a negative charge. The strength of this force is determined by Coulomb's inverse square law. When an electron is in motion, it generates a magnetic field. The Ampère-Maxwell law relates the magnetic field to the mass motion of electrons (the current) with respect to an observer. It is this property of induction which supplies the magnetic field that drives an electric motor. The electromagnetic field of an arbitrary moving charged particle is expressed by the Liénard–Wiechert potentials, which are valid even when the particle's speed is close to that of light (relativistic).



A particle with charge *q* (at left) is moving with velocity *v* through a magnetic field *B* that is oriented toward the viewer. For an electron, *q* is negative so it follows a curved trajectory toward the top.

When an electron is moving through a magnetic field, it is subject to the Lorentz force that exerts an influence in a direction perpendicular to the plane defined by the magnetic field and the electron velocity. This centripetal force causes the electron to follow a helical trajectory through the field at a radius called the gyroradius. The acceleration from this curving motion induces the electron to radiate energy in the form of synchrotron radiation. The energy emission in turn causes a recoil of the electron, known as the Abraham-Lorentz-Dirac force, which creates a friction that slows the electron. This force is caused by a back-reaction of the electron's own field upon itself.

In quantum electrodynamics the electromagnetic interaction between particles is mediated by photons. An isolated electron that is not undergoing acceleration is unable to emit or absorb a real photon; doing so would violate conservation of energy and momentum. Instead, virtual photons can transfer momentum between two charged particles. It is this exchange of virtual photons that, for example, generates the Coulomb force. Energy emission can occur when a moving electron is deflected by a charged particle, such as a proton. The acceleration of the electron results in the emission of Bremsstrahlung radiation.



Here, Bremsstrahlung is produced by an electron *e* deflected by the electric field of an atomic nucleus. The energy change $E_2 - E_1$ determines the frequency *f* of the emitted photon.

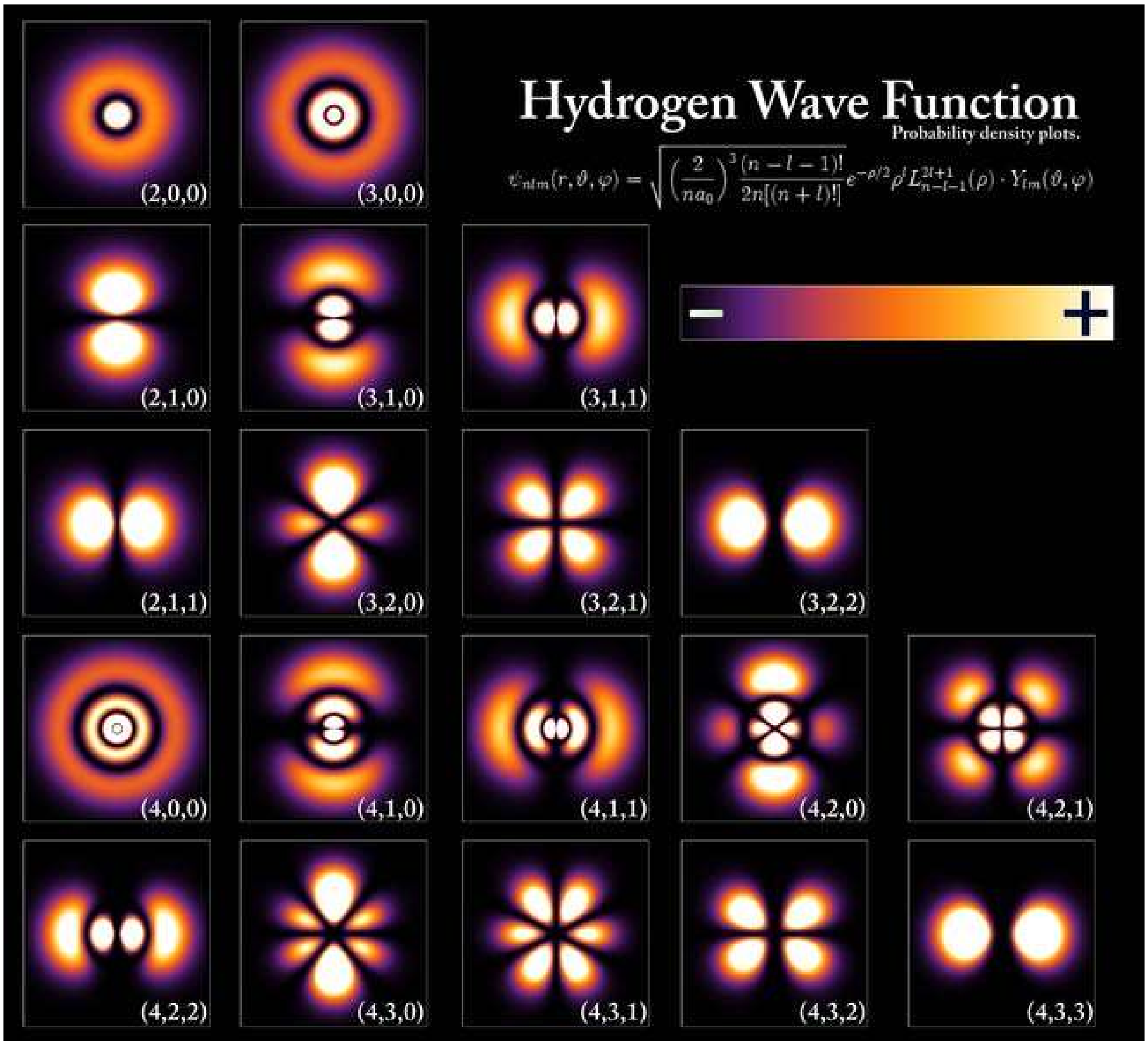
An inelastic collision between a photon (light) and a solitary (free) electron is called Compton scattering. This collision results in a transfer of momentum and energy between the particles, which modifies the wavelength of the photon by an amount called the Compton shift. The maximum magnitude of this wavelength shift is $h/m_e c$, which is known as the Compton wavelength. For an electron, it has a value of 2.43×10^{-12} m. When the wavelength of the light is long (for instance, the wavelength of the visible light is 0.4–0.7 μm) the wavelength shift becomes negligible. Such interaction between the light and free electrons is called Thomson scattering or Linear Thomson scattering.

The relative strength of the electromagnetic interaction between two charged particles, such as an electron and a proton, is given by the fine-structure constant. This value is a dimensionless quantity formed by the ratio of two energies: the electrostatic energy of attraction (or repulsion) at a separation of one Compton wavelength, and the rest energy of the charge. It is given by $\alpha \approx 7.297353\times10^{-3}$, which is approximately equal to $^1/_{137}$.

When electrons and positrons collide, they annihilate each other, giving rise to two or more gamma ray photons. If the electron and positron have negligible momentum, a positronium atom can form before annihilation results in two or three gamma ray photons totalling 1.022 MeV. On the other hand, high-energy photons may transform into an electron and a positron by a process called pair production, but only in the presence of a nearby charged particle, such as a nucleus.

In the theory of electroweak interaction, the left-handed component of electron's wavefunction forms a weak isospin doublet with the electron neutrino. This means that during weak interactions, electron neutrinos behave like electrons. Either member of this doublet can undergo a charged current interaction by emitting or absorbing a W and be converted into the other member. Charge is conserved during this reaction because the W boson also carries a charge, canceling out any net change during the transmutation. Charged current interactions are responsible for the phenomenon of beta decay in a radioactive atom. Both the electron and electron neutrino can undergo a neutral current interaction via a Z^0 exchange, and this is responsible for neutrino-electron elastic scattering.

Atoms and molecules



Probability densities for the first few hydrogen atom orbitals, seen in cross-section. The energy level of a bound electron determines the orbital it occupies, and the color reflects the probability to find the electron at a given position.

An electron can be *bound* to the nucleus of an atom by the attractive Coulomb force. A system of several electrons bound to a nucleus is called an atom. If the number of electrons is different from the nucleus' electrical charge, such an atom is called an ion. The wave-like behavior of a bound electron is described by a function called an atomic orbital. Each orbital has its own set of quantum numbers such as energy, angular momentum and projection of angular momentum, and only a discrete set of these orbitals exist around the nucleus. According to the Pauli exclusion principal each orbital can be occupied by up to two electrons, which must differ in their spin quantum number.

Electrons can transfer between different orbitals by the emission or absorption of photons with an energy that matches the difference in potential. Other methods of orbital transfer include collisions with particles, such as electrons, and the Auger effect. In order to escape the atom, the energy of the electron must be increased above its binding energy to the atom. This occurs, for example, with the photoelectric effect, where an incident photon exceeding the atom's ionization energy is absorbed by the electron.

The orbital angular momentum of electrons is quantized. Because the electron is charged, it produces an orbital magnetic moment that is proportional to the angular momentum. The net magnetic moment of an atom is equal to the vector sum of orbital and spin magnetic moments of all electrons and the nucleus. The nuclear magnetic moment is, however, negligible in comparison to the effect from the electrons. The magnetic moments of the electrons that occupy the same orbital (so called, paired electrons) cancel each other out.

The chemical bond between atoms occurs as a result of electromagnetic interactions, as described by the laws of quantum mechanics. The strongest bonds are formed by the sharing or transfer of electrons between atoms, allowing the formation of molecules. Within a molecule, electrons move under the influence of several nuclei, and occupy molecular orbitals; much as they can occupy atomic orbitals in isolated atoms. A fundamental factor in these molecular structures is the existence of electron pairs. These are electrons with opposed spins, allowing them to occupy the same molecular orbital without violating the Pauli exclusion principle (much like in atoms). Different molecular orbitals have different spatial distribution of the electron density. For instance, in bonded pairs (i.e. in the pairs that actually bind atoms together) electrons can be found with the maximal probability in a relatively small volume between the nuclei. On the contrary, in non-bonded pairs electrons are distributed in a large volume around nuclei.

Conductivity

A lightning discharge consists primarily of a flow of electrons. The electric potential needed for lightning may be generated by a triboelectric effect.

If a body has more or fewer electrons than are required to balance the positive charge of the nuclei, then that object has a net electric charge. When there is an excess of electrons, the object is said to be negatively charged. When there are fewer electrons than the number of protons in nuclei, the object is said to be positively charged. When the number of electrons and the number of protons are equal, their charges cancel each other and the object is said to be electrically neutral. A macroscopic body can develop an electric charge through rubbing, by the triboelectric effect.

Independent electrons moving in vacuum are termed *free* electrons. Electrons in metals also behave as if they were free. In reality the particles that are commonly termed electrons in metals and other solids are quasi-electrons—quasi-particles, which have the same electrical charge, spin and magnetic moment as real electrons but may have a different mass. When free electrons—both in vacuum and metals—move, they produce a net flow of charge called an electric current, which generates a magnetic field. Likewise a current can be created by a changing magnetic field. These interactions are described mathematically by Maxwell's equations.

At a given temperature, each material has an electrical conductivity that determines the value of electric current when an electric potential is applied. Examples of good conductors include metals such as copper and gold, whereas glass and Teflon are poor conductors. In any dielectric material, the electrons remain bound to their respective atoms and the material behaves as an insulator. Most semiconductors have a variable level of conductivity that lies between the extremes of conduction and insulation. On the other hand, metals have an electronic band structure containing partially filled electronic bands. The presence of such bands allows electrons in metals to behave as if they were free or delocalized electrons. These electrons are not associated with specific atoms, so when an electric field is applied, they are free to move like a gas (called Fermi gas) through the material much like free electrons.

Because of collisions between electrons and atoms, the drift velocity of electrons in a conductor is on the order of millimeters per second. However, the speed at which a change of current at one point in the material causes changes in currents in other parts of the material, the velocity of propagation, is typically about 75% of light speed. This occurs because electrical signals propagate as a wave, with the velocity dependent on the dielectric constant of the material.

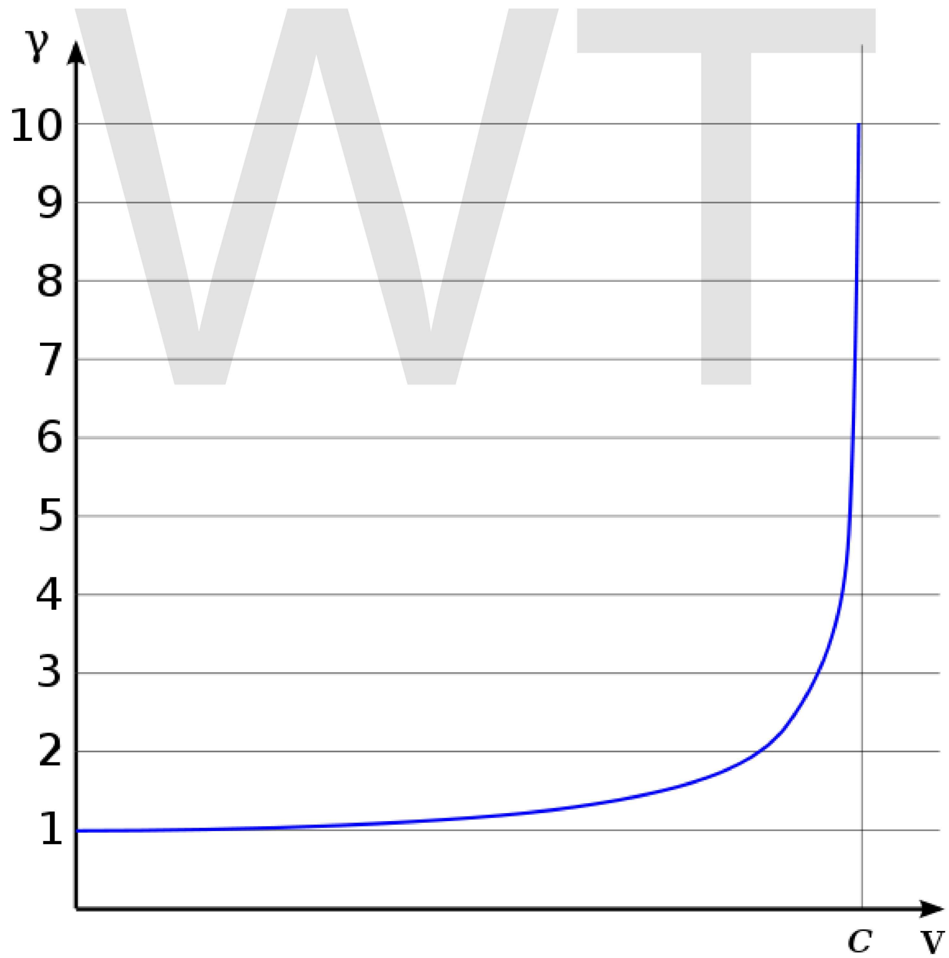
Metals make relatively good conductors of heat, primarily because the delocalized electrons are free to transport thermal energy between atoms. However, unlike electrical conductivity, the thermal conductivity of a metal is nearly independent of temperature. This is expressed mathematically by the Wiedemann-Franz law, which states that the ratio of thermal conductivity to the electrical conductivity is proportional to the temperature. The thermal disorder in the metallic lattice increases the electrical resistivity of the material, producing a temperature dependence for electrical current.

When cooled below a point called the critical temperature, materials can undergo a phase transition in which they lose all resistivity to electrical current, in a process known as superconductivity. In BCS theory, this behavior is modeled by pairs of electrons entering a quantum state known as a Bose–Einstein condensate. These Cooper pairs have their motion coupled to nearby matter via lattice vibrations called phonons, thereby avoiding the collisions with atoms that normally create electrical resistance. (Cooper pairs have a radius of roughly 100 nm, so they can overlap each other.) However, the mechanism by which higher temperature superconductors operate remains uncertain.

Electrons inside conducting solids, which are quasi-particles themselves, when tightly confined at temperatures close to absolute zero, behave as though they had split into two other quasiparticles: spinons and holons. The former carries spin and magnetic moment, while the latter electrical charge.

Motion and energy

According to Einstein's theory of special relativity, as an electron's speed approaches the speed of light, from an observer's point of view its relativistic mass increases, thereby making it more and more difficult to accelerate it from within the observer's frame of reference. The speed of an electron can approach, but never reach, the speed of light in a vacuum, *c*. However, when relativistic electrons—that is, electrons moving at a speed close to *c*—are injected into a dielectric medium such as water, where the local speed of light is significantly less than *c*, the electrons temporarily travel faster than light in the medium. As they interact with the medium, they generate a faint light called Cherenkov radiation.



Lorentz factor as a function of velocity. It starts at value 1 and goes to infinity as *v* approaches *c*.

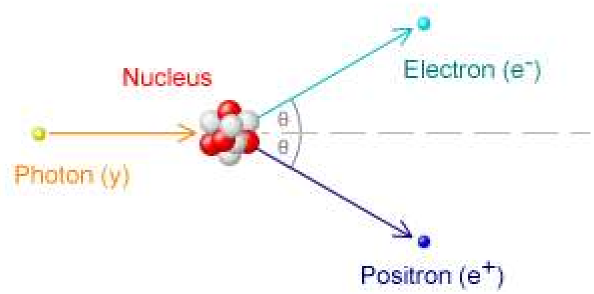
The effects of special relativity are based on a quantity known as the Lorentz factor, defined as $\gamma=1/\sqrt{1-v^2/c^2}$ where v is the speed of the particle. The kinetic energy K_e of an electron moving with velocity v is:

$$K_e = (\gamma - 1)m_e c^2,$$

where m_e is the mass of electron. For example, the Stanford linear accelerator can accelerate an electron to roughly 51 GeV. This gives a value of nearly 100,000 for γ, since the mass of an electron is 0.51 MeV/c^2. The relativistic momentum of this electron is 100,000 times the momentum that classical mechanics would predict for an electron at the same speed.

Since an electron behaves as a wave, at a given velocity it has a characteristic de Broglie wavelength. This is given by $\lambda_e = h/p$ where h is the Planck constant and p is the momentum. For the 51 GeV electron above, the wavelength is about 2.4×10^{-17} m, small enough to explore structures well below the size of an atomic nucleus.

Formation



Pair production caused by the collision of a photon with an atomic nucleus

The Big Bang theory is the most widely accepted scientific theory to explain the early stages in the evolution of the Universe. For the first millisecond of the Big Bang, the temperatures were over 10 billion kelvins and photons had mean energies over a million electronvolts. These photons were sufficiently energetic that they could react with each other to form pairs of electrons and positrons,

$$\gamma + \gamma \leftrightharpoons e^+ + e^-,$$

where γ is a photon, e^+ is a positron and e^- is an electron. Likewise, positron-electron pairs annihilated each other and emitted energetic photons. An equilibrium between electrons, positrons and photons was maintained during this phase of the evolution of the

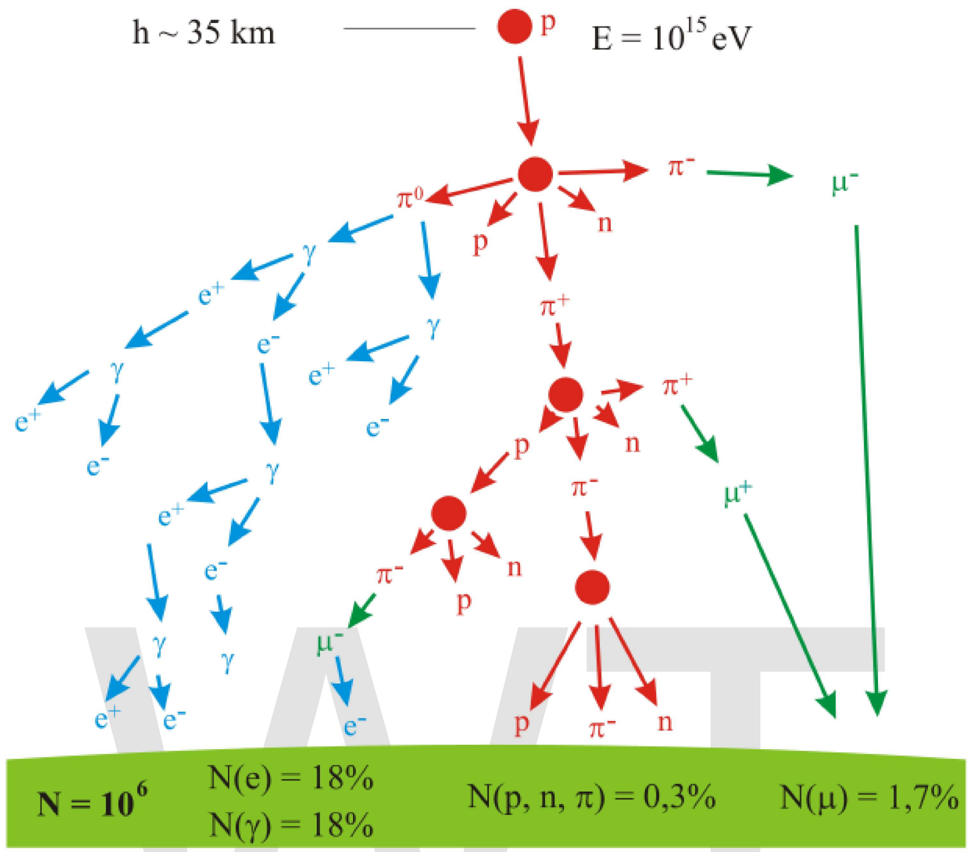
Universe. After 15 seconds had passed, however, the temperature of the universe dropped below the threshold where electron-positron formation could occur. Most of the surviving electrons and positrons annihilated each other, releasing gamma radiation that briefly reheated the universe.

For reasons that remain uncertain, during the process of leptogenesis there was an excess in the number of electrons over positrons. Hence, about one electron in every billion survived the annihilation process. This excess matched the excess of protons over anti-protons, in a condition known as baryon asymmetry, resulting in a net charge of zero for the universe. The surviving protons and neutrons began to participate in reactions with each other—in the process known as nucleosynthesis, forming isotopes of hydrogen and helium, with trace amounts of lithium. This process peaked after about five minutes. Any leftover neutrons underwent negative beta decay with a half-life of about a thousand seconds, releasing a proton and electron in the process,

$$\mathrm{n} \Rightarrow \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e},$$

where n is a neutron, p is a proton and ν_e is an electron antineutrino. For about the next 300,000–400,000 years, the excess electrons remained too energetic to bind with atomic nuclei. What followed is a period known as recombination, when neutral atoms were formed and the expanding universe became transparent to radiation.

Roughly one million years after the big bang, the first generation of stars began to form. Within a star, stellar nucleosynthesis results in the production of positrons from the fusion of atomic nuclei. These antimatter particles immediately annihilate with electrons, releasing gamma rays. The net result is a steady reduction in the number of electrons, and a matching increase in the number of neutrons. However, the process of stellar evolution can result in the synthesis of radioactive isotopes. Selected isotopes can subsequently undergo negative beta decay, emitting an electron and antineutrino from the nucleus. An example is the cobalt-60 (^{60}Co) isotope, which decays to form nickel-60 (^{60}Ni).



An extended air shower generated by an energetic cosmic ray striking the Earth's atmosphere

At the end of its lifetime, a star with more than about 20 solar masses can undergo gravitational collapse to form a black hole. According to classical physics, these massive stellar objects exert a gravitational attraction that is strong enough to prevent anything, even electromagnetic radiation, from escaping past the Schwarzschild radius. However, it is believed that quantum mechanical effects may allow Hawking radiation to be emitted at this distance. Electrons (and positrons) are thought to be created at the event horizon of these stellar remnants.

When pairs of virtual particles (such as an electron and positron) are created in the vicinity of the event horizon, the random spatial distribution of these particles may permit one of them to appear on the exterior; this process is called quantum tunneling. The gravitational potential of the black hole can then supply the energy that transforms this virtual particle into a real particle, allowing it to radiate away into space. In exchange, the other member of the pair is given negative energy, which results in a net loss of mass-energy by the black hole. The rate of Hawking radiation increases with decreasing mass, eventually causing the black hole to evaporate away until, finally, it explodes.

Cosmic rays are particles traveling through space with high energies. Energy events as high as 3.0×10^{20} eV have been recorded. When these particles collide with nucleons in the Earth's atmosphere, a shower of particles is generated, including pions. More than half of the cosmic radiation observed from the Earth's surface consists of muons. The particle called a muon is a lepton which is produced in the upper atmosphere by the decay of a pion. A muon, in turn, can decay to form an electron or positron. Thus, for the negatively charged pion π^-,

$$\begin{aligned} \pi^- &\to \mu^- + \bar{\nu}_\mu, \\ \mu^- &\to e^- + \bar{\nu}_e + \nu_\mu, \end{aligned}$$

where μ^- is a muon and ν_μ is a muon neutrino.

