Semiconductor
Engineering and Technologies

Jacqueline Roldan

Ellis Beyer

First Edition, 2012

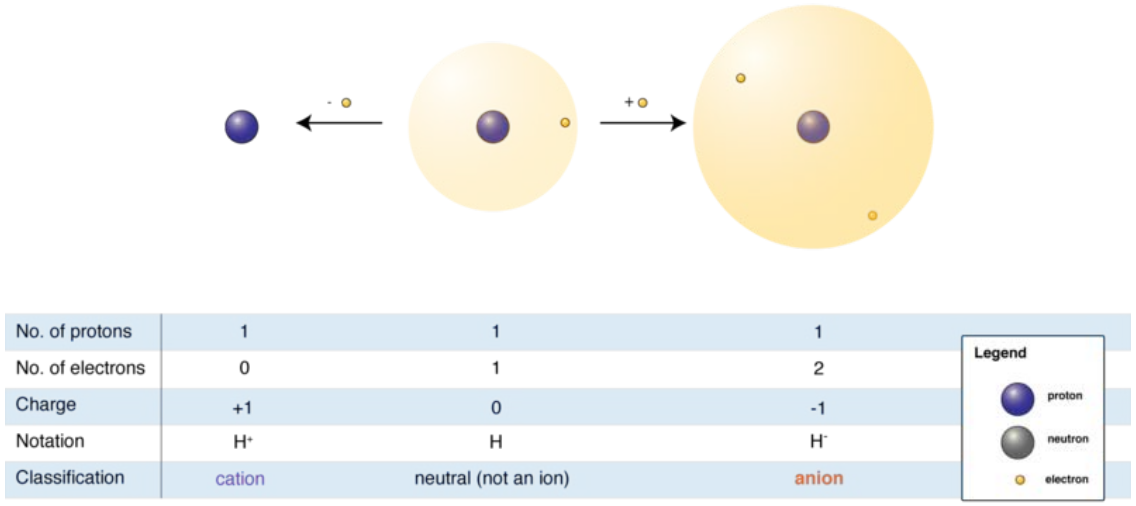
ISBN 978-81-323-0982-6

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

Table of Contents

Chapter-1

Ion



No. of protons	1	1	1
No. of electrons	0	1	2
Charge	+1	0	-1
Notation	H^+	H	H^-
Classification	cation	neutral (not an ion)	anion

Hydrogen atom (center) contains a single proton and a single electron. Removal of the electron gives a cation (left), whereas addition of an electron gives an anion (right). The hydrogen anion, with its loosely held two-electron cloud, has a larger radius than the neutral atom, which in turn is much larger than the bare proton of the cation. Hydrogen forms the only cation that has no electrons, but even cations that (unlike hydrogen) still retain one or more electrons, are still smaller than the neutral atoms or molecules from which they are derived.

An **ion** is an atom or molecule in which the total number of electrons is not equal to the total number of protons, giving it a net positive or negative electrical charge. The name was given by physicist Michael Faraday for the substances that allow a current to pass ("go") between electrodes in a solution, when an electric field is applied. It is from Greek *ιον*, meaning "going."

An **anion** from the Greek word ἄνω (*ánō*), meaning "up", is an ion with more electrons than protons, giving it a net negative charge (since electrons are negatively charged and protons are positively charged). Conversely, a **cation** from the Greek word κατά (*katá*), meaning "down", is an ion with fewer electrons than protons, giving it a positive charge. Since the charge on a proton is equal in magnitude to the charge on an electron, the net charge on an ion is equal to the number of protons in the ion minus the number of electrons.

An ion consisting of a single atom is an **atomic** or **monatomic ion**; if it consists of two or more atoms, it is a **molecular** or **polyatomic ion**.

General

History and Discovery

Etymologically the word *ion* is the Greek *ιον* (going), the present participle of *ιεναι*, *ienai*, "to go." This term was introduced by English physicist and chemist Michael Faraday in 1834 for the (then unknown) species that *goes* from one electrode to the other through an aqueous medium. Faraday did not know the nature of these species, but he knew that since metals dissolved into and entered solution at one electrode, and new metal came forth from solution at the other electrode, that some kind of substance moved through the solution in a current, conveying matter from one place to the other.

Faraday also introduced the words **anion** and **cation.** In Faraday's nomenclature, cations were named because they were attracted to the anode in a galvanic device and anions were named due to their attraction to the cathode.

Characteristics

Ions in their gas-like state are highly reactive, and do not occur in large amounts on Earth, except in flames, lightning, electrical sparks, and other plasmas. These gas-like ions rapidly interact with ions of opposite charge to give neutral molecules or ionic salts. Ions are also produced in the liquid or solid state when salts interact with solvents (for example, water) to produce "solvated ions," which are more stable, for reasons involving a combination of energy and entropy changes as the ions move away from each other to interact with the liquid. These stabilized species are more commonly found in the environment at low temperatures. A common example is the ions present in seawater, which are derived from the dissolved salts there.

All ions are charged, which means that like all charged objects they are:

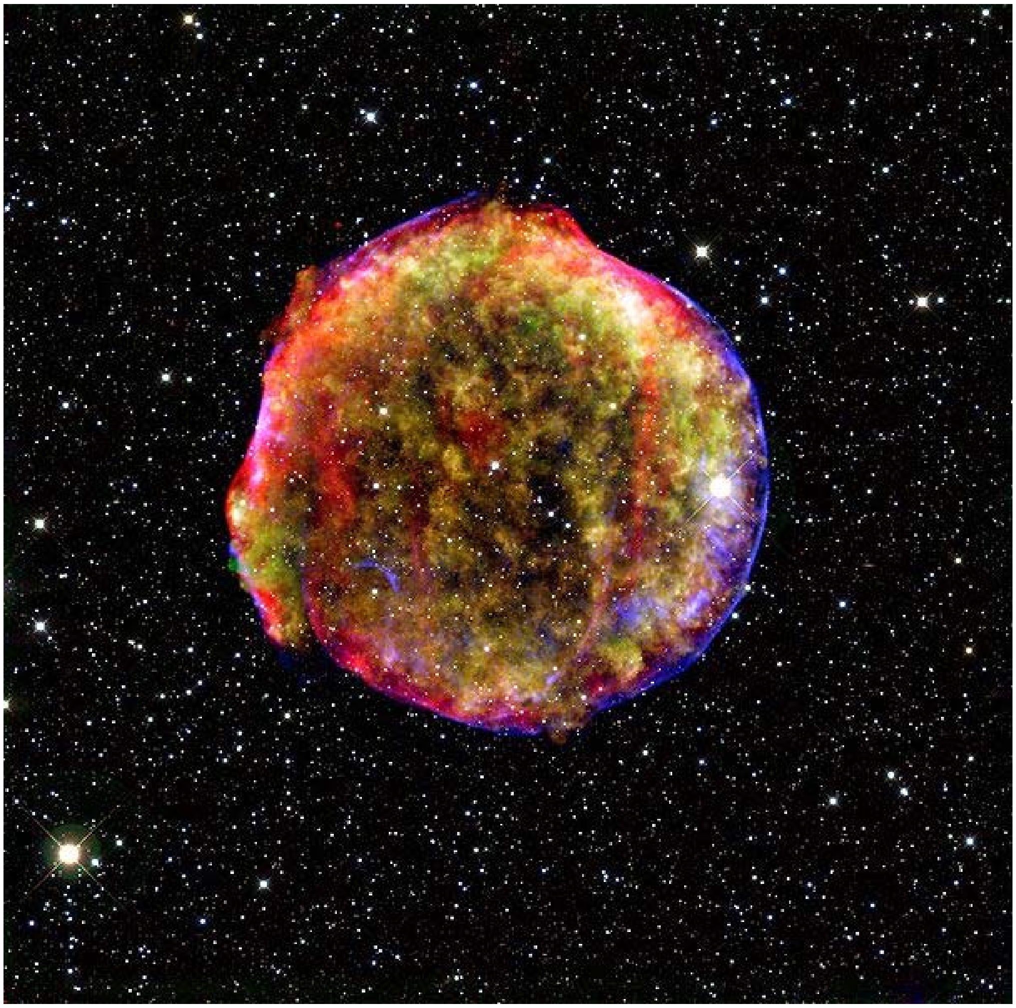
- attracted to opposite electric charges (positive to negative, and vice versa),
- repelled by like charges, and
- when moving, travel in trajectories that are deflected by a magnetic field.

Electrons, due to their smaller mass and thus larger space-filling properties as matter waves, determine the size of atoms and molecules that possess any electrons at all. Thus, anions (negatively charged ions) are larger than the parent molecule or atom, as the excess electron(s) repel each other, and add to the physical size of the ion, because its size is determined by its electron cloud. Conversely, cations are generally smaller than the corresponding parent atom or molecule, for the same reason. One particular cation (that of hydrogen) contains no electrons, and thus is *very much smaller* than the parent hydrogen atom.

Natural Occurrences

Ions are ubiquitous in nature and are responsible for diverse phenomena from the luminescence of the Sun, and the existence of ionosphere on Earth. Atoms in their ionic state may have a different color from neutral atoms, and thus light absorption by metal ions gives the color of gemstones. In both inorganic and organic chemistry (including biochemistry), the interaction of water and ions is extremely important (an example is the energy that drives breakdown of ATP. The following sections describe contexts in which ions feature prominently and are arranged in decreasing physical length-scale, from the astronomical to the microscopic.

Astronomical



The remnant of "Tycho's Supernova", a huge ball of expanding plasma. The outer shell shown in blue is X-ray emission by high-speed electrons.

A collection of non-aqueous gas-like ions, or even a gas containing a proportion of charged particles, is called a **plasma**. >99.9% of visible matter in the Universe may be in the form of plasmas. These include our Sun and other stars, the space between planets, as well as the space in between stars. Plasmas are often called the *fourth state of matter* because its properties are substantially different from solids, liquids, and gases. Astrophysical plasmas predominantly contain a mixture of electrons and protons (ionized hydrogen).

Related Technology

Ions can be non-chemically prepared using various ion sources, usually involving high voltage or temperature. These are used in a multitude of devices such as mass spectrometers, optical emission spectrometers, particle accelerators, ion implanters and ion engines.

As reactive charged particles, they are also used in air purification by disrupting microbes, and in household items such as smoke detectors.

As signaling and metabolism in organisms are controlled by a precise ionic gradient across membranes, the disruption of this gradient contributes to cell death. This is a common mechanism exploited by natural and artificial biocides, including the ion channels gramicidin and amphotericin (a fungicide).

Inorganic dissolved ions are a component of total dissolved solids, an indicator of water quality in the world.

Chemistry

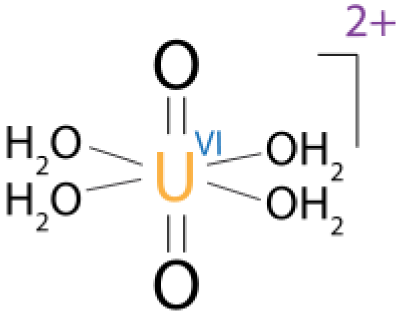
Notation

Denoting the charged state

Equivalent notations for an iron atom (Fe) that lost two electrons.

When writing the chemical formula for an ion, its net charge is written in superscript immediately after the chemical structure for the molecule/atom. The net charge is written with the magnitude *before* the sign; that is, a doubly charged cation is indicated as **2+** instead of **+2**. Conventionally the magnitude of the charge is omitted for singly charged molecules/atoms; for example, the sodium cation is indicated as Na^+ and *not* Na^{1+}.

An alternative (and acceptable) way of showing a molecule/atom with multiple charges is by drawing out the signs multiple times; this is often seen with transition metals. Chemists sometimes circle the sign; this is merely ornamental and does not alter the chemical meaning. All three representations of Fe^{2+} shown in the figure are thus equivalent.



Mixed Roman numerals and charge notations for the uranyl ion. The oxidation state of the metal is shown as superscripted Roman numerals, whereas the charge of the entire complex is shown by the angle symbol together with the magnitude and sign of the net charge.

Monatomic ions are sometimes also denoted with Roman numerals; for example, the Fe^{2+} example seen above is occasionally referred to as Fe(II) or Fe^{II}. The Roman numeral designates the *formal oxidation state* of an element, whereas the superscripted numerals denotes the net charge. The two notations are therefore exchangeable for monatomic ions, but the Roman numerals *cannot* be applied to polyatomic ions. It is however possible to mix the notations for the individual metal center with a polyatomic complex, as shown by the uranyl ion example.

Sub-classes

If an ion contains unpaired electrons, it is called a **radical** ion. Just like uncharged radicals, radical ions are very reactive. Polyatomic ions containing oxygen, such as carbonate and sulfate, are called **oxyanions**. Molecular ions that contain at least one carbon to hydrogen bond are called **organic ions**. If the charge in an organic ion is formally centered on a carbon, it is termed a **carbocation** (if positively charged) or **carboanion** (if negatively charged).

Formation

Formation of monatomic ions

Monatomic ions are formed by the addition of electrons to the valence shell of the atom, which is the outer-most electron shell in an atom, or the losing of electrons from this shell. The inner shells of an atom are filled with electrons that are tightly bound to the positively charged atomic nucleus, and so do not participate in this kind of chemical interaction. The process of gaining or losing electrons from a neutral atom or molecule is called **ionization**.

Atoms can be ionized by bombardment with radiation, but the more usual process of ionization encountered in chemistry is the transfer of electrons between atoms or molecules. This transfer is usually driven by the attaining of stable ("closed shell") electronic configurations. Atoms will gain or lose electrons depending on which action takes the least energy.

For example, a sodium atom, Na, has a single electron in its valence shell, surrounding 2 stable, filled inner shells of 2 and 8 electrons. Since these filled shells are very stable, a sodium atom tends to lose its extra electron and attain this stable configuration, becoming a sodium **cation** in the process

$$Na \rightarrow Na^+ + e-$$

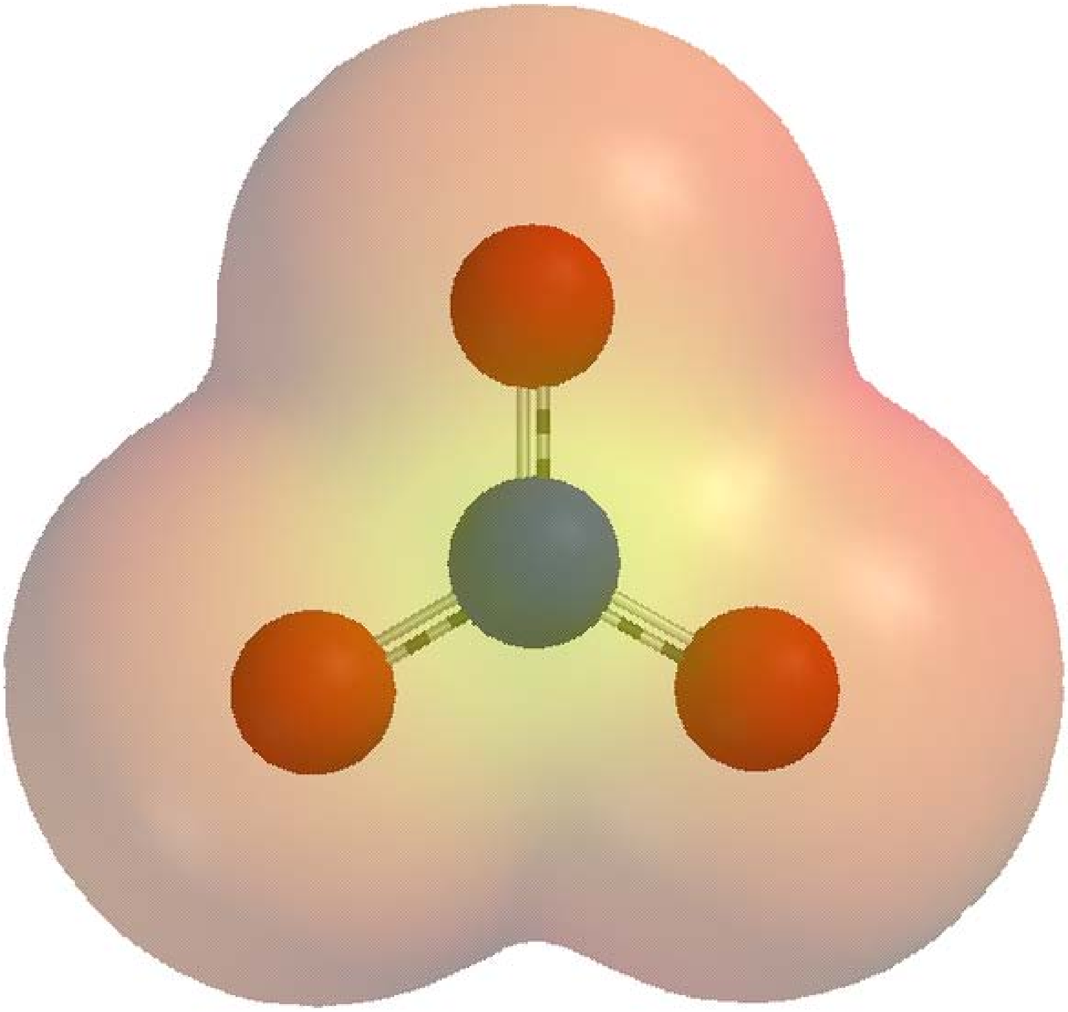
On the other hand, a chlorine atom, Cl, has 7 electrons in its valence shell, which is one short of the stable, filled shell with 8 electrons. Thus, a chlorine atom tends to *gain* an extra electron and attain a stable 8-electron configuration, becoming a chloride **anion** in the process:

$$Cl + e- \rightarrow Cl^-$$

This driving force is what causes sodium and chlorine to undergo a chemical reaction, where the "extra" electron is transferred from sodium to chlorine, forming sodium cations and chloride anions. Being oppositely charged, these cations and anions form ionic bonds and combine together to form sodium chloride, NaCl, more commonly known as rock salt.

$$Na^+ + Cl^- \rightarrow NaCl$$

Formation of polyatomic and molecular ions



An electrostatic potential map of the nitrate ion (NO_3^-). The 3-dimensional shell represents a single arbitrary isopotential.

Polyatomic and molecular ions are often formed by the gaining or losing of elemental ions such as H^+ in neutral molecules. For example, when ammonia, NH_3, accepts a proton, H^+, it forms the ammonium ion, NH_4^+. Ammonia and ammonium have the same number of electrons in essentially the same electronic configuration, but ammonium has an extra proton that gives it a net positive charge.

Ammonia can also lose an electron to gain a positive charge, forming the ion $\cdot NH_3^+$. However, this ion is unstable, because it has an incomplete valence shell around the nitrogen atom, making it a very reactive radical ion.

Due to the instability of radical ions, polyatomic and molecular ions are usually formed by gaining or losing elemental ions such as H^+, rather than gaining or losing electrons. This allows the molecule to preserve its stable electronic configuration while acquiring an electrical charge.

Ionization potential

The energy required to detach an electron in its lowest energy state from an atom or molecule of a gas with less net electric charge is called the *ionization potential*, or *ionization energy*. The *n*th ionization energy of an atom is the energy required to detach its *n*th electron after the first $n - 1$ electrons have already been detached.

Each successive ionization energy is markedly greater than the last. Particularly great increases occur after any given block of atomic orbitals is exhausted of electrons. For this reason, ions tend to form in ways that leave them with full orbital blocks. For example, sodium has one *valence electron* in its outermost shell, so in ionized form it is commonly found with one lost electron, as Na^+. On the other side of the periodic table, chlorine has seven valence electrons, so in ionized form it is commonly found with one gained electron, as Cl^-. Caesium has the lowest measured ionization energy of all the elements and helium has the greatest. The ionization energy of metals is generally much lower than the ionization energy of nonmetals, which is why metals will generally lose electrons to form positively charged ions while nonmetals will generally gain electrons to form negatively charged ions.

Ionic bonding

Ionic bonding is a kind of chemical bonding that arises from the mutual attraction of oppositely charged ions. Since ions of like charge repel each other, they do not usually exist on their own. Instead, many of them may form a crystal lattice, in which ions of opposite charge are bound to each other. The resulting compound is called an *ionic compound*, and is said to be held together by *ionic bonding*. In ionic compounds there arise characteristic distances between ion neighbors from which the spatial extension and the ionic radius of individual ions may be derived.

The most common type of ionic bonding is seen in compounds of metals and nonmetals (except noble gases, which rarely form chemical compounds). Metals are characterized by having a small number of electrons in excess of a stable, closed-shell electronic configuration. As such, they have the tendency to lose these extra electrons in order to attain a stable configuration. This property is known as *electropositivity*. Non-metals, on the other hand, are characterized by having an electron configuration just a few electrons short of a stable configuration. As such, they have the tendency to gain more electrons in order to achieve a stable configuration. This tendency is known as *electronegativity*. When a highly electropositive metal is combined with a highly electronegative nonmetal, the extra electrons from the metal atoms are transferred to the electron-deficient nonmetal atoms. This reaction produces metal cations and nonmetal anions, which are attracted to each other to form a *salt*.

Common ions

Common **Cations**

Common Name	**Formula**	**Historic Name**
Simple Cations		
Aluminium	Al^{3+}	
Calcium	Ca^{2+}	
Copper(II)	Cu^{2+}	cupric
Hydrogen	H^+	
Iron(II)	Fe^{2+}	ferrous
Iron(III)	Fe^{3+}	ferric
Magnesium	Mg^{2+}	
Mercury(II)	Hg^{2+}	mercuric
Potassium	K^+	kalic
Silver	Ag^+	
Sodium	Na^+	natric
Polyatomic Cations		
Ammonium	NH_4^+	
Oxonium	H_3O^+	hydronium
Mercury(I)	Hg_2^{2+}	mercurous

Common **Anions**

Formal Name	**Formula**	**Alt. Name**
Simple Anions		
Chloride	Cl^-	
Fluoride	F^-	
Oxide	O^{2-}	
Oxoanions		
Carbonate	CO_3^{2-}	
Hydrogen carbonate	HCO_3^-	bicarbonate
Hydroxide	OH^-	
Nitrate	NO_3^-	
Phosphate	PO_4^{3-}	
Sulfate	SO_4^{2-}	
Anions from Organic Acids		
Acetate	CH_3COO^-	ethanoate
Formate	$HCOO^-$	methanoate
Oxalate	$C_2O_4^{2-}$	ethandioate
Cyanide	CN^-	

Chapter-2

Charge Carriers

In physics, a **charge carrier** denotes a free (mobile, unbound) particle carrying an electric charge, especially the particles that carry electric currents in electrical conductors. Examples are electrons and ions.

In metals, the charge carriers are electrons. One or two of the outer valence electrons from each atom is able to move about freely within the crystal lattice of the metal. This cloud of free electrons is referred to as a Fermi gas.

In ionic solutions, such as salt water, the charge carriers are the dissolved cations and anions. Similarly, cations and anions of the dissociated liquid serve as charge carriers in liquids and melted ionic solids.

In plasma, such as an electric arc, the electrons and cations of ionized gas and vaporized material of electrodes act as charge carriers. (The electrode vaporization occurs in vacuum too, but then the arc is not technically occurring in vacuum, but in low-pressure electrode vapors.)

In vacuum, in an electric arc or in vacuum tubes free electrons act as charge carriers.