Observation

Aurorae are mostly caused by energetic electrons precipitating into the atmosphere

Remote observation of electrons requires detection of their radiated energy. For example, in high-energy environments such as the corona of a star, free electrons form a plasma that radiates energy due to Bremsstrahlung. Electron gas can undergo plasma oscillation, which is waves caused by synchronized variations in electron density, and these produce energy emissions that can be detected by using radio telescopes.

The frequency of a photon is proportional to its energy. As a bound electron transitions between different energy levels of an atom, it will absorb or emit photons at characteristic frequencies. For instance, when atoms are irradiated by a source with a broad spectrum, distinct absorption lines will appear in the spectrum of transmitted radiation. Each element or molecule displays a characteristic set of spectral lines, such as the hydrogen spectral series. Spectroscopic measurements of the strength and width of these lines allow the composition and physical properties of a substance to be determined.

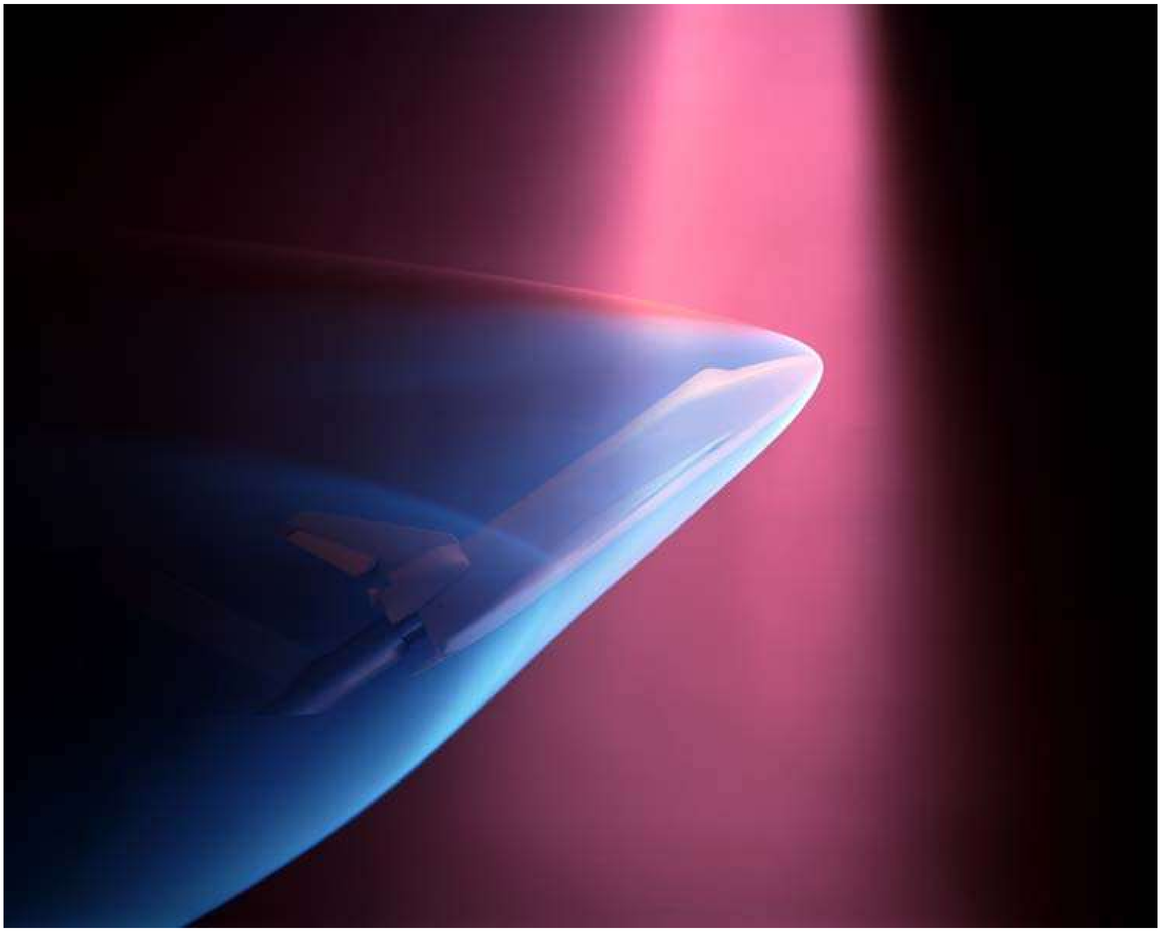
In laboratory conditions, the interactions of individual electrons can be observed by means of particle detectors, which allow measurement of specific properties such as energy, spin and charge. The development of the Paul trap and Penning trap allows charged particles to be contained within a small region for long durations. This enables precise measurements of the particle properties. For example, in one instance a Penning trap was used to contain a single electron for a period of 10 months. The magnetic moment of the electron was measured to a precision of eleven digits, which, in 1980, was a greater accuracy than for any other physical constant.

The first video images of an electron's energy distribution were captured by a team at Lund University in Sweden, February 2008. The scientists used extremely short flashes of light, called attosecond pulses, which allowed an electron's motion to be observed for the first time.

The distribution of the electrons in solid materials can be visualized by angle resolved photoemission spectroscopy (ARPES). This technique employs the photoelectric effect to measure the reciprocal space—a mathematical representation of periodic structures that is used to infer the original structure. ARPES can be used to determine the direction, speed and scattering of electrons within the material.

Plasma applications

Particle beams



During a NASA wind tunnel test, a model of the Space Shuttle is targeted by a beam of electrons, simulating the effect of ionizing gases during re-entry.

Electron beams are used in welding, which allows energy densities up to 10^7 W•cm^{-2} across a narrow focus diameter of 0.1–1.3 mm and usually does not require a filler material. This welding technique must be performed in a vacuum, so that the electron beam does not interact with the gas prior to reaching the target, and it can be used to join conductive materials that would otherwise be considered unsuitable for welding.

Electron beam lithography (EBL) is a method of etching semiconductors at resolutions smaller than a micron. This technique is limited by high costs, slow performance, the need to operate the beam in the vacuum and the tendency of the electrons to scatter in solids. The last problem limits the resolution to about 10 nm. For this reason, EBL is primarily used for the production of small numbers of specialized integrated circuits.

Electron beam processing is used to irradiate materials in order to change their physical properties or sterilize medical and food products. In radiation therapy, electron beams are generated by linear accelerators for treatment of superficial tumors. Because an electron beam only penetrates to a limited depth before being absorbed, typically up to 5 cm for electron energies in the range 5–20 MeV, electron therapy is useful for treating skin lesions such as basal cell carcinomas. An electron beam can be used to supplement the treatment of areas that have been irradiated by X-rays.

Particle accelerators use electric fields to propel electrons and their antiparticles to high energies. As these particles pass through magnetic fields, they emit synchrotron radiation. The intensity of this radiation is spin dependent, which causes polarization of the electron beam—a process known as the Sokolov–Ternov effect. The polarized electron beams can be useful for various experiments. Synchrotron radiation can also be used for cooling the electron beams, which reduces the momentum spread of the particles. Once the particles have accelerated to the required energies, separate electron and positron beams are brought into collision. The resulting energy emissions are observed with particle detectors and are studied in particle physics.

Imaging

Low-energy electron diffraction (LEED) is a method of bombarding a crystalline material with a collimated beam of electrons, then observing the resulting diffraction patterns to determine the structure of the material. The required energy of the electrons is typically in the range 20–200 eV. The reflection high energy electron diffraction (RHEED) technique uses the reflection of a beam of electrons fired at various low angles to characterize the surface of crystalline materials. The beam energy is typically in the range 8–20 keV and the angle of incidence is 1–4°.

The electron microscope directs a focused beam of electrons at a specimen. As the beam interacts with the material, some electrons change their properties, such as movement direction, angle, relative phase and energy. By recording these changes in the electron beam, microscopists can produce atomically resolved image of the material. In blue light, conventional optical microscopes have a diffraction-limited resolution of about 200 nm. By comparison, electron microscopes are limited by the de Broglie wavelength of the electron. This wavelength, for example, is equal to 0.0037 nm for electrons accelerated across a 100,000-volt potential. The Transmission Electron Aberration-corrected Microscope is capable of sub-0.05 nm resolution, which is more than enough to resolve individual atoms. This capability makes the electron microscope a useful laboratory instrument for high resolution imaging. However, electron microscopes are expensive instruments that are costly to maintain.

There are two main types of electron microscopes: transmission and scanning. Transmission electron microscopes function in a manner similar to overhead projector, with a beam of electrons passing through a slice of material then being projected by lenses on a photographic slide or a charge-coupled device. In scanning electron microscopes, the image is produced by rastering a finely focused electron beam, as in a

TV set, across the studied sample. The magnifications range from 100× to 1,000,000× or higher for both microscope types. The scanning tunneling microscope uses quantum tunneling of electrons from a sharp metal tip into the studied material and can produce atomically resolved images of its surface.

Other

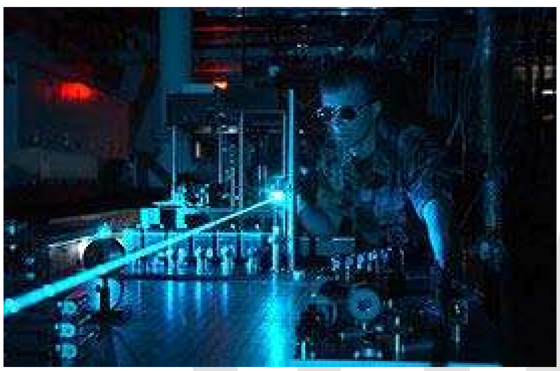
In the free electron laser (FEL), a relativistic electron beam is passed through a pair of undulators containing arrays of dipole magnets, whose fields are oriented in alternating directions. The electrons emit synchrotron radiation, which, in turn, coherently interacts with the same electrons. This leads to the strong amplification of the radiation field at the resonance frequency. FEL can emit a coherent high-brilliance electromagnetic radiation with a wide range of frequencies, from microwaves to soft X-rays. These devices can be used in the future for manufacturing, communication and various medical applications, such as soft tissue surgery.

Electrons are at the heart of cathode ray tubes, which are used extensively as display devices in laboratory instruments, computer monitors and television sets. In a photomultiplier tube, every photon striking the photocathode initiates an avalanche of electrons that produces a detectable current pulse. Vacuum tubes use the flow of electrons to manipulate electrical signals, and they played a critical role in the development of electronics technology. However, they have been largely supplanted by solid-state devices such as the transistor.

Chapter 5

Photon

Photon



Photons emitted in a coherent beam from a laser

Composition:	Elementary particle
Particle statistics:	Bosonic
Group:	Gauge boson
Interaction:	Electromagnetic
Symbol(s):	γ, hν, or ħω
Theorized:	Albert Einstein
Mass:	0 $<1\times10^{-18}$ eV
Mean lifetime:	Stable
Electric charge:	0 $<1\times10^{-35}$ *e*
Spin:	1
Parity:	-1
C parity:	-1
Condensed:	$I(J^{PC}) = 0,1(1^{--})$

In physics, a **photon** is an elementary particle, the quantum of the electromagnetic interaction and the basic unit of light and all other forms of electromagnetic radiation. It is also the force carrier for the electromagnetic force. The effects of this force are easily observable at both the microscopic and macroscopic level, because the photon has no rest mass; this allows for interactions at long distances. Like all elementary particles, photons are currently best explained by quantum mechanics and will exhibit wave–particle duality — they exhibit properties of both waves and particles. For example, a single photon may be refracted by a lens or exhibit wave interference with itself, but also act as a particle giving a definite result when quantitative momentum is measured.

The modern concept of the photon was developed gradually by Albert Einstein to explain experimental observations that did not fit the classical wave model of light. In particular, the photon model accounted for the frequency dependence of light's energy, and explained the ability of matter and radiation to be in thermal equilibrium. It also accounted for anomalous observations, including the properties of black body radiation, that other physicists, most notably Max Planck, had sought to explain using *semiclassical models*, in which light is still described by Maxwell's equations, but the material objects that emit and absorb light are quantized. Although these semiclassical models contributed to the development of quantum mechanics, further experiments validated Einstein's hypothesis that *light itself* is quantized; the quanta of light are photons.

In the modern Standard Model of particle physics, photons are described as a necessary consequence of physical laws having a certain symmetry at every point in spacetime. The intrinsic properties of photons, such as charge, mass and spin, are determined by the properties of this gauge symmetry. The neutrino theory of light, which attempts to describe the photon as a composite structure, has been unsuccessful so far.

The photon concept has led to momentous advances in experimental and theoretical physics, such as lasers, Bose–Einstein condensation, quantum field theory, and the probabilistic interpretation of quantum mechanics. It has been applied to photochemistry, high-resolution microscopy, and measurements of molecular distances. Recently, photons have been studied as elements of quantum computers and for sophisticated applications in optical communication such as quantum cryptography.

Nomenclature

In 1900, Max Planck was working on black-body radiation and suggested that the energy in electromagnetic waves could only be released in "packets" of energy. In his 1901 article in Annalen der Physik he called these packets "energy elements". The word *quanta* (singular *quantum*) was used even before 1900 to mean particles or amounts of different quantities, including electricity . Later, in 1905 Albert Einstein went further by suggesting that electromagnetic waves could only exist in these discrete wave-packets. He called such a wave-packet *the light quantum* (German: *das Lichtquant*). The name *photon* derives from the Greek word for light, *φως* (transliterated *phôs*), and was coined in 1926 by the physical chemist Gilbert Lewis, who published a speculative theory in which photons were "uncreatable and indestructible". Although Lewis' theory was never accepted as it was contradicted by many experiments, his new name, *photon*, was adopted immediately by most physicists. Isaac Asimov credits Arthur Compton with defining quanta of energy as photons in 1927.

In physics, a photon is usually denoted by the symbol γ (the Greek letter gamma). This symbol for the photon probably derives from gamma rays, which were discovered in 1900 by Paul Villard, named by Ernest Rutherford in 1903, and shown to be a form of electromagnetic radiation in 1914 by Rutherford and Edward Andrade. In chemistry and optical engineering, photons are usually symbolized by $h\nu$, the energy of a photon, where h is Planck's constant and the Greek letter ν (nu) is the photon's frequency. Much less commonly, the photon can be symbolized by hf, where its frequency is denoted by f.

Physical properties

The photon is massless, has no electric charge, and does not decay spontaneously in empty space. A photon has two possible polarization states and is described by exactly three continuous parameters: the components of its wave vector, which determine its wavelength λ and its direction of propagation. The photon is the gauge boson for electromagnetism, and therefore all other quantum numbers of the photon (such as lepton number, baryon number, and flavour quantum numbers) are zero.

Photons are emitted in many natural processes. For example, when a charge is accelerated it emits synchrotron radiation. During a molecular, atomic or nuclear transition to a lower energy level, photons of various energy will be emitted, from infrared light to gamma rays. A photon can also be emitted when a particle and its corresponding antiparticle are annihilated (for example, electron-positron annihilation).

In empty space, the photon moves at c (the speed of light) and its energy and momentum are related by $E = pc$, where p is the magnitude of the momentum vector $\mathbf{p}$. This derives from the following relativistic relation, with $m = 0$:

$$E^2 = p^2c^2 + m^2c^4.$$

The energy and momentum of a photon depend only on its frequency (ν) or inversely, its wavelength (λ):

$$E = \hbar\omega = h\nu = \frac{hc}{\lambda}$$
$$\mathbf{p} = \hbar\mathbf{k},$$

where $\mathbf{k}$ is the wave vector (where the wave number $k = |\mathbf{k}| = 2\pi/\lambda$), $\omega = 2\pi\nu$ is the angular frequency, and $\hbar = h/2\pi$ is the reduced Planck constant.

Since $\mathbf{p}$ points in the direction of the photon's propagation, the magnitude of the momentum is

$$p = \hbar k = \frac{h\nu}{c} = \frac{h}{\lambda}.$$

The photon also carries spin angular momentum that does not depend on its frequency. The magnitude of its spin is $\sqrt{2}\hbar$ and the component measured along its direction of motion, its helicity, must be $\pm\hbar$. These two possible helicities, called right-handed and left-handed, correspond to the two possible circular polarization states of the photon.

To illustrate the significance of these formulae, the annihilation of a particle with its antiparticle in free space must result in the creation of at least *two* photons for the following reason. In the center of mass frame, the colliding antiparticles have no net momentum, whereas a single photon always has momentum (since it is determined, as we have seen, only by the photon's frequency or wavelength—which cannot be zero). Hence, conservation of momentum (or equivalently, translational invariance) requires that at least two photons are created, with zero net momentum. (However it is possible if the system interacts with another particle or field for annihilation to produce one photon, as when a positron annihilates with a bound atomic electron, it is possible for only one photon to be emitted, as the nuclear Coulomb field breaks translational symmetry.) The energy of the two photons, or, equivalently, their frequency, may be determined from conservation of four-momentum. Seen another way, the photon can be considered as its own antiparticle. The reverse process, pair production, is the dominant mechanism by which high-energy photons such as gamma rays lose energy while passing through matter. That process is the reverse of "annihilation to one photon" allowed in the electric field of an atomic nucleus.

The classical formulae for the energy and momentum of electromagnetic radiation can be re-expressed in terms of photon events. For example, the pressure of electromagnetic radiation on an object derives from the transfer of photon momentum per unit time and unit area to that object, since pressure is force per unit area and force is the change in momentum per unit time.

Experimental checks on photon mass

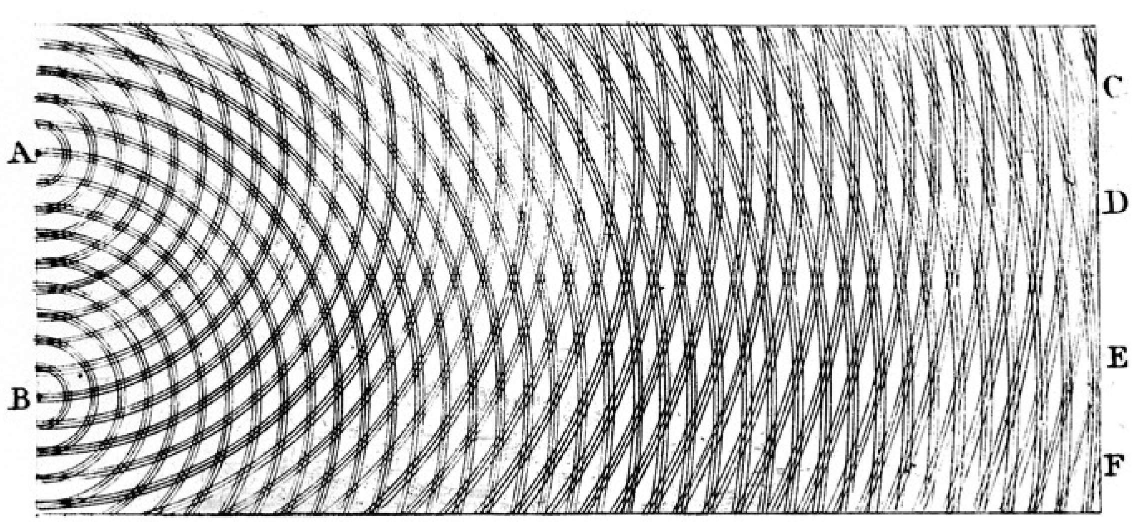
The photon is currently believed to be strictly massless, but this is an experimental question. If the photon is not a strictly massless particle, it would not move at the exact speed of light in vacuum, *c*. Its speed would be lower and depend on its frequency. Relativity would be unaffected by this; the so-called speed of light, *c*, would then not be the actual speed at which light moves, but a constant of nature which is the maximum speed that any object could theoretically attain in space-time. Thus, it would still be the speed of space-time ripples (gravitational waves and gravitons), but it would not be the speed of photons.

A massive photon would have other effects as well. Coulomb's law would be modified and the electromagnetic field would have an extra physical degree of freedom. These effects yield more sensitive experimental probes of the photon mass than the frequency dependence of the speed of light. If Coulomb's law is not exactly valid, then that would cause the presence of an electric field inside a hollow conductor when it is subjected to an external electric field. This thus allows one to test Coulomb's law to very high precision. A null result of such an experiment has set a limit of $m \lesssim 10^{-14}$ eV/c^2.

Sharper upper limits have been obtained in experiments designed to detect effects caused by the Galactic vector potential. Although the galactic vector potential is very large because the galactic magnetic field exists on very long length scales, only the magnetic field is observable if the photon is massless. In case of a massive photon, the mass term $\frac{1}{2}m^2 A_\mu A^\mu$ would affect the galactic plasma. The fact that no such effects are seen implies an upper bound on the photon mass of $m < 3\times10^{-27}$ eV/c^2. The galactic vector potential can also be probed directly by measuring the torque exerted on a magnetized ring. Such methods were used to obtain the sharper upper limit of 10^{-18}eV/c^2 given by the Particle Data Group.

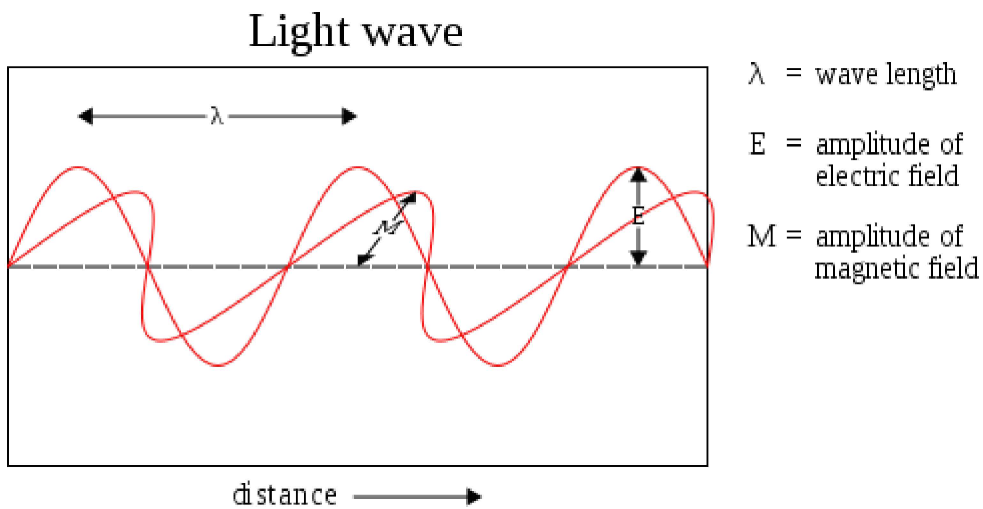
These sharp limits from the non-observation of the effects caused by the galactic vector potential have been shown to be model dependent. If the photon mass is generated via the Higgs mechanism then the upper limit of $m \lesssim 10^{-14}$ eV/c^2 from the test of Coulomb's law is valid.

Historical development



Thomas Young's double-slit experiment in 1805 showed that light can act as a wave, helping to defeat early particle theories of light.

In most theories up to the eighteenth century, light was pictured as being made up of particles. Since particle models cannot easily account for the refraction, diffraction and birefringence of light, wave theories of light were proposed by René Descartes (1637), Robert Hooke (1665), and Christian Huygens (1678); however, particle models remained dominant, chiefly due to the influence of Isaac Newton. In the early nineteenth century, Thomas Young and August Fresnel clearly demonstrated the interference and diffraction of light and by 1850 wave models were generally accepted. In 1865, James Clerk Maxwell's prediction that light was an electromagnetic wave—which was confirmed experimentally in 1888 by Heinrich Hertz's detection of radio waves—seemed to be the final blow to particle models of light.



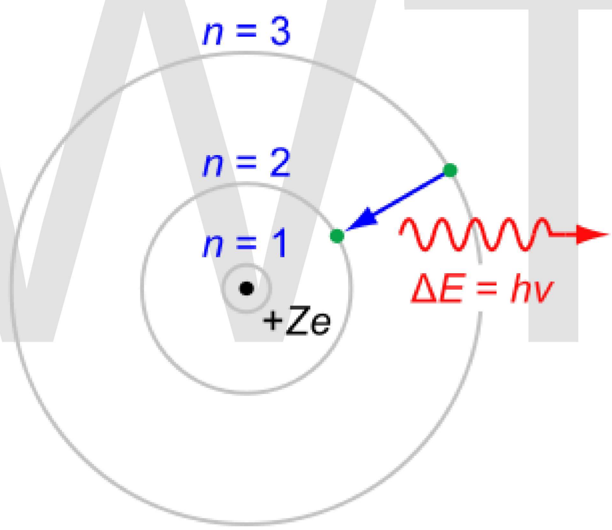
In 1900, Maxwell's theoretical model of light as oscillating electric and magnetic fields seemed complete. However, several observations could not be explained by any wave model of electromagnetic radiation, leading to the idea that light-energy was packaged into *quanta* described by E=hν. Later experiments showed that these light-quanta also carry momentum and, thus, can be considered particles: the *photon* concept was born, leading to a deeper understanding of the electric and magnetic fields themselves.

The Maxwell wave theory, however, does not account for *all* properties of light. The Maxwell theory predicts that the energy of a light wave depends only on its intensity, not on its frequency; nevertheless, several independent types of experiments show that the energy imparted by light to atoms depends only on the light's frequency, not on its intensity. For example, some chemical reactions are provoked only by light of frequency higher than a certain threshold; light of frequency lower than the threshold, no matter how intense, does not initiate the reaction. Similarly, electrons can be ejected from a metal plate by shining light of sufficiently high frequency on it (the photoelectric effect); the energy of the ejected electron is related only to the light's frequency, not to its intensity.

At the same time, investigations of blackbody radiation carried out over four decades (1860–1900) by various researchers culminated in Max Planck's hypothesis that the energy of *any* system that absorbs or emits electromagnetic radiation of frequency v is an integer multiple of an energy quantum $E=hv$. As shown by Albert Einstein, some form of energy quantization *must* be assumed to account for the thermal equilibrium observed between matter and electromagnetic radiation; for this explanation of the photoelectric effect, Einstein received the 1921 Nobel Prize in physics.

Since the Maxwell theory of light allows for all possible energies of electromagnetic radiation, most physicists assumed initially that the energy quantization resulted from some unknown constraint on the matter that absorbs or emits the radiation. In 1905, Einstein was the first to propose that energy quantization was a property of electromagnetic radiation itself. Although he accepted the validity of Maxwell's theory, Einstein pointed out that many anomalous experiments could be explained if the *energy* of a Maxwellian light wave were localized into point-like quanta that move independently of one another, even if the wave itself is spread continuously over space. In 1909 and 1916, Einstein showed that, if Planck's law of black-body radiation is accepted, the energy quanta must also carry momentum $p=h/\lambda$, making them full-fledged particles. This photon momentum was observed experimentally by Arthur Compton, for which he received the Nobel Prize in 1927. The pivotal question was then: how to unify Maxwell's wave theory of light with its experimentally observed particle nature? The answer to this question occupied Albert Einstein for the rest of his life, and was solved in quantum electrodynamics and its successor, the Standard Model.

Early objections



Up to 1923, most physicists were reluctant to accept that light itself was quantized. Instead, they tried to explain photon behavior by quantizing only *matter*, as in the Bohr model of the hydrogen atom (shown here). Even though these semiclassical models were only a first approximation, they were accurate for simple systems and they led to quantum mechanics.

Einstein's 1905 predictions were verified experimentally in several ways in the first two decades of the 20th century, as recounted in Robert Millikan's Nobel lecture. However, before Compton's experiment showing that photons carried momentum proportional to their wave number (or frequency) (1922), most physicists were reluctant to believe that electromagnetic radiation itself might be particulate. (See, for example, the Nobel lectures of Wien, Planck and Millikan.). Instead, there was a widespread belief that energy quantization resulted from some unknown constraint on the matter that absorbs or emits radiation. Attitudes changed over time. In part, the change can be traced to experiments such as Compton scattering, where it was much more difficult not to ascribe quantization to light itself to explain the observed results.

Even after Compton's experiment, Bohr, Hendrik Kramers and John Slater made one last attempt to preserve the Maxwellian continuous electromagnetic field model of light, the so-called BKS model. To account for the then-available data, two drastic hypotheses had to be made:

1. **Energy and momentum are conserved only on the average in interactions between matter and radiation, not in elementary processes such as absorption and emission.** This allows one to reconcile the (at the time believed to be) discontinuously changing energy of the atom (jump between energy states) with the continuous release of energy into radiation. (It is now known that this is actually a continuous process, the combined atom-field system evolving in time according to Schroedinger's equation.)
2. **Causality is abandoned**. For example, spontaneous emissions are merely emissions induced by a "virtual" electromagnetic field.

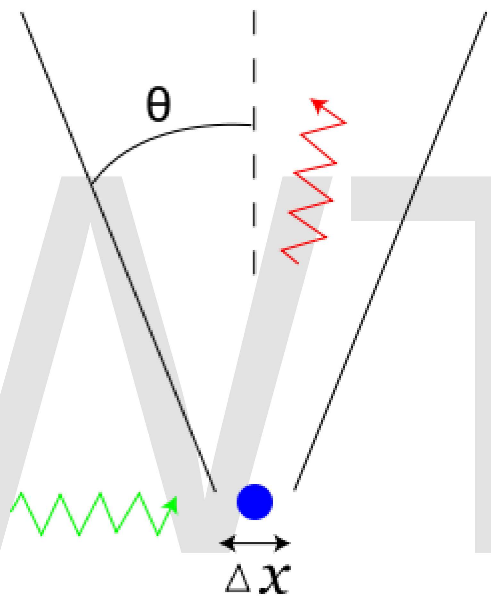
However, refined Compton experiments showed that energy-momentum is conserved extraordinarily well in elementary processes; and also that the jolting of the electron and the generation of a new photon in Compton scattering obey causality to within 10 ps. Accordingly, Bohr and his co-workers gave their model "as honorable a funeral as possible". Nevertheless, the failures of the BKS model inspired Werner Heisenberg in his development of matrix mechanics.

A few physicists persisted in developing semiclassical models in which electromagnetic radiation is not quantized, but matter appears to obey the laws of quantum mechanics. Although the evidence for photons from chemical and physical experiments was overwhelming by the 1970s, this evidence could not be considered as *absolutely* definitive; since it relied on the interaction of light with matter, a sufficiently complicated theory of matter could in principle account for the evidence. Nevertheless, *all* semiclassical theories were refuted definitively in the 1970s and 1980s by photon-correlation experiments. Hence, Einstein's hypothesis that quantization is a property of light itself is considered to be proven.

Wave–particle duality and uncertainty principles

Photons, like all quantum objects, exhibit both wave-like and particle-like properties. Their dual wave–particle nature can be difficult to visualize. The photon displays clearly wave-like phenomena such as diffraction and interference on the length scale of its wavelength. For example, a single photon passing through a double-slit experiment lands on the screen exhibiting interference phenomena but only if no measure was made on the actual slit being run across. To account for the particle interpretation that phenomena is called probability distribution but behaves according to the Maxwell's equations.

However, experiments confirm that the photon is *not* a short pulse of electromagnetic radiation; it does not spread out as it propagates, nor does it divide when it encounters a beam splitter". Rather, the photon seems to be a point-like particle since it is absorbed or emitted *as a whole* by arbitrarily small systems, systems much smaller than its wavelength, such as an atomic nucleus ($\approx 10^{-15}$ m across) or even the point-like electron. Nevertheless, the photon is *not* a point-like particle whose trajectory is shaped probabilistically by the electromagnetic field, as conceived by Einstein and others; that hypothesis was also refuted by the photon-correlation experiments cited above. According to our present understanding, the electromagnetic field itself is produced by photons, which in turn result from a local gauge symmetry and the laws of quantum field theory.



Heisenberg's thought experiment for locating an electron (shown in blue) with a high-resolution gamma-ray microscope. The incoming gamma ray (shown in green) is scattered by the electron up into the microscope's aperture angle θ. The scattered gamma ray is shown in red. Classical optics shows that the electron position can be resolved only up to an uncertainty Δx that depends on θ and the wavelength λ of the incoming light.

A key element of quantum mechanics is Heisenberg's uncertainty principle, which forbids the simultaneous measurement of the position and momentum of a particle along the same direction. Remarkably, the uncertainty principle for charged, material particles *requires* the quantization of light into photons, and even the frequency dependence of the photon's energy and momentum. An elegant illustration is Heisenberg's thought experiment for locating an electron with an ideal microscope. The position of the electron can be determined to within the resolving power of the microscope, which is given by a formula from classical optics

$$\Delta x \sim \frac{\lambda}{\sin\theta}$$

where θ is the aperture angle of the microscope. Thus, the position uncertainty Δx can be made arbitrarily small by reducing the wavelength λ. The momentum of the electron is uncertain, since it received a "kick" Δp from the light scattering from it into the microscope. If light were *not* quantized into photons, the uncertainty Δp could be made arbitrarily small by reducing the light's intensity. In that case, since the wavelength and intensity of light can be varied independently, one could simultaneously determine the position and momentum to arbitrarily high accuracy, violating the uncertainty principle. By contrast, Einstein's formula for photon momentum preserves the uncertainty principle; since the photon is scattered anywhere within the aperture, the uncertainty of momentum transferred equals

$$\Delta p \sim p_{\text{photon}} \sin\theta = \frac{h}{\lambda}\sin\theta$$

giving the product $\Delta x \Delta p \ \sim \ h$, which is Heisenberg's uncertainty principle. Thus, the entire world is quantized; both matter and fields must obey a consistent set of quantum laws, if either one is to be quantized.

The analogous uncertainty principle for photons forbids the simultaneous measurement of the number n of photons in an electromagnetic wave and the phase φ of that wave

$$\Delta n \Delta \varphi > 1$$

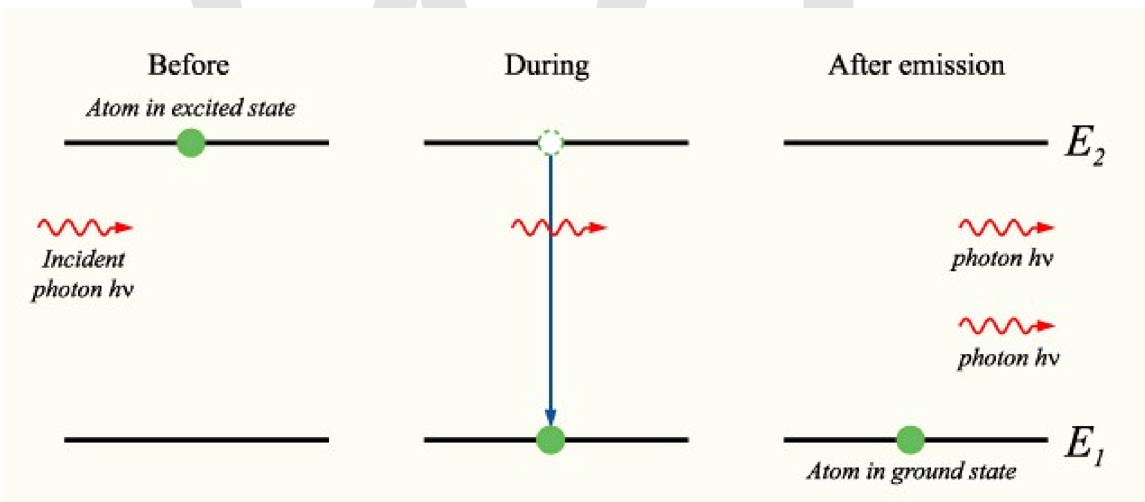
Both photons and material particles such as electrons create analogous interference patterns when passing through a double-slit experiment. For photons, this corresponds to the interference of a Maxwell light wave whereas, for material particles, this corresponds to the interference of the Schrödinger wave equation. Although this similarity might suggest that Maxwell's equations are simply Schrödinger's equation for photons, most physicists do not agree. For one thing, they are mathematically different; most obviously, Schrödinger's one equation solves for a complex field, whereas Maxwell's four equations solve for real fields. More generally, the normal concept of a Schrödinger probability wave function cannot be applied to photons. Being massless, they cannot be localized without being destroyed; technically, photons cannot have a position eigenstate $|\mathbf{r}\rangle$, and, thus, the normal Heisenberg uncertainty principle $\Delta x \Delta p > h / 2$ does not pertain to photons. A few substitute wave functions have been suggested for the photon, but they have not come into general use. Instead, physicists generally accept the second-quantized theory of photons described below, quantum electrodynamics, in which photons are quantized excitations of electromagnetic modes.

Bose–Einstein model of a photon gas

In 1924, Satyendra Nath Bose derived Planck's law of black-body radiation without using any electromagnetism, but rather a modification of coarse-grained counting of phase space. Einstein showed that this modification is equivalent to assuming that photons are rigorously identical and that it implied a "mysterious non-local interaction", now understood as the requirement for a symmetric quantum mechanical state. This work led to the concept of coherent states and the development of the laser. In the same papers, Einstein extended Bose's formalism to material particles (bosons) and predicted that they would condense into their lowest quantum state at low enough temperatures; this Bose–Einstein condensation was observed experimentally in 1995.

The modern view on this is that photons are, by virtue of their integer spin, bosons (as opposed to fermions with half-integer spin). By the spin-statistics theorem, all bosons obey Bose–Einstein statistics (whereas all fermions obey Fermi-Dirac statistics).

Stimulated and spontaneous emission



Stimulated emission (in which photons "clone" themselves) was predicted by Einstein in his kinetic analysis, and led to the development of the laser. Einstein's derivation inspired further developments in the quantum treatment of light, which led to the statistical interpretation of quantum mechanics.

In 1916, Einstein showed that Planck's radiation law could be derived from a semi-classical, statistical treatment of photons and atoms, which implies a relation between the rates at which atoms emit and absorb photons. The condition follows from the assumption that light is emitted and absorbed by atoms independently, and that the thermal equilibrium is preserved by interaction with atoms. Consider a cavity in thermal equilibrium and filled with electromagnetic radiation and atoms that can emit and absorb that radiation. Thermal equilibrium requires that the energy density $\rho(\nu)$ of photons with frequency ν (which is proportional to their number density) is, on average, constant in time; hence, the rate at which photons of any particular frequency are *emitted* must equal the rate of *absorbing* them.

Einstein began by postulating simple proportionality relations for the different reaction rates involved. In his model, the rate R_{ji} for a system to *absorb* a photon of frequency ν and transition from a lower energy E_j to a higher energy E_i is proportional to the number N_j of atoms with energy E_j and to the energy density ρ(ν) of ambient photons with that frequency,

$$R_{ji} = N_j B_{ji} \rho(\nu)$$

where B_{ji} is the rate constant for absorption. For the reverse process, there are two possibilities: spontaneous emission of a photon, and a return to the lower-energy state that is initiated by the interaction with a passing photon. Following Einstein's approach, the corresponding rate R_{ij} for the emission of photons of frequency ν and transition from a higher energy E_i to a lower energy E_j is

$$R_{ij} = N_i A_{ij} + N_i B_{ij} \rho(\nu)$$

where A_{ij} is the rate constant for emitting a photon spontaneously, and B_{ij} is the rate constant for emitting it in response to ambient photons (induced or stimulated emission). In thermodynamic equilibrium, the number of atoms in state i and that of atoms in state j must, on average, be constant; hence, the rates R_{ji} and R_{ij} must be equal. Also, by arguments analogous to the derivation of Boltzmann statistics, the ratio of N_i and N_j is g_i / $g_j \exp(E_j - E_i) / kT)$, where $g_{i,j}$ are the degeneracy of the state i and that of j, respectively, $E_{i,j}$ their energies, k the Boltzmann constant and T the system's temperature. From this, it is readily derived that $g_i B_{ij} = g_j B_{ji}$ and

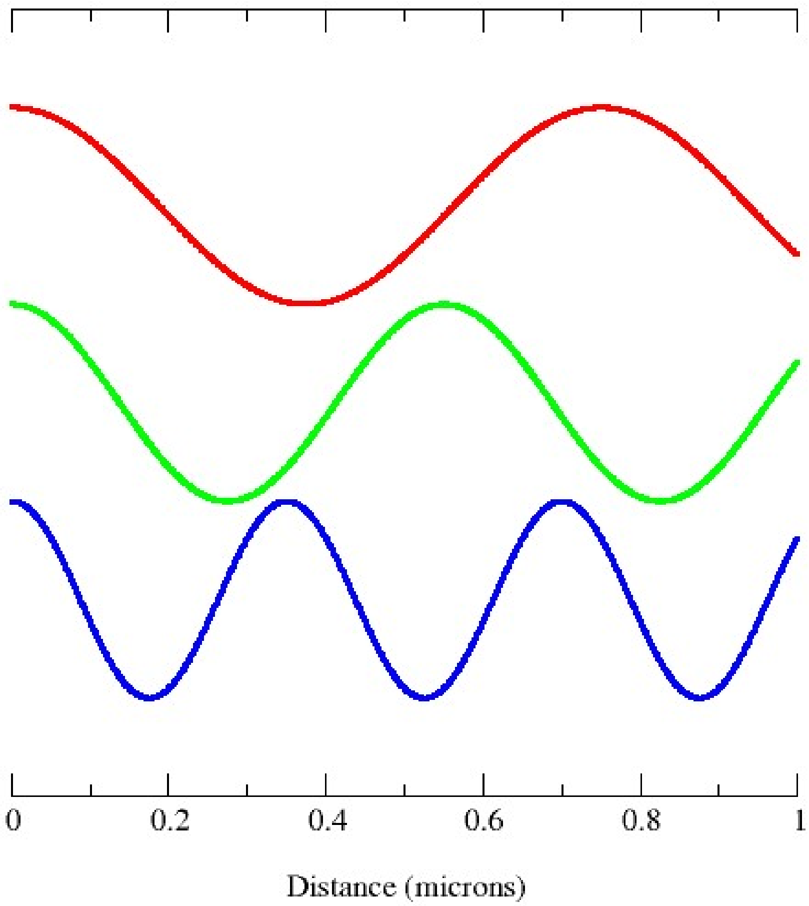
$$A_{ij} = \frac{8\pi h \nu^3}{c^3} B_{ij}.$$

The A and Bs are collectively known as the *Einstein coefficients*.

Einstein could not fully justify his rate equations, but claimed that it should be possible to calculate the coefficients A_{ij}, B_{ji} and B_{ij} once physicists had obtained "mechanics and electrodynamics modified to accommodate the quantum hypothesis". In fact, in 1926, Paul Dirac derived the B_{ij} rate constants in using a semiclassical approach, and, in 1927, succeeded in deriving *all* the rate constants from first principles within the framework of quantum theory. Dirac's work was the foundation of quantum electrodynamics, i.e., the quantization of the electromagnetic field itself. Dirac's approach is also called *second quantization* or quantum field theory; earlier quantum mechanical treatments only treat material particles as quantum mechanical, not the electromagnetic field.

Einstein was troubled by the fact that his theory seemed incomplete, since it did not determine the *direction* of a spontaneously emitted photon. A probabilistic nature of light-particle motion was first considered by Newton in his treatment of birefringence and, more generally, of the splitting of light beams at interfaces into a transmitted beam and a reflected beam. Newton hypothesized that hidden variables in the light particle determined which path it would follow. Similarly, Einstein hoped for a more complete theory that would leave nothing to chance, beginning his separation from quantum mechanics. Ironically, Max Born's probabilistic interpretation of the wave function was inspired by Einstein's later work searching for a more complete theory.

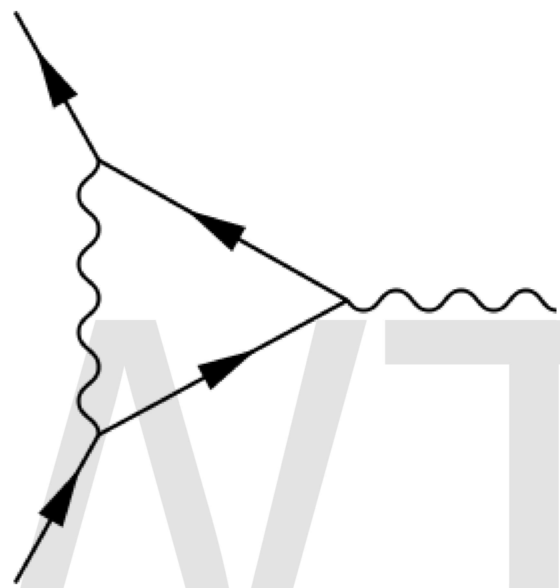
Second quantization



Different *electromagnetic modes* (such as those depicted here) can be treated as independent simple harmonic oscillators. A photon corresponds to a unit of energy E=hν in its electromagnetic mode.

In 1910, Peter Debye derived Planck's law of black-body radiation from a relatively simple assumption. He correctly decomposed the electromagnetic field in a cavity into its Fourier modes, and assumed that the energy in any mode was an integer multiple of *hν*, where ν is the frequency of the electromagnetic mode. Planck's law of black-body radiation follows immediately as a geometric sum. However, Debye's approach failed to give the correct formula for the energy fluctuations of blackbody radiation, which were derived by Einstein in 1909.

In 1925, Born, Heisenberg and Jordan reinterpreted Debye's concept in a key way. As may be shown classically, the Fourier modes of the electromagnetic field—a complete set of electromagnetic plane waves indexed by their wave vector **k** and polarization state—are equivalent to a set of uncoupled simple harmonic oscillators. Treated quantum mechanically, the energy levels of such oscillators are known to be $E = nh\nu$, where ν is the oscillator frequency. The key new step was to identify an electromagnetic mode with energy $E = nh\nu$ as a state with n photons, each of energy $h\nu$. This approach gives the correct energy fluctuation formula.



In quantum field theory, the probability of an event is computed by summing the probability amplitude (a complex number) for all possible ways in which the event can occur, as in the Feynman diagram shown here; the probability equals the square of the modulus of the total amplitude.

Dirac took this one step further. He treated the interaction between a charge and an electromagnetic field as a small perturbation that induces transitions in the photon states, changing the numbers of photons in the modes, while conserving energy and momentum overall. Dirac was able to derive Einstein's A_{ij} and B_{ij} coefficients from first principles, and showed that the Bose–Einstein statistics of photons is a natural consequence of quantizing the electromagnetic field correctly (Bose's reasoning went in the opposite direction; he derived Planck's law of black body radiation by *assuming* BE statistics). In Dirac's time, it was not yet known that all bosons, including photons, must obey BE statistics.

Dirac's second-order perturbation theory can involve virtual photons, transient intermediate states of the electromagnetic field; the static electric and magnetic interactions are mediated by such virtual photons. In such quantum field theories, the probability amplitude of observable events is calculated by summing over *all* possible intermediate steps, even ones that are unphysical; hence, virtual photons are not constrained to satisfy $E = pc$, and may have extra polarization states; depending on the gauge used, virtual photons may have three or four polarization states, instead of the two states of real photons. Although these transient virtual photons can never be observed, they contribute measurably to the probabilities of observable events. Indeed, such second-order and higher-order perturbation calculations can give apparently infinite contributions to the sum. Such unphysical results are corrected for using the technique of renormalization. Other virtual particles may contribute to the summation as well; for example, two photons may interact indirectly through virtual electron-positron pairs. In fact, such photon-photon scattering, as well as electron-photon scattering, is meant to be one of the modes of operations of the planned particle accelerator, the International Linear Collider.

In modern physics notation, the quantum state of the electromagnetic field is written as a Fock state, a tensor product of the states for each electromagnetic mode

$$|n_{k_0}\rangle \otimes |n_{k_1}\rangle \otimes \dots \otimes |n_{k_n}\rangle \dots$$

where $|n_{k_i}\rangle$ represents the state in which n_{k_i} photons are in the mode k_i. In this notation, the creation of a new photon in mode k_i (e.g., emitted from an atomic transition) is written as $|n_{k_i}\rangle \rightarrow |n_{k_i} + 1\rangle$. This notation merely expresses the concept of Born, Heisenberg and Jordan described above, and does not add any physics.

The photon as a gauge boson

The electromagnetic field can be understood as a gauge field, i.e., as a field that results from requiring a gauge symmetry holds independently at every position in spacetime. For the electromagnetic field, this gauge symmetry is the Abelian U(1) symmetry of a complex number, which reflects the ability to vary the phase of a complex number without affecting observables or real valued functions made from it, such as the energy or the Lagrangian.