In semiconductor physics, two types of charge carriers are recognized. One of them is electrons. In addition, it is convenient to treat the traveling vacancies in the valence-band electron population (holes) as the second type of charge carriers which carry the positive charge.

Free Carrier Concentration

Free carrier concentration is the concentration of free carriers in a doped semiconductior. It is similar to the carrier concentration in a metal and for the purposes of calculating currents or drift velocities can be used in the same way. Free carriers are electrons (or holes) which have been introduced directly into the conduction (valence) band by doping and are not promoted thermally. For this reason electrons (holes) will not act as double carriers by leaving behind holes (electrons) in the other band.

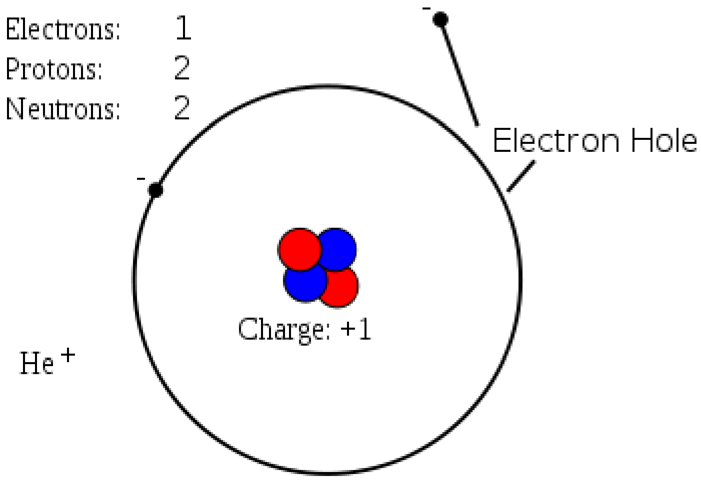
Electron hole

An **electron hole** is the conceptual and mathematical opposite of an electron, useful in the study of physics, chemistry, and electrical engineering. The concept describes the lack of an electron at a position where one could exist in an atom or atomic lattice. It is different from the positron, which is the antimatter analogue of the electron.

The electron hole was introduced into calculations for the following two situations:

- If an electron is excited into a higher state it leaves a hole in its old state. This meaning is used in Auger electron spectroscopy (and other x-ray techniques), in computational chemistry, and to explain the low electron-electron scattering-rate in crystals (metals, semiconductors).
- In crystals, band structure calculations lead to an effective mass for the charge carriers, which can be negative. Inspired by the Hall effect, Newton's law is used to attach the negative sign onto the charge.

Solid state physics



When an electron leaves a helium atom, it leaves an electron hole in its place. This causes the helium atom to become positively charged.

In solid state physics, an **electron hole** (usually referred to simply as a **hole**) is the absence of an electron from an otherwise full electron shell. A hole is essentially a way to conceptualise the interactions of the electrons within a nearly *full* system, which is *missing* just a few electrons. In some ways, the behaviour of a hole within a semiconductor crystal lattice is comparable to that of the bubble in an otherwise full bottle of water.

Hole conduction in a valence band can be explained by the following analogy. Imagine a row of people seated in an auditorium, where there are no spare chairs. Someone in the middle of the row wants to leave, so he jumps over the back of the seat into an empty row, and walks out. The empty row is analogous to the conduction band, and the person walking out is analogous to a free electron.

Now imagine someone else comes along and wants to sit down. The empty row has a poor view; so he does not want to sit there. Instead, a person in the crowded row moves into the empty seat the first person left behind. The empty seat moves one spot closer to the edge and the person waiting to sit down. The next person follows, and the next, etcetera. One could say that the empty seat moves towards the edge of the row. Once the empty seat reaches the edge, the new person can sit down.

In the process everyone in the row has moved along. If those people were negatively charged (like electrons), this movement would constitute conduction. If the seats themselves were positively charged, then only the vacant seat would be positive. This is a very simple model of how hole conduction works.

In reality, due to the crystal structure properties, the hole is not localized to a single position as described in the previous example. Rather, the hole spans an area in the crystal lattice covering many hundreds of unit cells. This is equivalent to being unable to tell which broken bond corresponds to the "missing" electron.

Instead of analyzing the movement of an empty state in the valence band as the movement of billions of separate electrons, a single equivalent imaginary particle called a "hole" is considered. In an applied electric field, the electrons move in one direction, corresponding to the hole moving in the other. If a hole associates itself with a neutral atom, that atom loses an electron and becomes positive. Therefore the hole is taken to have positive charge of +e, precisely the opposite of the electron charge.

Coulomb's law allows the force on the "hole" due to an electric field to be calculated. An effective mass can then be derived which relates the (imaginary) force on the (imaginary) hole to the acceleration of that hole. In some semiconductors, such as silicon, the hole's effective mass is dependent on direction (anisotropic), however a value averaged over all directions can be used for some macroscopic calculations.

In most semiconductors, the effective mass of a hole is much larger than that of an electron. This results in lower mobility for holes under the influence of an electric field and this may slow down the speed of the electronic device made of that semiconductor. This is one major reason for adopting electrons as the primary charge carriers, whenever possible in semiconductor devices instead of holes.

Holes in quantum chemistry

An alternate meaning for the term **electron hole** is used in computational chemistry. In coupled cluster methods, the ground (or lowest energy) state of a molecule is interpreted as the "vacuum state"—conceptually, in this state there are no electrons. In this scheme, the absence of an electron from a normally-filled state is called a "hole" and is treated as a particle, and the presence of an electron in a normally-empty state is simply called an "electron". This terminology is almost identical to that used in solid-state physics.

Non-radiative recombination

Non-radiative recombination is a process in phosphors and semiconductors, whereby charge carriers recombine without releasing photons. A phonon is released instead.

Non-radiative recombination in optoelectronics and phosphors is an unwanted process, lowering the light generation efficiency and increasing heat losses.

Shockley–Read–Hall (SRH) process

The electron in transition between bands passes through a new energy state created within the band gap by an impurity in the crystal lattice. The impurity state can absorb differences in momentum between the carriers, and so this process is the dominant generation and recombination process in silicon and other indirect bandgap materials. It can also dominate in direct bandgap materials under conditions of very low carrier densities (very low level injection). The energy is exchanged in the form of lattice vibration, or a phonon exchanging thermal energy with the material.

Various impurities and dislocations create energy levels within the band gap corresponding to neither donor nor acceptor levels, forming deep-level traps. Non-radiative recombination occurs primarily at such sites.

Charge carrier density

The **charge carrier density** denotes the number of charge carriers per volume. It is measured in m^{-3}. As any density it can depend on position. It should not be confused with the charge density which is the number of charges per volume at a given energy.

The carrier density is obtained by integrating the charge density over the energy that the charges are allowed to have.

The charge carrier density is a particle density, so integrating it over a volume *V* gives the number of charge carriers *N* in that volume

$$N = \int_V n(\mathbf{r})\,\mathrm{d}V .$$

where

$n(\mathbf{r})$ is the position-dependent charge carrier density.

If the density does not depend on position and is instead equal to a constant n_0 this equation simplifies to

$$N = V \cdot n_0.$$

The charge carrier densities enters equations concerning the electrical conductivity and related phenomena like the thermal conductivity.

Charge carriers in semiconductors

There are two recognized types of **charge carriers in semiconductors**. One of them is electrons, which carry negative electric charge. In addition, it is convenient to treat the traveling vacancies in the valence-band electron population (holes) as the second type of charge carriers, which carry a positive charge equal in magnitude to that of an electron.

Carrier generation and recombination

When an electron meets with a hole, they recombine and these free carriers effectively vanish. The energy released can be either thermal, heating up the semiconductor (*thermal recombination*, one of the sources of waste heat in semiconductors), or released as photons (*optical recombination*, used in LEDs and semiconductor lasers).

Majority and minority carriers

The more abundant charge carriers are called **majority carriers**. In n-type semiconductors they are electrons, while in p-type semiconductors they are holes. The less abundant charge carriers are called **minority carriers**; in n-type semiconductors they are holes, while in p-type semiconductors they are electrons.

In an intrinsic semiconductor the concentrations of both types of carriers are ideally equal.

Minority carriers play an important role in bipolar transistors and solar cells. However, their role in field-effect transistors (FETs) is a bit more complex: for example, a MOSFET has both p-type and n-type regions. The transistor action involves the majority carriers of the source and drain regions, but these carriers traverse the body of the opposite type, where they are minority carriers. However, the traversing carriers hugely outnumber their opposite type in the transfer region (in fact, the opposite type carriers are removed by an applied electric field that creates an inversion layer), so conventionally the source and drain designation for the carriers is adopted, and FETs are called "majority carrier" devices.

Saturation velocity

Saturation velocity is the maximum velocity a charge carrier in a semiconductor, generally an electron, attains in the presence of very high electric fields. Charge carriers normally move at an average drift speed proportional to the electric field strength they experience temporally. The proportionality constant is known as mobility of the carrier, which is a material property. A good conductor would have a high mobility value for its charge carrier, which means higher velocity, and consequently higher current values for a given electric field strength. There is a limit though to this process and at some high field value, a charge carrier can not move any faster, having reached its saturation velocity, due to mechanisms that eventually limit the movement of the carriers in the material.

Saturation velocity is a very important parameter in the design of semiconductor devices, especially field effect transistors, which are basic building blocks of almost all modern integrated circuits. Typical values of saturation velocity may vary greatly for different materials, for example for Si it is in the order of 1×10^7 cm/s, for GaAs 1.2×10^7 cm/s, while for 6H-SiC, it is near 2×10^7 cm/s. Typical electric field strengths at which carrier velocity saturates is usually on the order of 10-100 kV/cm. Both saturation field and the saturation velocity of a semiconductor material are typically strong function of impurities, crystal defects and temperature.

For extremely small scale devices, where the high-field regions may be comparable or smaller than the average mean free path of the charge carrier, one can observe velocity overshoot, or hot electron effects which has become more important as the transistor geometries continually decrease to enable design of faster, larger and more dense integrated circuits. The regime where the two terminals between which the elctron moves is much smaller than the mean free path, is sometimes referred as ballistic transport. There has been numerous attempts in the past to build transistors based on this principle without much success. Nevertheless, developing field of nanotechnology, and new materials such as Carbon nanotubes and graphene, offers new hope.

Though in a semiconductor such as Si saturation velocity of a carrier is same as the peak velocity of the carrier, for some other materials with more complex energy band structures, this is not true. In GaAs or InP for example the carrier drift velocity reaches to a maximum as a function of field and then it begins to actually decrease as the electric field applied is increased further. Carriers which have gained enough energy are kicked up to a different conduction band which presents a lower drift velocity and eventually a lower saturation velocity in these materials. This results in an overall decrease of current for higher voltage until all electrons are in the "slow" band and this is the principle behind operation of a Gunn diode, which can display negative differential resistivity. Due to the transfer of electrons to a different conduction band involved, such devices, usually single terminal, are referred to as transferred electron devices, or TEDs.

Spontaneous emission

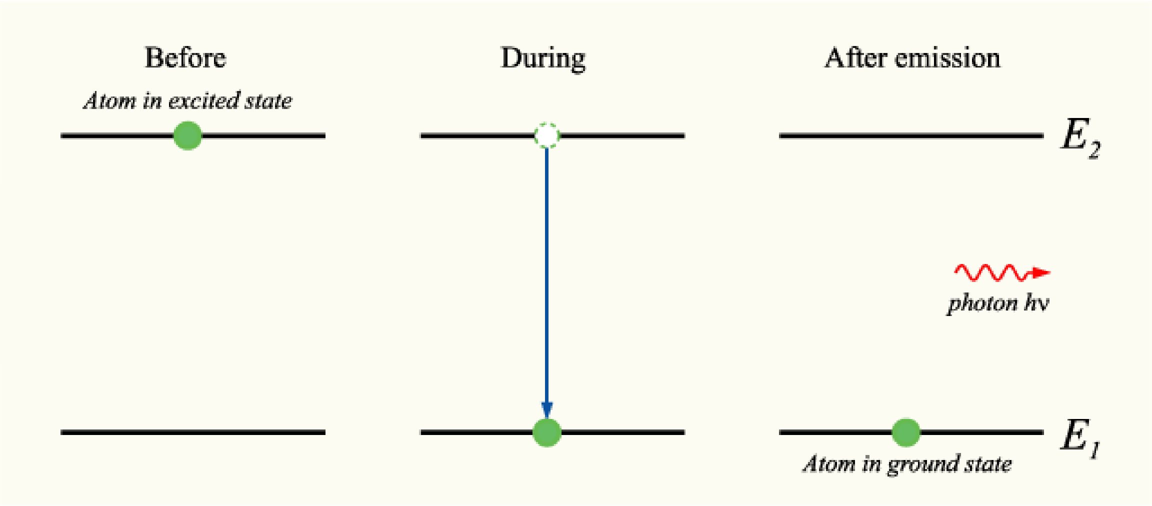
Spontaneous emission is the process by which a light source such as an atom, molecule, nanocrystal or nucleus in an excited state undergoes a transition to a state with a lower energy, e.g., the ground state and emits a photon. Spontaneous emission of light or luminescence is a fundamental process that plays an essential role in many phenomena in nature and forms the basis of many applications, such as fluorescent tubes, older television screens (cathode ray tubes), plasma display panels, lasers (for startup - normal continuous operation works by stimulated emission instead) and light emitting diodes.

Introduction

If a light source ('the atom') is in the excited state with energy E_2, it may spontaneously decay to a lower lying level (e.g., the ground state) with energy E_1, releasing the difference in energy between the two states as a photon. The photon will have angular frequency ω and energy $\hbar\omega$(= $h\nu$, where h is the Planck constant and ν is the frequency):

$$E_2 - E_1 = \hbar\omega,$$

where $\hbar$is the reduced Planck constant. The phase of the photon in spontaneous emission is random as is the direction the photon propagates in. This is not true for stimulated emission. An energy level diagram illustrating the process of spontaneous emission is shown below:



If the number of light sources in the excited state is given by N, the rate at which N decays is:

$$\frac{\partial N}{\partial t} = -A_{21}N,$$

where A_{21} is the rate of spontaneous emission. In the rate-equation A_{21} is a proportionality constant for this particular transition in this particular light source. The constant is referred to as the *Einstein A coefficient*, and has units s^{-1}. The above equation can be solved to give:

$$N(t) = N(0)e^{-A_{21}t} = N(0)e^{-\Gamma_{rad}t},$$

where $N(0)$ is the initial number of light sources in the excited state, t is the time and Γ_{rad} is the radiative decay rate of the transition. The number of excited states N thus decays exponentially with time, similar to radioactive decay. After one lifetime, the number of excited states decays to 36.8% of its original value ($\frac{1}{e}$-time). The radiative decay rate Γ_{rad} is inversely proportional to the lifetime τ_{12}:

$$A_{21} = \Gamma_{12} = \frac{1}{\tau_{21}}.$$

Theory

Spontaneous transitions was not explainable within the framework of the old quantum theory, that is a theory in which the atomic levels are quantized, but the electromagnetic field is not. In fact, using the machinery of the usually called "first-quantized" quantum mechanics and one computes the probability of spontaneous transitions from one stationary state to another, one finds that it is zero. In order to explain spontaneous transitions, quantum mechanics must be extended to a "second-quantized" theory, wherein the electromagnetic field is quantized at every point in space. Such a theory is known as a quantum field theory; the quantum field theory of electrons and electromagnetic fields is known as quantum electrodynamics.

In quantum electrodynamics (or QED), the electromagnetic field has a ground state, the vacuum state, which can mix with the excited stationary states of the atom. As a result of this interaction, the "stationary state" of the atom is no longer a true eigenstate of the combined system of the atom plus electromagnetic field. In particular, the electron transition from the excited state to the electronic ground state mixes with the transition of the electromagnetic field from the ground state to an excited state, a field state with one photon in it. Spontaneous emission in free space depends upon vacuum fluctuations to get started.

Although there is only one electronic transition from the excited state to ground state, there are many ways in which the electromagnetic field may go from the ground state to a one-photon state. That is, the electromagnetic field has infinitely more degrees of freedom, corresponding to the different directions in which the photon can be emitted. Equivalently, one might say that the phase space offered by the electromagnetic field is infinitely larger than that offered by the atom. Since one must consider probabilities that occupy all of phase space equally, the combined system of atom plus electromagnetic field must undergo a transition from electronic excitation to a photonic excitation; the atom must decay by spontaneous emission. The time the light source remains in the excited state thus depends on the light source itself as well as its environment. Imagine trying to hold a pencil upright on the end of your finger. It will stay there if your hand is perfectly stable and nothing perturbs the equilibrium. But the slightest perturbation will make the pencil fall into a more stable equilibrium position. Similarly, vacuum fluctuations cause an excited atom to fall into its ground state.

In spectroscopy one can frequently find that atoms or molecules in the excited states dissipate their energy in the absence of any external source of photons. This is not spontaneous emission, but is actually nonradiative relaxation of the atoms or molecules caused by the fluctuation of the surrounding molecules present inside the bulk.

Rate of spontaneous emission

The rate of spontaneous emission (i.e., the radiative rate) can be described by Fermi's golden rule. The rate of emission depends on two factors: an 'atomic part', which describes the internal structure of the light source and a 'field part', which describes the density of electromagnetic modes of the environment. The atomic part describes the strength of a transition between two states in terms of transition moments. In a homogeneous medium, such as free space, the rate of spontaneous emission in the dipole approximation is given by:

$$\Gamma_{rad}(\omega) = \frac{\omega^3 n |\mu_{12}|^2}{3\pi\varepsilon_0 \hbar {c_0}^3}$$

where ω is the emission frequency, n is the index of refraction, μ_{12} is the transition dipole moment, $\varepsilon_0$is the vacuum permittivity, $\hbar$is the reduced Planck constant and c_0 is the vacuum speed of light. (This approximation breaks down in the case of inner shell electrons in high-Z atoms.) Clearly, the rate of spontaneous emission in free space increases with ω^3. In contrast with atoms, which have a discrete emission spectrum, quantum dots can be tuned continuously by changing their size. This property has been used to check the ω^3-frequency dependence of the spontaneous emission rate as described by Fermi's golden rule.

Radiative and nonradiative decay: the quantum efficiency

In the rate-equation above, it is assumed that decay of the number of excited states *N* only occurs under emission of light. In this case one speaks of full radiative decay and this means that the quantum efficiency is 100%. Besides radiative decay, which occurs under the emission of light, there is a second decay mechanism; nonradiative decay. To determine the total decay rate Γ_{tot}, radiative and nonradiative rates should be summed:

$$\Gamma_{tot} = \Gamma_{rad} + \Gamma_{nrad}$$

where Γ_{tot} is the total decay rate, Γ_{rad} is the radiative decay rate and Γ_{nrad} the nonradiative decay rate. The quantum efficiency (QE) is defined as the fraction of emission processes in which emission of light is involved:

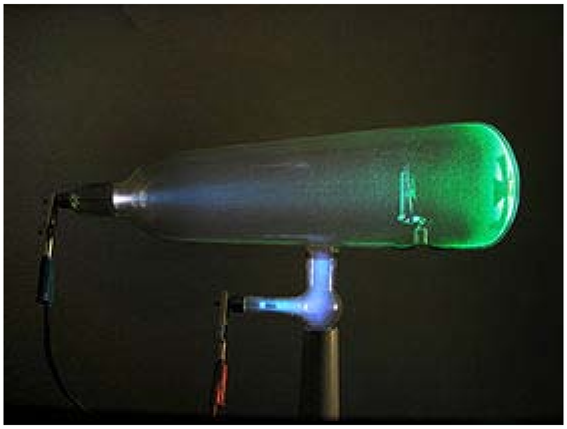
$$QE = \frac{\Gamma_{rad}}{\Gamma_{nrad} + \Gamma_{rad}}.$$

In nonradiative relaxation, the energy is released as phonons, more commonly known as heat. Nonradiative relaxation occurs when the energy difference between the levels is very small, and these typically occur on a much faster time scale than radiative transitions. For many materials (for instance, semiconductors), electrons move quickly from a high energy level to a meta-stable level via small nonradiative transitions and then make the final move down to the bottom level via an optical or radiative transition. This final transition is the transition over the bandgap in semiconductors. Large nonradiative transitions do not occur frequently because the crystal structure generally can not support large vibrations without destroying bonds (which generally doesn't happen for relaxation). Meta-stable states form a very important feature that is exploited in the construction of lasers. Specifically, since electrons decay slowly from them, they can be piled up in this state without too much loss and then stimulated emission can be used to boost an optical signal.

Chapter-3

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

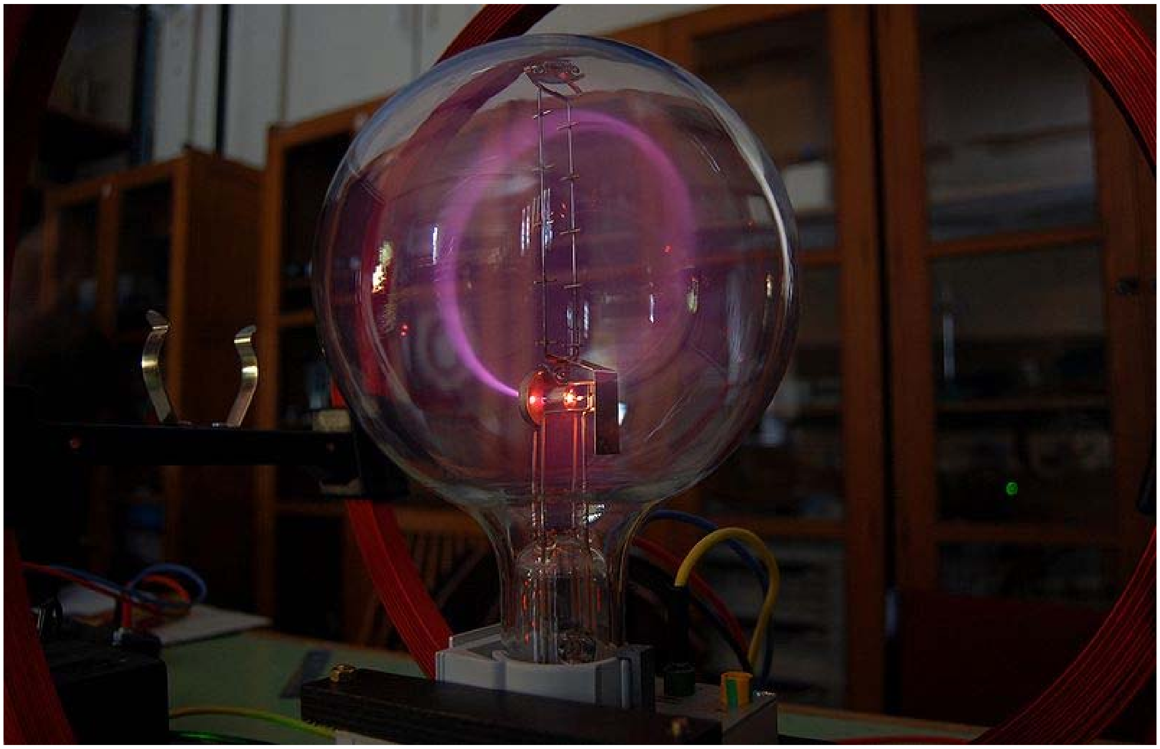
The ancient Greeks noticed that amber attracted small objects when rubbed with fur. Apart from lightning, this phenomenon is humanity's earliest recorded experience with electricity. In his 1600 treatise *De Magnete*, the English scientist William Gilbert coined the New Latin term *electricus*, to refer 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 the rays 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.