The quanta of an Abelian gauge field must be massless, uncharged bosons, as long as the symmetry is not broken; hence, the photon is predicted to be massless, and to have zero electric charge and integer spin. The particular form of the electromagnetic interaction specifies that the photon must have spin ±1; thus, its helicity must be $\pm\hbar$. These two spin components correspond to the classical concepts of right-handed and left-handed circularly polarized light. However, the transient virtual photons of quantum electrodynamics may also adopt unphysical polarization states.

In the prevailing Standard Model of physics, the photon is one of four gauge bosons in the electroweak interaction; the other three are denoted W^+, W^- and Z^0 and are responsible for the weak interaction. Unlike the photon, these gauge bosons have mass, owing to a mechanism that breaks their SU(2) gauge symmetry. The unification of the photon with W and Z gauge bosons in the electroweak interaction was accomplished by Sheldon Glashow, Abdus Salam and Steven Weinberg, for which they were awarded the 1979 Nobel Prize in physics. Physicists continue to hypothesize grand unified theories that connect these four gauge bosons with the eight gluon gauge bosons of quantum chromodynamics; however, key predictions of these theories, such as proton decay, have not been observed experimentally.

Contributions to the mass of a system

The energy of a system that emits a photon is *decreased* by the energy E of the photon as measured in the rest frame of the emitting system, which may result in a reduction in mass in the amount E / c^2. Similarly, the mass of a system that absorbs a photon is *increased* by a corresponding amount. As an application, the energy balance of nuclear reactions involving photons is commonly written in terms of the masses of the nuclei involved, and terms of the form E / c^2 for the gamma photons (and for other relevant energies, such as the recoil energy of nuclei).

This concept is applied in key predictions of quantum electrodynamics. In that theory, the mass of electrons (or, more generally, leptons) is modified by including the mass contributions of virtual photons, in a technique known as renormalization. Such "radiative corrections" contribute to a number of predictions of QED, such as the magnetic dipole moment of leptons, the Lamb shift, and the hyperfine structure of bound lepton pairs, such as muonium and positronium.

Since photons contribute to the stress-energy tensor, they exert a gravitational attraction on other objects, according to the theory of general relativity. Conversely, photons are themselves affected by gravity; their normally straight trajectories may be bent by warped spacetime, as in gravitational lensing, and their frequencies may be lowered by moving to a higher gravitational potential, as in the Pound-Rebka experiment. However, these effects are not specific to photons; exactly the same effects would be predicted for classical electromagnetic waves.

Photons in matter

(Visible) light that travels through transparent matter does so at a lower speed than c, the speed of light in a vacuum. X-rays, on the other hand, usually have a phase velocity above c, as evidenced by total external reflection. In addition, light can also undergo scattering and absorption. There are circumstances in which heat transfer through a material is mostly radiative, involving emission and absorption of photons within it. An example would be in the core of the sun. Energy can take about a million years to reach the surface. However, this phenomenon is distinct from scattered radiation passing diffusely through matter, as it involves local equilibration between the radiation and the temperature. Thus, the time is how long it takes the *energy* to be transferred, not the *photons* themselves. Once in open space, a photon from the Sun takes only 8.3 minutes to reach Earth. The factor by which the speed of light is decreased in a material is called the refractive index of the material. In a classical wave picture, the slowing can be explained by the light inducing electric polarization in the matter, the polarized matter radiating new light, and the new light interfering with the original light wave to form a delayed wave. In a particle picture, the slowing can instead be described as a blending of the photon with quantum excitations of the matter (quasi-particles such as phonons and excitons) to form a polariton; this polariton has a nonzero effective mass, which means that it cannot travel at *c*.

Alternatively, photons may be viewed as *always* traveling at *c*, even in matter, but they have their phase shifted (delayed or advanced) upon interaction with atomic scatters: this modifies their wavelength and momentum, but not speed. A light wave made up of these photons does travel slower than the speed of light. In this view the photons are "bare", and are scattered and phase shifted, while in the view of the preceding paragraph the photons are "dressed" by their interaction with matter, and move without scattering or phase shifting, but at a lower speed.

Light of different frequencies may travel through matter at different speeds; this is called dispersion. In some cases, it can result in extremely slow speeds of light in matter. The effects of photon interactions with other quasi-particles may be observed directly in Raman scattering and Brillouin scattering.

Photons can also be absorbed by nuclei, atoms or molecules, provoking transitions between their energy levels. A classic example is the molecular transition of retinal $C_{20}H_{28}O$, which is responsible for vision, as discovered in 1958 by Nobel laureate biochemist George Wald and co-workers. The absorption provokes a cis-trans isomerization that, in combination with other such transitions, is transduced into nerve impulses. The absorption of photons can even break chemical bonds, as in the photodissociation of chlorine; this is the subject of photochemistry. Analogously, gamma rays can in some circumstances dissociate atomic nuclei in a process called photodisintegration.

Technological applications

Photons have many applications in technology. These examples are chosen to illustrate applications of photons *per se*, rather than general optical devices such as lenses, etc. that could operate under a classical theory of light. The laser is an extremely important application and is discussed above under stimulated emission.

Individual photons can be detected by several methods. The classic photomultiplier tube exploits the photoelectric effect: a photon landing on a metal plate ejects an electron, initiating an ever-amplifying avalanche of electrons. Charge-coupled device chips use a similar effect in semiconductors: an incident photon generates a charge on a microscopic capacitor that can be detected. Other detectors such as Geiger counters use the ability of photons to ionize gas molecules, causing a detectable change in conductivity.

Planck's energy formula $E = h\nu$ is often used by engineers and chemists in design, both to compute the change in energy resulting from a photon absorption and to predict the frequency of the light emitted for a given energy transition. For example, the emission spectrum of a fluorescent light bulb can be designed using gas molecules with different electronic energy levels and adjusting the typical energy with which an electron hits the gas molecules within the bulb.

Under some conditions, an energy transition can be excited by "two" photons that individually would be insufficient. This allows for higher resolution microscopy, because the sample absorbs energy only in the region where two beams of different colors overlap significantly, which can be made much smaller than the excitation volume of a single beam. Moreover, these photons cause less damage to the sample, since they are of lower energy.

In some cases, two energy transitions can be coupled so that, as one system absorbs a photon, another nearby system "steals" its energy and re-emits a photon of a different frequency. This is the basis of fluorescence resonance energy transfer, a technique that is used in molecular biology to study the interaction of suitable proteins.

Several different kinds of hardware random number generator involve the detection of single photons. In one example, for each bit in the random sequence that is to be produced, a photon is sent to a beam-splitter. In such a situation, there are two possible outcomes of equal probability. The actual outcome is used to determine whether the next bit in the sequence is "0" or "1".

Recent research

Much research has been devoted to applications of photons in the field of quantum optics. Photons seem well-suited to be elements of an extremely fast quantum computer, and the quantum entanglement of photons is a focus of research. Nonlinear optical processes are another active research area, with topics such as two-photon absorption, self-phase modulation, modulational instability and optical parametric oscillators. However, such processes generally do not require the assumption of photons *per se*; they may often be modeled by treating atoms as nonlinear oscillators. The nonlinear process of spontaneous parametric down conversion is often used to produce single-photon states. Finally, photons are essential in some aspects of optical communication, especially for quantum cryptography.

Chapter 6

Spintronics

Spintronics (a neologism meaning "spin transport electronics"), also known as magnetoelectronics, is an emerging technology that exploits both the intrinsic spin of the electron and its associated magnetic moment, in addition to its fundamental electronic charge, in solid-state devices.

History

Spintronics emerged from discoveries in the 1980s concerning spin-dependent electron transport phenomena in solid-state devices. This includes the observation of spin-polarized electron injection from a ferromagnetic metal to a normal metal by Johnson and Silsbee (1985), and the discovery of giant magnetoresistance independently by Albert Fert et al. and Peter Grünberg et al. (1988). The origins of spintronics can be traced back even further to the ferromagnet/superconductor tunneling experiments pioneered by Meservey and Tedrow, and initial experiments on magnetic tunnel junctions by Julliere in the 1970s. The use of semiconductors for spintronics can be traced back at least as far as the theoretical proposal of a spin field-effect-transistor by Datta and Das in 1990.

Theory

Electrons are spin-1/2 fermions and therefore constitute a two-state system with spin "up" and spin "down". To make a spintronic device, the primary requirements are, first, a system that can generate a current of spin-polarized electrons comprising more of one spin species—up or down—than the other (called a spin injector), and, secondly, a separate system sensitive to the spin polarization of the electrons (spin detector). Manipulation of the electron spin during transport between injector and detector (especially in semiconductors) via spin precession can be accomplished using real external magnetic fields or effective fields caused by spin-orbit interaction.

Spin polarization in non-magnetic materials can be achieved either through the Zeeman effect in large magnetic fields and low temperatures, or by non-equilibrium methods. In the latter case, the non-equilibrium polarization will decay over a timescale called the "spin lifetime". Spin lifetimes of conduction electrons in metals are relatively short (typically less than 1 nanosecond). However in semiconductors the lifetimes can be very long (microseconds at low temperatures), especially when the electrons are isolated in local trapping potentials (for instance, at impurities, where lifetimes can be milliseconds).

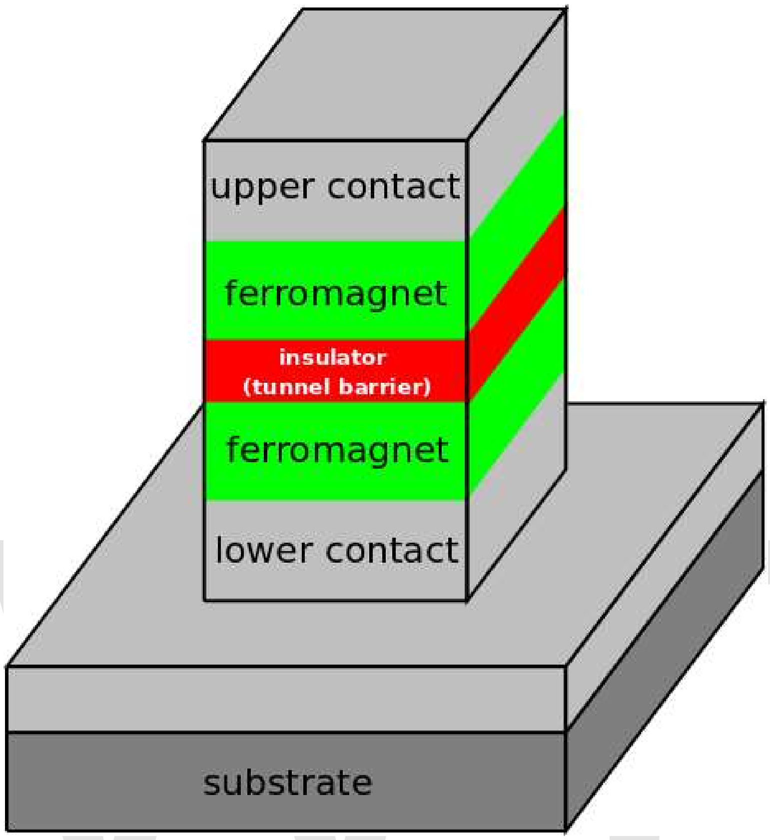
Metal-based spintronic devices

The simplest method of generating a spin-polarised current in a metal is to pass the current through a ferromagnetic material. The most common application of this effect is a giant magnetoresistance (GMR) device. A typical GMR device consists of at least two layers of ferromagnetic materials separated by a spacer layer. When the two magnetization vectors of the ferromagnetic layers are aligned, the electrical resistance will be lower (so a higher current flows at constant voltage) than if the ferromagnetic layers are anti-aligned. This constitutes a magnetic field sensor.

Two variants of GMR have been applied in devices: (1) current-in-plane (CIP), where the electric current flows parallel to the layers and (2) current-perpendicular-to-plane (CPP), where the electric current flows in a direction perpendicular to the layers.

Other metals-based spintronics devices:

Tunnel magnetoresistance



Magnetic tunnel junction (schematic)

The **Tunnel magnetoresistance** (TMR) is a magnetoresistive effect that occurs in magnetic tunnel junctions (MTJs). This is a component consisting of two ferromagnets separated by a thin insulator. If the insulating layer is thin enough (typically a few nanometers), electrons can tunnel from one ferromagnet into the other. Since this process is forbidden in classical physics, the tunnel magnetoresistance is a strictly quantum mechanical phenomenon.

Magnetic tunnel junctions are manufactured in thin film technology. On an industrial scale the film deposition is done by magnetron sputter deposition, on a laboratory scale molecular beam epitaxy, pulsed laser deposition and electron beam physical vapor deposition are also utilized. The junctions are prepared by photolithography.

Phenomenological description

The direction of the two magnetizations of the ferromagnetic films can be switched individually by an external magnetic field. If the magnetizations are in a parallel orientation it is more likely that electrons will tunnel through the insulating film than if they are in the oppositional (antiparallel) orientation. Consequently, such a junction can be switched between two states of electrical resistance, one with low and one with very high resistance.

History

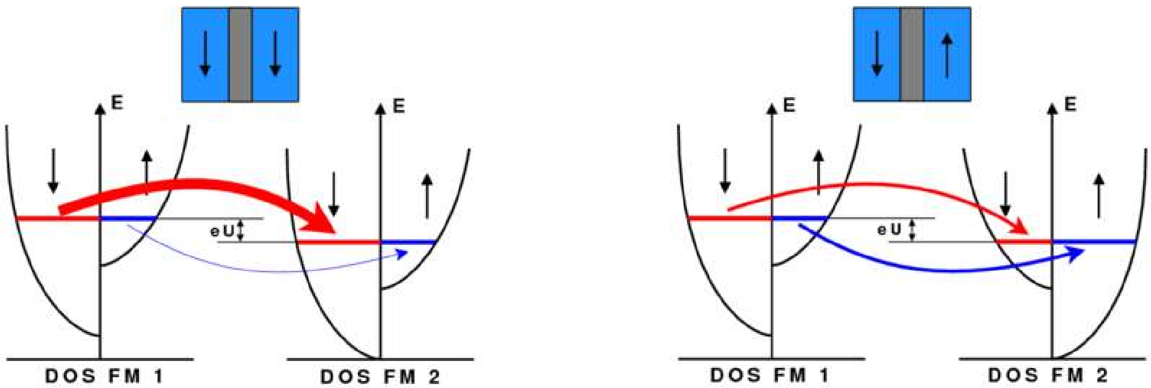
The effect was originally discovered in 1975 by M. Jullière (University of Rennes, France) in Fe/Ge-O/Co-junctions at 4.2 K. The relative change of resistance was around 14%, and did not attract much attention. In 1991 T. Miyazaki (University Tohoku, Japan) found an effect of 2.7% at room temperature. Later, in 1994, Miyazaki found 18% in junctions of iron separated by an amorphous aluminum oxide insulator and J. Moodera found 11.8% in junctions with electrodes of CoFe and Co . The highest effects observed to date with aluminum oxide insulators are around 70% at room temperature.

Since the year 2000, tunnel barriers of crystalline magnesium oxide (MgO) are under development. In 2001 Butler and Mathon independently made the theoretical prediction that using iron as the ferromagnet and MgO as the insulator, the tunnel magnetoresistance can reach several thousand percent . The same year, Bowen et al. were the first to report experiments showing a significant TMR in a MgO based magnetic tunnel junction [Fe/MgO/FeCo(001)] . In 2004, Parkin and Yuasa were able to make Fe/MgO/Fe junctions that reach over 200% TMR at room temperature . Today (2009) effects of up to 600% at room temperature and more than 1100% at 4.2 K are observed in junctions of CoFeB/MgO/CoFeB.

Applications

The read-heads of modern hard disk drives work on the basis of magnetic tunnel junctions. TMR, or more specifically the magnetic tunnel junction, is also the basis of MRAM, a new type of non-volatile memory. The 1st generation technologies relied on creating cross-point magnetic fields on each bit to write the data on it, although this approach has a scaling limit at around 90-130 nm. There are two 2nd generation techniques currently being developed: Thermal Assisted Switching (TAS) which is being developed by Crocus Technology, and Spin Torque Transfer (STT) on which Crocus, Hynix, IBM, and several other companies are working Further, magnetic tunnel junctions are also used for sensing applications.

Physical explanation



Two-current model for parallel and anti-parallel alignment of the magnetizations

The relative resistance change—or effect amplitude—is defined as

$$TMR := \frac{R_{ap} - R_p}{R_p}$$

where R_{ap} is the electrical resistance in the anti-parallel state, whereas R_{p} is the resistance in the parallel state.

The TMR effect was explained by Jullière with the spin polarizations of the ferromagnetic electrodes. The spin polarization P is calculated from the spin dependent density of states (DOS) $\mathcal{D}$ at the Fermi energy:

$$P = \frac{\mathcal{D}_\uparrow(E_{\mathrm{F}}) - \mathcal{D}_\downarrow(E_{\mathrm{F}})}{\mathcal{D}_\uparrow(E_{\mathrm{F}}) + \mathcal{D}_\downarrow(E_{\mathrm{F}})}$$

The spin-up electrons are those with spin orientation parallel to the external magnetic field, whereas the spin-down electrons have anti-parallel alignment with the external field. The relative resistance change is now given by the spin polarizations of the two ferromagnets, P_1 and P_2:

$$TMR = \frac{2P_1P_2}{1 - P_1P_2}$$

If no voltage is applied to the junction, electrons tunnel in both directions with equal rates. With a bias voltage U, electrons tunnel preferentially to the positive electrode. With the assumption that spin is conserved during tunneling, the current can be described in a two-current model. The total current is split in two partial currents, one for the spin-up electrons and another for the spin-down electrons. These vary depending on the magnetic state of the junctions.

There are two possibilities to obtain a defined anti-parallel state. First, one can use ferromagnets with different coercivities (by using different materials or different film thicknesses). And second, one of the ferromagnets can be coupled with an antiferromagnet (exchange bias). In this case the magnetization of the uncoupled electrode remains "free".

The TMR decreases with both increasing temperature and increasing bias voltage. Both can be understood in principle by magnon excitations and interactions with magnons.

It is obvious that the TMR becomes infinite if P_1 and P_2 equal 1, i.e. if both electrodes have 100% spin polarization. In this case the magnetic tunnel junction becomes a switch, that switches magnetically between low resistance and infinite resistance. Materials that come into consideration for this are called *ferromagnetic half-metals*. Their conduction electrons are fully spin polarized. This property is theoretically predicted for a number of materials (e.g. CrO_2, various Heusler alloys) but could not be experimentally confirmed to date.

Spin torque transfer

Spin torque transfer is a writing technology in which data is written by reorienting the magnetization of a thin magnetic layer in a tunnel magnetoresistance (TMR) element using a spin-polarized current. Charge carriers (such as electrons) have a property known as spin (physics) which is a small quantity of angular momentum intrinsic to the carrier. An electrical current is generally unpolarized (consisting of 50% spin-up and 50% spin-down electrons); a spin polarized current is one with more electrons of either spin. By passing a current through a thick magnetic layer, one can produce a spin-polarized current.

At very small device scales it is possible that a spin-polarized current can transfer its spin angular momentum to a small magnetic element. Spin torque transfer magnetic RAM (STT-MRAM or SPRAM) has the advantages of lower power consumption and better scalability over conventional MRAM. Spin torque transfer technology has the potential to make possible MRAM devices combining low current requirements and reduced cost; however, the amount of current needed to reorient the magnetization is at present too high for most commercial applications, and the reduction of this current density alone is the basis for current academic research in spin electronics.

Hynix Semiconductor and Grandis formed a partnership in April 2008 to explore commercial development of STT-RAM technology.

Hitachi and Tohoku University demonstrated a 32-Mbit spin-transfer torque RAM in June 2009.

Other companies working on STT-MRAM include Crocus Technology and Spin Transfer Technologies.

Applications

Motorola has developed a 1st generation 256 kb MRAM based on a single magnetic tunnel junction and a single transistor and which has a read/write cycle of under 50 nanoseconds (Everspin, Motorola's spin-off, has since developed a 4 Mbit version). There are two 2nd generation MRAM techniques currently in development: Thermal Assisted Switching (TAS) which is being developed by Crocus Technology, and Spin Torque Transfer (STT) on which Crocus, Hynix, IBM, and several other companies are working.

Another design in development, called Racetrack memory, encodes information in the direction of magnetization between domain walls of a ferromagnetic metal wire.

Semiconductor-based spintronic devices

Ferromagnetic semiconductor sources (like manganese-doped gallium arsenide GaMnAs), increase the interface resistance with a tunnel barrier, or using hot-electron injection.

Spin detection in semiconductors is another challenge, which has been met with the following techniques:

- Faraday/Kerr rotation of transmitted/reflected photons
- Circular polarization analysis of electroluminescence
- Nonlocal spin valve (adapted from Johnson and Silsbee's work with metals)
- Ballistic spin filtering

The latter technique was used to overcome the lack of spin-orbit interaction and materials issues to achieve spin transport in silicon, the most important semiconductor for electronics.

Because external magnetic fields (and stray fields from magnetic contacts) can cause large Hall effects and magnetoresistance in semiconductors (which mimic spin-valve effects), the only conclusive evidence of spin transport in semiconductors is demonstration of spin precession and dephasing in a magnetic field non-collinear to the injected spin orientation. This is called the Hanle effect.

Applications

Applications such as semiconductor lasers using spin-polarized electrical injection have shown threshold current reduction and controllable circularly polarized coherent light output. Future applications may include a spin-based transistor having advantages over MOSFET devices such as steeper sub-threshold slope.

Chapter 7

Spinmechatronics & Biexciton

Spinmechatronics

Spinmechatronics is neologism referring to an emerging field of research concerned with the exploitation of spin-dependent phenomena and established spintronic methodologies and technologies in conjunction with electro-mechanical, magno-mechanical, acousto-mechanical and opto-mechanical systems. Most especially, spinmechatronics (or spin mechatronics) concerns the integration of micro- and nano-mechatronic systems with spin physics and spintronics.

History and origins

Whilst spinmechatronics has been recognised only recently (2008) as an independent field, hybrid spin-mechanical system development dates back to the early nineteen-nineties, with devices combining spintronics and micromechanics emerging at the turn of the twenty-first century.

One of the longest established spinmechatronic systems is the Magnetic Resonance Force Microscope or MRFM. First proposed by J. A. Sidles in a seminal paper of 1991 – and since extensively developed both theoretically and experimentally by a number of international research groups – the MRFM operates by coupling a magnetically loaded micro-mechanical cantilever to an excited nuclear, proton or electron spin system. The MRFM concept effectively combines scanning atomic force microscopy (AFM) with magnetic resonance spectroscopy to provide a spectroscopic tool of unparalleled sensitivity. Nanometre resolution is possible, and the technique potentially forms the basis for ultra-high sensitivity, ultra-high resolution magnetic, biochemical, biomedical, and clinical diagnostics.

The synergy of micromechanics and established spintronic technologies for sensing applications is one of the most significant spinmechatronic developments of the last decade. At the beginning of this century, strain sensors incorporating magnetoresistive technologies emerged and a wide range of devices exploiting similar principles are likely to realize research and commercial potential by 2015.

Contemporary innovation in spinmechatronics drives forward the independent advancement of cutting-edge science in spin physics, spintronics and micro- and nano-mechatronics and catalyses the development of wholly new instrumentation, control and fabrication techniques to facilitate and exploit their integration.

Key constitutive technologies

Micro- and nano- mechatronics

MEMS: micro-electromechanical systems are the key ingredient of micro-mechatronics. Micro-electromechanical systems are – as the name suggests – devices with significant dimensions in the micrometre regime or less . Highly suited to integration with electronic and microwave circuitry, they provide the key to electro-mechanical functionalities unachievable with classical precision mechatronics. Commercialisation of mass produced Microelectromechanical systems products is rapidly picking up pace and includes printer ink-jet technology, 3D accelerometers, integrated pressure sensors, and Digital Light Processing (DLP) displays. At the cutting edge of Microelectromechanical systems fabrication and integration technologies are nano- electromechanical systems (NEMS). Typical examples are micrometres long, tens of nanometres thick, and have mechanical resonance frequencies approaching 100 MHz. Their small physical dimensions and mass (of order pico-grams) makes them highly sensitive to changes in stiffness; this, their synergy with mechanical and data processing systems, and the option of attaching chemical/ biological molecules, makes them ideal for ultra high-performance mechanical, chemical and biological sensing applications.

Spin physics

Spin physics is a broad and active area of condensed-matter physics research. 'Spin' in this context refers to a quantum mechanical property of certain elementary particles and nuclei, and should not be confused with the classical (and better-known) concept of rotation. Spin physics spans studies of nuclear, electron and proton magnetic resonance, magnetism, and certain areas of optics. Spintronics is a branch of spin physics. Perhaps the two best known applications of spin physics are Magnetic Resonance Imaging (or MRI) and the spintronic giant-magnetoresistive (GMR) hard disk read head.

Spintronics

Spintronic magnetoresistance is a major scientific and commercial success story. Today, most families own a spintronic device: the giant-magnetoresistive (GMR) hard disk read head in their computer. The science that gave rise to this phenomenal business opportunity – and earned the 2007 Nobel Prize for Physics – was the recognition that electrical carriers are characterized by both charge and spin. Today, tunnelling-magnetoresistance (TMR) – which uses the electron spin as a label to allow or forbid electron tunnelling – dominates the hard disk market and is rapidly establishing itself in areas as diverse as magnetic logic devices and biosensors . Ongoing development is pushing the frontiers of TMR devices towards the nanoscale.

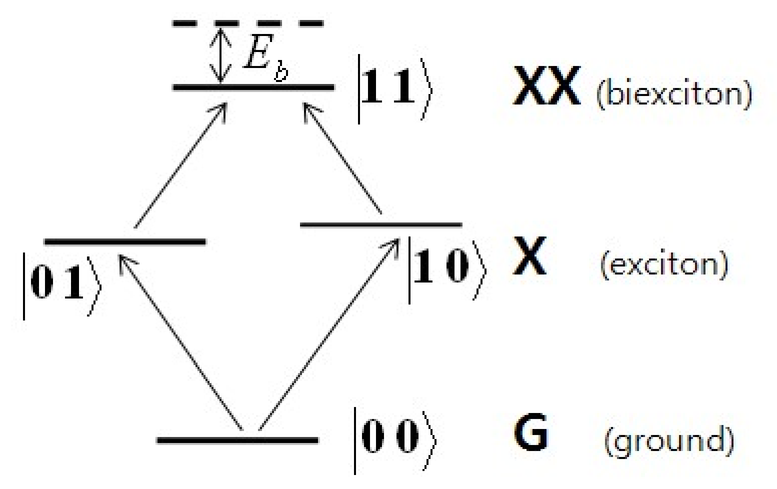
Biexciton

Biexcitons are created from two free excitons.

Formation of biexcitons

In quantum information and computation, it is essential to construct coherent combinations of quantum states. The basic quantum operations can be performed on a sequence of pairs of physically distinguishable quantum bits and, therefore, can be illustrated by a simple four-level system.

In an optically driven system where the $|01\rangle$ and $|10\rangle$ states can be directly excited, direct excitation of the upper $|11\rangle$ level from the ground state $|00\rangle$ is usually forbidden and the most efficient alternative is coherent nondegenerate two-photon excitation, using $|01\rangle$ or $|10\rangle$ as an intermediate state.



Model for a single quantum dots. E_b is the biexciton binding energy

At high exciton densities in some materials, the formation of exciton molecules called biexcitons is also observed. This is the equivalent process to the formation of an H_2 molecule from two isolated hydrogen atoms.

Biexcitons have been observed in a number of compound semiconductors, including CdS, ZnSe, ZnO and especially copper chloride. CuCl has a band gap at 3.40 eV, and the ground state exciton is observed at 3.20 eV, implying that Rx = 0.2 eV. At high densities, a new feature is observed in the absorption spectrum at 3.18 eV. This is attributed to biexciton formation. The energy difference between the two features tells us that the binding energy of the biexciton is 0.02 eV.

Attempts to observe biexcitons in materials like GaAs have been hindered by the nonlinear saturation effects described.

Observation of biexcitons

Three possibilities of observing biexcitons exist:

(a) excitation from the one-exciton band to the biexciton band (pump-probe experiments);

(b) two-photon absorption of light from the ground state to the biexciton state;

(c) luminescence from a biexciton state made up from two free excitons in a dense exciton system.

Binding energy of biexcitons

The biexciton is a quasi-particle formed from two excitons, and its energy is expressed as

$$E_b = 2E_X - E_{XX}$$

where E_{XX} is the biexciton energy, E_X is the exciton energy, and

E_b is the biexciton binding energy.

When a biexciton is annihilated, it disintegrates into a free exciton and a photon. The energy of the photon is smaller than that of the biexciton by the biexciton binding energy, so the biexciton luminescence peak appears on the low-energy side of the exciton peak.

The biexciton binding energy in semiconductor quantum dots has been the subject of extensive theoretical study. Because a biexciton is a composite of two electrons and two holes, we must solve a four-body problem under spatially restricted conditions. The biexciton binding energies for CuCl quantum dots, as measured by the site selective luminescence method, increased with decreasing quantum dot size. The data were well fitted by the function

$$B_{XX} = \frac{c_1}{a^2} + \frac{c_2}{a} + B_{bulk}$$

where B_{XX} is biexciton binding energy, a is the radius of the quantum dots, B_{bulk} is the binding energy of bulk crystal, and c_1 and c_2 are fitting parameters.

A simple model for describing binding energy of biexcitons

In the effective-mass approximation, the Hamiltonian of the system consisting of two electrons (1, 2) and two holes (a, b) is given by

$$H_{XX} = -\frac{\hbar^2}{2m_e^*}(\nabla_1{}^2 + \nabla_2{}^2) - \frac{\hbar^2}{2m_h^*}(\nabla_a{}^2 + \nabla_b{}^2) + V$$

where m_e^* and m_h^* are the effective masses of electrons and holes, respectively, and

$$V = V_{12} - V_{1a} - V_{1b} - V_{2a} - V_{2b} + V_{ab}$$

where V_{ij} denotes the Coulomb interaction between the charged particles i and j (i,j = 1,2,a,b denote the two electrons and two holes in the biexciton) given by

$$V_{ij} = \frac{e^2}{\epsilon|\mathbf{r}_i - \mathbf{r}_j|}$$

where ε is the dielectric constant of the material.

Denoting $\mathbf{R}$ and $\mathbf{r}$ are the c.m. coordinate and the relative coordinate of the biexciton, respectively, and $M = m_e^* + m_h^*$ is the effective mass of the exciton, the Hamiltonian becomes

$$H_{XX} = -\frac{\hbar^2}{4M}\nabla_R{}^2 - \frac{\hbar^2}{M}\nabla_r{}^2 - \frac{\hbar^2}{2\mu}(\nabla_{1a}{}^2 + \nabla_{2b}{}^2) + V$$

where $1/\mu = 1/m_e^* + 1/m_h^*$; $\nabla_{1a}{}^2$ and $\nabla_{2b}{}^2$ are the Laplacians with respect to relative coordinates between electron and hole, respectively. And $\nabla_r{}^2$ is that with respect to relative coordinate between the c. m. of excitons, and $\nabla_R{}^2$ is that with respect to the c. m. coordinate $\mathbf{R}$ of the system.

In the units of the exciton Rydberg and Bohr radius, the Hamiltonian can be written in dimensionless form

$$H_{XX} = -(\nabla_{1a}{}^2 + \nabla_{2b}{}^2) - 2\sigma(1+\sigma)^2\nabla_r{}^2 + V$$

where $\sigma = m_e^*/m_h^*$ with neglecting kinetic energy operator of c. m. motion. And V can be written as

$$V = 2(\frac{1}{r_{12}} - \frac{1}{r_{1a}} - \frac{1}{r_{1b}} - \frac{1}{r_{2a}} - \frac{1}{r_{2b}} + \frac{1}{r_{ab}})$$

To solve the problem of the bound states of the biexciton complex, it is required to find the wave functions ψ satisfying the wave equation

$$H_{XX}\psi = E_{XX}\psi$$

If the eigenvalue E_{XX} can be obtained, the binding energy of the biexciton can be also acquired

$$E_b = 2E_X - E_{XX}$$

where E_b is the binding energy of the biexciton and E_X is the energy of exciton.

Binding energy in nanotubes

Biexcitons with bound complexes formed by two excitons are surprisingly stable for carbon nanotube in a wide diameter range. Thus, a biexciton binding energy exceeding the inhomogeneous exciton line width is predicted for a wide range of nanotubes.

The biexciton binding energy in carbon nanotube is quite accurately approximated by an inverse dependence on *r*, except perhaps for the smallest values of *r*.

$$E_{XX} \approx \frac{0.195eV}{r}$$

The actual biexciton binding energy is inversely proportional to the physical nanotube radius.

Binding energy in CuCl QDs

The binding energy of biexcitons increase with the decrease in their size and its size dependence and bulk value are well represented by the expression

$$\frac{78}{a^{*2}} + \frac{52}{a^*} + 33 \text{ (meV)}$$

where a^* is the effective radius of microcrystallites in a unit of nm. The enhanced Coulomb interaction in microcrystallites still increase the biexciton binding energy in the large-size regime, where the quantum confinement energy of excitons is not considerable.

Lifetime of biexcitons and excitons

In solid state physics, the recombination of biexcitons (XX), which can be regarded as an entity of two weakly bound excitons (X), is often discussed in a simple way of argumentation: The recombination of any of the two excitons forming the biexciton will result in its disappearance. Therefore one would expect that the biexciton lifetime, τ_2, is about half of that of the exciton, τ_1.

However, some elementary points of view should be taken into account. For example, a direct radiative recombination into the ground state of the crystal, allowed for the exciton state, is optically forbidden for singlet biexcitons. Furthermore, the spin structure and the Coulomb interaction of the electrons and the holes forming the excitons and the biexcitons has to be considered.

In a CdSe/ZnSe quantum dot, the lifetime of biexciton is measured as 310 ps, and that of exciton is about 290 ps. The ratio of the two lifetimes is not about 2, but about 1.

The final state of the biexciton recombination is an optically allowed (bright) exciton state. Because of exchange interaction, the excitons are split into optically allowed states (bright, X^*) and optically forbidden states (dark, X^O). The spin flip time between X^* and X^O states, τ_{1s}, is expected to be in the range of several ns, i.e., much longer than the bright exciton lifetime.

The time evolution of the probabilities for the population of the biexciton and exciton states, w_2 and w_1, respectively:

$$\frac{dw_2}{dt} = -\frac{w_2}{\tau_2}, \frac{dw_1}{dt} = -\frac{w_1}{\tau_{1t}} + \frac{w_2}{\tau_2}$$

with $1 / \tau_{1t} = 1 / \tau_1 + 1 / \tau_{1s}$, where τ_1 and τ_2 are the X^* and the XX radiative recombination times, respectively.

Solving this equation system, the X and XX dynamics can be described quantitatively. As a result, while decay of XX is monoexponential, the onset of the X line is significantly delayed, resulting in a plateaulike characteristics of the exciton decay curve.

Chapter 8

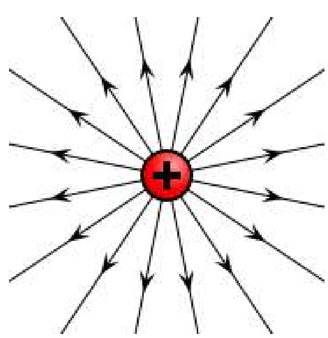
Electric Charge

Electric charge is a physical property of matter which causes it to experience a force when near other electrically charged matter. Classically, electric charge comes in two types. These are called positive and negative. Two positively charged substances, or objects, experience a mutual repulsive force, as do two negatively charged objects. Positively charged objects and negatively charged objects experience an attractive force. The SI unit of electric charge is the coulomb (C). The study of how charged substances interact is classical electrodynamics, which is accurate insofar as quantum effects can be ignored.

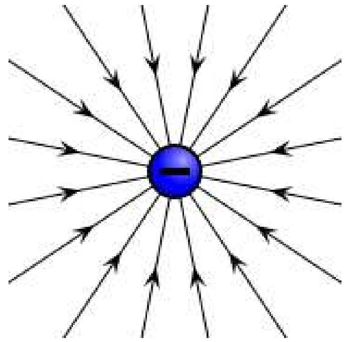
The *electric charge* is a fundamental conserved property of some subatomic particles, which determines their electromagnetic interaction. Electrically charged matter is influenced by, and produces, electromagnetic fields. The interaction between a moving charge and an electromagnetic field is the source of the electromagnetic force, which is one of the four fundamental forces.

Twentieth-century experiments demonstrated that electric charge is *quantized*; that is, it comes in multiples of individual small units called the elementary charge, *e*, approximately equal to 1.602×10^{-19} coulombs (except for particles called quarks which have charges that are multiples of ⅓*e*). The proton has a charge of *e*, and the electron has a charge of −*e*. The study of charged particles, and how their interactions are mediated by photons, is quantum electrodynamics.

Overview



Electric field induced by a positive electric charge



Electric field induced by a negative electric charge

Charge is the fundamental property of forms of matter that exhibit electrostatic attraction or repulsion in the presence of other matter. Electric charge is a characteristic property of many subatomic particles. The charges of free-standing particles are integer multiples of the elementary charge e; we say that electric charge is *quantized*. Michael Faraday, in his electrolysis experiments, was the first to note the discrete nature of electric charge. Robert Millikan's oil-drop experiment demonstrated this fact directly, and measured the elementary charge.

By convention, the charge of an electron is −1, while that of a proton is +1. Charged particles whose charges have the same sign repel one another, and particles whose charges have different signs attract. Coulomb's law quantifies the electrostatic force between two particles by asserting that the force is proportional to the product of their charges, and inversely proportional to the square of the distance between them.

The charge of an antiparticle equals that of the corresponding particle, but with opposite sign. Quarks have fractional charges of either $-\frac{1}{3}$ or $+\frac{2}{3}$, but free-standing quarks have never been observed (the theoretical reason for this fact is asymptotic freedom).

The electric charge of a macroscopic object is the sum of the electric charges of the particles that make it up. This charge is often small, because matter is made of atoms, and atoms typically have equal numbers of protons and electrons, in which case their charges cancel out, yielding a net charge of zero, thus making the atom neutral.

An *ion* is an atom (or group of atoms) that has lost one or more electrons, giving it a net positive charge (cation), or that has gained one or more electrons, giving it a net negative charge (anion). *Monatomic ions* are formed from single atoms, while *polyatomic ions* are formed from two or more atoms that have been bonded together, in each case yielding an ion with a positive or negative net charge.

During the formation of macroscopic objects, usually the constituent atoms and ions will combine in such a manner that they form structures composed of neutral *ionic compounds* electrically bound to neutral atoms. Thus macroscopic objects tend toward being neutral overall, but macroscopic objects are rarely perfectly net neutral.

There are times when macroscopic objects contain ions distributed throughout the material, rigidly bound in place, giving an overall net positive or negative charge to the object. Also, macroscopic objects made of conductive elements, can more or less easily (depending on the element) take on or give off electrons, and then maintain a net negative or positive charge indefinitely. When the net electric charge of an object is non-zero and motionless, the phenomenon is known as static electricity. This can easily be produced by rubbing two dissimilar materials together, such as rubbing amber with fur or glass with silk. In this way non-conductive materials can be charged to a significant degree, either positively or negatively. Of course, charge taken from one material is simply moved to the other material, leaving an opposite charge of the same magnitude behind. The law of *conservation of charge* always applies, giving the object from which a negative charge has been taken a positive charge of the same magnitude, and vice-versa.

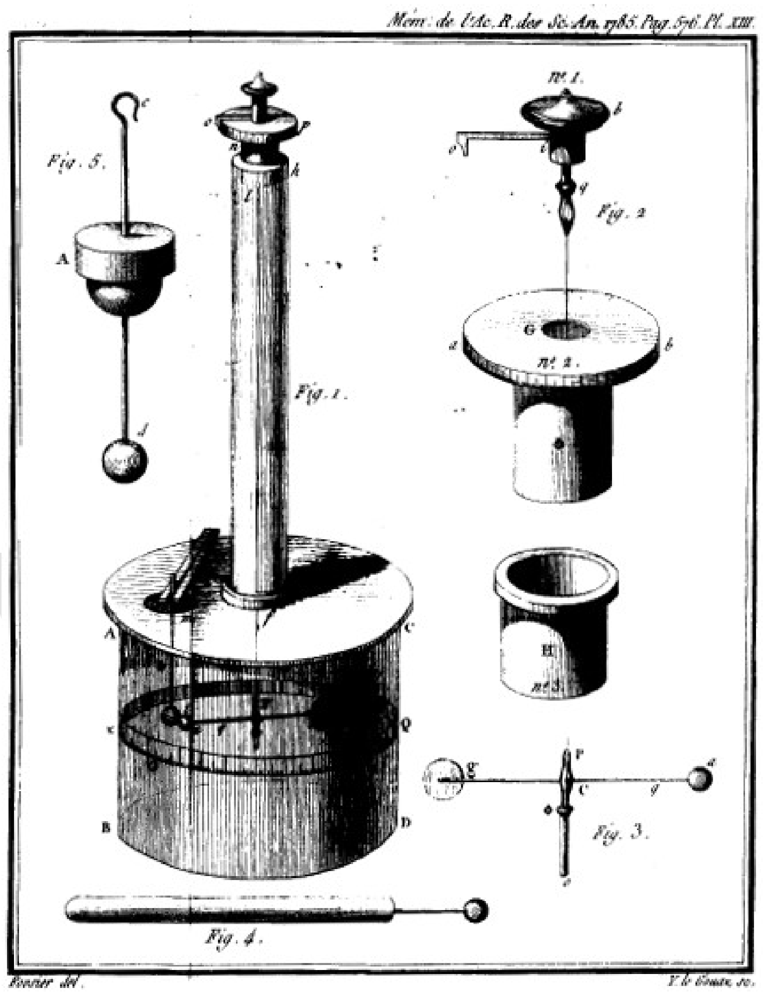
Even when an object's net charge is zero, charge can be distributed non-uniformly in the object (e.g., due to an external electromagnetic field, or bound polar molecules). In such cases the object is said to be polarized. The charge due to polarization is known as bound charge, while charge on an object produced by electrons gained or lost from outside the object is called *free charge*. The motion of electrons in conductive metals in a specific direction is known as electric current.

Units

The SI unit of quantity of electric charge is the coulomb, which is equivalent to about 6.242×10^{18} *e* (*e* is the charge of a proton). Hence, the charge of an electron is approximately -1.602×10^{-19} C. The coulomb is defined as the quantity of charge that has passed through the cross section of an electrical conductor carrying one ampere within one second. The symbol *Q* is often used to denote a quantity of electricity or charge. The quantity of electric charge can be directly measured with an electrometer, or indirectly measured with a ballistic galvanometer.

After finding the quantized character of charge, in 1891 George Stoney proposed the unit 'electron' for this fundamental unit of electrical charge. This was before the discovery of the particle by J.J. Thomson in 1897. The unit is today treated as nameless, referred to as "elementary charge", "fundamental unit of charge", or simply as "e". A measure of charge should be a multiple of the elementary charge *e*, even if at large scales, charge seems to behave as a real quantity. In some contexts it is meaningful to speak of fractions of a charge; for example in the charging of a capacitor, or in the fractional quantum Hall effect.

History



Coulomb's torsion balance

As reported by the ancient Greek philosopher Thales of Miletus around 600 BC, charge (or *electricity*) could be accumulated by rubbing fur on various substances, such as amber. The Greeks noted that the charged amber buttons could attract light objects such as hair. They also noted that if they rubbed the amber for long enough, they could even get a spark to jump. This property derives from the triboelectric effect.

In 1600, the English scientist William Gilbert returned to the subject in *De Magnete*, and coined the New Latin word *electricus* from *ηλεκτρον* (*elektron*), the Greek word for "amber", which soon gave rise to the English words "electric" and "electricity." He was followed in 1660 by Otto von Guericke, who invented what was probably the first electrostatic generator. Other European pioneers were Robert Boyle, who in 1675 stated that electric attraction and repulsion can act across a vacuum; Stephen Gray, who in 1729 classified materials as conductors and insulators; and C. F. du Fay, who proposed in 1733 that electricity came in two varieties which cancelled each other, and expressed this in terms of a two-fluid theory. When glass was rubbed with silk, du Fay said that the glass was charged with *vitreous electricity*, and when amber was rubbed with fur, the amber was said to be charged with *resinous electricity*. In 1839, Michael Faraday showed that the apparent division between static electricity, current electricity and bioelectricity was incorrect, and all were a consequence of the behavior of a single kind of electricity appearing in opposite polarities. It is arbitrary which polarity you call positive and which you call negative. Positive charge can be defined as the charge left on a glass rod after being rubbed with silk.

One of the foremost experts on electricity in the 18th century was Benjamin Franklin, who argued in favour of a one-fluid theory of electricity. Franklin imagined electricity as being a type of invisible fluid present in all matter; for example he believed that it was the glass in a Leyden jar that held the accumulated charge. He posited that rubbing insulating surfaces together caused this fluid to change location, and that a flow of this fluid constitutes an electric current. He also posited that when matter contained too little of the fluid it was "negatively" charged, and when it had an excess it was "positively" charged. Arbitrarily (or for a reason that was not recorded) he identified the term "positive" with vitreous electricity and "negative" with resinous electricity. William Watson arrived at the same explanation at about the same time.

Static electricity and electric current

Static electricity and electric current are two separate phenomena, both involving electric charge, and may occur simultaneously in the same object. Static electricity is a reference to the electric charge of an object and the related electrostatic discharge when two objects are brought together that are not at equilibrium. An electrostatic discharge creates a change in the charge of each of the two objects. In contrast, electric current is the flow of electric charge through an object, which produces no net loss or gain of electric charge. Although charge flows between two objects during an electrostatic discharge, time is too short for current to be maintained.

Electrification by friction

Experiment I

Let a piece of glass and a piece of resin, neither of which exhibits any electrical properties, be rubbed together and left with the rubbed surfaces in contact. They will still exhibit no electrical properties. Let them be separated. They will now attract each other.

If a second piece of glass be rubbed with a second piece of resin, and if the piece be then separated and suspended in the neighbourhood of the former pieces of glass and resin, it may be observed:

1. that the two pieces of glass repel each other.
2. that each piece of glass attracts each piece of resin.
3. that the two pieces of resin repel each other.

These phenomena of attraction and repulsion are called electrical phenomena and the bodies which exhibit them are said to be 'electrified', or to be 'charged with electricity'.

Bodies may be electrified in many other ways, as well as by friction.

The electrical properties of the two pieces of glass are similar to each other but opposite to those of the two pieces of resin: the glass attracts what the resin repels and repels what the resin attracts.

If a body electrified in any manner whatever behaves as the glass does, that is, if it repels the glass and attracts the resin, the body is said to be 'vitreously' electrified, and if it attracts the glass and repels the resin it is said to be 'resinously' electrified. All electrified bodies are found to be either vitreously or resinously electrified.

It is the established convention of the scientific community to define the vitreous electrification as positive, and the resinous electrification as negative. The exactly opposite properties of the two kinds of electrification justify us in indicating them by opposite signs, but the application of the positive sign to one rather than to the other kind must be considered as a matter of arbitrary convention, just as it is a matter of convention in mathematical diagram to reckon positive distances towards the right hand.

No force, either of attraction or of repulsion, can be observed between an electrified body and a body not electrified.

We now know that the Franklin/Watson model was fundamentally correct. There is only one kind of electrical charge, and only one variable is required to keep track of the amount of charge. On the other hand, just knowing the charge is not a complete description of the situation. Matter is composed of several kinds of electrically charged particles, and these particles have many properties, not just charge.

The most common charge carriers are the positively charged proton and the negatively charged electron. The movement of any of these charged particles constitutes an electric current. In many situations, it suffices to speak of the *conventional current* without regard to whether it is carried by positive charges moving in the direction of the conventional current and/or by negative charges moving in the opposite direction. This macroscopic viewpoint is an approximation that simplifies electromagnetic concepts and calculations.