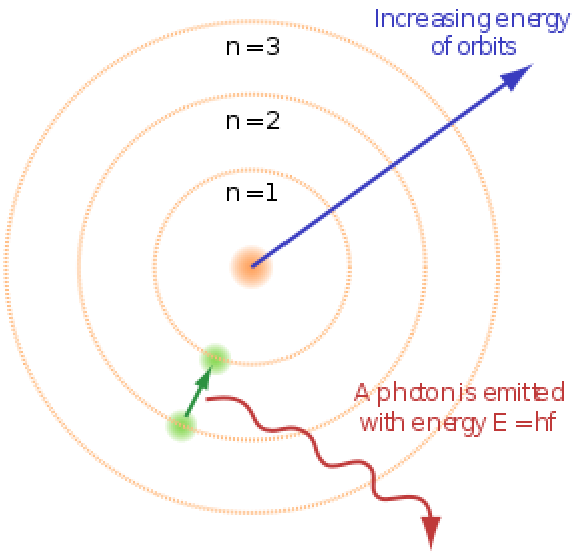
In 1896, the British physicist J. J. Thomson, with his colleagues John S. Townsend and H. A. Wilson, performed experiments indicating that cathode rays really were unique particles, rather than waves, atoms or molecules as was believed earlier. Thomson made good estimates of both the charge *e* and the 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.

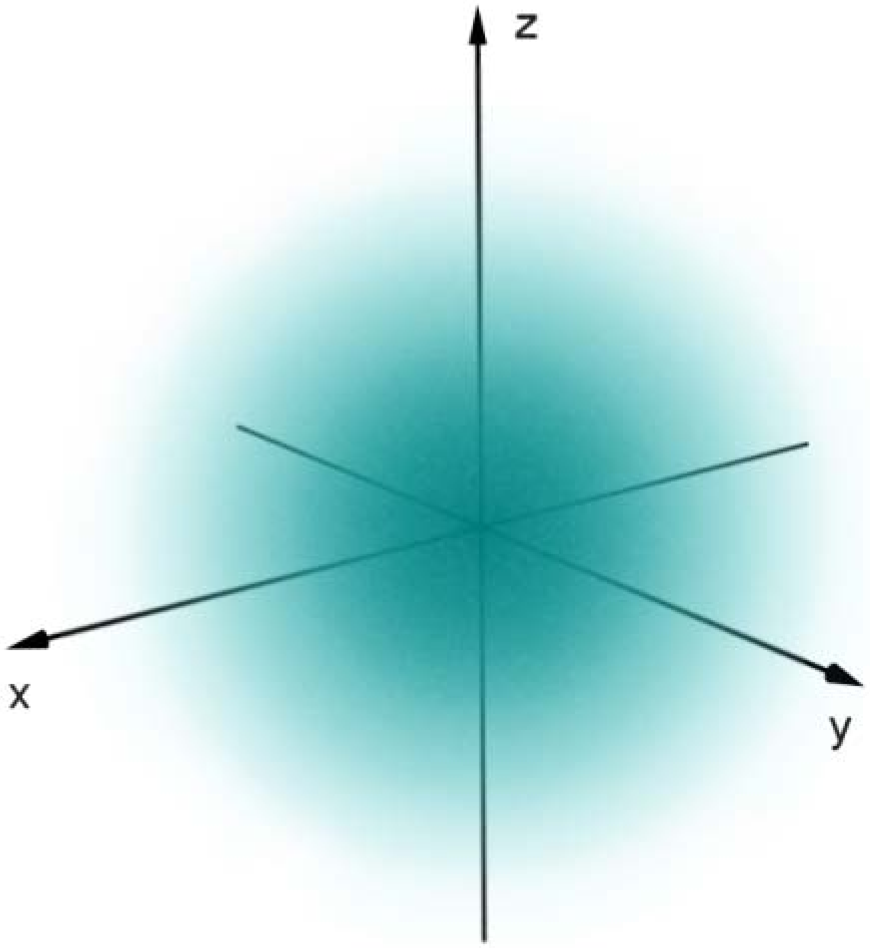
By 1914, experiments by physicists Ernest Rutherford, Henry Moseley, James Franck and Gustav Hertz had largely established the structure of an atom as a dense nucleus of positive charge surrounded by lower-mass electrons. In 1913, Danish physicist Niels Bohr postulated that electrons resided in quantized 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.

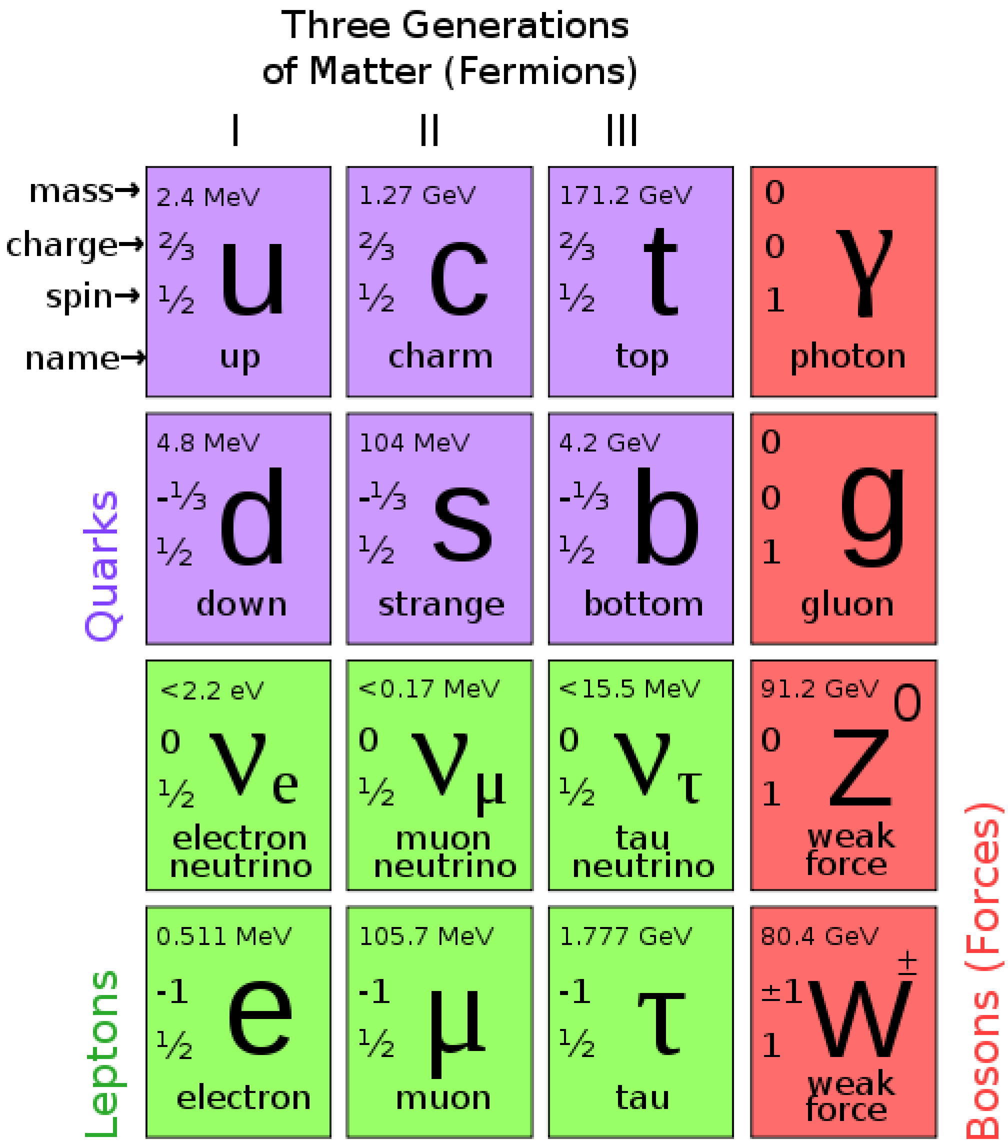
Particle accelerators

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



	I	II	III	
Quarks	2.4 MeV ⅔ ½ u up	1.27 GeV ⅔ ½ c charm	171.2 GeV ⅔ ½ t top	0 0 1 γ photon
	4.8 MeV -⅓ ½ d down	104 MeV -⅓ ½ s strange	4.2 GeV -⅓ ½ b bottom	0 0 1 g gluon
Leptons	<2.2 eV 0 ½ ν_e electron neutrino	<0.17 MeV 0 ½ ν_μ muon neutrino	<15.5 MeV 0 ½ ν_τ tau neutrino	91.2 GeV 0 1 Z^0 weak force
	0.511 MeV -1 ½ e electron	105.7 MeV -1 ½ μ muon	1.777 GeV -1 ½ τ tau	80.4 GeV ±1 1 $W^{\pm}$ weak force

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 $^1/_2$.

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.

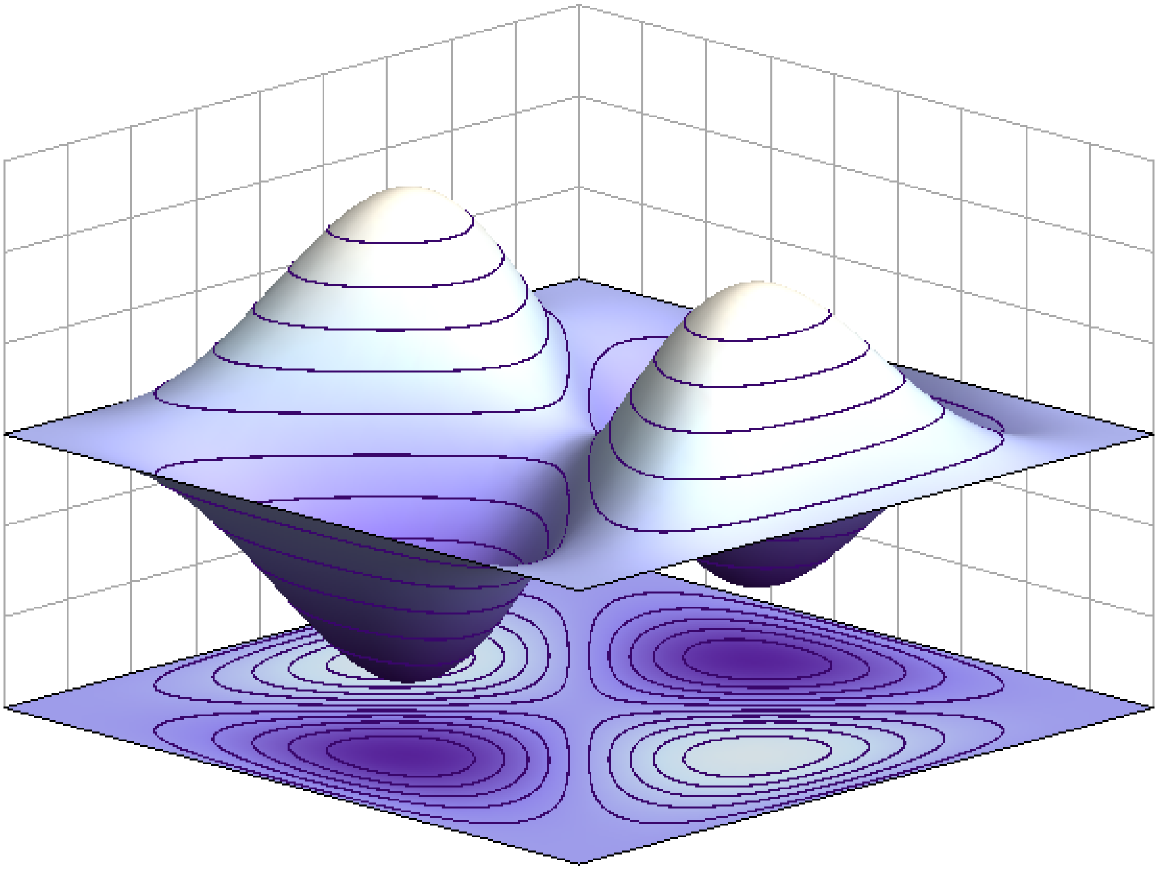
The electron has an intrinsic angular momentum or spin of $^1/_2$. This property is usually stated by referring to the electron as a spin-$^1/_2$ particle. For such particles the spin magnitude is $^{\sqrt{3}}/_2\ \hbar$. while the result of the measurement of a projection of the spin on any axis can only be $\pm^{\hbar}/_2$. 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 $9.27400915(23)\times10^{-24}$ 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.



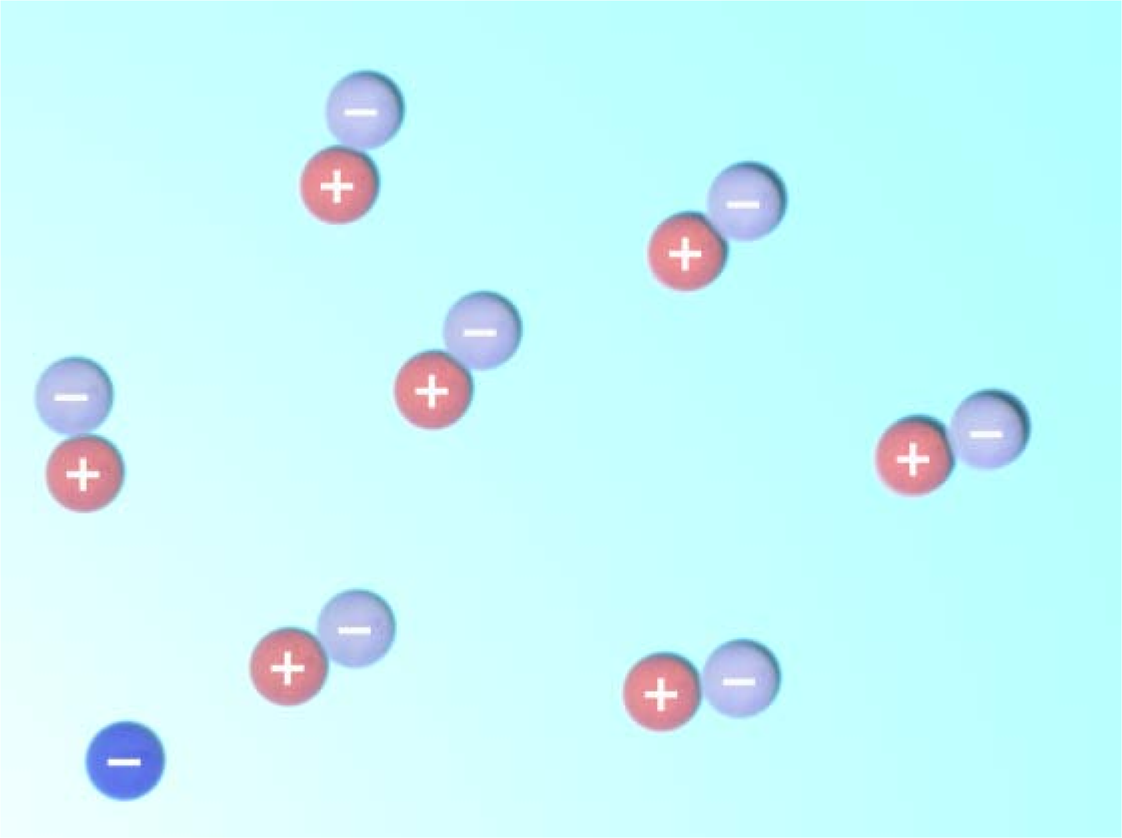
Example of an antisymmetric wave function for a quantum state of two identical fermions in a 2-dimensional box. If the particles swap position, the wave function inverts its sign.

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, $\Delta E \cdot \Delta t \geq \hbar$. 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)

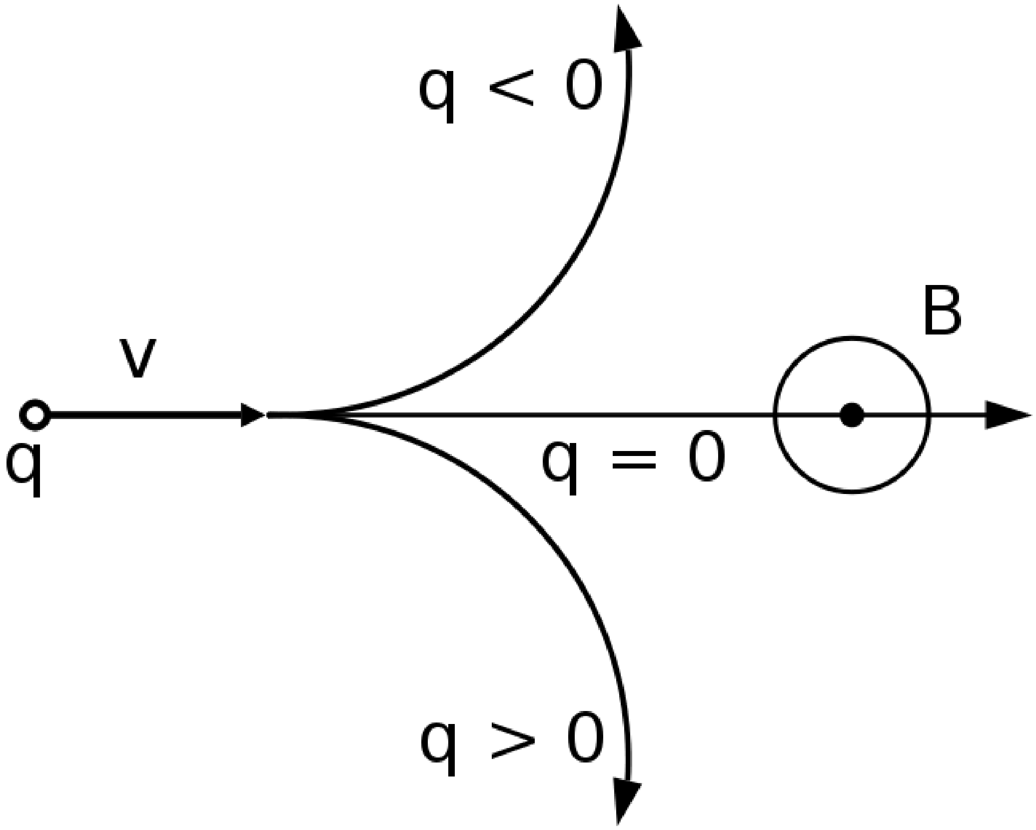
While an electron–positron virtual pair is in existence, the coulomb force from the ambient electric field surrounding an electron causes a created positron to be attracted to the original electron, while a created electron experiences a repulsion. This causes what is called vacuum polarization. In effect, 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

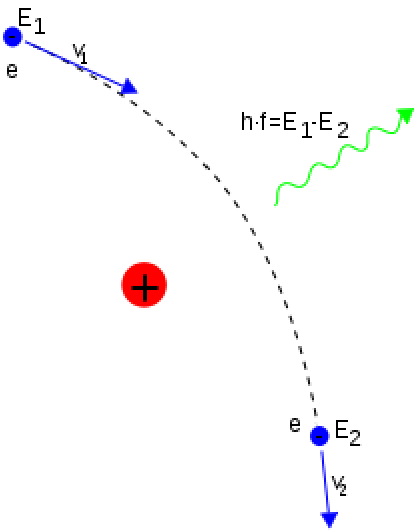
An electron generates an electric field that exerts an attractive force on a particle with a positive charge, such as the proton, and a repulsive force on a particle with a negative charge. The strength of this force is determined by Coulomb's inverse square law. When an electron is in motion, it generates 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_ec, which is known as the Compton wavelength. For an electron, it has a value of $2.43\times10{-}12$ 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.

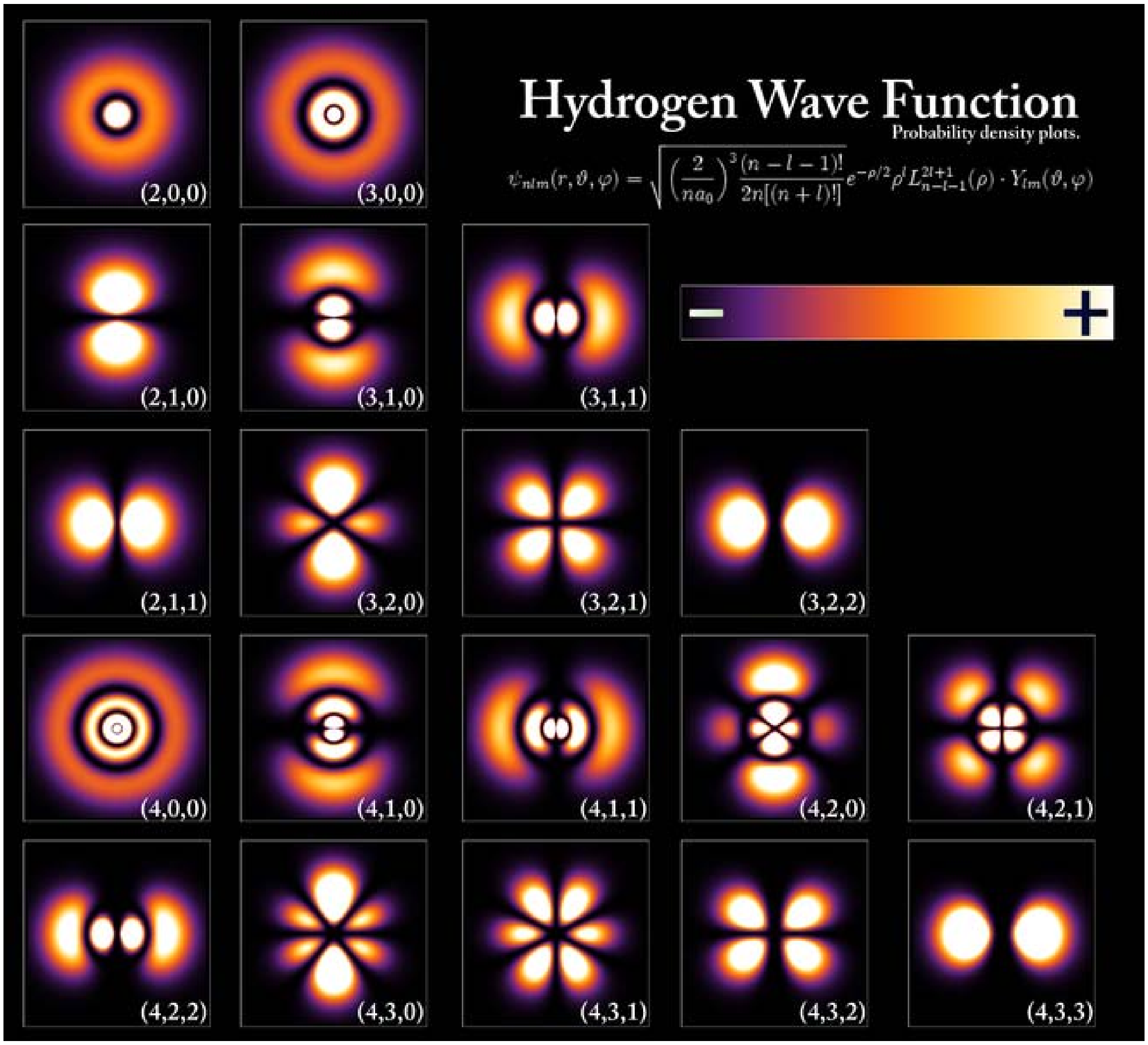
The relative strength of the electromagnetic interaction between two charged particles, such as an electron and a proton, is given by the fine-structure constant. This value is a dimensionless quantity formed by the ratio of two energies: the electrostatic energy of attraction (or repulsion) at a separation 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 Z0
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.