At the opposite extreme, if one looks at the microscopic situation, one sees there are many ways of carrying an electric current, including: a flow of electrons; a flow of electron "holes" which act like positive particles; and both negative and positive particles (ions or other charged particles) flowing in opposite directions in an electrolytic solution or a plasma).

Beware that in the common and important case of metallic wires, the direction of the conventional current is opposite to the drift velocity of the actual charge carriers, i.e. the electrons. This is a source of confusion for beginners.

Properties

Aside from the properties described in articles about electromagnetism, charge is a relativistic invariant. This means that any particle that has charge *Q*, no matter how fast it goes, always has charge *Q*. This property has been experimentally verified by showing that the charge of *one* helium nucleus (two protons and two neutrons bound together in a nucleus and moving around at high speeds) is the same as *two* deuterium nuclei (one proton and one neutron bound together, but moving much more slowly than they would if they were in a helium nucleus).

Conservation of electric charge

The total electric charge of an isolated system remains constant regardless of changes within the system itself. This law is inherent to all processes known to physics and can be derived in a local form from gauge invariance of the wave function. The conservation of charge results in the charge-current continuity equation. More generally, the net change in charge density ρ within a volume of integration V is equal to the area integral over the current density **J** through the closed surface $S = \partial V$, which is in turn equal to the net current I:

$$-\frac{d}{dt}\int_V \rho \, \mathrm{d}V = \oiint_{\partial V} \mathbf{J} \cdot \mathrm{d}\mathbf{S} = \int J dS \cos\theta = I.$$

Thus, the conservation of electric charge, as expressed by the continuity equation, gives the result:

$$I = \frac{dQ}{dt}.$$

The charge transferred between times t_i and t_f is obtained by integrating both sides:

$$Q = \int_{t_\mathrm{i}}^{t_\mathrm{f}} I \, \mathrm{d}t$$

where I is the net outward current through a closed surface and Q is the electric charge contained within the volume defined by the surface.

Chapter 9

Magnetoresistive Random Access Memory

Magnetoresistive Random Access Memory is a non-volatile computer memory (NVRAM) technology that has been under development since the 1990s. Continued increases in density of existing memory technologies – notably flash RAM and DRAM – kept it in a niche role in the market, but its proponents believe that the advantages are so overwhelming that magnetoresistive RAM will eventually become dominant for all types of memory, becoming a true universal memory.

Description

Unlike conventional **RAM** chip technologies, in MRAM data is not stored as electric charge or current flows, but by magnetic storage elements. The elements are formed from two ferromagnetic plates, each of which can hold a magnetic field, separated by a thin insulating layer. One of the two plates is a permanent magnet set to a particular polarity, the other's field can be changed to match that of an external field to store memory. This configuration is known as a spin valve and is the simplest structure for a MRAM bit. A memory device is built from a grid of such "cells".

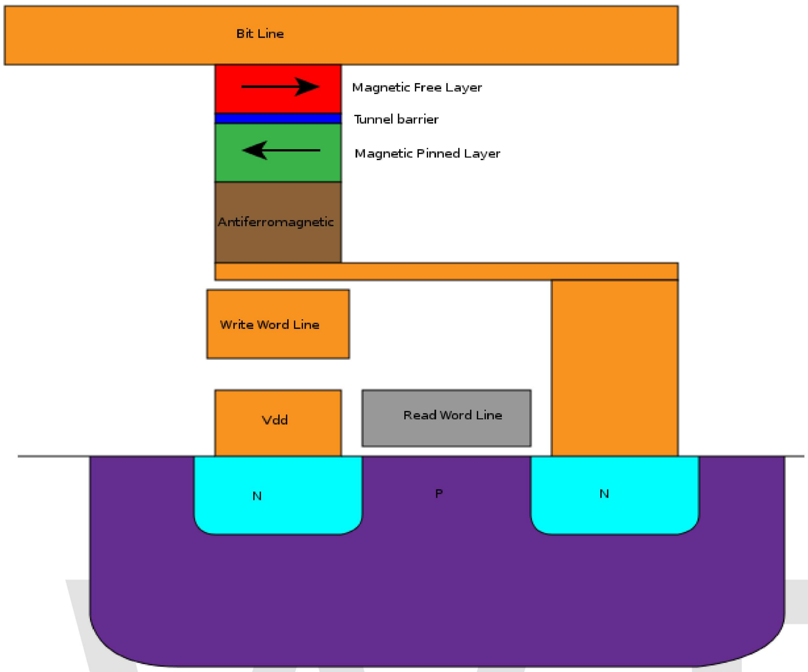
The simplest method of reading is accomplished by measuring the electrical resistance of the cell. A particular cell is (typically) selected by powering an associated transistor which switches current from a supply line through the cell to ground. Due to the magnetic tunnel effect, the electrical resistance of the cell changes due to the orientation of the fields in the two plates. By measuring the resulting current, the resistance inside any particular cell can be determined, and from this the polarity of the writable plate. Typically if the two plates have the same polarity this is considered to mean "1", while if the two plates are of opposite polarity the resistance will be higher and this means "0".

Data is written to the cells using a variety of means. In the simplest, each cell lies between a pair of write lines arranged at right angles to each other, above and below the cell. When current is passed through them, an induced magnetic field is created at the junction, which the writable plate picks up. This pattern of operation is similar to core memory, a system commonly used in the 1960s. This approach requires a fairly substantial current to generate the field, however, which makes it less interesting for low-power uses, one of MRAM's primary disadvantages. Additionally, as the device is scaled down in size, there comes a time when the induced field overlaps adjacent cells over a small area, leading to potential false writes. This problem, the half-select (or write disturb) problem, appears to set a fairly large size for this type of cell. One experimental solution to this problem was to use circular domains written and read using the giant magnetoresistive effect, but it appears this line of research is no longer active.

Another approach, the **toggle mode**, uses a multi-step write with a modified multi-layer cell. The cell is modified to contain an "artificial antiferromagnet" where the magnetic orientation alternates back and forth across the surface, with both the pinned and free layers consisting of multi-layer stacks isolated by a thin "coupling layer". The resulting layers have only two stable states, which can be toggled from one to the other by timing the write current in the two lines so one is slightly delayed, thereby "rotating" the field. Any voltage less than the full write level actually increases its resistance to flipping. That means that other cells located along one of the write lines will not suffer from the half-select problem, allowing for smaller cell sizes.

A newer technique, **spin-torque-transfer (STT)** or **Spin Transfer Switching**, uses spin-aligned ("polarized") electrons to directly torque the domains. Specifically, if the electrons flowing into a layer have to change their spin, this will develop a torque that will be transferred to the nearby layer. This lowers the amount of current needed to write the cells, making it about the same as the read process. There are concerns that the "classic" type of MRAM cell will have difficulty at high densities due to the amount of current needed during writes, a problem STT avoids. For this reason, the STT proponents expect the technique to be used for devices of 65 nm and smaller. The downside is the need to maintain the spin coherence. Overall, the STT requires much less write current than conventional or toggle MRAM. However, higher speed operation still requires higher current.

Other potential arrangements include "Thermal Assisted Switching" (TAS-MRAM) which briefly heats up (reminiscent of phase-change memory) the magnetic tunnel junctions during the write process and keeps the MTJ's stable at a colder temperature the rest of the time; and "vertical transport MRAM" (VMRAM), which uses current through a vertical column to change magnetic orientation, a geometric arrangement that reduces the write disturb problem and so can be used at higher density.



Simplified structure of an MRAM cell

Comparison with other systems

Density

The main determinant of a memory system's cost is the density of the components used to make it up. Smaller components, and fewer of them, mean that more "cells" can be packed onto a single chip, which in turn means more can be produced at once from a single silicon wafer. This improves yield, which is directly related to cost.

DRAM uses a small capacitor as a memory element, wires to carry current to and from it, and a transistor to control it – referred to as a "1T1C" cell. This makes DRAM the highest density RAM currently available, and thus the least expensive, which is why it is used for the majority of RAM found in a computer.

MRAM is physically similar to DRAM in makeup, although often does not require a transistor for the write operation. However, as mentioned above, the most basic MRAM cell suffers from the half-select problem, which limits cell sizes to around 180 nm or more. Toggle-mode MRAM offers a much smaller size before this becomes a problem, apparently around 90 nm, the same size as most current DRAM products. To be worth putting into wide production, however, it is generally believed that MRAM will have to move to the 65 nm size of the most advanced memory devices, which will require the use of STT.

Power consumption

Since the capacitors used in DRAM lose their charge over time, memory assemblies using them must periodically *refresh* all the cells in their chips approximately 20 times a second, reading each one and re-writing its contents. This demands a constant power supply, which is why DRAM loses its memory when power is turned off on the computer. As DRAM cells decrease in size, the refresh cycles become shorter, and the power draw more continuous.

In contrast, MRAM requires no refresh at any time. Not only does this mean it retains its memory with the power turned off, but also that there is no constant power draw. While the read process theoretically requires more power than the same process in a DRAM, in practice the difference appears to be very close to zero. However, the write process requires more power in order to overcome the existing field stored in the junction, varying from three to eight times the power required during reading. Although the exact amount of power savings depends on the nature of the work – more frequent writing will require more power – in general MRAM proponents expect much lower power consumption (up to 99% less) compared to DRAM. STT-based MRAMs eliminate the difference between reading and writing, further reducing power requirements.

It is also worth comparing MRAM with another common memory system, flash RAM. Like MRAM, flash does not lose its memory when power is removed, which makes it very common as a "hard disk replacement" in small devices such as digital audio players or digital cameras. When used for reading, flash and MRAM are very similar in power requirements. However, flash is re-written using a large pulse of voltage (about 10 V) that is stored up over time in a charge pump, which is both power-hungry and time consuming. Additionally the current pulse physically degrades the flash cells, which means flash can only be written to some finite number of times before it must be replaced.

In contrast, MRAM requires only slightly more power to write than read, and no change in the voltage, eliminating the need for a charge pump. This leads to much faster operation, lower power consumption, and an indefinitely long "lifetime".

Performance

DRAM performance is limited by the rate at which the charge stored in the cells can be drained (for reading) or stored (for writing). MRAM operation is based on measuring voltages rather than charges or currents, so there is less "settling time" needed. IBM researchers have demonstrated MRAM devices with access times on the order of 2 ns, somewhat better than even the most advanced DRAMs built on much newer processes. A team at the German Physikalisch-Technische Bundesanstalt have demonstrated MRAM devices with 1 ns settling times, better than the currently accepted theoretical limits for DRAM, although the demonstration was a single cell. The differences compared to flash are far more significant, with write times as much as thousands of times faster.

The only current memory technology that easily competes with MRAM in terms of performance is static RAM, or SRAM. SRAM consists of a series of transistors arranged in a flip-flop, which will hold one of two states as long as power is applied. Since the transistors have a very low power requirement, their switching time is very low. However, since an SRAM cell consists of several transistors, typically four or six, its density is much lower than DRAM. This makes it expensive, which is why it is used only for small amounts of high performance memory, notably the CPU cache in almost all modern CPU designs.

Although MRAM is not quite as fast as SRAM, it is close enough to be interesting even in this role. Given its much higher density, a CPU designer may be inclined to use MRAM to offer a much larger but somewhat slower cache, rather than a smaller but faster one. It remains to be seen how this trade off will play out in the future.

Overall

MRAM has similar performance to SRAM, similar density of DRAM but much lower power consumption than DRAM, and is much faster and suffers no degradation over time in comparison to flash memory. It is this combination of features that some suggest make it the "universal memory", able to replace SRAM, DRAM, EEPROM and flash. This also explains the huge amount of research being carried out into developing it.

However, to date, MRAM has not been widely adopted in the market. It may be that vendors are not prepared to take the risk of allocating a modern fab to MRAM production when such fabs cost upwards of a few billion dollars to build and can instead generate revenue by serving developed markets producing flash and DRAM memories.

The very latest fabs seem to be used for flash, for example producing 16 Gbit parts produced by Samsung on a 50 nm process. Slightly older fabs are being used to produce most DDR2 DRAM, most of which is produced on a one-generation-old 90 nm process rather than using up scarce leading-edge capacity.

In comparison, MRAM is still largely "in development", and being produced on older non-critical fabs. The only commercial product widely available at this point is Everspin's 4 Mbit part, produced on a several-generations-old 180 nm process. As demand for flash continues to outstrip supply, it appears it will be some time before a company can afford to "give up" one of their latest fabs for MRAM production. Even then, MRAM designs currently do not come close to flash in terms of cell size, even using the same fab.

History

Most of the following has been taken from mram-info web site:

- 1955 - Magnetic core memory had the same reading writing principle as MRAM
- 1988 - European scientists (Albert Fert and Peter Grünberg) discovered the "giant magnetoresistive effect" in thin-film structures.

- 1995 - Motorola (later to become Freescale) initiates work on MRAM development
- 2000 - IBM and Infineon established a joint MRAM development program.
- 2000 - Spintec laboratory's first Spin Torque Transfer patent.
- 2002 - NVE Announces Technology Exchange with Cypress Semiconductor.
- 2003 - A 128 kbit MRAM chip was introduced, manufactured with a 180 nm lithographic process

2004

- June - Infineon unveiled a 16-Mbit prototype, manufactured with a 180 nm lithographic process
- September - MRAM becomes a standard product offering at Freescale.
- October - Taiwan developers of MRAM tape out 1 Mbit parts at TSMC.
- October - Micron drops MRAM, mulls other memories.
- December - TSMC, NEC, Toshiba describe novel MRAM cells.
- December - Renesas Technology trumpets a high performance, high-reliability MRAM technology.
- Spintech laboratory's first observation of Thermal Assisted Switching (TAS) as MRAM approach.
- Crocus Technology is founded; the company is a developer of 2nd generation MRAM

2005

- January - Cypress Semiconductor samples MRAM, using NVE IP.
- March - Cypress to Sell MRAM Subsidiary.
- June - Honeywell posts data sheet for 1-Mbit rad-hard MRAM using a 150 nm lithographic process
- August - MRAM record: memory cell runs at 2 GHz.
- November - Renesas Technology and Grandis collaborate on development of 65 nm MRAM employing spin torque transfer (STT).
- November - NVE receives an SBIR grant to research cryptographic tamper-responsive memory.
- December - Sony announced the first lab-produced spin-torque-transfer MRAM, which utilizes a spin-polarized current through the tunneling magnetoresistance layer to write data. This method consumes less power and is more scalable than conventional MRAM. With further advances in materials, this process should allow for densities higher than those possible in DRAM.
- December - Freescale Semiconductor Inc. demonstrates an MRAM that uses magnesium oxide, rather than an aluminum oxide, allowing for a thinner insulating tunnel barrier and improved bit resistance during the write cycle, thereby reducing the required write current.
- Spintec laboratory gives Crocus Technology exclusive license on its patents.

Current status

2006

- February - Toshiba and NEC announced a 16 Mbit MRAM chip with a new "power-forking" design. It achieves a transfer rate of 200 MB/s, with a 34 ns cycle time - the best performance of any MRAM chip. It also boasts the smallest physical size in its class—78.5 square millimeters—and a low voltage requirement of 1.8 volts.
- July - On July 10, Austin Texas - Freescale Semiconductor begins marketing a 4-Mbit MRAM chip, which sells for approximately $25.00 per chip.

2007

- R&D moving to spin transfer torque RAM (SPRAM)
- February - Tohoku University and Hitachi developed a prototype 2 Mbit Non-Volatile RAM Chip employing spin-transfer torque switching.
- August - "IBM, TDK Partner In Magnetic Memory Research on Spin Transfer Torque Switching" IBM and TDK to lower the cost and boost performance of MRAM to hopefully release a product to market.
- November - Toshiba applied and proved the spin transfer torque switching with perpendicular magnetic anisotropy MTJ device.
- November - NEC Develops World's Fastest SRAM-Compatible MRAM With Operation Speed of 250 MHz.

2008

- Japanese satellite, SpriteSat, to use Freescale MRAM to replace SRAM and FLASH components
- June - Samsung and Hynix become partner on STT-MRAM
- June - Freescale spins off MRAM operations as new company Everspin
- August - Scientists in Germany have developed next-generation MRAM that is said to operate as fast as fundamental performance limits allow, with write cycles under 1 nanosecond.

2009

- June - Hitachi and Tohoku University demonstrated a 32-Mbit spin-transfer torque RAM (SPRAM).
- June - Crocus Technology and Tower Semiconductor announce deal port Crocus' MRAM process technology to Tower's manufacturing environment

2010

- June - Hitachi and Tohoku Univ announced Multi-level SPRAM

Applications

Proposed uses for MRAM include devices such as:

- Aerospace and military systems
- Digital cameras
- Notebooks
- Smart cards
- Mobile telephones
- Cellular base stations
- Personal Computers
- Battery-Backed SRAM replacement
- Datalogging specialty memories (black box solutions)
- Media players
- Book readers

Chapter 10

Racetrack Memory

Racetrack Memory is an experimental non-volatile memory device under development at IBM's Almaden Research Center by a team led by Stuart Parkin. In early 2008, a 3-bit version was successfully demonstrated. If it is developed successfully, racetrack would offer storage density higher than comparable solid-state memory devices like flash RAM and similar to conventional disk drives, and also have much higher read/write performance. It is one of a number of new technologies trying to become a universal memory in the future.

Description

Racetrack Memory uses a spin-coherent electric current to move magnetic domains along a nanoscopic permalloy wire about 200 nm across and 100 nm thick. As current is passed through the wire, the domains pass by magnetic read/write heads positioned near the wire, which alter the domains to record patterns of bits. A Racetrack Memory device is made up of many such wires and read/write elements. In general operational concept, Racetrack Memory is similar to the earlier twistor memory or bubble memory of the 1960s and 70s. Both of these used electrical currents to "push" a magnetic pattern through a substrate. Dramatic improvements in magnetic detection capabilities, based on the development of spintronic magnetoresistive sensing materials and devices, allow the use of much smaller magnetic domains to provide far higher areal densities.

In production, it is expected that the wires can be scaled down to around 50 nm. There are two ways to arrange Racetrack Memory. The simplest is a series of flat wires arranged in a grid with read and write heads arranged nearby. A more widely studied arrangement uses U-shaped wires arranged vertically over a grid of read/write heads on an underlying substrate. This allows the wires to be much longer without increasing its 2D area, although the need to move individual domains further along the wires before they reach the read/write heads results in slower random access times. This does not present a real performance bottleneck; both arrangements offer about the same throughput. Thus the primary concern in terms of construction is practical; whether or not the 3D vertical arrangement is feasible to mass produce.

Comparison to other memory devices

Current projections suggest that Racetrack Memory will offer performance on the order of 20-32 ns to read or write a random bit. This compares to about 3,000,000 ns for a hard drive, or 6-40 ns for conventional DRAM. The authors of the primary work also discuss ways to improve the access times with the use of a "reservoir," improving to about 9.5 ns. Aggregate throughput, with or without the reservoir, is on the order of 250-670 Mbit/s for Racetrack Memory, compared to 102400 for dual channel DDR2 DRAM, 1000 for high-performance hard drives, and much slower performance on the order of 30 to 100 Mbit/s for Flash devices. The only current technology that offers a clear performance benefit over Racetrack Memory is SRAM, on the order of 2 ns, but is much more expensive and far lower density.

Flash, in particular, is a highly asymmetrical device. Although read performance is fairly fast, especially compared to a hard drive, writing is much slower. Flash works by "trapping" electrons in the chip surface, and requires a burst of high voltage to remove this charge and reset the cell. In order to do this, charge is accumulated in a device known as a charge pump, which takes a relatively long time to charge up. In the case of "NOR" flash, which allows random bit-wise access like Racetrack Memory, read times are on the order of 70 ns, while write times are much slower, about 2,500 ns. To address this concern, "NAND" flash allows reading and writing only in large blocks, but this means that the time to access any random *bit* is greatly increased, to about 1,000 ns. Additionally, the use of the burst of high voltage physically degrades the cell, so most flash devices allow on the order of 100,000 writes to any particular bit before their operation becomes unpredictable. Wear leveling and other techniques can spread this out, but only if the underlying data can be re-arranged.

The key determinant of the cost of any memory device is the physical size of the storage medium. The reason for this is due to the way memory devices are fabricated. In the case of solid-state devices like Flash or DRAM, a large "wafer" of silicon is processed into many individual devices, which are then cut apart and packaged. The cost of packaging is about $1 per device, so as the density increases and the number of bits per devices increases with it, the *cost per bit* falls by an equal amount. In the case of hard drives, data is stored on a number of rotating platters, and the cost of the device is strongly related to the number of platters. Increasing the density allows the number of platters to be reduced for any given amount of storage.

In most cases memory devices store one bit in any given location, so they are typically compared in terms of "cell size", a cell storing one bit. Cell size itself is given in units of F^2, where F is the design rule, representing usually the metal line width. Flash and racetrack both store multiple bits per cell, but the comparison can still be made. For instance, modern hard drives appear to be rapidly reaching their current theoretical limits around 650 nm^2/bit, which is defined primarily by our capability to read and write to tiny patches of the magnetic surface. DRAM has a cell size of about 6 F^2, SRAM is much worse at 120 F^2. NAND flash is currently the densest form of non-volatile memory in widespread use, with a cell size of about 4.5 F^2, but storing three bits per cell for an effective size of 1.5 F^2. NOR is slightly less dense, at an effective 4.75 F^2, accounting for 2-bit operation on a 9.5 F^2 cell size. In the vertical orientation (U-shaped) racetrack, about 10-20 bits are stored per cell, which itself can have a physical size of at least about 20 F^2. In addition, bits at different positions on the "track" would take different times (from ~10 ns to nearly a microsecond, or 10 ns/bit) to be accessed by the read/write sensor, because the "track" is moved at fixed rate (~100 m/s) past the read/write sensor.

Racetrack Memory is one of a number of new technologies aiming to replace Flash, and potentially offer a "universal" memory device applicable to a wide variety of roles. Other leading contenders include MRAM, PCRAM and FeRAM. Most of these technologies offer densities similar to Flash, in most cases worse, and their primary advantage is the lack of write endurance limits like those in Flash. Field-MRAM offers excellent performance as high as 3 ns access time, but requires a large 25-40 F^2 cell size. It might see use as a SRAM replacement, but not as a mass storage device. The highest densities from any of these devices is offered by PCRAM, which has a cell size of about 5.8 F^2, similar to Flash, as well as fairly good performance around 50 ns. Nevertheless, none of these can come close to competing with Racetrack Memory in overall terms, especially density. For example, 50 ns allows about five bits to be operated in a Racetrack Memory device, resulting in an effective cell size of 20/5=4 F^2, easily exceeding the performance-density product of PCM. On the other hand, without sacrificing bit density, the same 20 F^2 area can also fit 2.5 2-bit 8 F^2 alternative memory cells (such as RRAM or spin-torque transfer MRAM), each of which could individually operated much faster (~10 ns).

A difficulty for this technology arises from the need for high current density (>10^8 A/cm²); a 30 nm x 100 nm cross-section would require >3 mA. The resulting power draw would be higher than, for example, spin-torque transfer memory or Flash.

Development difficulties

One limitation of the early experimental devices was that the magnetic domains could only be pushed slowly through the wires, requiring current pulses on the orders of microseconds to move them successfully. This was unexpected, and led to performance roughly equal to hard drives, as much as 1000 times slower than predicted. Recent research at the University of Hamburg has traced this problem to microscopic imperfections in the crystal structure of the wires which led to the domains becoming "stuck" at these imperfections. Using an x-ray microscope to directly image the boundaries between the domains, their research found that domain walls would be moved by pulses as short as a few nanoseconds when these imperfections were absent. This corresponds to a macroscopic performance of about 110 m/s.

The voltage required to drive the domains along the racetrack would be proportional to the length of the wire. The current density must be sufficiently high to push the domain walls (as in electromigration). For example, a permalloy racetrack of resistivity $5*10^{-7}$ ohm-m, that is 1 cm long to cover an entire chip array, and uses a current density of $3*10^8$ A/cm², would require a driving voltage of 15 kV along the racetrack.