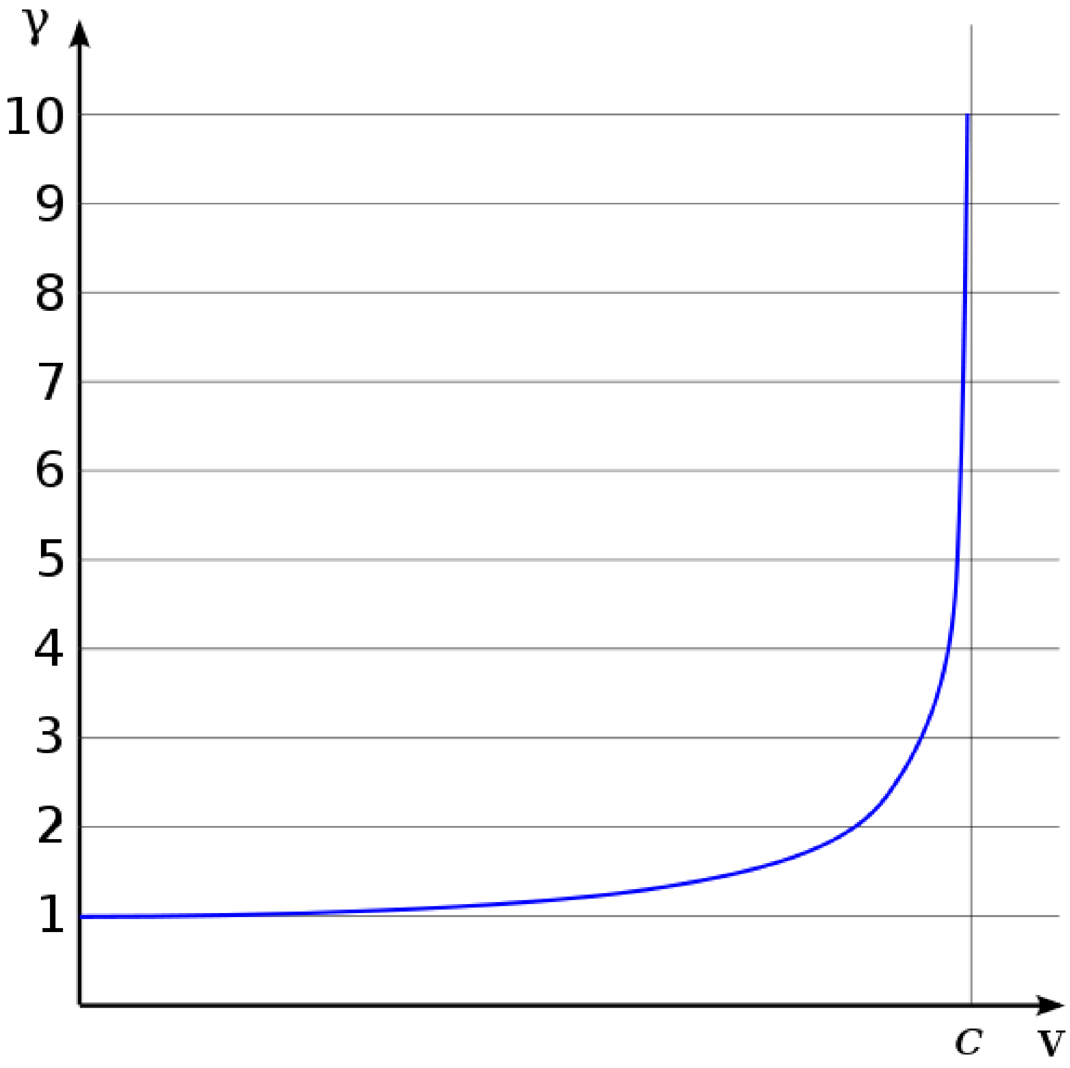
Metals make relatively good conductors of heat, primarily because the delocalized electrons are free to transport thermal energy between atoms. However, unlike electrical conductivity, the thermal conductivity of a metal is nearly independent of temperature. This is expressed mathematically by the Wiedemann-Franz 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.



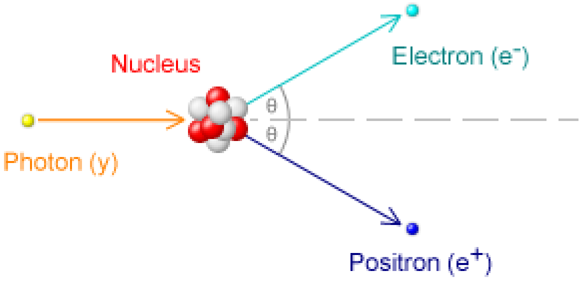
Lorentz factor as a function of velocity. It starts at value 1 and goes to infinity as *v* approaches *c*.

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. 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,

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

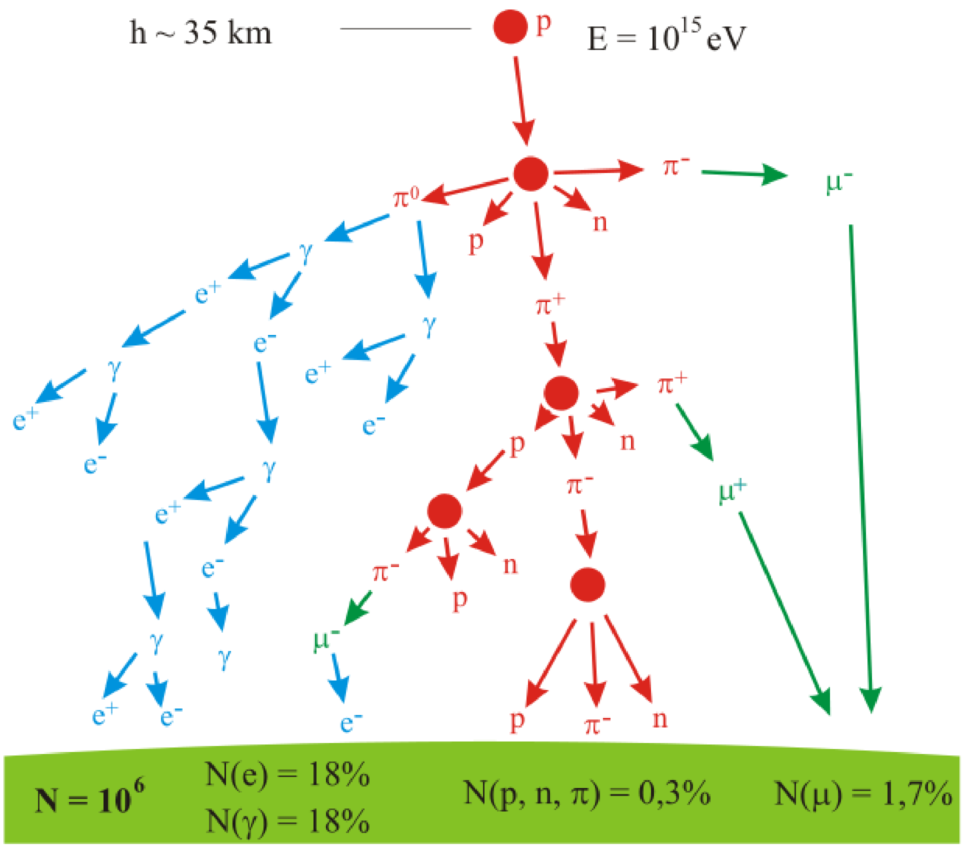
where γ is a photon, e^{+} is a positron and e^{-} is an electron. Likewise, positron-electron pairs annihilated each other and emitted energetic photons. An equilibrium between electrons, positrons and photons was maintained during this phase of the evolution of the Universe. After 15 seconds had passed, 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,

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

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^-,

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

where $\mu^-$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.

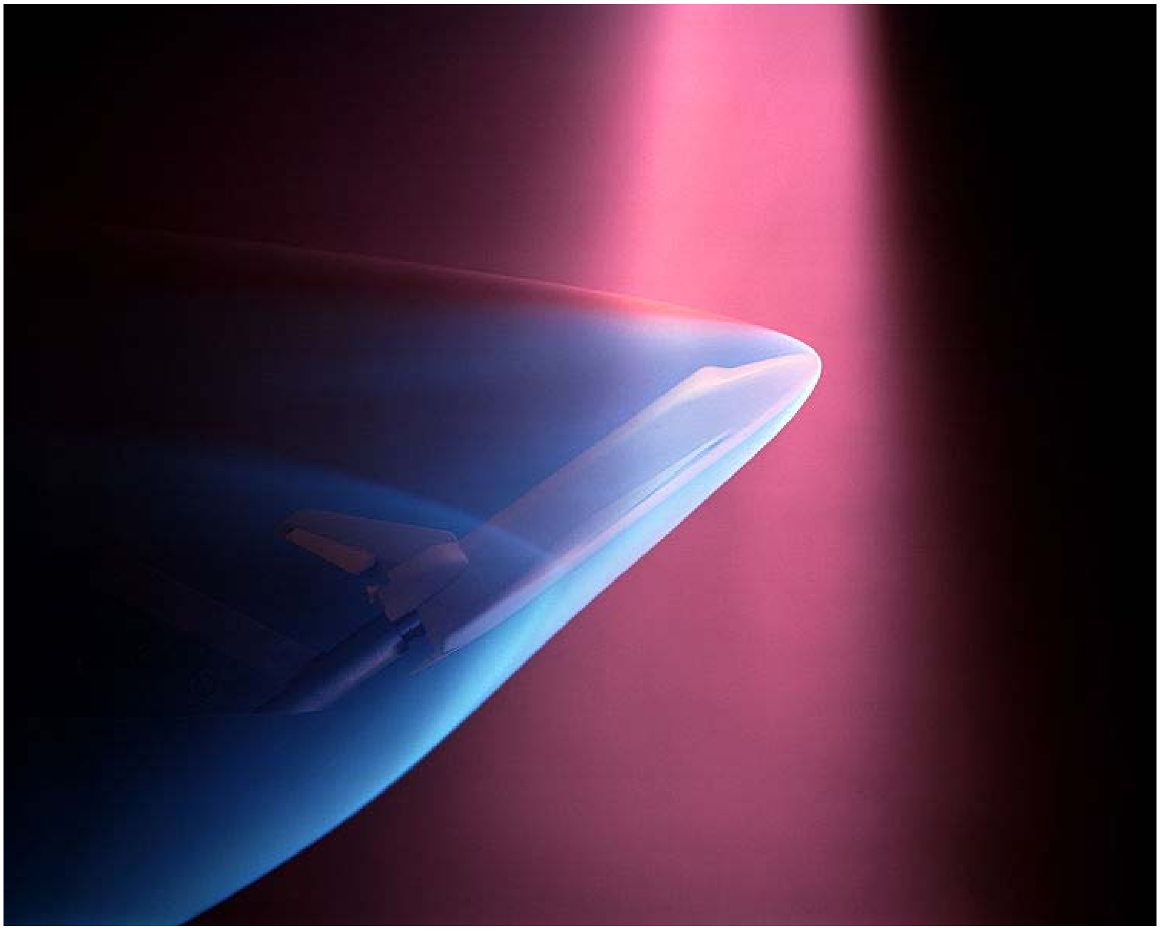
In laboratory conditions, the interactions of individual electrons can be observed by means of particle detectors, which allow measurement of specific properties such as energy, spin and charge. The development of the Paul trap and Penning trap allows charged particles to be contained within a small 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-4

High-k Dielectric and Low-k Dielectric

High-k dielectric

The term **high-κ dielectric** refers to a material with a high dielectric constant κ (as compared to silicon dioxide) used in semiconductor manufacturing processes which replaces the silicon dioxide gate dielectric. The implementation of high-κ gate dielectrics is one of several strategies developed to allow further miniaturization of microelectronic components, colloquially referred to as extending Moore's Law.

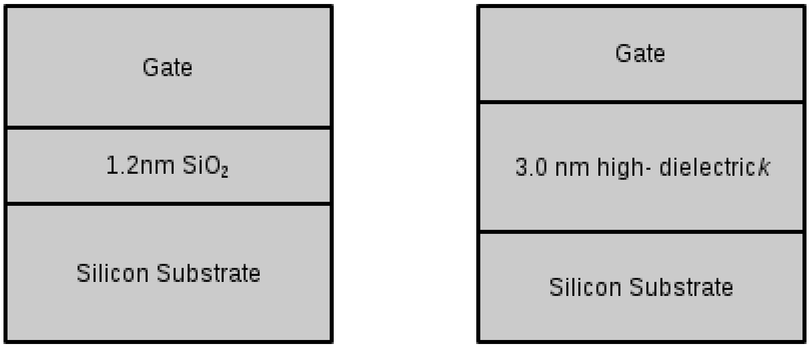
Need for high-κ materials

Silicon dioxide has been used as a gate oxide material for decades. As transistors have decreased in size, the thickness of the silicon dioxide gate dielectric has steadily decreased to increase the gate capacitance and thereby drive current and device performance. As the thickness scales below 2 nm, leakage currents due to tunneling increase drastically, leading to unwieldy power consumption and reduced device reliability. Replacing the silicon dioxide gate dielectric with a high-κ material allows increased gate capacitance without the concomitant leakage effects.

First principles

The gate oxide in a MOSFET can be modeled as a parallel plate capacitor. Ignoring quantum mechanical and depletion effects from the Si substrate and gate, the capacitance *C* of this parallel plate capacitor is given by

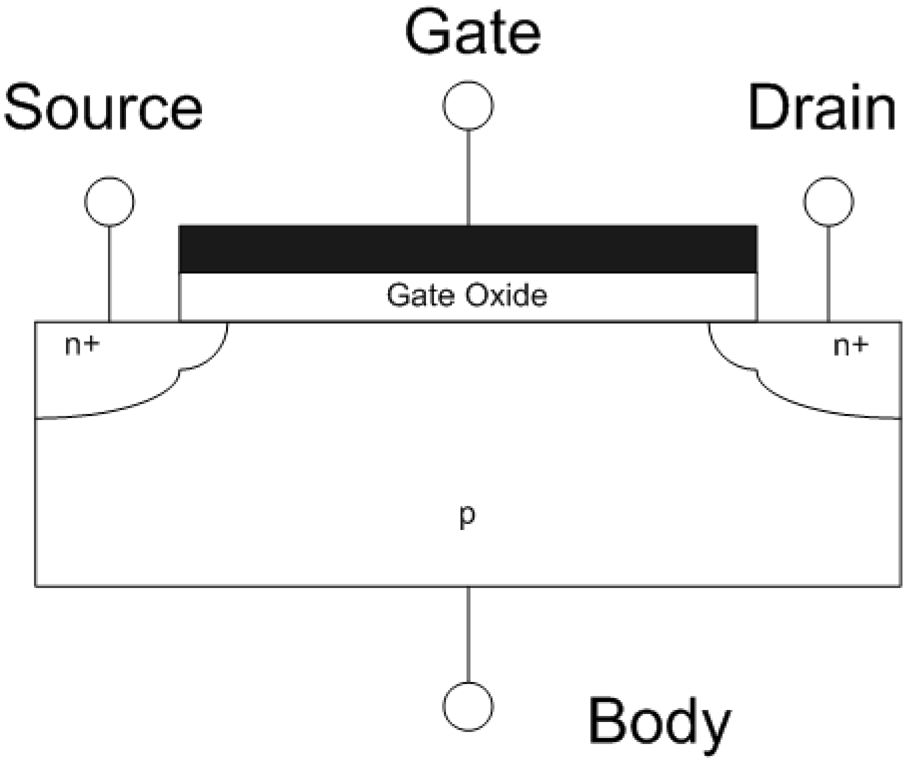
$$C = \frac{\kappa \varepsilon_0 A}{t}$$



Existing 90nm Process
Capacitance = 1x
Leakage Current = 1x

A potential high- process*k*
Capacitance = 1.6x
Leakage Current = 0.01x

Conventional silicon dioxide gate dielectric structure compared to a potential high-k dielectric structure



Cross-section of an NMOS transistor showing the gate oxide dielectric

Where

- A is the capacitor area
- κ is the relative dielectric constant of the material (3.9 for silicon dioxide)
- ε_0 is the permittivity of free space
- t is the thickness of the capacitor oxide insulator

Since leakage limitation constrains further reduction of t, an alternative method to increase gate capacitance is alter κ by replacing silicon dioxide with a high-κ material. In such a scenario, a thicker gate layer might be used which can reduce the leakage current flowing through the structure as well as improving the gate dielectric reliability.

Gate capacitance impact on drive current

The drive current I_D for a MOSFET can be written (using the gradual channel approximation) as

$$I_{D,Sat} = \frac{W}{L} \mu C_{inv} \frac{(V_G - V_{th})^2}{2}$$

Where

- W is the width of the transistor channel
- L is the channel length
- μ is the channel carrier mobility (assumed constant here)
- C_{inv} is the capacitance density associated with the gate dielectric when the underlying channel is in the inverted state
- V_G is the voltage applied to the transistor gate
- V_D is the voltage applied to the transistor drain
- V_{th} is the threshold voltage

The term $V_G - V_{th}$ is limited in range due to reliability and room temperature operation constraints, since a too large V_G would create an undesirable, high electric field across the oxide. Furthermore, V_{th} cannot easily be reduced below about 200 mV, because leakage currents due to increased oxide leakage (that is, assuming high-κ dielectrics are not available) and subthreshold conduction raise stand-by power consumption to unacceptable levels. Thus, according to this simplified list of factors, an increased $I_{D,sat}$ requires a reduction in the channel length or an increase in the gate dielectric capacitance.

Materials and considerations

Replacing the silicon dioxide gate dielectric with another material adds complexity to the manufacturing process. Silicon dioxide can be formed by oxidizing the underlying silicon, ensuring a uniform, conformal oxide and high interface quality. As a consequence, development efforts have focused on finding a material with a requisitely high dielectric constant that can be easily integrated into a manufacturing process. Other key considerations include band alignment to silicon (which may alter leakage current), film morphology, thermal stability, maintenance of a high mobility of charge carriers in the channel and minimization of electrical defects in the film/interface. Materials which have received considerable attention are hafnium silicate, zirconium silicate, hafnium dioxide and zirconium dioxide, typically deposited using atomic layer deposition.

It is expected that defect states in the high-k dielectric can influence its electrical properties. Defect states can be measured for example by using zero-bias thermally stimulated current, zero-temperature-gradient zero-bias thermally stimulated current spectroscopy, or inelastic electron tunneling spectroscopy (IETS).

Use in industry

The industry has employed oxynitride gate dielectrics since the 1990s, wherein a conventionally formed silicon oxide dielectric is infused with a small amount of nitrogen. The nitride content subtly raises the dielectric constant and is thought to offer other advantages, such as resistance against dopant diffusion through the gate dielectric.

In early 2007, Intel announced the deployment of hafnium-based high-k dielectrics in conjunction with a metallic gate for components built on 45 nanometer technologies, and has shipped it in the 2007 processor series codenamed Penryn. At the same time, IBM announced plans to transition to high-k materials, also hafnium-based, for some products in 2008. While not identified, it is most likely the dielectrics used by these companies are some form of nitrided hafnium silicates (HfSiON). HfO_2 and HfSiO are susceptible to crystallization during dopant activation annealing. NEC Electronics has also announced the use of a HfSiON dielectric in their 55 nm *UltimateLowPower* technology. However, even HfSiON is susceptible to trap-related leakage currents, which tend to increase with stress over device lifetime. The higher the hafnium concentration, the more severe the issue. However, there is no absolute guarantee that hafnium will be the basis of future high-k dielectrics. The 2006 ITRS roadmap predicts the implementation of high-k materials to be commonplace in the industry by 2010.

Low-k dielectric

In semiconductor manufacturing, a **low-κ dielectric** is a material with a small dielectric constant relative to silicon dioxide. Although the proper symbol for the dielectric constant is the Greek letter κ (kappa), in conversation such materials are referred to as being "low-k" (low-kay) rather than "low-κ" (low-kappa). Low-κ dielectric material implementation is one of several strategies used to allow continued scaling of microelectronic devices, colloquially referred to as extending Moore's law. In digital circuits, insulating dielectrics separate the conducting parts (wire interconnects and transistors) from one another. As components have scaled and transistors have got closer together, the insulating dielectrics have thinned to the point where charge build up and crosstalk adversely affect the performance of the device. Replacing the silicon dioxide with a low-κ dielectric of the same thickness reduces parasitic capacitance, enabling faster switching speeds and lower heat dissipation.

Low-κ materials

The dielectric constant of SiO_2, the insulating material used in silicon chips, is 3.9. This number is the ratio of the permittivity of SiO_2 divided by permittivity of vacuum,

$\varepsilon_{SiO2}/\varepsilon_0$,where $\varepsilon_0 = 8.854\times10^{-6}$ pF/μm . There are many materials with lower dielectric constants but few of them can be suitably integrated into a manufacturing process. Development efforts have focused primarily on three classes of materials:

Fluorine-doped silicon dioxide

By doping SiO_2 with fluorine to produce fluorinated silica glass, the dielectric constant is lowered from 3.9 to 3.5.

Carbon-doped silicon dioxide

By doping SiO_2 with carbon, one can lower the dielectric constant to 3.0. Major products of carbon doped silicon dioxide include **Black Diamond** from Applied Materials, **Aurora** from ASM International N.V. The Aurora is the low-κ material used in Intel 90 nm, 65 nm and 45 nm lines, while the Black Diamond controlled about 80% of low-κ material market. Novellus Systems' **Coral** also falls in this category.

Porous silicon dioxide

Various methods may be employed to create large voids or pores in a silicon dioxide dielectric. Air has a dielectric constant of roughly 1.0005, thus the dielectric constant of the porous material may be reduced by increasing the porosity of the film. Dielectric constants lower than 2.0 have been reported. Integration difficulties related to porous silicon dioxide implementation include low mechanical strength and difficult integration with etch and polish processes.

Porous carbon-doped silicon dioxide

By UV curing, floating methyl group in carbon doped silicon dioxide can be eliminated and pores can be introduced to the carbon doped silicon dioxide low-κ materials. The technology for producing porous dielectric layers with very low dielectric constant was developed by A. Grill, V.V. Patel, S.M. Gates in the late 1990's at IBM (US6312793-International Business Machines]. While all materials used for the deposition of low-k dielectric layers are in the public domain and therefore not patentable, the technique of making such layers was rapidly adopted by the semiconductor manufacturing equipment industry and commercialised as an essential application supplied with the equipment. Examples include Applied Materials' **Black Diamond II** technology and ASM International's **Aurora 2.7** and **Aurora ULK** process.

Spin-on organic polymeric dielectrics

Polymeric dielectrics are generally deposited by a spin-on approach, such as those traditionally used to deposit photoresist, rather than chemical vapor deposition. Integration difficulties include low mechanical strength and thermal stability. **SiLK** from Dow Chemical is a well known example of low-κ material in this category. Other spin-on organic low-κ include polyimide, polynorbornenes, Benzocyclobutene, PTFE, SU-8 2000.

Porous SiLK

By introducing pores into the SiLK resin, the dielectric constant value can be lowered to 2.2.