Chapter 11

Spin

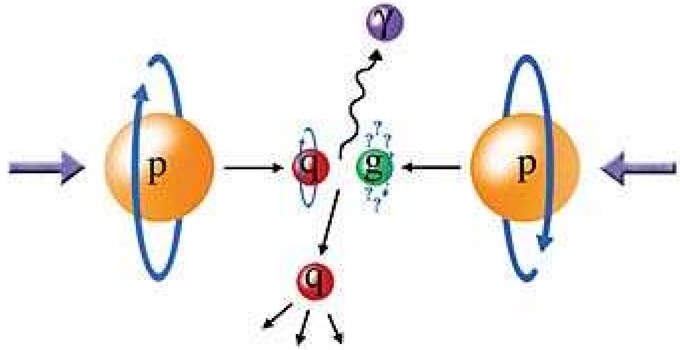
spin is a fundamental characteristic property of elementary particles, composite particles (hadrons), and atomic nuclei.

All elementary particles of a given kind have the same **spin quantum number**, an important part of a particle's quantum state. When combined with the spin-statistics theorem, the spin of electrons results in the Pauli exclusion principle, which in turn underlies the periodic table of chemical elements. The **spin direction** (also called *spin* for short) of a particle is an important intrinsic degree of freedom.

Wolfgang Pauli was the first to propose the concept of spin, but he did not name it. In 1925, Ralph Kronig, George Uhlenbeck, and Samuel Goudsmit suggested a physical interpretation of particles spinning around their own axis. The mathematical theory was worked out in depth by Pauli in 1927. When Paul Dirac derived his relativistic quantum mechanics in 1928, electron spin was an essential part thereof.

Spin is a type of angular momentum, where angular momentum is defined in the modern way. This modern definition of angular momentum is *not* the same as the historical classical mechanics definition, **L**=**r**×**p**. (The latter definition, which does not include spin, is more specifically called "orbital angular momentum".)

Since spin is a type of angular momentum, it has the same units: J·s in SI units. In practice, however, SI units are almost never used to describe spin: Instead, it is written as a multiple of the reduced Planck's constant $\hbar$. In natural units, the $\hbar$ is omitted, so the units of spin are implied. However, by definition the "spin quantum number" is always unitless.



The head-on collision of a quark (the red ball) from one proton (the orange ball) with a gluon (the green ball) from another proton with opposite spin; spin is represented by the blue arrows circling the protons and the quark. The blue question marks circling the gluon represent the question: Are gluons polarized? The particles ejected from the collision are a shower of quarks and one photon of light (the purple ball).

Spin quantum number

As the name suggests, spin was originally conceived as the rotation of a particle around some axis. This picture is correct in so far as spins obey the same mathematical laws as do quantized angular momenta. On the other hand, spins have some peculiar properties that distinguish them from orbital angular momenta:

- Spin quantum numbers may take on half-integer values;
- The spin of a charged particle is associated with a magnetic dipole moment with a g-factor differing from 1. This is incompatible with classical physics, which assumes that the charge and mass of the particle are distributed evenly in spheres of equal radius.

Elementary particles

Elementary particles are particles for which there is no known way of dividing them into smaller units. Theoretical and experimental studies have shown that the spin possessed by such particles cannot be explained by postulating that they are made up of even smaller particles rotating about a common center of mass; as far as can be determined, these elementary particles are true point particles. The spin of an elementary particle is a truly intrinsic physical property, akin to the particle's electric charge and rest mass.

Let the **spin quantum number** s be $n/2$, where n can be any non-negative integer. Hence the allowed values of s are 0, 1/2, 1, 3/2, 2, etc. The value of s for an elementary particle depends only on the type of particle, and cannot be altered in any known way (in contrast to the *spin direction* described below). The spin angular momentum S of any physical system is quantized. The allowed values of S are:

$$S = \hbar \sqrt{s(s+1)},$$

where $\hbar$ is the reduced Planck's constant. In contrast, orbital angular momentum can only take on integer quantum numbers.

All known matter is ultimately composed of elementary particles called fermions, and all elementary fermions have s=1/2. Examples of fermions are the electron and positron, the quarks making up protons and neutrons, and the neutrinos. Elementary particles emit and receive one or more particles called bosons. This boson exchange gives rise to the three fundamental interactions ("forces") of the Standard model of particle physics; hence bosons are also called *force carriers*. These bosons have s=1. The best understood boson is the photon. Electromagnetism is the force that results when charged particles exchange photons.

Theory predicts the existence of two bosons whose s differs from 1. The force carrier for gravity is the hypothetical graviton; theory suggests that it has s=2. The Higgs mechanism predicts that elementary particles acquire nonzero rest mass by exchanging hypothetical Higgs bosons with an all-pervasive Higgs field. Theory predicts that the Higgs boson has s=0. If so, it would be the only elementary particle for which this is the case.

Composite particles

The spin of composite particles, such as protons, neutrons, and atomic nuclei is usually understood to mean the total angular momentum. This is the sum of the spins and orbital angular momenta of the constituent particles. Such a composite spin is subject to the same quantization condition as any other angular momentum.

Composite particles are often referred to as having a definite spin, just like elementary particles; for example, the proton is a spin-1/2 particle. This is understood to refer to the spin of the lowest-energy internal state of the composite particle (i.e., a given spin and orbital configuration of the constituents).

It is not always easy to deduce the spin of a composite particle from first principles; for example, even though we know that the proton is a spin-1/2 particle, the question of how this spin is distributed among the three internal valence quarks and the surrounding sea quarks and gluons is an active area of research.

Delta baryons, which decay into protons and neutrons, have spin 3/2. All the three quarks inside a Δ particle have their spin axis pointing in the same direction, unlike the nearly identical proton and neutron (called "nucleons") in which the intrinsic spin of one of the three constituent quarks is always opposite the spin of the other two. This difference in spin alignment is the only quantum number distinction between the Δ^+ and Δ^0 and ordinary nucleons.

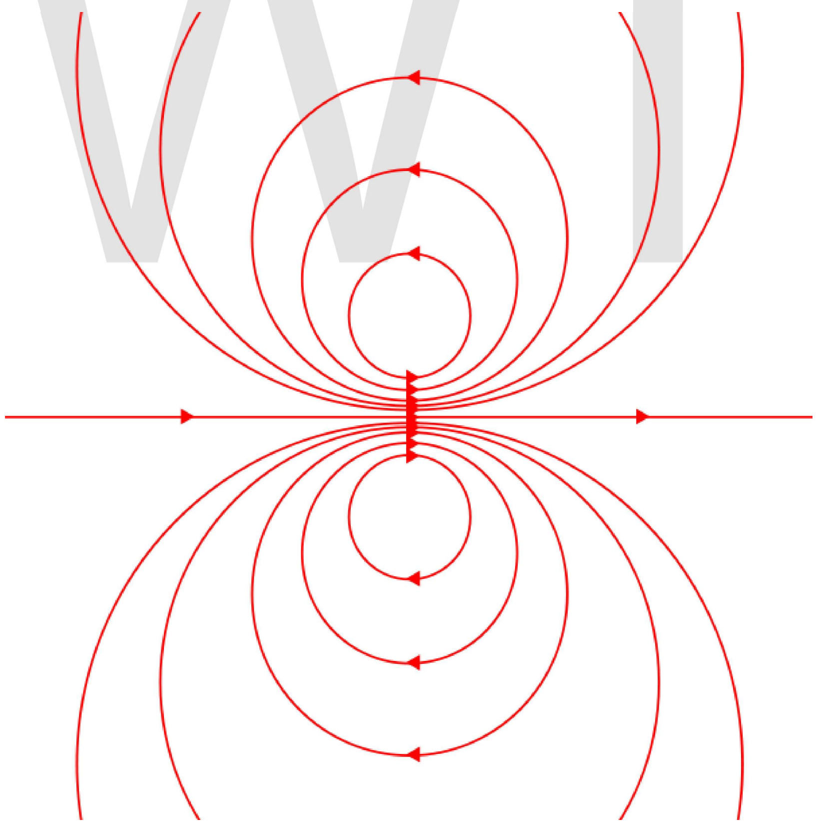
Atoms and molecules

The spin of atoms and molecules is the sum of the spins of unpaired electrons, which may be parallel or antiparallel. It is responsible for paramagnetism.

The spin-statistics theorem

The spin of a particle has crucial consequences for its properties in statistical mechanics. Particles with half-integer spin obey Fermi-Dirac statistics, and are known as fermions. They are required to occupy antisymmetric quantum states. This property forbids fermions from sharing quantum states – a restriction known as the Pauli exclusion principle. Particles with integer spin, on the other hand, obey Bose-Einstein statistics, and are known as bosons. These particles occupy "symmetric states", and can therefore share quantum states. The proof of this is known as the spin-statistics theorem, which relies on both quantum mechanics and the theory of special relativity. In fact, "the connection between spin and statistics is one of the most important applications of the special relativity theory".

Magnetic moments



Magnetic field lines around a *magnetostatic dipole*; the magnetic dipole itself is in the center and is seen from the side.

Particles with spin can possess a magnetic dipole moment, just like a rotating electrically charged body in classical electrodynamics. These magnetic moments can be experimentally observed in several ways, e.g. by the deflection of particles by inhomogeneous magnetic fields in a Stern–Gerlach experiment, or by measuring the magnetic fields generated by the particles themselves.

The intrinsic magnetic moment μ of an elementary particle with charge q, mass m, and spin angular momentum S, is

$$\mu = g \frac{q}{2m} S$$

where the dimensionless quantity g is called the g-factor. For exclusively orbital rotations it would be 1 (assuming that the mass and the charge occupy spheres of equal radius).

The electron, being a charged elementary particle, possesses a nonzero magnetic moment. One of the triumphs of the theory of quantum electrodynamics is its accurate prediction of the electron g-factor, which has been experimentally determined to have the value −2.002 319 304 3622(15), with the digits in parentheses denoting measurement uncertainty in the last two digits at one standard deviation. The value of 2 arises from the Dirac equation, a fundamental equation connecting the electron's spin with its electromagnetic properties, and the correction of 0.002 319 304… arises from the electron's interaction with the surrounding electromagnetic field, including its own field. Composite particles also possess magnetic moments associated with their spin. In particular, the neutron possesses a non-zero magnetic moment despite being electrically neutral. This fact was an early indication that the neutron is not an elementary particle. In fact, it is made up of quarks, which are electrically charged particles. The magnetic moment of the neutron comes from the spins of the individual quarks and their orbital motions.

Neutrinos are both elementary and electrically neutral. The minimally extended Standard Model that takes into account finite neutrino masses predicts neutrino magnetic moments of:

$$\mu_\nu \approx 3 \times 10^{-19} \mu_B \frac{m_\nu}{\text{eV}}$$

where the μ_ν are the neutrino magnetic moments, m_ν are the neutrino masses, and μ_B is the Bohr magneton. New physics above the electroweak scale could, however, lead to significantly higher neutrino magnetic moments. It can be shown in a model independent way that neutrino magnetic moments larger than about $10^{-14}\mu_B$ are unnatural, because they would also lead to large radiative contributions to the neutrino mass. Since the neutrino masses cannot exceed about 1 eV, these radiative corrections must then be assumed to be fine tuned to cancel out to a large degree.

The measurement of neutrino magnetic moments is an active area of research. As of 2001, the latest experimental results have put the neutrino magnetic moment at less than 1.2×10^{-10} times the electron's magnetic moment.

In ordinary materials, the magnetic dipole moments of individual atoms produce magnetic fields that cancel one another, because each dipole points in a random direction. Ferromagnetic materials below their Curie temperature, however, exhibit magnetic domains in which the atomic dipole moments are locally aligned, producing a macroscopic, non-zero magnetic field from the domain. These are the ordinary "magnets" with which we are all familiar.

In paramagnetic materials, the magnetic dipole moments of individual atoms spontaneously align with an externally applied magnetic field. In diamagnetic materials, on the other hand, the magnetic dipole moments of individual atoms spontaneously align oppositely to any externally applied magnetic field, even if it requires energy to do so.

The study of the behavior of such "spin models" is a thriving area of research in condensed matter physics. For instance, the Ising model describes spins (dipoles) that have only two possible states, up and down, whereas in the Heisenberg model the spin vector is allowed to point in any direction. These models have many interesting properties, which have led to interesting results in the theory of phase transitions.

Spin direction

Spin projection quantum number and spin multiplicity

In classical mechanics, the angular momentum of a particle possesses not only a magnitude (how fast the body is rotating), but also a direction (either up or down on the axis of rotation of the particle). Quantum mechanical spin also contains information about direction, but in a more subtle form. Quantum mechanics states that the component of angular momentum measured along any direction (say along the *z*-axis) can only take on the values

$$\hbar s_z, \qquad s_z = -s, -s+1, \cdots, s-1, s$$

where *s* is the principal spin quantum number discussed in the previous section. One can see that there are $2s+1$ possible values of s_z. The number $2s+1$ is called the multiplicity of the spin system. For example, there are only two possible values for a spin-1/2 particle: $s_z = +1/2$ and $s_z = -1/2$. These correspond to quantum states in which the spin is pointing in the $+z$ or $-z$ directions respectively, and are often referred to as "spin up" and "spin down".

Spin vector

For a given quantum state, it is possible to describe a spin vector $\langle S \rangle$ whose components are the expectation values of the spin components along each axis, i.e., $\langle S \rangle = [\langle s_x \rangle, \langle s_y \rangle, \langle s_z \rangle]$. This vector describes the "direction" in which the spin is pointing, corresponding to the classical concept of the axis of rotation. It turns out that the spin vector is not very useful in actual quantum mechanical calculations, because it cannot be measured directly — s_x, s_y and s_z cannot possess simultaneous definite values, because of a quantum uncertainty relation between them. However, for statistically large collections of particles that have been placed in the same pure quantum state, such as through the use of a Stern-Gerlach apparatus, the spin vector does have a well-defined experimental meaning: It specifies the direction in ordinary space in which a subsequent detector must be oriented in order to achieve the maximum possible probability (100%) of detecting every particle in the collection. For spin-1/2 particles, this maximum probability drops off smoothly as the angle between the spin vector and the detector increases, until at an angle of 180 degrees — that is, for detectors oriented in the opposite direction to the spin vector—the expectation of detecting particles from the collection reaches a minimum of 0%.

As a qualitative concept, the spin vector is often handy because it is easy to picture classically. For instance, quantum mechanical spin can exhibit phenomena analogous to classical gyroscopic effects. For example, one can exert a kind of "torque" on an electron by putting it in a magnetic field (the field acts upon the electron's intrinsic magnetic dipole moment). The result is that the spin vector undergoes precession, just like a classical gyroscope. This phenomenon is used in nuclear magnetic resonance sensing.

Mathematically, quantum mechanical spin is not described by vectors as in classical angular momentum, but by objects known as spinors. There are subtle differences between the behavior of spinors and vectors under coordinate rotations. For example, rotating a spin-1/2 particle by 360 degrees does not bring it back to the same quantum state, but to the state with the opposite quantum phase; this is detectable, in principle, with interference experiments. To return the particle to its exact original state, one needs a 720 degree rotation. A spin-zero particle can only have a single quantum state, even after torque is applied. Rotating a spin-2 particle 180 degrees can bring it back to the same quantum state and a spin-4 particle should be rotated 90 degrees to bring it back to the same quantum state. The spin 2 particle can be analogous to a straight stick that looks the same even after it is rotated 180 degrees and a spin 0 particle can be imagined as sphere which looks the same after whatever angle it is turned through.

Mathematical formulation of spin

Spin operator

Spin obeys commutation relations analogous to those of the orbital angular momentum:

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k$$

where ε_{ijk} is the Levi-Civita symbol. It follows (as with angular momentum) that the eigenvectors of S^2 and S_z (expressed as kets in the total S basis) are:

$$S^2|s,m\rangle = \hbar^2 s(s+1)|s,m\rangle$$
$$S_z|s,m\rangle = \hbar m|s,m\rangle.$$

The spin raising and lowering operators acting on these eigenvectors give:

$$S_\pm|s,m\rangle = \hbar\sqrt{s(s+1) - m(m\pm 1)}|s,m\pm 1\rangle$$, where
$$S_\pm = S_x \pm iS_y.$$

But unlike orbital angular momentum the eigenvectors are not spherical harmonics. They are not functions of θ and φ. There is also no reason to exclude half integer values of s and m.

In addition to their other properties, all quantum mechanical particles possess an intrinsic spin (though it may have the intrinsic spin 0, too). The spin is quantized in units of the reduced action constant, such that the state function of the particle is, e.g., not $\psi = \psi(\mathbf{r})$, but $\psi = \psi(\mathbf{r}, \sigma)$,where σ is out of the following discrete set of values:

$$\sigma \in \{-s\cdot\hbar, -(s-1)\cdot\hbar, ..., +(s-1)\cdot\hbar, +s\cdot\hbar\}.$$

One distinguishes bosons (s=0 or 1 or 2 or ...) and fermions (s=1/2 or 3/2 or 5/2 or ...). The total angular momentum conserved in interaction processes is then the *sum* of the orbital angular momentum and the spin.

Pauli matrices and spin operators

The quantum mechanical operators associated with spin observables are:

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y, \quad S_z = \frac{\hbar}{2}\sigma_z.$$

In the special case of spin-1/2 σ_x, σ_y and σ_z are the three Pauli matrices, given by:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Spin and the Pauli exclusion principle

For systems of *N* identical particles this is related to the Pauli exclusion principle, which states that by interchanges of any two of the N particles one must have

$$\psi(\,\dots;\,\mathbf{r}_i,\sigma_i\,;\,\dots;\mathbf{r}_j,\sigma_j\,;\,\dots)\stackrel{!}{=}(-1)^{2s}\cdot\psi(\,\dots;\,\mathbf{r}_j,\sigma_j\,;\,\dots;\mathbf{r}_i,\sigma_i\,;\,\dots)\,.$$

Thus, for bosons the prefactor $(-1)^{2s}$ will reduce to +1, for fermions to –1. In quantum mechanics all particles are either bosons or fermions. In relativistic quantum field theories also "supersymmetric" particles exist, where linear combinations of bosonic and fermionic components appear. In two dimensions the prefactor $(-1)^{2s}$ can be replaced by any complex number of magnitude 1.

Electrons are fermions with s=1/2; quanta of light ("photons") are bosons with s=1. This shows also explicitly that the property *spin* cannot be fully explained as a classical intrinsic orbital angular momentum, e.g., similar to that of a "spinning top", since orbital angular rotations would lead to integer values of *s*. Instead one is dealing with an essential legacy of relativity. The photon, in contrast, is always relativistic (velocity $v\equiv c)$, and the corresponding classical theory, that of Maxwell, is also relativistic.

The above permutation postulate for *N*-particle state functions has most-important consequences in daily life, e.g. the periodic table of the chemists or biologists.

Spin and rotations

As described above, quantum mechanics states that component of angular momentum measured along any direction can only take a number of discrete values. The most convenient quantum mechanical description of particle's spin is therefore with a set of complex numbers corresponding to amplitudes of finding a given value of projection of its intrinsic angular momentum on a given axis. For instance, for a spin 1/2 particle, we would need two numbers $a_{\pm 1/2}$, giving amplitudes of finding it with projection of angular momentum equal to $\hbar/2$ and $-\hbar/2$, satisfying the requirement

$$|a_{1/2}|^2+|a_{-1/2}|^2\;=1.$$

Since these numbers depend on the choice of the axis, they transform into each other non-trivially when this axis is rotated. It's clear that the transformation law must be linear, so we can represent it by associating a matrix with each rotation, and the product of two transformation matrices corresponding to rotations A and B must be equal (up to phase) to the matrix representing rotation AB. Further, rotations preserve quantum mechanical inner product, and so should our transformation matrices:

$$\sum_{m=-j}^{j} a_m^* b_m = \sum_{m=-j}^{j} (\sum_{n=-j}^{j} U_{nm} a_n)^* (\sum_{k=-j}^{j} U_{km} b_k)$$

$$\sum_{n=-j}^{j} \sum_{k=-j}^{j} U_{np}^* U_{kq} = \delta_{pq}.$$

Mathematically speaking, these matrices furnish a unitary projective representation of the rotation group SO(3). Each such representation corresponds to a representation of the covering group of SO(3), which is SU(2). There is one n-dimensional irreducible representation of SU(2) for each dimension, though this representation is n-dimensional real for odd n and n-dimensional complex for even n (hence of real dimension 2n). For a rotation by angle θ in the plane with normal vector $\hat{\theta}$, *U* can be written

$$U = e^{\frac{i}{\hbar}\vec{\theta}\cdot\vec{S}},$$

where $\vec{\theta} = \theta\hat{\theta}$and $\vec{S}$ is the vector of spin operators.

More generic rotations can be built by compounding operators of this type. For example, a 3-D rotation of spin 1/2 particles can be written

$$\begin{pmatrix} a'_{1/2} \\ a'_{-1/2} \end{pmatrix} = \exp\left(i\sigma_z\gamma/2\right)\exp\left(i\sigma_x\beta/2\right)\exp\left(i\sigma_z\alpha/2\right)\begin{pmatrix} a_{1/2} \\ a_{-1/2} \end{pmatrix}$$

where α,β,γ are Euler angles.

It is interesting to calculate a rotation by 2π on a spin state. One finds that states with half-integral spin pick up a minus sign, while integral spin states are rotated into themselves.

This fact is crucial to the proof of the spin-statistics theorem.

Spin and Lorentz transformations

We could try the same approach to determine the behavior of spin under general Lorentz transformations, but we'd immediately discover a major obstacle. Unlike SO(3), the group of Lorentz transformations SO(3,1) is non-compact and therefore does not have any faithful unitary finite-dimensional representations.

In case of spin 1/2 particles, it is possible to find a construction that includes both a finite-dimensional representation and a scalar product that is preserved by this representation. We associate a 4-component Dirac spinor ψ with each particle. These spinors transform under Lorentz transformations according to the law

$$\psi' = \exp\left(\frac{1}{8}\omega_{\mu\nu}[\gamma_\mu, \gamma_\nu]\right)\psi$$

where γ_μ are gamma matrices and $\omega_{\mu\nu}$ is an antisymmetric 4x4 matrix parametrizing the transformation. It can be shown that the scalar product

$$\langle\psi|\phi\rangle = \bar{\psi}\phi = \psi^\dagger\gamma_0\phi$$

is preserved. (It is not, however, positive definite, so the representation is not unitary.)

Measuring spin along the x, y, and z axes

Each of the (hermitian) Pauli matrices has two eigenvalues, +1 and -1. The corresponding normalized eigenvectors are:

$$\begin{aligned}
\psi_{x+} &= \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \psi_{x-} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix},\\
\psi_{y+} &= \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}, \psi_{y-} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix},\\
\psi_{z+} &= \begin{pmatrix}1\\0\end{pmatrix}, \psi_{z-} = \begin{pmatrix}0\\1\end{pmatrix}.
\end{aligned}$$

By the postulates of quantum mechanics, an experiment designed to measure the electron spin on the x, y or z axis can only yield an eigenvalue of the spin operator (S_x, S_y, S_z) on that axis, $\frac{\hbar}{2}$ and $\frac{-\hbar}{2}$. The quantum state of a particle (with respect to spin), can be represented by a two component spinor:

$$\psi = \begin{pmatrix}a+bi\\c+di\end{pmatrix}.$$

When the spin of this particle is measured with respect to a given axis (in this example, the x-axis), the probability that its spin will be measured as $\frac{\hbar}{2}$ is just $|\langle\psi \mid \psi_{x+}\rangle|^2$. Correspondingly, the probability that its spin will be measured as $\frac{-\hbar}{2}$ is just $|\langle\psi \mid \psi_{x-}\rangle|^2$. Following the measurement, the spin state of the particle will collapse into the corresponding eigenstate. As a result, if the particle's spin along a given axis has been measured to have a given eigenvalue, all measurements will yield the same eigenvalue (since $|\langle\psi_{x+} \mid \psi_{x+}\rangle|^2 = 1$, etc), provided that no measurements of the spin are made along other axes.

Measuring spin along an arbitrary axis

The operator to measure spin along an arbitrary axis direction is easily obtained from the Pauli spin matrices. Let $u = (u_x, u_y, u_z)$ be an arbitrary unit vector. Then the operator for spin in this direction is simply $S_u = \hbar(u_x\sigma_x + u_y\sigma_y + u_z\sigma_z)/2$. The operator S_u has eigenvalues of $\pm\hbar/2$, just like the usual spin matrices. This method of finding the operator for spin in an arbitrary direction generalizes to higher spin states, one takes the dot product of the direction with a vector of the three operators for the three x,y,z axis directions.