Spin-on silicone based polymeric dielectric

There are two kinds of silicone based polymeric dielectric materials, hydrogen silsesquioxane (HSQ) and methylsilsesquioxane (MSQ).

Chapter-5

Semiconductor Device Defects

Catastrophic optical damage

Catastrophic optical damage (**COD**), or **catastrophic optical mirror damage** (**COMD**), is a failure mode of high-power semiconductor lasers. It occurs when the semiconductor junction is overloaded by exceeding its power density and absorbs too much of the produced light energy, leading to melting and recrystallization of the semiconductor material at the facets of the laser. This is often colloquially referred to as "blowing the diode." The affected area contains a large number of lattice defects, negatively affecting its performance. If the affected area is sufficiently large, it can be observable under optical microscope as darkening of the laser facet, and/or as presence of cracks and grooves. The damage can occur within a single laser pulse, in less than a millisecond. The time to COD is inversely proportional to the power density.

Catastrophic optical damage is one of the limiting factors in increasing performance of semiconductor lasers. It is the primary failure mode for AlGaInP/AlGaAs red lasers.

Short-wavelength lasers are more susceptible to COD than long-wavelength ones.

The typical values for COD in industrial products range between 12 and 20 MW/cm^2.

Causes and mechanisms

At the edge of a diode laser, where light is emitted, a mirror is traditionally formed by cleaving the semiconductor wafer to form a specularly reflecting plane. This approach is facilitated by the weakness of the [110] crystallographic plane in III-V semiconductor crystals (such as GaAs, InP, GaSb, etc.) compared to other planes. A scratch made at the edge of the wafer and a slight bending force causes a nearly atomically perfect mirror-like cleavage plane to form and propagate in a straight line across the wafer.

But it so happens that the atomic states at the cleavage plane are altered (compared to their bulk properties within the crystal) by the termination of the perfectly periodic lattice at that plane. Surface states at the cleaved plane have energy levels within the (otherwise forbidden) band gap of the semiconductor.

The absorbed light causes generation of electron-hole pairs. These can lead to breaking of chemical bonds on the crystal surface followed by oxidation, or to release of heat by nonradiative recombination. The oxidized surface then shows increased absorption of the laser light, which further accelerates its degradation. The oxidation is especially problematic for semiconductor layers containing aluminium.

Essentially, as a result when light propagates through the cleavage plane and transits to free space from within the semiconductor crystal, a fraction of the light energy is absorbed by the surface states whence it is converted to heat by phonon-electron interactions. This heats the cleaved mirror. In addition the mirror may heat simply because the edge of the diode laser—which is electrically pumped—is in less-than-perfect contact with the mount that provides a path for heat removal. The heating of the mirror causes the band gap of the semiconductor to shrink in the warmer areas. The band gap shrinkage brings more electronic band-to-band transitions into alignment with the photon energy causing yet more absorption. This is thermal runaway, a form of positive feedback, and the result can be melting of the facet, known as *catastrophic optical damage*, or COD.

Deterioration of the laser facets with aging and effects of the environment (erosion by water, oxygen, etc.) increases light absorption by the surface, and decreases the COD threshold. A sudden catastrophic failure of the laser due to COD then can occur after many thousands hours in service.

Improvements

One of the methods of increasing the COD threshold in AlGaInP laser structures is the sulfur treatment, which replaces the oxides at the laser facet with chalcogenide glasses. This decreases the recombination velocity of the surface states.

Reduction of recombination velocity of surface states can be also achieved by cleaving the crystals in ultrahigh vacuum and immediate deposition of a suitable passivation layer.

A thin layer of aluminium can be deposited over the surface, for gettering the oxygen.

Another approach is doping of the surface, increasing the band gap and decreasing absorption of the lasing wavelength, shifting the absorption maximum several nanometers up.

Current crowding near the mirror area can be avoided by prevention of injecting charge carriers near the mirror region. This is achieved by depositing the electrodes away from the mirror, at least several carrier diffusion distances.

Energy density on the surface can be reduced by employing a waveguide broadening the optical cavity, so the same amount of energy exits through a larger area. Energy density of 15–20 MW/cm^2 corresponding to 100 mW per micrometer of stripe width are now achievable. A wider laser stripe can be used for higher output power, for the cost of transverse mode oscillations and therefore worsening of spectral and spatial beam quality.

In the 1970s, this problem, which is particularly nettlesome for GaAs-based lasers emitting between 1 µm and 0.630 µm wavelengths (less so for InP based lasers used for long-haul telecommunications which emit between 1.3 µm and 2 µm), was identified. Michael Ettenberg, a researcher and later Vice President at RCA Laboratories' David Sarnoff Research Center in Princeton, New Jersey, devised a solution. A thin layer of aluminum oxide was deposited on the facet. If the aluminum oxide thickness is chosen correctly, it functions as an anti-reflective coating, reducing reflection at the surface. This alleviated the heating and COD at the facet.

Since then, various other refinements have been employed. One approach is to create a so-called non-absorbing mirror (NAM) such that the final 10 µm or so before the light emits from the cleaved facet are rendered non-absorbing at the wavelength of interest. Such lasers are called **window lasers**.

In the very early 1990s, SDL, Inc. began supplying high power diode lasers with good reliability characteristics. CEO Donald Scifres and CTO David Welch presented new reliability performance data at, e.g., SPIE Photonics West conferences of the era. The methods used by SDL to defeat COD were considered to be highly proprietary and have still not been disclosed publicly as of June, 2006.

In the mid-1990s IBM Research (Ruschlikon, Switzerland) announced that it had devised its so-called "E2 process" which conferred extraordinary resistance to COD in GaAs-based lasers. This process, too, has never been disclosed as of June, 2006.

Hot carriers injection

Hot carriers injection (**HCI**) is the phenomenon in solid-state or semiconductor electronic devices where either an electron or a “hole” gains sufficient kinetic energy to overcome a potential barrier necessary to break an interface state. The term hot electron comes from the effective temperature used to model carrier density (with fermi dirac function). One should not think that the term "hot" refers to the actual temperature of the MOS transistor. At higher temperatures, the mean free path (distance between two collisions with atoms in the substrate) is shorter, which decreases the energy gain by the carrier. As a result, Hot Carrier Degradation is more important at low temperature.

HCI and CMOS Semiconductor Technology

Semiconductor Physics of HCI

The term “hot carrier injection” usually refers to the effect in MOSFETs, where a carrier is injected from the conducting channel in the silicon substrate to the gate dielectric, which usually is made of silicon dioxide (SiO_2).

To become “hot” and enter the conduction band of SiO_2, an electron must gain a kinetic energy of 3.3 eV. For holes, the valence band offset in this case dictates they must have a kinetic energy of 4.6 eV.

When electrons are accelerated in the channel, they gain energy along the mean free path. This energy is lost in two different ways:

1. The carrier hit an atom in the substrate. Then the collision create a cold carrier and an additional electron-hole pair. In the case of nMOS transistors, additional electrons are collected by the channel and additional holes are evacuated by the substrate.
2. The carrier hit a Si-H bond and break the bond. An interface state is created and the Hydrogen atom is released in the substrate.

The probability to hit either an atom or a Si-H bond is random. And the average energy involved in each process is the same in both case.

This is the reason why the substrate current is monitored during HCI stress. A high substrate current means a large number of created electron-hole pairs and thus an efficient Si-H bond breakage mechanism.

When interface states are created, the threshold voltage is modified and the subthreshold slope is degraded. This leads to lower current, and degrades the operating frequency of integrated circuit.

Scaling and HCI

Advances in semiconductor manufacturing techniques and ever increasing demand for faster and more complex integrated circuits (ICs) have driven the associated Metal–Oxide–Semiconductor field-effect transistor (MOSFET) to scale to smaller dimensions.

However, it has not been possible to scale the supply voltage used to operate these ICs proportionately due to factors such as compatibility with previous generation circuits, noise margin, power and delay requirements, and non-scaling of threshold voltage, subthreshold slope, and parasitic capacitance.

As a result internal electric fields increase in aggressively scaled MOSFETs, which comes with the additional benefit of increased carrier velocities (up to velocity saturation), and hence increased switching speed, but also presents a major reliability problem for the long term operation of these devices, as high fields induce hot carrier injection which affects device reliability.

Large electric fields in MOSFETs imply the presence of high-energy carriers, referred to as “**hot carriers**”. These hot carriers that have sufficiently high energies and momenta to allow them to be injected from the semiconductor into the surrounding dielectric films such as the gate and sidewall oxides as well as the buried oxide in the case of silicon on insulator (SOI) MOSFETs.

CMOS Reliability Impact of HCI

The presence of such mobile carriers in the oxides triggers numerous physical damage processes that can drastically change the device characteristics over prolonged periods. The accumulation of damage can eventually cause the circuit to fail as key parameters such as threshold voltage shift due to such damage. The accumulation of damage resulting degradation in device behavior due to hot carrier injection is called “**hot carrier degradation**”.

The useful life-time of circuits and integrated circuits based on such a MOS device are thus affected by the life-time of the MOS device itself. To assure that integrated circuits manufactured with minimal geometry devices will not have their useful life impaired, the life-time of the component MOS devices must have their HCI degradation well understood. Failure to accurately characterize HCI life-time effects can ultimately affect business costs such as warranty and support costs and impact marketing and sales promises for a foundry or IC manufacturer.

Relationship to Radiation Effects

Hot carrier degradation is fundamentally same as the ionization radiation effect known as the total dose damage to semiconductors, as experienced in space systems due to solar proton, electron, X-ray and gamma ray exposure.

HCI and NOR Flash Memory Cells

HCI is the basis of operation for a number of non-volatile memory technologies such as Electrically Erasable Programmable Read-Only Memory (EEPROM) cells. As soon as the potential detrimental influence of HC injection on the circuit reliability was recognized, several fabrication strategies were devised to reduce it without compromising the circuit performance.

NOR flash memory exploits the principle of hot carriers injection by deliberately injecting carriers across the gate oxide to charge the floating gate. This charge alters the MOS transistor threshold voltage to represent a logic '0' state. An uncharged floating gate represents a '1' state. Erasing the NOR Flash memory cell removes stored charge through the process of Fowler–Nordheim tunneling.

Because of the damage to the oxide caused by normal NOR Flash operation, HCI damage is one of the factors that cause the number of write-erase cycles to be limited. Because the ability to hold charge and the formation of damage traps in the oxide affects the ability to have distinct '1' and '0' charge states, HCI damage results in the closing of the non-volatile memory logic margin window over time. The number of write-erase cycles at which '1' and '0' can no longer be distinguished defines the endurance of a non-volatile memory.

Latchup

Latchup is a term used in the realm of integrated circuits (ICs) to describe a particular type of short circuit which can occur in an improperly designed circuit. More specifically it is the inadvertent creation of a low-impedance path between the power supply rails of a MOSFET circuit, triggering a parasitic structure which disrupts proper functioning of the part and possibly even leading to its destruction due to overcurrent. A power cycle is required to correct this situation.

The parasitic structure is usually equivalent to a thyristor (or SCR), a PNPN structure which acts as a PNP and an NPN transistor stacked next to each other. During a latchup when one of the transistors is conducting, the other one begins conducting too. They both keep each other in saturation for as long as the structure is forward-biased and some current flows through it - which usually means until a power-down. The SCR parasitic structure is formed as a part of the totem-pole PMOS and NMOS transistor pair on the output drivers of the gates.

The latchup does not have to happen between the power rails; it can happen at any place where the required parasitic structure exists. A spike of positive or negative voltage on an input or output pin of a digital chip, exceeding the rail voltage by more than a diode drop, is a common cause of latchup. Another cause is the supply voltage exceeding the absolute maximum rating, often from a transient spike in the power supply, leading to a breakdown of some internal junction. This frequently happens in circuits which use multiple supply voltages that do not come up in the proper order after a power-up, leading to voltages on data lines exceeding the input rating of parts that have not yet reached a nominal supply voltage.

Yet another common cause of latchups is ionizing radiation which makes this a significant issue in electronic products designed for space (or very high-altitude) applications.

Latchup prevention

It is possible to design chips that are latchup-resistant, where a layer of insulating oxide (called a *trench*) surrounds both the NMOS and the PMOS transistors. This breaks the parasitic SCR structure between these transistors. Such parts are important in the cases where the proper sequencing of power and signals cannot be guaranteed (e.g., in hot swap devices).

Devices fabricated in lightly doped epitaxial layers grown on heavily doped substrates are also less susceptible to latchup. The heavily doped layer acts as a current sink where excess minority carriers can quickly recombine.

Another possibility for a latchup prevention is the *Latchup Protection Technology* circuit. When a latchup is detected, the LPT circuit shuts down the chip and holds it powered-down for a preset time.

Most silicon-on-insulator devices are inherently latchup-resistant. Latchup is the low resistance connection between tub and power supply rails.

Chapter-6

Radiation Hardening

Radiation hardening is a method of designing and testing electronic components and systems to make them resistant to damage or malfunctions caused by ionizing radiation (particle radiation and high-energy electromagnetic radiation), such as would be encountered in outer space, high-altitude flight, around nuclear reactors, particle accelerators, or during nuclear accidents or nuclear warfare.

Most radiation-hardened chips are based on their commercial equivalents, with some manufacturing and design variations that reduce the susceptibility to interference from electromagnetic radiation. Due to the extensive development and testing required to produce a radiation-tolerant design of a microelectronic chip, radiation-hardened chips tend to lag behind the cutting-edge of developments.

Problems caused by radiation

Environments with high levels of ionizing radiation create special design challenges. A single charged particle can knock thousands of electrons loose, causing electronic noise and signal spikes. In the case of digital circuits, this can cause results which are inaccurate or unintelligible. This is a particularly serious problem in the design of artificial satellites, spacecraft, military aircraft, nuclear power stations, and nuclear weapons. In order to ensure the proper operation of such systems, manufacturers of integrated circuits and sensors intended for the (military) aerospace markets employ various methods of radiation hardening. The resulting systems are said to be **rad(iation)-hardened**, **rad-hard**, or (within context) **hardened**.

Major radiation damage sources

Typical sources of exposure of electronics to ionizing radiation are the Van Allen radiation belts for satellites, nuclear reactors in power plants for sensors and control circuits, residual radiation from isotopes in chip packaging materials, cosmic radiation for spacecraft and high-altitude aircraft, and nuclear explosions for potentially all military and civilian electronics.

- Cosmic rays come from all directions and consist of approx. 85% protons, 14% alpha particles, and 1% heavy ions, together with x-ray and gamma-ray radiation. Most effects are caused by particles with energies between 10^8 and $2*10^{10}$ eV. The atmosphere filters most of these, so they are primarily a concern for spacecraft and high-altitude aircraft.

- Solar particle events come from the direction of the sun and consist of a large flux of high-energy (several GeV) protons and heavy ions, again accompanied by x-ray radiation.

- Van Allen radiation belts contain electrons (up to about 10 MeV) and protons (up to 100s MeV) trapped in the geomagnetic field. The particle flux in the regions farther from the Earth can vary wildly depending on the actual conditions of the sun and the magnetosphere. Due to their position they pose a concern for satellites.

- Secondary particles result from interaction of other kinds of radiation with structures around the electronic devices.

- Nuclear reactors produce gamma radiation and neutron radiation which can affect sensor and control circuits in nuclear power plants.

- Nuclear explosions produce a short and extremely intense surge through a wide spectrum of electromagnetic radiation, an electromagnetic pulse (EMP), neutron radiation, and a flux of both primary and secondary charged particles. In case of a nuclear war they pose a potential concern for all civilian and military electronics.

- Chip packaging materials were an insidious source of radiation that was found to be causing soft errors in new DRAM chips in the 1970s. Traces of radioactive elements in the packaging of the chips were producing alpha particles, which were then occasionally discharging some of the capacitors used to store the DRAM data bits. These effects have been reduced today by using purer packaging materials, and employing error-correcting codes to detect and often correct DRAM errors.

Radiation effects on electronics

Fundamental mechanisms

Two fundamental damage mechanisms take place:

- *Lattice displacement*, caused by neutrons, protons, alpha particles, heavy ions, and very high energy gamma photons. They change the arrangement of the atoms in the crystal lattice, creating lasting damage, and increasing the number of recombination centers, depleting the minority carriers and worsening the analog properties of the affected semiconductor junctions. Counterintuitively, higher doses over short time cause partial annealing ("healing") of the damaged lattice, leading to a lower degree of damage than with the same doses delivered in low intensity over a long time. This type of damage is especially important for bipolar transistors, which are dependent on minority carriers in their base regions; increased losses caused by recombination cause loss of the transistor gain.

- *Ionization effects* are caused by charged particles, including the ones with energy too low to cause lattice effects. The ionization effects are usually transient, creating glitches and soft errors, but can lead to destruction of the device if they trigger other damage mechanisms, e.g. a latchup. Photocurrent caused by ultraviolet and x-ray radiation may belong to this category as well. Gradual accumulation of holes in the oxide layer in MOSFET transistors leads to worsening of their performance, up to device failure when the dose is high enough.

The effects can vary wildly depending on all the parameters - the type of radiation, total dose and the radiation flux, combination of types of radiation, and even the kind of the device load (operating frequency, operating voltage, actual state of the transistor during the instant it is struck by the particle), which makes thorough testing difficult, time consuming, and requiring a lot of test samples.

Resultant effects

The "end-user" effects can be characterized in several groups:

- **Neutron effects**: A neutron interacting with the semiconductor lattice will displace its atoms. This leads to an increase in the count of recombination centers and deep-level defects, reducing the lifetime of minority carriers, thus affecting bipolar devices more than CMOS ones. Bipolar devices on silicon tend to show changes in electrical parameters at levels of 10^{10} to 10^{11} neutrons/cm², CMOS devices aren't affected until 10^{15} neutrons/cm². The sensitivity of the devices may increase together with increasing level of integration and decreasing size of individual structures. There is also the risk of induced radioactivity caused by neutron activation, which is a major source of noise in high energy astrophysics instruments. Induced radiation, together with residual radiation from impurities in used materials, can cause all sorts of single-event problems during the device's lifetime. GaAs LEDs, common in optocouplers, are very sensitive to neutrons. The lattice damage influences the frequency of crystal oscillators. Kinetic energy effects (namely lattice displacement) of charged particles belong here too.

- **Total ionizing dose effects**: The cumulative damage of the semiconductor lattice (*lattice displacement* damage) caused by ionizing radiation over the exposition time. It is measured in rads and causes slow gradual degradation of the device's performance; total dose greater than 5000 rads delivered to silicon-based devices in seconds to minutes will cause long-term degradation. In CMOS devices, the radiation creates electron–hole pairs in the gate insulation layers, which cause photocurrents during their recombination, and the holes trapped in the lattice defects in the insulator create a persistent gate biasing and influence the transistors' threshold voltage, making the N-type MOSFET transistors easier and the P-type ones more difficult to switch on. The accumulated charge can be high enough to keep the transistors permanently open (or closed), leading to device failure. Some self-healing takes place over time, but this effect is not too significant. This effect is the same as hot carrier degradation in high-integration high-speed electronics. Crystal oscillators are somewhat sensitive to radiation dose, which alters their frequency; the sensitivity can be greatly reduced by using swept quartz. Natural quartz crystals are especially sensitive.

- **Transient dose effects**: The short-time high-intensity pulse of radiation, typically occurring during a nuclear explosion. The high radiation flux creates photocurrents in the entire body of the semiconductor, causing transistors to randomly open, changing logical states of flip-flops and memory cells. Permanent damage may occur if the duration of the pulse is too long, or if the pulse causes junction damage or causes a latchup. Latchups are commonly caused by the x-rays and gamma radiation flash of a nuclear explosion. Crystal oscillators may stop oscillating for the duration of the flash due to prompt photoconductivity induced in quartz.

- **Systems-generated EMP effects (SGEMP)** are caused by the radiation flash traveling through the equipment and causing local ionization and electric currents in the material of the chips, circuit boards, cables and cases.

- **Single-event effects (SEE)** are phenomena affecting mostly digital devices.

Digital damage: SEE

Single-event effects (SEE), mostly affecting only digital devices, were not studied extensively until relatively recently. When a high-energy particle travels through a semiconductor, it leaves an ionized track behind. This ionization may cause a highly localized effect similar to the transient dose one - a benign glitch in output, a less benign bit flip in memory or a register, or, especially in high-power transistors, a destructive latchup and burnout. Single event effects have importance for electronics in satellites, aircraft, and other both civilian and military aerospace applications. Sometimes in circuits not involving latches it is helpful to introduce RC time constant circuits, slowing down the circuit's reaction time beyond the duration of an SEE.

- **Single-event upsets (SEU)**, or **transient radiation effects in electronics**, are state changes of memory or register bits caused by a single ion interacting with the chip. They do not cause lasting damage to the device, but may cause lasting problems to a system which cannot recover from such an error. In very sensitive devices, a single ion can cause a multiple-bit upset (MBU) in several adjacent memory cells. SEUs can become **Single-event Functional Interrupts** (**SEFI**) when they upset control circuits, such as state machines, placing the device into an undefined state, a test mode, or a halt, which would then need a reset or a power cycle to recover.

- **Single-event latchup (SEL)** can occur in any chip with a parasitic PNPN structure. A heavy ion or a high-energy proton passing through one of the two inner-transistor junctions can turn on the thyristor-like structure, which then stays "shorted" (an effect known as latchup) until the device is power-cycled. As the effect can happen between the power source and substrate, destructively high current can be involved and the part may fail. Bulk CMOS devices are most susceptible.

- **Single-event transient (SET)** happens when the charge collected from an ionization event discharges in the form of a spurious signal traveling through the circuit. This is de facto the effect of an electrostatic discharge.

- **Single-event snapback**, similar to SEL but not requiring the PNPN structure, can be induced in N-channel MOS transistors switching large currents, when an ion hits near the drain junction and causes avalanche multiplication of the charge carriers. The transistor then opens and stays opened.

- **Single-event induced burnout (SEB)** may occur in power MOSFETs when the substrate right under the source region gets forward-biased and the drain-source voltage is higher than the breakdown voltage of the parasitic structures. The resulting high current and local overheating then may destroy the device.

- **Single-event gate rupture (SEGR)** was observed in power MOSFETs when a heavy ion hits the gate region while a high voltage is applied to the gate. A local breakdown then happens in the insulating layer of silicon dioxide, causing local overheat and destruction (looking like a microscopic explosion) of the gate region. It can occur even in EEPROM cells during write or erase, when the cells are subjected to a comparatively high voltage.

SEE Testing

While proton beams are widely used for SEE testing due to availability, at lower energies proton irradiation can often underestimate SEE susceptibility. Furthermore, proton beams expose devices to risk of total ionizing dose (TID) failure which can cloud proton testing results or result in pre-mature device failure. White neutron beams—while ostensibly the most representative SEE test method—are usually derived from solid target-based sources, resulting in flux non-uniformity and small beam areas. White neutron beams also have some measure of uncertainty in their energy spectrum, often with high thermal neutron content.

The disadvantages of both proton and spallation neutron sources can be avoided by using mono-energetic 14 MeV neutrons for SEE testing. A potential concern is that mono-energetic neutron-induced single event effects will not accurately represent the real-world effects of broad-spectrum atmospheric neutrons. However, recent studies have indicated that, to the contrary, mono-energetic neutrons—particularly 14 MeV neutrons—can be used to quite accurately understand SEE cross-sections in modern microelectronics.

A particular study of interest, performed in 2010 by Normand and Dominik, powerfully demonstrates the effectiveness of 14 MeV neutrons.

The first devoted SEE testing Laboratory in Canada is currently being established in Southern Ontario under the name Radiation Effects Laboratories.

Radiation-hardening techniques

- Physical:
    - Hardened chips are often manufactured on insulating substrates instead of the usual semiconductor wafers. Silicon oxide (SOI) and sapphire (SOS) are commonly used. While normal commercial-grade chips can withstand between 50 and 100 gray(5 and 10 krad), space-grade SOI and SOS chips can survive doses many orders of magnitude greater. At one time many 4000 series chips were available in radiation-hardened versions (RadHard).
    - Bipolar integrated circuits generally have higher radiation tolerance than CMOS circuits. The low-power Schottky (LS) 5400 series can withstand 1000 krad, and many ECL devices can withstand 10 000 krad
    - Magnetoresistive RAM, or MRAM, is considered a likely candidate to provide radiation hardened, rewritable, non-volatile conductor memory. Physical principles and early tests suggest that MRAM is not susceptible to ionization induced data loss.
    - Shielding the package against radioactivity, to reduce exposure of the bare device.
    - Capacitor-based DRAM is often replaced by more rugged (but larger, and more expensive) SRAM.
    - Choice of substrate with wide band gap, which gives it higher tolerance to deep-level defects; e.g. silicon carbide or gallium nitride.
    - Shielding the chips themselves by use of depleted boron (consisting only of isotope Boron-11) in the borophosphosilicate glass passivation layer protecting the chips, as boron-10 readily captures neutrons and undergoes alpha decay.
- Logical:
    - Error correcting memory uses additional parity bits to check for and possibly correct corrupted data. Since radiation effects damage the memory content even when the system is not accessing the RAM, a "scrubber" circuit must continuously sweep the RAM; reading out the data, checking the parity for data errors, then writing back any corrections to the RAM.
    - Redundant elements can be used at the system level. Three separate microprocessor boards may independently compute an answer to a

calculation and compare their answers. Any system that produces a minority result will recalculate. Logic may be added such that if repeated errors occur from the same system, that board is shut down.

- Redundant elements may be used at the circuit level. A single bit may be replaced with three bits and separate "voting logic" for each bit to continuously determine its result. This increases area of a chip design by a factor of 5, so must be reserved for smaller designs. But it has the secondary advantage of also being "fail-safe" in real time. In the event of a single-bit failure (which may be unrelated to radiation), the voting logic will continue to produce the correct result without resorting to a watchdog timer. System level voting between three separate processor systems will generally need to use some circuit-level voting logic to perform the votes between the three processor systems.
- A watchdog timer will perform a hard reset of a system unless some sequence is performed that generally indicates the system is alive, such as a write operation from an onboard processor. During normal operation, software schedules a write to the watchdog timer at regular intervals to prevent the timer from running out. If radiation causes the processor to operate incorrectly, it is unlikely the software will work correctly enough to clear the watchdog timer. The watchdog eventually times out and forces a hard reset to the system. This is considered a last resort to other methods of radiation hardening.

Chapter-7

Etching (Microfabrication)

Etching tanks used to perform Piranha, Hydrofluoric acid or RCA clean on 4-inch wafer batches at LAAS technological facility in Toulouse, France.

Etching is used in microfabrication to chemically remove layers from the surface of a wafer during manufacturing. Etching is a critically important process module, and every wafer undergoes many etching steps before it is complete.

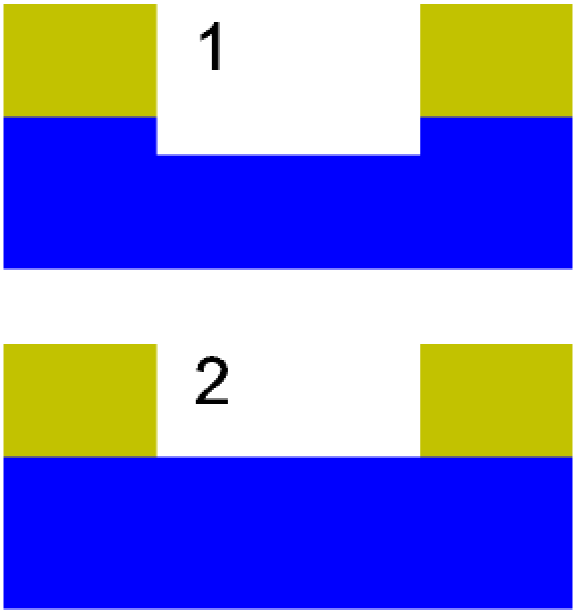
For many etch steps, part of the wafer is protected from the etchant by a "masking" material which resists etching. In some cases, the masking material is a photoresist which has been patterned using photolithography. Other situations require a more durable mask, such as silicon nitride.

Figures of merit

If the etch is intended to make a cavity in a material, the depth of the cavity may be controlled approximately using the etching time and the known etch rate. More often, though, etching must entirely remove the top layer of a multilayer structure, without damaging the underlying or masking layers. The etching system's ability to do this depends on the ratio of etch rates in the two materials (*selectivity*).

Some etches undercut the masking layer and form cavities with sloping sidewalls. The distance of undercutting is called *bias*. Etchants with large bias are called *isotropic*, because they erode the substrate equally in all directions. Modern processes greatly prefer anisotropic etches, because they produce sharp, well-controlled features.

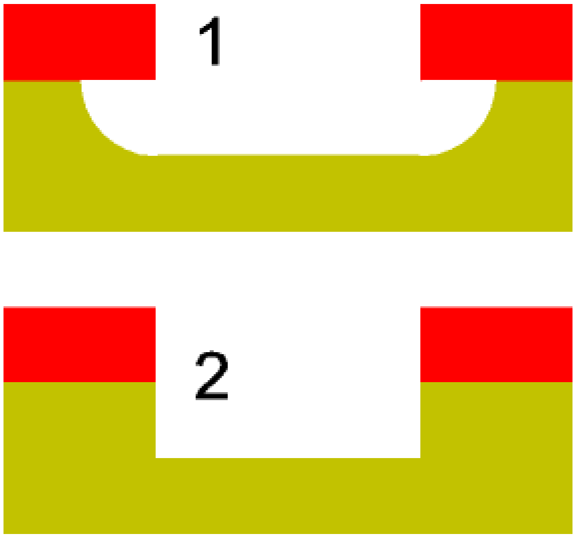
Selectivity



Yellow: layer to be removed; blue: layer to remain

1. A poorly selective etch removes the top layer, but also attacks the underlying material.
2. A highly selective etch leaves the underlying material unharmed.

Isotropy



Red: masking layer; yellow: layer to be removed

1. A perfectly isotropic etch produces round sidewalls.
2. A perfectly anisotropic etch produces vertical sidewalls.

Etching media and technology

The two fundamental types of etchants are liquid-phase ("wet") and plasma-phase ("dry"). Each of these exists in several varieties.

Wet etching

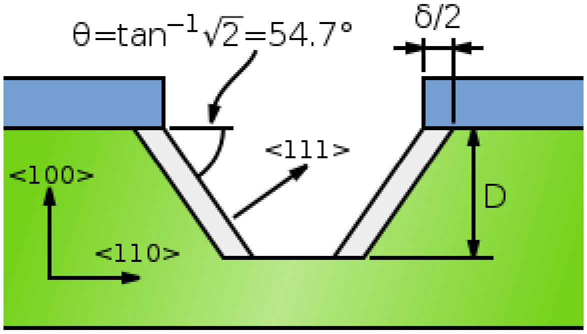
The first etching processes used liquid-phase ("wet") etchants. The wafer can be immersed in a bath of etchant, which must be agitated to achieve good process control. For instance, buffered hydrofluoric acid (BHF) is used commonly to etch silicon dioxide over a silicon substrate.

Different specialised etchants can be used to characterise the surface etched.

Wet etchants are usually isotropic, which leads to large bias when etching thick films. They also require the disposal of large amounts of toxic waste. For these reasons, they are seldom used in state-of-the-art processes. However, the photographic developer used for photoresist resembles wet etching.

As an alternative to immersion, single wafer machines use the Bernoulli principle to employ a gas (usually, pure nitrogen) to cushion and protect one side of the wafer while etchant is applied to the other side. It can be done to either the front side or back side. The etch chemistry is dispensed on the top side when in the machine and the bottom side is not affected. This etch method is particularly effective just before "backend" processing (BEOL), where wafers are normally very much thinner after wafer backgrinding, and very sensitive to thermal or mechanical stress. Etching a thin layer of even a few micrometres will remove microcracks produced during backgrinding resulting in a the wafer having dramatically increased strength and flexibility without breaking.

Anisotropic wet etching



An anisotropic wet etch on a silicon wafer creates a cavity with a trapezoidal cross-section. The bottom of the cavity is a <100> plane, and the sides are <111> planes. The blue material is an etch mask, and the green material is silicon.

Some wet etchants etch crystalline materials at very different rates depending upon which crystal face is exposed. In single-crystal materials (e.g. silicon wafers), this effect can allow very high anisotropy, as shown in the figure.

Several anisotropic wet etchants are available for silicon. For instance, potassium hydroxide (KOH) can achieve selectivity of 400 between <100> and <111> planes. Another option is EDP (an aqueous solution of ethylene diamine and pyrocatechol), which also displays high selectivity for p-type doping. Neither of these etchants may be used on wafers that contain CMOS integrated circuits. Both of them etch aluminum, commonly used as a metallization (wiring) material. KOH introduces mobile potassium ions into silicon dioxide, and EDP is highly corrosive and carcinogenic. Tetramethylammonium hydroxide (TMAH) presents a safer alternative, although it has even worse selectivity between <100> and <111> planes in silicon than does EDP.

Etching a rectangular hole in a (100)-Si wafer will result in a pyramid shaped etch pit. The wall will be flat and angled (as opposed to curved in isotropic etching), and have an angle to the surface of the wafer of:

$$\tan^{-1}\sqrt{2} = 54.7^\circ$$

If the etching is stopped before the pyramid is formed, a frustum will be formed. The undercut, δ, under the resist mask is given by:

$$\delta = \frac{\sqrt{6}D}{S} = \frac{\sqrt{6}R_{100}T}{R_{100}/R_{111}} = \sqrt{6}TR_{111},$$

where R_{xxx} is the etch rate in the <xxx> direction, T is the etch time, D is the etch depth and S is the anisotropy of the material and etchant.

Different etchants have different anisotropies. Below is a table of common anisotropic etchants for silicon:

Etchant	Operating temp (°C)	R_{100} (µm/min)	$S=R_{100}/R_{111}$	Mask materials
Ethylenediamine pyrocatechol (EDP)	110	0.47	17	SiO_2, Si_3N_4, Au, Cr, Ag, Cu
Potassium hydroxide/Isopropyl alcohol (KOH/IPA)	50	1.0	400	Si_3N_4, SiO_2 (etches at 140nm/min)
Tetramethylammonium hydroxide (TMAH)	80	0.6	37	Si_3N_4, SiO_2

Plasma etching

Modern VLSI processes avoid wet etching, and use *plasma etching* instead. Plasma etchers can operate in several modes by adjusting the parameters of the plasma. Ordinary plasma etching operates between 0.1 and 5 Torr. (This unit of pressure, commonly used in vacuum engineering, equals approximately 133.3 pascals.) The plasma produces energetic free radicals, neutrally charged, that react at the surface of the wafer. Since neutral particles attack the wafer from all angles, this process is isotropic.

The source gas for the plasma usually contains small molecules rich in chlorine or fluorine. For instance, carbon tetrachloride (CCl_4) etches silicon and aluminium, and trifluoromethane etches silicon dioxide and silicon nitride. A plasma containing oxygen is used to oxidize ("ash") photoresist and facilitate its removal.

Ion milling, or *sputter etching*, uses lower pressures, often as low as 10^{-4} Torr (10 mPa). It bombards the wafer with energetic ions of noble gases, often Ar^+, which knock atoms from the substrate by transferring momentum. Because the etching is performed by ions, which approach the wafer approximately from one direction, this process is highly anisotropic. On the other hand, it tends to display poor selectivity. *Reactive-ion etching*

(RIE) operates under conditions intermediate between sputter and plasma etching (between 10^{-3} and 10^{-1} Torr). *Deep reactive-ion etching* (DRIE) modifies the RIE technique to produce deep, narrow features.

Common etch processes used in microfabrication

Etchants for common microfabrication materials

Material to be etched	Wet etchants	Plasma etchants
Aluminium (Al)	80% phosphoric acid (H_3PO_4) + 5% acetic acid + 5% nitric acid (HNO_3) + 10% water (H_2O) at 35–45 °C	Cl_2, CCl_4, $SiCl_4$, BCl_3
Indium tin oxide [ITO] (In_2O_3:SnO_2)	Hydrochloric acid (HCl) + nitric acid (HNO_3) + water (H_2O) (1:0.1:1) at 40 °C	
Chromium (Cr)	- "Chrome etch": ceric ammonium nitrate ($(NH_4)_2Ce(NO_3)_6$) + nitric acid (HNO_3) - Hydrochloric acid (HCl)	
Gold (Au)	Aqua regia	
Molybdenum (Mo)		CF_4
Organic residues and photoresist	Piranha etch: sulfuric acid (H_2SO_4) + hydrogen peroxide (H_2O_2)	O_2 (ashing)
Platinum (Pt)	Aqua regia	
Silicon (Si)	Nitric acid (HNO_3) + hydrofluoric acid (HF)	- CF_4, SF_6, NF_3 - Cl_2, CCl_2F_2
Silicon dioxide (SiO_2)	- Hydrofluoric acid (HF) - Buffered oxide etch [BOE]: ammonium fluoride (NH_4F) and hydrofluoric acid (HF)	CF_4, SF_6, NF_3
Silicon nitride (Si_3N_4)	- 85% Phosphoric acid (H_3PO_4) at 180 °C (Requires SiO_2 etch mask)	CF_4, SF_6, NF_3

Tantalum (Ta)		CF_4
Titanium (Ti)	Hydrofluoric acid (HF)	BCl_3
Titanium nitride (TiN)	- Nitric acid (HNO_3) + hydrofluoric acid (HF) - SC1	
Tungsten (W)	- Nitric acid (HNO_3) + hydrofluoric acid (HF) - Hydrogen Peroxide (H_2O_2)	- CF_4 - SF_6

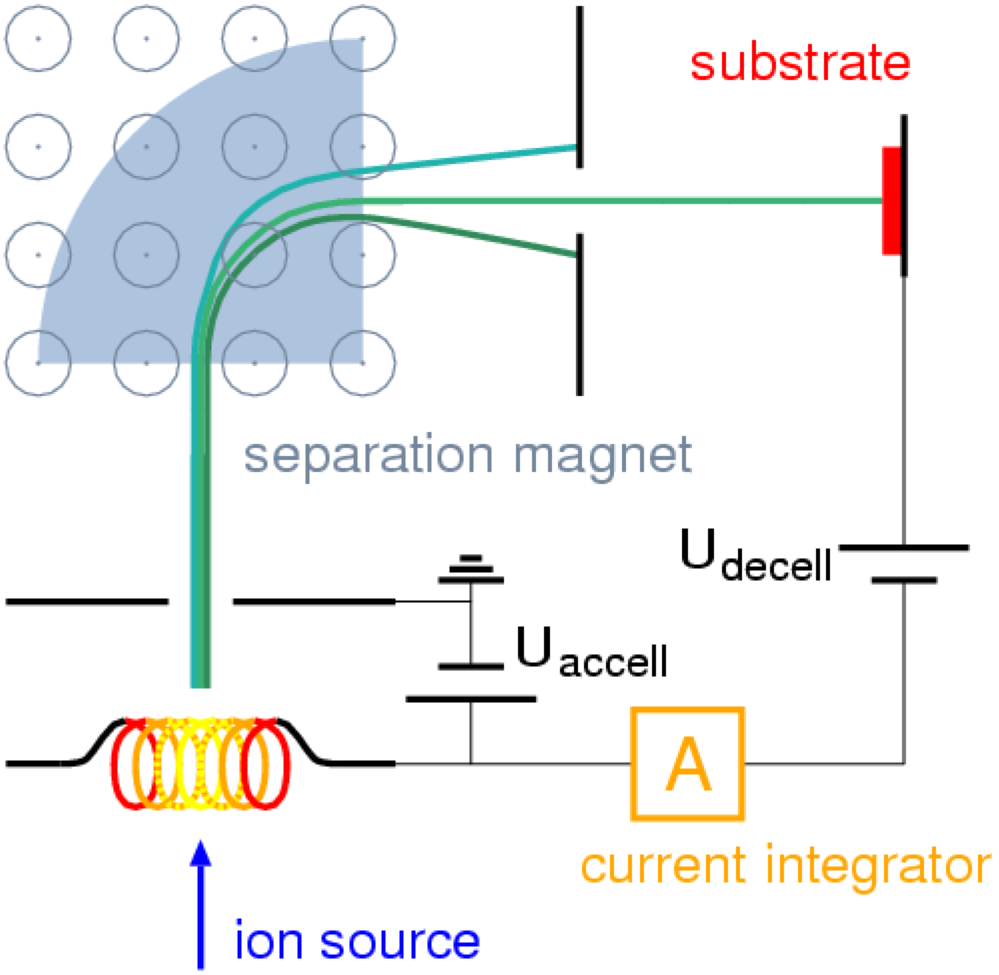
Chapter-8

Ion Implantation

An ion implantation system at LAAS technological facility in Toulouse, France.

Ion implantation is a materials engineering process by which ions of a material can be implanted into another solid, thereby changing the physical properties of the solid. Ion implantation is used in semiconductor device fabrication and in metal finishing, as well as various applications in materials science research. The ions introduce both a chemical change in the target, in that they can be a different element than the target or induce a nuclear transmutation, and a structural change, in that the crystal structure of the target can be damaged or even destroyed by the energetic collision cascades.

General principle



Ion implantation setup with mass separator

Ion implantation equipment typically consists of an ion source, where ions of the desired element are produced, an accelerator, where the ions are electrostatically accelerated to a high energy, and a target chamber, where the ions impinge on a target, which is the material to be implanted. Thus ion implantation is a special case of particle radiation. Each ion is typically a single atom or molecule, and thus the actual amount of material implanted in the target is the integral over time of the ion current. This amount is called the dose. The currents supplied by implanters are typically small (microamperes), and thus the dose which can be implanted in a reasonable amount of time is small. Therefore, ion implantation finds application in cases where the amount of chemical change required is small.

Typical ion energies are in the range of 10 to 500 keV (1,600 to 80,000 aJ). Energies in the range 1 to 10 keV (160 to 1,600 aJ) can be used, but result in a penetration of only a few nanometers or less. Energies lower than this result in very little damage to the target, and fall under the designation ion beam deposition. Higher energies can also be used: accelerators capable of 5 MeV (800,000 aJ) are common. However, there is often great structural damage to the target, and because the depth distribution is broad, the net composition change at any point in the target will be small.

The energy of the ions, as well as the ion species and the composition of the target determine the depth of penetration of the ions in the solid: A monoenergetic ion beam will generally have a broad depth distribution. The average penetration depth is called the range of the ions. Under typical circumstances ion ranges will be between 10 nanometers and 1 micrometer. Thus, ion implantation is especially useful in cases where the chemical or structural change is desired to be near the surface of the target. Ions gradually lose their energy as they travel through the solid, both from occasional collisions with target atoms (which cause abrupt energy transfers) and from a mild drag from overlap of electron orbitals, which is a continuous process. The loss of ion energy in the target is called stopping and can be simulated with the binary collision approximation method.

Application in semiconductor device fabrication

Doping

The introduction of dopants in a semiconductor is the most common application of ion implantation. Dopant ions such as boron, phosphorus or arsenic are generally created from a gas source, so that the purity of the source can be very high. These gases tend to be very hazardous. When implanted in a semiconductor, each dopant atom can create a charge carrier in the semiconductor after annealing. A hole can be created for a p-type dopant, and an electron for an n-type dopant. This modifies the conductivity of the semiconductor in its vicinity. The technique is used, for example, for adjusting the threshold of a MOSFET.

Ion implantation was developed as a method of producing the p-n junction of photovoltaic devices in the late 1970s and early 1980s, along with the use of pulsed-electron beam for rapid annealing, although it has not to date been used for commercial production.

Silicon on insulator

One prominent method for preparing silicon on insulator (SOI) substrates from conventional silicon substrates is the *SIMOX* (**S**eparation by **IM**plantation of **OX**ygen) process, wherein a buried high dose oxygen implant is converted to silicon oxide by a high temperature annealing process.

Mesotaxy

Mesotaxy is the term for the growth of a crystallographically matching phase underneath the surface of the host crystal (compare to epitaxy, which is the growth of the matching phase on the surface of a substrate). In this process, ions are implanted at a high enough energy and dose into a material to create a layer of a second phase, and the temperature is controlled so that the crystal structure of the target is not destroyed. The crystal orientation of the layer can be engineered to match that of the target, even though the exact crystal structure and lattice constant may be very different. For example, after the implantation of nickel ions into a silicon wafer, a layer of nickel silicide can be grown in which the crystal orientation of the silicide matches that of the silicon.

Application in metal finishing

Tool steel toughening

Nitrogen or other ions can be implanted into a tool steel target (drill bits, for example). The structural change caused by the implantation produces a surface compression in the steel, which prevents crack propagation and thus makes the material more resistant to fracture. The chemical change can also make the tool more resistant to corrosion.

Surface finishing

In some applications, for example prosthetic devices such as artificial joints, it is desired to have surfaces very resistant to both chemical corrosion and wear due to friction. Ion implantation is used in such cases to engineer the surfaces of such devices for more reliable performance. As in the case of tool steels, the surface modification caused by ion implantation includes both a surface compression which prevents crack propagation and an alloying of the surface to make it more chemically resistant to corrosion.

Other applications

Ion beam mixing

Ion implantation can be used to achieve ion beam mixing, i.e. mixing up atoms of different elements at an interface. This may be useful for achieving graded interfaces or strengthening adhesion between layers of immiscible materials.

Problems with ion implantation

Crystallographic damage

Each individual ion produces many point defects in the target crystal on impact such as vacancies and interstitials. Vacancies are crystal lattice points unoccupied by an atom: in this case the ion collides with a target atom, resulting in transfer of a significant amount of energy to the target atom such that it leaves its crystal site. This target atom then itself becomes a projectile in the solid, and can cause successive collision events. Interstitials result when such atoms (or the original ion itself) come to rest in the solid, but find no vacant space in the lattice to reside. These point defects can migrate and cluster with each other, resulting in dislocation loops and other defects.

Damage recovery

Because ion implantation causes damage to the crystal structure of the target which is often unwanted, ion implantation processing is often followed by a thermal annealing. This can be referred to as damage recovery.

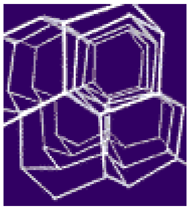
Amorphization

The amount of crystallographic damage can be enough to completely amorphize the surface of the target: i.e. it can become an amorphous solid (such a solid produced from a melt is called a glass). In some cases, complete amorphization of a target is preferable to a highly defective crystal: An amorphized film can be regrown at a lower temperature than required to anneal a highly damaged crystal.

Sputtering

Some of the collision events result in atoms being ejected (sputtered) from the surface, and thus ion implantation will slowly etch away a surface. The effect is only appreciable for very large doses.