A normalized spinor for spin-1/2 in the (u_x, u_y, u_z) direction (which works for all spin states except spin down where it will give 0/0), is:

$$\frac{1}{\sqrt{2+2u_z}}\begin{bmatrix} 1+u_z \\ u_x+iu_y \end{bmatrix}.$$

The above spinor is obtained in the usual way by diagonalizing the σ_u matrix and finding the eigenstates corresponding to the eigenvalues. In quantum mechanics, vectors are termed "normalized" when multiplied by a normalizing factor, which results in the vector having a length of unity.

Compatibility of spin measurements

Since the Pauli matrices do not commute, measurements of spin along the different axes are incompatible. This means that if, for example, we know the spin along the x-axis, and we then measure the spin along the y-axis, we have invalidated our previous knowledge of the x-axis spin. This can be seen from the property of the eigenvectors (i.e. eigenstates) of the Pauli matrices that:

$$|\langle\psi_{x+/-}|\psi_{y+/-}\rangle|^2 = |\langle\psi_{x+/-}|\psi_{z+/-}\rangle|^2 = |\langle\psi_{y+/-}|\psi_{z+/-}\rangle|^2 = \frac{1}{2}.$$

So when we measure the spin of a particle along the x-axis as, for example, $\frac{\hbar}{2}$, the particle's spin state collapses into the eigenstate $|\psi_{x+}\rangle$. When we then subsequently measure the particle's spin along the y-axis, the spin state will now collapse into either $|\psi_{y+}\rangle$ or $|\psi_{y-}\rangle$, each with probability $\frac{1}{2}$. Let us say, in our example, that we measure $\frac{-\hbar}{2}$. When we now return to measure the particle's spin along the x-axis again, the probabilities that we will measure $\frac{\hbar}{2}$ or $\frac{-\hbar}{2}$ are each $\frac{1}{2}$ (i.e. they are $|\langle\psi_{x+}|\psi_{y-}\rangle|^2$ and

$| \langle \psi_{x-} \mid \psi_{y-} \rangle |^2$). This implies that our original measurement of the spin along the x-axis is no longer valid, since the spin along the x-axis will now be measured to have either eigenvalue with equal probability.

Applications

Spin has important theoretical implications and practical applications. Well-established *direct* applications of spin include:

- Nuclear magnetic resonance spectroscopy in chemistry;
- Electron spin resonance spectroscopy in chemistry and physics;
- Magnetic resonance imaging (MRI) in medicine, which relies on proton spin density;
- Giant magnetoresistive (GMR) drive head technology in modern hard disks.

Electron spin plays an important role in magnetism, with applications for instance in computer memories. The manipulation of *nuclear spin* by radiofrequency waves (nuclear magnetic resonance) is important in chemical spectroscopy and medical imaging.

Spin-orbit coupling leads to the fine structure of atomic spectra, which is used in atomic clocks and in the modern definition of the second. Precise measurements of the g-factor of the electron have played an important role in the development and verification of quantum electrodynamics. *Photon spin* is associated with the polarization of light.

A possible future direct application of spin is as a binary information carrier in spin transistors. Original concept proposed in 1990 is known as Datta-Das spin transistor. Electronics based on spin transistors is called spintronics, which includes the manipulation of spins in semiconductor devices.

There are many *indirect* applications and manifestations of spin and the associated Pauli exclusion principle, starting with the periodic table of chemistry.

History

Spin was first discovered in the context of the emission spectrum of alkali metals. In 1924 Wolfgang Pauli introduced what he called a "two-valued quantum degree of freedom" associated with the electron in the outermost shell. This allowed him to formulate the Pauli exclusion principle, stating that no two electrons can share the same quantum state at the same time.

The physical interpretation of Pauli's "degree of freedom" was initially unknown. Ralph Kronig, one of Landé's assistants, suggested in early 1925 that it was produced by the self-rotation of the electron. When Pauli heard about the idea, he criticized it severely, noting that the electron's hypothetical surface would have to be moving faster than the speed of light in order for it to rotate quickly enough to produce the necessary angular momentum. This would violate the theory of relativity. Largely due to Pauli's criticism, Kronig decided not to publish his idea.

In the autumn of 1925, the same thought came to two Dutch physicists, George Uhlenbeck and Samuel Goudsmit. Under the advice of Paul Ehrenfest, they published their results. It met a favorable response, especially after Llewellyn Thomas managed to resolve a factor-of-two discrepancy between experimental results and Uhlenbeck and Goudsmit's calculations (and Kronig's unpublished ones). This discrepancy was due to the orientation of the electron's tangent frame, in addition to its position.

Mathematically speaking, a fiber bundle description is needed. The tangent bundle effect is additive and relativistic; that is, it vanishes if c goes to infinity. It is one half of the value obtained without regard for the tangent space orientation, but with opposite sign. Thus the combined effect differs from the latter by a factor two (Thomas precession).

Despite his initial objections, Pauli formalized the theory of spin in 1927, using the modern theory of quantum mechanics discovered by Schrödinger and Heisenberg. He pioneered the use of Pauli matrices as a representation of the spin operators, and introduced a two-component spinor wave-function.

Pauli's theory of spin was non-relativistic. However, in 1928, Paul Dirac published the Dirac equation, which described the relativistic electron. In the Dirac equation, a four-component spinor (known as a "Dirac spinor") was used for the electron wave-function. In 1940, Pauli proved the *spin-statistics theorem*, which states that fermions have half-integer spin and bosons integer spin.

In retrospect, the first direct experimental evidence of the electron spin was the Stern-Gerlach experiment of 1922. However, the correct explanation of this experiment was only given in 1927.

Chapter 12

Spin–Orbit Interaction

The spin-orbit interaction (also called *spin-orbit effect* or *spin-orbit coupling*) is any interaction of a particle's spin with its motion. The first and best known example of this is that spin-orbit interaction causes shifts in an electron's atomic energy levels due to electromagnetic interaction between the electron's spin and the nucleus's magnetic field. This is detectable as a splitting of spectral lines). A similar effect, due to the relationship between angular momentum and the strong nuclear force, occurs for protons and neutrons moving inside the nucleus, leading to a shift in their energy levels in the nucleus shell model. In the field of spintronics, spin-orbit effects for electrons in semiconductors and other materials are explored and put to useful work.

Spin–orbit interaction in atomic energy levels

Using some semiclassical electrodynamics and non-relativistic quantum mechanics, in this section we present a relatively simple and quantitative description of the spin-orbit interaction for an electron bound to an atom, up to first order in perturbation theory. This gives results that agree reasonably well with observations. A more rigorous derivation of the same result would start with the Dirac equation, and achieving a more precise result would involve calculating small corrections from quantum electrodynamics.

Energy of a magnetic moment

The energy of a magnetic moment in a magnetic field is given by:

$$\Delta H = -\boldsymbol{\mu} \cdot \boldsymbol{B},$$

where **μ** is the magnetic moment of the particle and **B** is the magnetic field it experiences.

Magnetic field

We shall deal with the magnetic field first. Although in the rest frame of the nucleus, there is no magnetic field, there *is* one in the rest frame of the electron. Ignoring for now that this frame is not inertial, in SI units we end up with the equation

$$\boldsymbol{B} = -\frac{\boldsymbol{v} \times \boldsymbol{E}}{c^2},$$

where **v** is the velocity of the electron and **E** the electric field it travels through. Now we know that E is radial so we can rewrite $\boldsymbol{E} = \left|\frac{E}{r}\right| \boldsymbol{r}$. Also we know that the momentum of the electron $\boldsymbol{p} = m_e \boldsymbol{v}$. Substituting this in and changing the order of the cross product gives:

$$\boldsymbol{B} = \frac{\boldsymbol{r} \times \boldsymbol{p}}{m_e c^2} \left|\frac{E}{r}\right|.$$

Next, we express the electric field as the gradient of the electric potential $\boldsymbol{E} = -\boldsymbol{\nabla} V$. Here we make the central field approximation, that is, that the electrostatic potential is spherically symmetric, so is only a function of radius. This approximation is exact for hydrogen, and indeed hydrogen-like systems. Now we can say

$$|E| = \frac{\partial V}{\partial r} = \frac{1}{e} \frac{\partial U(r)}{\partial r},$$

where $U = Ve$ is the potential energy of the electron in the central field, and e is the elementary charge. Now we remember from classical mechanics that the angular momentum of a particle $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$. Putting it all together we get

$$\boldsymbol{B} = \frac{1}{m_e e c^2} \frac{1}{r} \frac{\partial U(r)}{\partial r} \boldsymbol{L}.$$

It is important to note at this point that B is a positive number multiplied by L, meaning that the magnetic field is parallel to the orbital angular momentum of the particle.

Magnetic Moment of the Electron

The magnetic moment of the electron is

$$\boldsymbol{\mu} = -g_s \mu_B \boldsymbol{S}/\hbar,$$

where $\boldsymbol{S}$ is the spin angular momentum vector, μ_B is the Bohr magneton and $g_s \approx 2$ is the electron spin g-factor. Here, $\boldsymbol{\mu}$ is a negative constant multiplied by the spin, so the magnetic moment is antiparallel to the spin angular momentum.

The spin-orbit potential consists of two parts. The Larmor part is connected to interaction of the magnetic moment of electron with magnetic field of nucleus in the co-moving frame of electron. The second contribution is related to Thomas precession.

Larmor interaction energy

The Larmor interaction energy is

$$\Delta H_L = -\boldsymbol{\mu} \cdot \boldsymbol{B}.$$

Substituting in this equation expressions for the magnetic moment and the magnetic field, one gets

$$\Delta H_L = \frac{2\mu_B}{\hbar m_e e c^2} \frac{1}{r} \frac{\partial U(r)}{\partial r} \boldsymbol{L} \cdot \boldsymbol{S}.$$

Now, we have to take into account Thomas precession correction for the electron's curved trajectory.

Thomas interaction energy

In 1926 Llewellyn Thomas relativistically recomputed the doublet separation in the fine structure of the atom.. Thomas precession rate, $\boldsymbol{\Omega}_T$, is related to the angular frequency of the orbital motion, $\boldsymbol{\omega}$, of a spinning particle as follows

$$\boldsymbol{\Omega}_T = \boldsymbol{\omega}(1 - \gamma),$$

where γ is the Lorentz factor of moving particle. The Hamiltonian producing the spin precession $\boldsymbol{\Omega}_T$ is given by

$$\Delta H_T = \boldsymbol{\Omega}_T \cdot \boldsymbol{S}.$$

To the first order in $(v / c)^2$, we obtain

$$\Delta H_T = -\frac{\mu_B}{\hbar m_e e c^2} \frac{1}{r} \frac{\partial U(r)}{\partial r} \boldsymbol{L} \cdot \boldsymbol{S}.$$

Total interaction energy

The total spin-orbit potential in an external electrostatic potential takes the form

$$\Delta H \equiv \Delta H_L + \Delta H_T = \frac{\mu_B}{\hbar m_e e c^2} \frac{1}{r} \frac{\partial U(r)}{\partial r} (\boldsymbol{L} \cdot \boldsymbol{S}).$$

The net effect of Thomas precession is the reduction of the Larmor interaction energy by factor 1/2 which came to be known as the *Thomas half*.

Evaluating the energy shift

Thanks to all the above approximations, we can now evaluate the detailed energy shift in this model. In particular, we wish to find a basis that diagonalizes both H_0 (the non-perturbed Hamiltonian) and ΔH. To find out what basis this is, we first define the total angular momentum operator

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}.$$

Taking the dot product of this with itself, we get

$$J^2 = L^2 + S^2 + 2\boldsymbol{L}\cdot\boldsymbol{S}$$

(since **L** and **S** commute), and therefore

$$\boldsymbol{L}\cdot\boldsymbol{S} = \frac{1}{2}(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2)$$

It can be shown that the five operators H_0, J^2, L^2, S^2, and J_z all commute with each other and with ΔH. Therefore, the basis we were looking for is the simultaneous eigenbasis of these five operators (i.e., the basis where all five are diagonal). Elements of this basis have the five quantum numbers: n (the "principal quantum number") j (the "total angular momentum quantum number"), l (the "orbital angular momentum quantum number"), s (the "spin quantum number"), and j_z (the "z-component of total angular momentum").

To evaluate the energies, we note that

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{2}{a^3 n^3 l(l+1)(2l+1)}$$

for hydrogenic wavefunctions (here $a = \hbar / Z\alpha m_e c$ is the Bohr radius divided by the nuclear charge Z); and

$$\langle \boldsymbol{L}\cdot\boldsymbol{S}\rangle = \frac{1}{2}(\langle \boldsymbol{J}^2\rangle - \langle \boldsymbol{L}^2\rangle - \langle \boldsymbol{S}^2\rangle)$$
$$= \frac{\hbar^2}{2}(j(j+1) - l(l+1) - s(s+1))$$

Final Energy Shift

We can now say

$$\Delta E = \frac{\beta}{2}(j(j+1) - l(l+1) - s(s+1))$$

where

$$\beta = \frac{-\mu_B}{m_e e c^2} \left\langle \frac{1}{r} \frac{\partial U(r)}{\partial r} \right\rangle$$

For hydrogen, we can write the explicit result

$$\beta(n,l) = \frac{\mu_0}{4\pi} g_s \mu_B^2 \frac{1}{n^3 a_0^3 l(l+1/2)(l+1)}$$

For any hydrogen-like atom with Z protons

$$\beta(n,l) = Z^4 \frac{\mu_0}{4\pi} g_s \mu_B^2 \frac{1}{n^3 a_0^3 l(l+1/2)(l+1)}$$

Chapter 13

Spin Polarization, Spin Pumping & Spin Transistor

Spin polarization

Spin polarization is the degree to which the spin, i.e. the intrinsic angular momentum of elementary particles, is aligned with a given direction. This property may pertain to the spin, hence to the magnetic moment, of conduction electrons in ferromagnetic metals, such as iron, giving rise to spin polarized currents. It may refer to (static) spin waves, preferential correlation of spin orientation with ordered lattices (semiconductors or insulators).

It may also pertain to beams of particles, produced for particular aims, such as polarized neutron scattering or muon spin spectroscopy. Spin polarization of electrons or of nuclei, often called simply magnetization, is also produced by the application of a magnetic field, thanks to the Curie law and it is used to produce an induction signal in Electron spin resonance (ESR or EPR) and in Nuclear magnetic resonance (NMR).

Spin polarization is also important for spintronics, a branch of electronics. Magnetic semiconductors are being researched as possible spintronics materials.

The spin of free electrons is measured either by a LEED image from a clean wolfram-crystal (SPLEED) or by an electron microscope composed purely of electrostatic lenses and a gold foil as a sample. Back scattered electrons are decelerated by annular optics and focused onto a ring shaped electron mulitplier at about 15°. The position on the ring is recorded. This whole device is called a Mott-detector. Depending on their spin the electrons have the chance to hit the ring at different positions. 1% of the electrons are scattered in the foil. Of these 1% are collected by the detector and then about 30% of the electrons hit the detector at the wrong position. Both devices work due to spin orbit coupling.

The circular polarization of electromagnetic fields is due to spin polarization of their constituent photons.

In the most generic context, spin polarization is any alignment of the components of a non-scalar (vectorial, tensorial, spinor) field with its arguments, i.e., with the nonrelativistic three spatial or relativistic four spatiotemporal regions over which it is defined. In this sense, it also includes gravitational waves and any field theory that couples its constituents with the differential operators of vector analysis.

Spin pumping

Spin pumping is a method of generating a spin current, the spintronic analog of a battery in conventional electronics.

In order to make a spintronic device, the primary requirement is to have a system that can generate a current of spin-polarized electrons, as well as a system that is sensitive to the spin polarization. Most spintronic devices also have a unit in between these two that changes the current of electrons depending on the spin states. Candidates for such devices include injection schemes based on magnetic semiconductors and ferromagnetic metals, ferromagnetic resonance devices, and a variety of spin-dependent pumps. Optical, microwave and electrical methods are also being explored.

Spin transistor

The magnetically-sensitive transistor (also known as the **spin transistor** or spintronic transistor—named for spintronics, the technology which this development spawned), originally proposed in 1990 and currently still being developed, is an improved design on the common transistor invented in the 1940s. The spin transistor comes about as a result of research on the ability of electrons (and other fermions) to naturally exhibit one of two (and only two) states of spin: known as "spin up" and "spin down". Unlike its namesake predecessor, which operates on an electric current, spin transistors operate on electrons on a more fundamental level; it is essentially the application of electrons set in particular states of spin to store information.

One advantage over regular transistors is that these spin states can be detected and altered without necessarily requiring the application of an electric current. This allows for detection hardware (such as hard drive heads) that are much smaller but even more sensitive than today's devices, which rely on noisy amplifiers to detect the minute charges used on today's data storage devices. The potential end result is devices that can store more data in less space and consume less power, using less costly materials. The increased sensitivity of spin transistors is also being researched in creating more sensitive automotive sensors, a move being encouraged by a push for more environmentally-friendly vehicles.

A second advantage of a spin transistor is that the spin of an electron is semi-permanent and can be used as means of creating cost-effective non-volatile solid state storage that does not require the constant application of current to sustain. It is one of the technologies being explored for Magnetic Random Access Memory (MRAM).

Because of its high potential for practical use in the computer world, spin transistors are currently being researched in various firms throughout the world, such as in England and in Sweden. Recent breakthroughs have allowed the production of spin transistors, using readily-available substances, that can operate at room temperature: a precursor to commercial viability.

Chapter 14

Spin Magnetic Moment & Spin Valve

Spin Magnetic Moment

Basis for spin magnetic moments

A **spin magnetic moment** is induced by all charged particles, but the particle most usable by modern technology is a certain lepton, the electron. Specifically, a spin magnetic moment is created because a particle has physical properties known as spin and electric charge. The spin within classical physics would be an object that rotates axially around its center of mass. In quantum mechanics, elementary particles are points, which have no axis to revolve around. This means these particles do not have spin in a classical sense, as angular momentum is defined by $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, but have the physical property of angular momentum. Maxwell's theory of magnetic fields dictates that any moving charged particle creates a magnetic moment, and by definition, angular momentum designates movement. This is where the magnetic moment emerges in classical electromagnetism.

Calculation

We can calculate the observable spin magnetic moment (a vector), $\vec{\mu}_S$, for a sub-atomic particle with charge q, mass m, and spin angular momentum (also a vector), $\vec{s}$, via:

$$\vec{\mu}_S = g \frac{q}{2m} \vec{s} \qquad (1)$$

where g is a dimensionless number, called the g-factor. This number depends on the particle: it is $g = 2.0023$ for the electron, $g = 5.586$ for the proton, and $g = -3.826$ for the neutron. The proton and neutron are composed of quarks, which have a non-zero charge and a spin of $\hbar/2$, and this must be taken into account when calculating their g-factors. Even though the neutron has a charge $q = 0$, its quarks give it a magnetic moment. The proton and electron's spin magnetic moments can be calculated by simply setting $q = e$.

The intrinsic electron magnetic dipole moment is approximately equal to the Bohr magneton because $g \approx 2$ and the electron's spin is also $\hbar/2$:

$$\mu_S \approx 2\frac{e}{2m_e}\frac{\hbar}{2} = \mu_B \quad \text{...} \quad \textbf{(2)}$$

Eq. (2) is normally written as

$$\vec{\mu}_S = g\mu_B\vec{s} \quad \text{...} \quad \textbf{(3)}$$

where μ_B is the Bohr magneton.

Just like the *total spin angular momentum* cannot be measured, neither can the *total spin magnetic moment* be measured. Equations (1) - (3) give the physical observable, that component of the magnetic moment measured along an axis, relative to or along the applied field direction. Conventionally, the *z*-axis is chosen but the observable values of the component of spin angular momentum along all three axes (assuming a Cartesian coordinate system) are each $\pm\frac{\hbar}{2}$. However, in order to obtain the magnitude of the total spin angular momentum, wave mechanical corrections dictate that $\vec{s}$ be replaced by its eigenvalue, $\sqrt{S(S+1)}$. In turn, calculation of the magnitude of the total spin magnetic moment requires that Eq. (3) be replaced by:

$$|\vec{\mu}_S| = g\mu_B\sqrt{S(S+1)} \quad \text{...} \quad \textbf{(4)}$$

Thus, for a single electron, with spin quantum number $s = \frac{1}{2}$, the component of the magnetic moment along the field direction is, from Eq. (3), $\vec{\mu}_{S,z} = \mu_B$, or one BM, while the [magnitude of the] total spin magnetic moment is, from Eq. (4), $|\vec{\mu}_S| = \sqrt{3}\mu_B$, or 1.73 BM.

The analysis is readily extended to the spin-only magnetic moment of an atom. For example, the total spin magnetic moment (sometimes referred to as the *effective magnetic moment* when the orbital moment contribution to the total magnetic moment is neglected) of a transition metal ion with a single *d* electron outside of closed shells (e.g. Ti^{3+}) is 1.73 μ_B since $S = 1/2$, while an atom with two unpaired electrons (e.g. V^{3+}) with $S = 1$ would have an effective magnetic moment of 2.83 μ_B.

Spin in chemistry

Spin magnetic moments create a basis for one of the most important principles in chemistry, the Pauli exclusion principle. This principle, first suggested by Wolfgang Pauli, governs most of modern-day chemistry. The theory plays further roles than just the explanations of doublets within electromagnetic spectrum. This additional quantum number, spin, became the basis for the modern standard model used today, which includes the use of Hund's rules, and an explanation of beta decay.

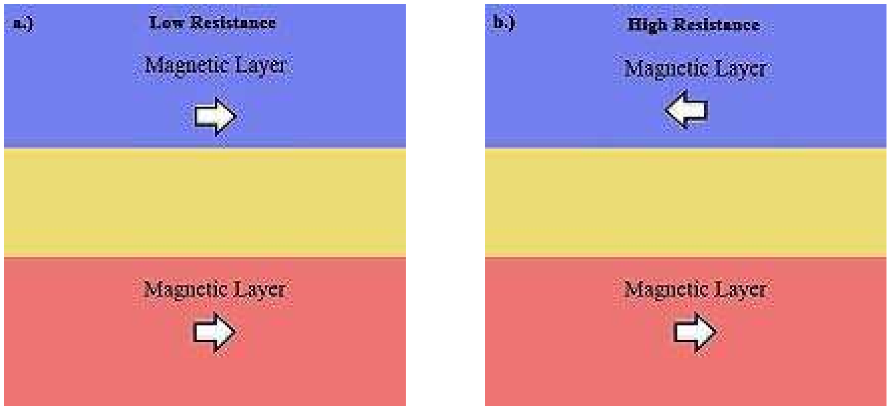
History of spin magnetic moments

The idea of a spin angular momentum was first proposed in a 1925 publication by George Uhlenbeck and Samuel Goudsmit to explain hyperfine splitting in atomic spectra. In 1928, Paul Dirac provided a rigorous theoretical foundation for the concept with his relativistic equation of motion for the wavefunction of the electron.

Spin Valve

A **spin valve** is a device consisting of two or more conducting magnetic materials, that alternates its electrical resistance (from low to high or high to low) depending on the alignment of the magnetic layers, in order to exploit the Giant Magnetoresistive effect. The magnetic layers of the device align "up" or "down" depending on an external magnetic field. Layers are made of two materials with different magnetic coercivity, which can be seen in the layers' hysteresis curves. Due to the different coercivities one layer ("soft" layer) changes polarity at small magnetic fields while the other ("hard" layer) changes polarity at a higher magnetic field. As the magnetic field across the sample is swept two distinct states can exist, one with the magnetisations of the layers parallel, and one with the magnetisations of the layers antiparallel. In the figures below, the top layer is soft and the bottom layer is hard.

How it works



Spin valves work because of a quantum property of electrons (and other particles) called spin. When a magnetic layer is polarized, the unpaired carrier electrons align their spins to the external magnetic field. When a potential exists across a spin valve, the spin-polarized electrons keep their spin alignment as they move through the device. If these electrons encounter a material with a magnetic field pointing in the opposite direction, they have to flip spins to find an empty energy state in the new material. This flip requires extra energy which causes the device to have a higher resistance than when the magnetic materials are polarized in the same direction.

Applications

Spin valves are used in magnetic sensors and hard disk read heads. They are also used in magnetic random access memories (MRAM).