Ion channelling



A diamond cubic crystal viewed from the <110> direction, showing hexagonal ion channels.

If there is a crystallographic structure to the target, and especially in semiconductor substrates where the crystal structure is more open, particular crystallographic directions offer much lower stopping than other directions. The result is that the range of an ion can be much longer if the ion travels exactly along a particular direction, for example the <110> direction in silicon and other diamond cubic materials. This effect is called *ion channelling*, and, like all the channelling effects, is highly nonlinear, with small variations from perfect orientation resulting in extreme differences in implantation depth. For this reason, most implantation is carried out a few degrees off-axis, where tiny alignment errors will have more predictable effects.

Ion channelling can be used directly in Rutherford backscattering and related techniques as an analytical method to determine the amount and depth profile of damage in crystalline thin film materials.

Hazardous Materials Note

In the ion implantation semiconductor fabrication process of wafers, it is important for the workers to minimize their exposure to the toxic materials used in the ion implanter process. Such hazardous elements, solid source and gasses are used, such as Arsine and Phosphine. For this reason, the semiconductor fabrication facilities are highly automated, and may feature negative pressure gas bottles safe delivery system (SDS). Other elements may include Antimony, Arsenic, Phosphorus, and Boron. Residue of these elements show up when the machine is opened to atmosphere, and can also be accumulated and found concentrated in the vacuum pumps hardware. It is important not to expose yourself to these carcinogenic, corrosive, flammable, and toxic elements. Many overlapping safety protocols must be used when handling these deadly compounds. Use safety, and read MSDS's.

High Voltage Safety

High voltage power supplies in ion implantation equipment can pose a risk of electrocution. In addition, high-energy atomic collisions can, in some cases, generate radionuclides. Operators and Maintenance personnel should learn and follow the safety advice of the manufacturer and/or the institution responsible for the equipment. Prior to entry to high voltage area, terminal components must be grounded using a grounding stick. Next, power supplies should be locked in the off state and tagged to prevent unauthorized energizing.

Chapter-9

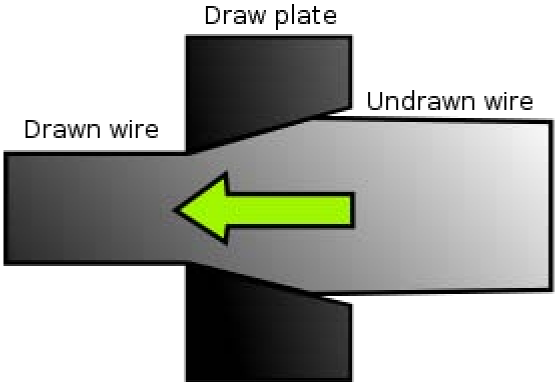
Wire Drawing

Drawing silver wire by hand pulling.

Drawing thicker silver wire by cranked pulling.

Wire drawing is a metalworking process used to reduce the cross-section of a wire by pulling the wire through a single, or series of, drawing die(s). There are many applications for wire drawing, including electrical wiring, cables, tension-loaded structural components, springs, paper clips, spokes for wheels, and stringed musical instruments. Although similar in process, drawing is different from extrusion, because in drawing the wire is pulled, rather than pushed, through the die. Drawing is usually performed at room temperature, thus classified a cold working process, but it may be performed at elevated temperatures for large wires to reduce forces. More recently drawing has been used with molten glass to produce high quality optical fibers.

Process



Wire drawing concept

The wire drawing process is quite simple in concept. The wire is prepared by shrinking the beginning of it, by hammering, filing, rolling or swaging, so that it will fit through the die; the wire is then pulled through the die. As the wire is pulled through the die, its volume remains the same, so as the diameter decreases, the length increases. Usually the wire will require more than one draw, through successively smaller dies, to reach the desired size. The American wire gauge scale is based on this. This can be done on a small scale with a draw plate, or on a large commercial scale using automated machinery. The process of wire drawing improves material properties due to cold working.

The areal reduction of small wires are 15–25% and larger wires are 20–45%. Very fine wires are usually drawn in bundles. In a bundle, the wires are separated by a metal with similar properties, but with lower chemical resistance so that it can be removed after drawing. If the reduction in diameter is greater than 50%, the process may require annealing between the process of drawing the wire through the dies. Commercial wire drawing usually starts with a coil of hot rolled 9 mm (0.35 in) diameter wire. The surface is first treated to remove scales. It is then fed into either a single block or continuous wire drawing machine.

Single block wire drawing machines include means for holding the dies accurately in position and for drawing the wire steadily through the holes. The usual design consists of a cast-iron bench or table having a bracket standing up to hold the die, and a vertical drum which rotates and by coiling the wire around its surface pulls it through the die, the coil of wire being stored upon another drum or "swift" which lies behind the die and reels off the wire as fast as required. The wire drum or "block" is provided with means for rapidly coupling or uncoupling it to its vertical shaft, so that the motion of the wire may be stopped or started instantly. The block is also tapered, so that the coil of wire may be easily slipped off upwards when finished. Before the wire can be attached to the block, a sufficient length of it must be pulled through the die; this is effected by a pair of gripping pincers on the end of a chain which is wound around a revolving drum, so drawing the wire until enough can be coiled two or three times on the block, where the end is secured by a small screw clamp or vice. When the wire is on the block, it is set in motion and the wire is drawn steadily through the die; it is very important that the block rotates evenly and that it runs true and pulls the wire at a constant velocity, otherwise "snatching" occurs which will weaken or even break the wire. The speeds at which wire is drawn vary greatly, according to the material and the amount of reduction.

Continuous wire drawing machines differ from the single block machines in having a series of dies through which the wire passes in a continuous manner. The difficulty of feeding between each die is solved by introducing a block between each die. The speeds of the blocks are increased successively, so that the elongation is taken up and any slip compensated for. One of these machines may contain 3 to 12 dies. The operation of threading the wire through all the dies and around the blocks is termed "stringing-up". The arrangements for lubrication include a pump which floods the dies, and in many cases also the bottom portions of the blocks run in lubricant.

Often intermediate anneals are required to counter the effects of cold working, and to allow more further drawing. A final anneal may also be used on the finished product to maximize ductility and electrical conductivity.

An example of product produced in a continuous wire drawing machine is telephone wire. It is drawn 20 to 30 times from hot rolled rod stock.

While round cross-sections dominate most drawing processes, non-circular cross-sections are drawn. They are usually drawn when the cross-section is small and quantities are too low to justify rolling. In these processes, a block or Turk's-head machine are used.

Lubrication

Lubrication in the drawing process is essential for maintaining good surface finish and long die life. The following are different methods of lubrication:

- Wet drawing: the dies and wire or rod are completely immersed in lubricant
- Dry drawing: the wire or rod passes through a container of lubricant which coats the surface of the wire or rod
- Metal coating: the wire or rod is coated with a soft metal which acts as a solid lubricant
- Ultrasonic vibration: the dies and mandrels are vibrated, which helps to reduce forces and allow larger reductions per pass

Various lubricants, such as oil, are employed. Another lubrication method is to immerse the wire in a copper (II) sulfate solution, such that a film of copper is deposited which forms a kind of lubricant. In some classes of wire the copper is left after the final drawing to serve as a preventive of rust or to allow easy soldering.

Mechanical Properties

The strength-enhancing effect of wire drawing can be substantial. The highest tensile strengths available on any steel have been recorded on small-diameter cold-drawn austenitic stainless wire. Tensile strength can be as high as 400 ksi (3760 MPa).

Drawing dies

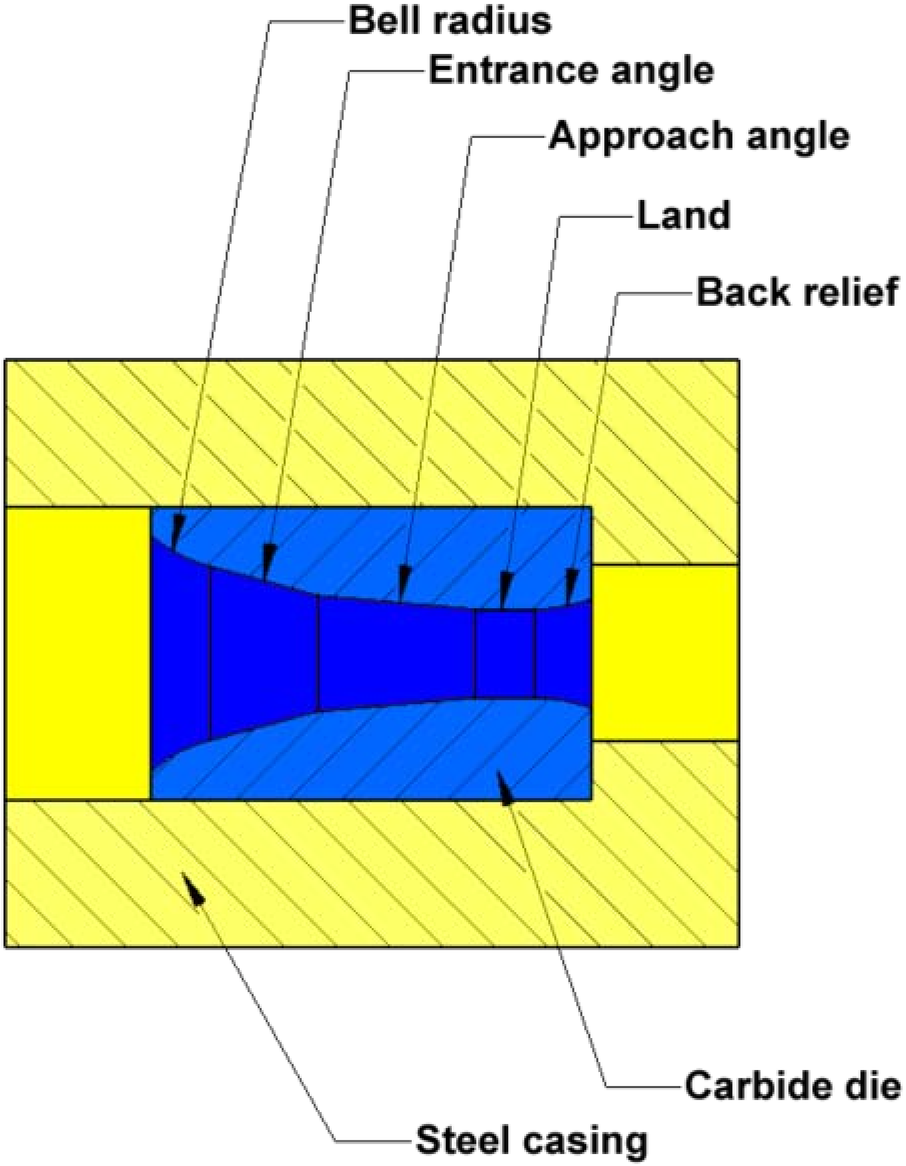


Diagram of a carbide wire drawing die

Drawing dies are typically made of tool steel, tungsten carbide, or diamond, with tungsten carbide and manufactured diamond being the most common. Synthetic diamond is usually used in the early stages of the drawing process, whereas natural diamond dies are used in the final stages. For drawing very fine wire a single crystal diamond die is used. For hot drawing, cast-steel dies are used. For steel wire drawing, a tungsten carbide die is used. The dies are placed in a steel casing, which backs the die and allow for easy die changes. Die angles usually range from 6–15°, and each die has at least 2 different angles: the entering angle and approach angle. Wire dies usually are used with power as to pull the wire through them. There are coils of wire on either end of the die which pull and roll up the wire with a reduced diameter.

Chapter-10

Thermal Oxidation & Thermal Management of High-Power LEDs

Thermal oxidation

Furnaces used for diffusion and thermal oxidation at LAAS technological facility in Toulouse, France.

In microfabrication, **thermal oxidation** is a way to produce a thin layer of oxide (usually silicon dioxide) on the surface of a wafer (semiconductor). The technique forces an oxidizing agent to diffuse into the wafer at high temperature and react with it. The rate of oxide growth is often predicted by the Deal-Grove model.

The chemical reaction

Thermal oxidation of silicon is usually performed at a temperature between 800 and 1200°C, resulting in so called **High Temperature Oxide** layer (HTO). It may use either water vapor (steam) or molecular oxygen as the oxidant; it is consequently called either *wet* or *dry* oxidation. The reaction is one of the following:

$$\mathrm{Si} + 2\mathrm{H_2O} \rightarrow \mathrm{SiO_2} + 2\mathrm{H_2}\ _{(g)}$$
$$\mathrm{Si} + \mathrm{O_2} \rightarrow \mathrm{SiO_2}$$

The oxidizing ambient may also contain several percent of hydrochloric acid (HCl). The chlorine removes metal ions that may occur in the oxide.

Thermal oxide incorporates silicon consumed from the substrate and oxygen supplied from the ambient. Thus, it grows both down into the wafer and up out of it. For every unit thickness of silicon consumed, 2.27 unit thicknesses of oxide will appear . Conversely, if a bare silicon surface is oxidized, 46% of the oxide thickness will lie below the original surface, and 54% above it.

Deal-Grove model

According to the commonly-used Deal-Grove model, the time *t* required to grow an oxide of thickness X_o, at a constant temperature, on a bare silicon surface, is:

$$t = \frac{X_o^2}{B} + \frac{X_o}{B/A}$$

where the constants A and B encapsulate the properties of the reaction and the oxide layer, respectively.

If a wafer that already contains oxide is placed in an oxidizing ambient, this equation must be modified by adding a corrective term τ, the time that would have been required to grow the pre-existing oxide under current conditions. This term may be found using the equation for *t* above.

Solving the quadratic equation for X_o yields:

$$X_o(t) = A/2 \cdot \left[\sqrt{1 + \frac{4B}{A^2}(t+\tau)} - 1\right]$$

Oxidation technology

Most thermal oxidation is performed in furnaces, at temperatures between 800 and 1200°C. A single furnace accepts many wafers at the same time, in a specially designed quartz rack (called a "boat"). Historically, the boat entered the oxidation chamber from the side (this design is called "horizontal"), and held the wafers vertically, beside each other. However, many modern designs hold the wafers horizontally, above and below each other, and load them into the oxidation chamber from below.

Vertical furnaces stand higher than horizontal furnaces, so they may not fit into some microfabrication facilities. However, they help to prevent dust contamination. Unlike horizontal furnaces, in which falling dust can contaminate any wafer, vertical furnaces only allow it to fall on the top wafer in the boat.

Vertical furnaces also eliminate an issue that plagued horizontal furnaces: non-uniformity of grown oxide across the wafer. Horizontal furnaces typically have convection currents inside the tube which causes the bottom of the tube to be slightly colder than the top of the tube. As the wafers lie vertically in the tube the convection and the temperature gradient with it causes the top of the wafer to have a thicker oxide than the bottom of the wafer. Vertical furnaces solve this problem by having wafer sitting horizontally, and then having the gas flow in the furnace flowing from top to bottom, significantly dampening any thermal convections.

Vertical furnaces also allow the use of load locks to purge the wafers with nitrogen before oxidation to limit the growth of native oxide on the Si surface.

Oxide quality

Wet oxidation is preferred to dry oxidation for growing thick oxides, because of the higher growth rate. However, fast oxidation leaves more dangling bonds at the silicon interface, which produce quantum states for electrons and allow current to leak along the interface. (This is called a "dirty" interface.) Wet oxidation also yields a lower-density oxide, with lower dielectric strength.

The long time required to grow a thick oxide in dry oxidation makes this process impractical. Thick oxides are usually grown with a long wet oxidation bracketed by short dry ones (a *dry-wet-dry* cycle). The beginning and ending dry oxidations produce films of high-quality oxide at the outer and inner surfaces of the oxide layer, respectively.

Mobile metal ions can degrade performance of MOSFETs (sodium is of particular concern). However, chlorine can immobilize sodium by forming sodium chloride. Chlorine is often introduced by adding hydrogen chloride or trichloroethylene to the oxidizing medium. Its presence also increases the rate of oxidation.

Other notes

- Thermal oxidation can be performed on selected areas of a wafer, and blocked on others. Areas which are not to be oxidized are covered with a film of silicon nitride, which blocks diffusion of oxygen and water vapor. The nitride is removed after oxidation is complete. This process cannot produce sharp features, because lateral (parallel to the surface) diffusion of oxidant molecules under the nitride mask causes the oxide to protrude into the masked area.

- Because impurities dissolve differently in silicon and oxide, a growing oxide will selectively take up or reject dopants. This redistribution is governed by the segregation coefficient, which determines how strongly the oxide absorbs or rejects the dopant, and the diffusivity.

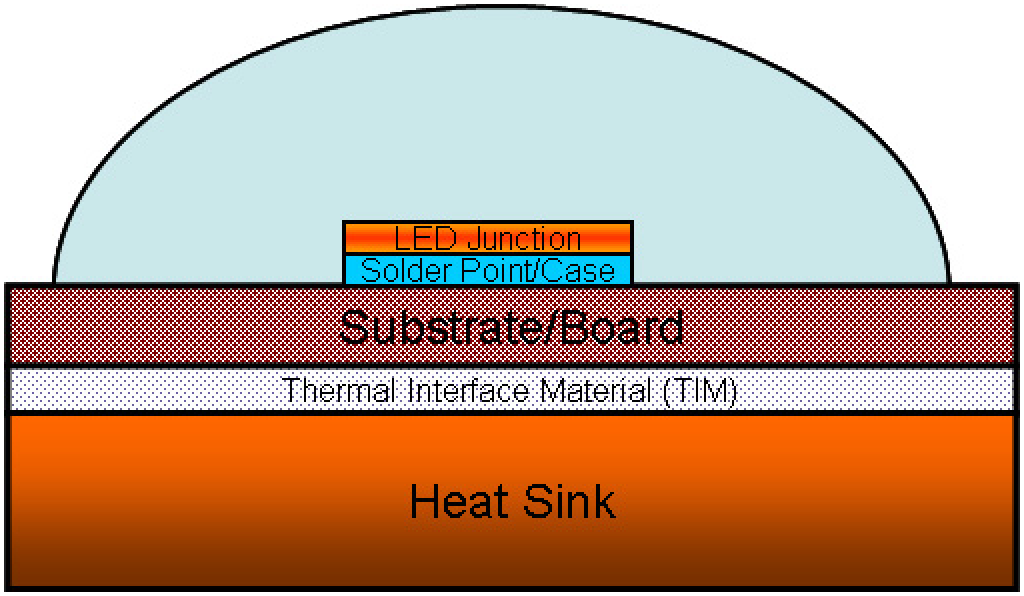
- The orientation of the silicon crystal affects oxidation. A <100> wafer oxidizes more slowly than a <111> wafer, but produces an electrically cleaner oxide interface.

- Thermal oxidation of any variety produces a higher-quality oxide, with a much cleaner interface, than chemical vapor deposition of oxide resulting in **Low Temperature Oxide** layer (reaction of TEOS at about 600 °C). However, the high temperatures required to produce High Temperature Oxide (HTO) restrict its usability. For instance, in MOSFET processes, thermal oxidation is never performed after the doping for the source and drain terminals is performed, because it would disturb the placement of the dopants.

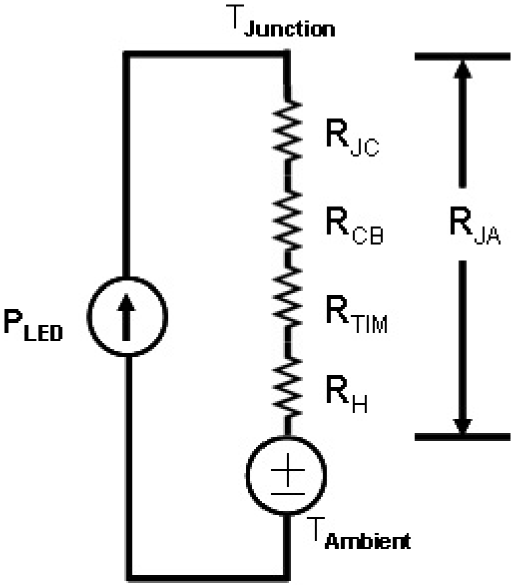
Thermal management of high-power LEDs

High power light-emitting diodes (LEDs) are likely to replace other technologies such as incandescent and fluorescent bulbs in signaling, solid state lighting, and vehicle headlights because they save energy and extend the light's lifetime. LEDs that use from 500 milliwatts to as much as 10 watts in a single package have become standard, and researchers expect to use even more power in the future. Some of the electricity in an LED becomes heat rather than light. If that heat is not removed, the LEDs run at high temperatures, which not only lowers their efficiency, but also makes the LED more dangerous and less reliable. Thus, thermal management of high power LEDs is a crucial area of research and development.

Heat transfer procedure



Typical LED package including thermal management design



Typical thermal model of LED package. LED power dissipation is modeled as a current source; thermal resistance is modeled as a resistor; and the ambient temperature is modeled as a voltage source.

In order to maintain a low junction temperature to keep good performance of an LED, every method of releasing heat from LEDs should be considered. Conduction, convection, and radiation are the three means of heat transfer. Typically, LEDs are encapsulated in a transparent resin, which is a poor thermal conductor. Nearly all heat produced is conducted through the back side of the chip. Heat is generated from the PN junction by electrical energy that was not converted to useful light, and conducted to outside ambience through a long and extensive path, from junction to solder point, solder point to board, and board to the heat sink and then to the atmosphere. The heat path of tungsten light bulbs is almost all straight into the atmosphere, starting from filament to the glass and ending with the thermal resistance from glass to the atmosphere. A typical LED side view and its thermal model are shown in the figures.

The thermal resistance between two points is defined as the ratio of the difference in temperature to the power dissipated; the unit is ^{o}C/W. From the LED junction to the thermal contact at the bottom of package, the thermal resistance is governed by the package design. It is referred to as the thermal resistance between junction and ambient (R_{JA}). Different components in the heat conduction path can be modeled as different thermal resistances. The total power dissipated by the LED (P_{LED}) is the product of the forward voltage and the forward current of the LED, which can be modeled as a current source. The ambient temperature is modeled as a voltage source. Therefore, the junction temperature (T_J) is the sum of the ambient temperature (T_A) and the product of the thermal resistance from junction to ambient and the power dissipated. By "thermic Ohm's Law", we have the equation as follows: $T_J = T_A + (R_{JA} \times P_{LED})$, and $R_{JA} = R_{JC} + R_{CB} + R_{TIM} + R_H$

Intuitively, you can see that the junction temperature will be lower if the thermal impedance is smaller and likewise, with a lower ambient temperature. To maximize the useful ambient temperature range for a given power dissipation, the total thermal resistance from junction to ambient must be minimized. The values for the thermal resistance vary widely depending on the material or component supplier. For example, R_{JC} will range from 2.6^{o}C/W to 18^{o}C/W, depending on the LED manufacturer. The thermal interface material's (TIM) thermal resistance will also vary depending on the type of material selected. Common TIMs are epoxy, thermal grease, pressure sensitive adhesive and solder. In the most cases, power LEDs will be mounted on metal-core printed circuit boards (MCPCB), which will be attached to a heat sink. Heat flows from the LED junction through the MCPCB to the heat sink by way of conduction, and the heat sink diffuses heat to the ambient surroundings by convection. So, we can also add convection to the thermal model at the end of the heat transmission path. In the package design, the surface flatness and quality of each component, applied mounting pressure, contact area, the type of interface material and its thickness are all important parameters to thermal resistance design.

Passive thermal designs

Some considerations for passive thermal designs to ensure good thermal management for high power LED operation include:

Adhesive

Adhesive is commonly used to bond LED and board, and board and heat sinks. Using a thermal conductive adhesive can further optimize the thermal performance.

Heat sink

Heat sinks provide a path for heat from the LED source to outside medium. Heat sinks can dissipate power in three ways: conduction (heat transfer from one solid to another), convection (heat transfer from a solid to a moving fluid, for most LED applications the fluid will be air), or radiation (heat transfer from two bodies of different surface temperatures through electromagnetic waves).

- **Material** – The thermal conductivity of the material that the heat sink is made from directly affects the dissipation efficiency through conduction. Normally this is aluminum, although copper may be used with an advantage for flat-sheet heat sinks.
- **Shape** - Thermal transfer takes place at the surface of the heat sink. Therefore, heat sinks should be designed to have a large surface area. This goal can be reached by using a large number of fine fins or by increasing the size of the heat sink itself.
- **Surface Finish** - Thermal radiation of heat sinks is a function of surface finish, especially at higher temperatures. A painted surface will have a greater emissivity than a bright, unpainted one. The effect is most remarkable with flat-plate heat sinks, where about one-third of the heat is dissipated by radiation. Moreover, a perfectly flat contact area allows the use of a thinner layer of thermal compound, which will reduce the thermal resistance between the heat sink and LED source. On the other hand, anodizing or etching will also decrease the thermal resistance.
- **Mounting method**- Heat-sink mountings with screws or springs are often better than regular clips, thermal conductive glue or sticky tape.

PCB (Printed Circuit Board)

- **MCPCB** - MCPCB (Metal Core PCB) are those boards which incorporate a base metal material as heat spreader as an integral part of the circuit board. The metal core usually consists of aluminum alloy. Furthermore MCPCB can take advantage of incorporating a dielectric polymer layer with high thermal conductivity for lower thermal resistance.
- **Separation** - Separating the LED drive circuitry from the LED board prevents the heat generated by the driver from raising the LED junction temperature.

Package type

- **Flip chip** - The concept is similar to flip-chip in package configuration widely used in the silicon integrated circuit industry. Briefly speaking, the LED die is assembled face down on the sub-mount, which is usually silicon or ceramic, acting as the heat spreader and supporting substrate. The flip-chip joint can be eutectic, high-lead, lead-free solder or gold stub. The primary source of light comes from the back side of the LED chip, and there is usually a built-in reflective layer between the light emitter and the solder joints to reflect the light emitted downwards up. Several companies have adopted flip-chip packages for their high-power LED, achieving bout 60% reduction in the thermal resistance of the LED while keeping its thermal reliability.

Chapter-11

AAL1gator & Silicon on Insulator

AAL1gator

The **AAL1gator** is a semiconductor device that implements the Circuit Emulation Service. It was developed between 1994 and 1998 and became a run-away success. It also played a role in the acquisition of four companies. The name was based on the fact that the AAL1gator implements the ATM AAL-1 standard.

Development of the AAL1gator

The AAL1gator was developed by Network Synthesis, Inc. under contract from Integrated Telecom Technology (IgT). It was the first semiconductor solution to implement the Circuit Emulation Service standard from the ATM Forum. It implemented 8 DS1/E1 lines worth of CES and had 256 channels. It flexibly converted the PDH DS1 signal into Asynchronous Transfer Mode cells.

The AAL1gator was principally designed by the Network Synthesis CEO, Brian Holden and a consultant, Ed Lennox. Brian Holden was also involved in the ATM Forum standardization effort for the Circuit Emulation Service. Additional design efforts came from Andy Annudurai, Ravi Sajwan, and Imran Chaudhri (who also came up with the name). Chee Hu did most of the work on getting the "C" version to work at speed and to be manufacturable. Denis Smetana did most of the work on the "D" version and on the later 32 DS1 version. Jim Jacobson of OnStream Networks was the Beta Customer.

Patents on the AAL1gator

Two U.S. patents were issued on the AAL1gator's calendar-based transmit scheduler, one on the original product and an even better one on the "D" version enhancements designed by Denis Smetana . The scheduler implemented several intricate methods of minimizing the jitter caused by the scheduling of the 256 channels. The AAL1gator also could have gotten another patent on its method of queuing the SRTS samples, but the designers were too busy to get the application in.

Functions of the AAL1gator

The AAL1gator could flexibly map individual DS0s or groups of DS0s into 256 ATM VCs. It also had a high speed mode which mapped a single DS-3 into ATM. Additionally, it had a high performance implementation of the SRTS clock recovery algorithm. The original AAL1gator also was delivered along with the code for an external digital frequency synthesizer. A later version incorporated that synthesizer.

The AAL1gator has been used in several applications that were completely different than the application it was designed for. The designers knew they had a hit product when reports of these uses came in. One use was to provide fractional T1 service over microwave radio. Another use was to move DS1's around within a box.

The AAL1gator and acquisitions

The AAL1gator played a significant role in the following four technology acquisitions:

- Network Synthesis by Integrated Telecom Technology in June 1996
- Integrated Telecom Technology by PMC-Sierra in April 1998
- OnStream Networks by 3Com in 1996
- Sentient Networks by Cisco Systems in 1998

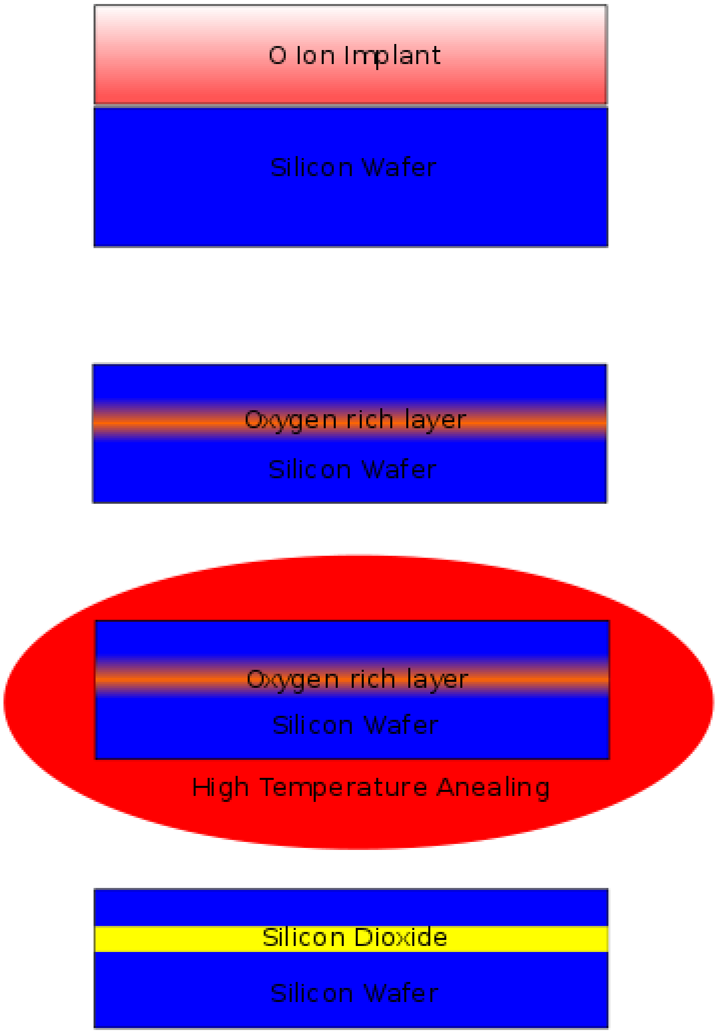
The IgT acquisition happened after the development was showing signs of success. The PMC-Sierra acquisition happened after IgT's AAL1gator began to drive the sales of T1/E1 framers from PMC-Sierra. The Onstream and Sentient acquisitions happened after successful Circuit Emulation Service (CES) product developments based on the AAL1gator.

An anecdote from the AAL1gator's development is that Brian Holden and Ed Lennox made a quick, off-handed decision in a hallway to add the "PMC Mode" to the "B" version of the device in mid-1996 to provide a glueless interface to PMC-Sierra's TQUAD and EQUAD T1 and E1 framers. Little did they know then that this off-handed decision would be a key enabler of PMC-Sierra's acquisition of the company two years later.

The AAL1gator-32

PMC-Sierra developed a 32 line version of the AAL1gator known as the AAL1gator32.

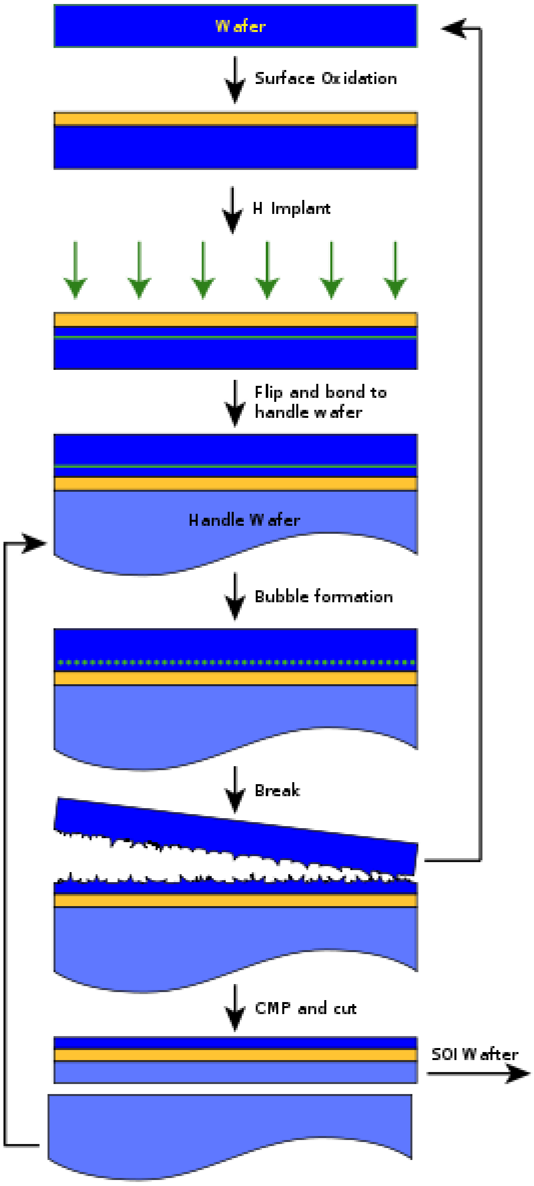
Silicon on insulator



SIMOX process

Silicon on insulator technology (**SOI**) refers to the use of a layered silicon-insulator-silicon substrate in place of conventional silicon substrates in semiconductor manufacturing, especially microelectronics, to reduce parasitic device capacitance and thereby improving performance. SOI-based devices differ from conventional silicon-built devices in that the silicon junction is above an electrical insulator, typically silicon dioxide or (less commonly) sapphire. (These types of devices are called silicon on sapphire, or SOS). The choice of insulator depends largely on intended application, with sapphire being used for radiation-sensitive applications and silicon dioxide preferred for improved performance and diminished short channel effects in microelectronics devices . The insulating layer and topmost silicon layer also vary widely with application . The first industrial implementation of SOI was announced by IBM in August 1998.

Industry need



Smart Cut process

The implementation of SOI technology is one of several manufacturing strategies employed to allow the continued miniaturization of microelectronic devices, colloquially referred to as extending Moore's Law. Reported benefits of SOI technology relative to conventional silicon (bulk CMOS) processing include :

- *Lower parasitic capacitance* due to isolation from the bulk silicon, which improves power consumption at matched performance.
- *Resistance to latchup* due to complete isolation of the n- and p-well structures.

From a manufacturing perspective, SOI substrates are compatible with most conventional fabrication processes. In general, an SOI-based process may be implemented without special equipment or significant retooling of an existing factory. Among challenges unique to SOI are novel metrology requirements to account for the buried oxide layer and concerns about differential stress in the topmost silicon layer. The primary barrier to SOI implementation is the drastic increase in substrate cost, which contributes an estimated 10–15% increase to total manufacturing costs.

Manufacture of SOI wafers

SiO_2-based SOI wafers can be produced by several methods:

- *SIMOX* - **S**eparation by **IM**plantation of **OX**ygen – uses an oxygen ion beam implantation process followed by high temperature annealing to create a buried SiO_2 layer.

- Wafer bonding – the insulating layer is formed by directly bonding oxidized silicon with a second substrate. The majority of the second substrate is subsequently removed, the remnants forming the topmost Si layer.
    - One prominent example of a wafer bonding process is the *Smart Cut* method developed by the French firm Soitec which uses ion implantation followed by controlled exfoliation to determine the thickness of the uppermost silicon layer.
    - *NanoCleave* is a technology developed by Silicon Genesis Corporation that separates the silicon via stress at the interface of silicon and silicon-germanium alloy.
    - *ELTRAN* is a technology developed by Canon which is based on porous silicon and water cut.
- Seed methods - wherein the topmost Si layer is grown directly on the insulator. Seed methods require some sort of template for homoepitaxy, which may be achieved by chemical treatment of the insulator, an appropriately oriented crystalline insulator, or vias through the insulator from the underlying substrate.

An exhaustive review of these various manufacturing processes may be found in reference

Use in the microelectronics industry

IBM began to use SOI in the high-end RS64-IV "Istar" PowerPC-AS microprocessor in 2000. Other examples of microprocessors built on SOI technology include AMD's 130 nm, 90 nm, 65 nm and 45 nm single, dual, quad and six core processors since 2001. Freescale adopted SOI in their PowerPC 7455 CPU in late 2001, currently Freescale is shipping SOI products in 180 nm, 130 nm, 90 nm and 45 nm lines. The 90 nm Power Architecture based processors used in the Xbox 360, PlayStation 3 and Wii use SOI technology as well. Competitive offerings from Intel, however, such as the 65 nm Core 2 and Core 2 Duo microprocessors, are built using conventional bulk CMOS technology. Intel's new 45 nm process will continue to use conventional technology. However, Intel made a claim of single-chip silicon laser based on SOI.

On the foundry side, TSMC claimed no customer wanted SOI, but Chartered Semiconductor devoted a whole fab to SOI.

Use in Photonics

SOI wafers are widely used in silicon photonics. The crystalline silicon layer on insulator can be used to fabricate optical waveguides and other passive optical devices for integrated optics. The crystalline silicon layer in sandwiched between the buried insulator (Silicon oxide, Sapphire etc.) and top cladding of air (or Silicon oxide or any other low refractive index material). This enables propagation of electromagnetic waves in the waveguides on the basis of total internal reflection.

Chapter-12

Applications & Examples of Semiconductor Technology

Technology aware design

Technology Aware Design (**TAD**) is a research program that started in 2001 at IMEC, Leuven, Belgium. It anticipates the end of the traditional "happy scaling" paradigm, where CMOS technology and CMOS design evolved on formally separate tracks, the interface between the two being standard cell, or SPICE compact models.

Today, both sides (design and technology) are confronted with the need to understand the other in order to overcome new scaling induced issues. The TAD program pursues analysis and solutions for these scaling induced problems.

Wafer-level optics

Wafer-level optics (WLO) enables the design and manufacture of miniaturized optics at the wafer level using advanced semiconductor-like techniques. The end product is cost effective, miniaturized optics that enable the reduced form factor of camera modules for mobile devices. . The technology is scalable from a single-element CIF/VGA lens to a multi-element mega pixel lens structure, where the lens wafers are precision aligned, bonded together and diced to form multi-element lens stacks.

LOCOS

Typical LOCOS structure.
1) Silicon 2) Silicon dioxide

LOCOS, short for **LOCal Oxidation of Silicon**, is a microfabrication process where silicon dioxide is formed in selected areas on a silicon wafer having the Si-SiO_2 interface at a lower point than the rest of the silicon surface.

This technology was developed to insulate MOS transistors from each other. The main goal is to create a silicon oxide insulating structure that penetrates under the surface of the wafer, so that the Si-SiO_2 interface occurs at a lower point than the rest of the silicon surface. This cannot be easily achieved by etching field oxide. Thermal oxidation of selected regions surrounding transistors is used instead. The oxygen penetrates in depth of the wafer, reacts with silicon and transforms it in to silicon oxide. In this way, an immersed structure is formed.

The immersed insulating barrier limits the transistor cross-talk.

Process

Typical process steps are the following:

I. Preparation of silicon substrate (layer 1)
II. CVD deposition of SiO_2, pad/buffer oxide (layer 2)
III. CVD deposition of Si_3N_4, nitride mask (layer 3)
IV. Etching of nitride layer (layer 3) and silicon oxide layer (layer 2)
V. Thermal growth of silicon oxide (structure 4)
VI. Further growth of thermal silicon oxide (structure 4)
VII. Removal of nitride mask (layer 3)

There are 4 basic layers/structures:

1. Si, silicon substrate, wafer
2. SiO_2, buffer oxide (pad oxide), chemical vapor deposition silicon oxide
3. Si_3N_4, nitride mask
4. SiO_2, insulation oxide, thermal o

Function of layers and structures

The silicon wafer (layer 1) is used as a basis for building electronic structures (such as MOS transistors).

To perform local oxidation, the areas not meant to be oxidized will be coated in a material that does not permit the diffusion of oxygen at high temperatures (thermal oxidation is performed in temperatures between 800 and 1200°C), such as silicon nitride (layer 3, step III).

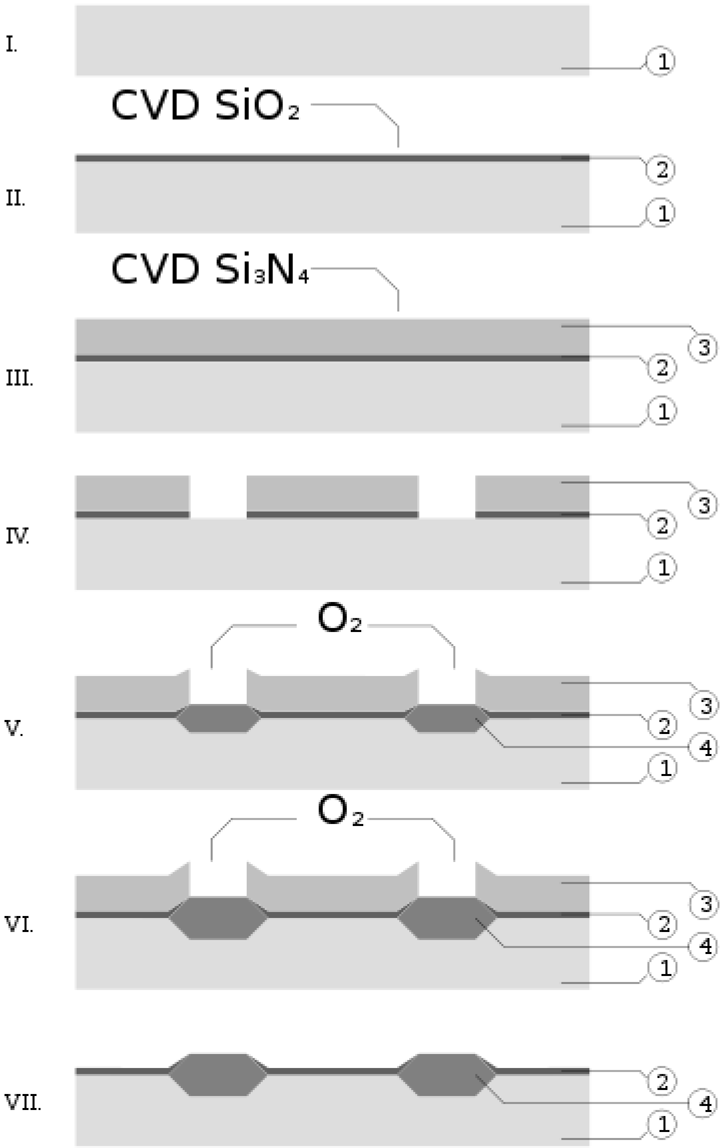
During the growth of the immersed insulating thermal oxide structures (steps V and VI), the silicon nitride layer (layer 3) is pushed upwards. Without the buffer oxide (layer 2, also known as pad oxide), this would create too much tension in the Si substrate (layer 1), the plastic deformation would occur and the electronic devices would be damaged.

Therefore a buffer oxide (layer 2) is deposed by the CVD (step II) between the Si substrate (layer 1) and the silicon nitride (layer 3). At high temperatures, the viscosity of silicon oxide decreases and the stress created between the silicon substrate (layer 1) and nitride layer (layer 3), by the growth of the thermal oxide (steps V and VI), is relieved.

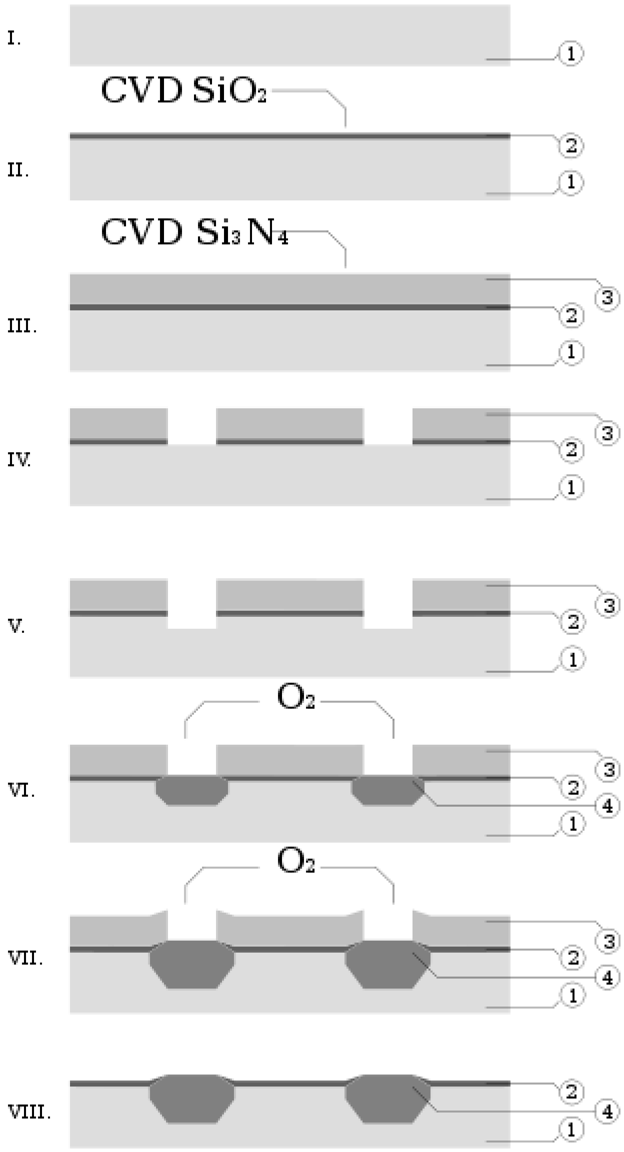
The insulating structures (structure 4) are formed by thermal oxidation of silicon. During this process, the silicon wafer is "consumed" and "replaced" by silicon oxide. The volume of silicon oxide to silicon is about 2.4:1, which explains the growth of the insulation structures and the created tension.

The disadvantage of this technology is that the insulating structures are rather large, and therefore, less MOS transistors can be formed on one wafer.

Reduction of dimensions of insulating structures is solved by the STI (Shallow Trench Isolation, also known as Box Isolation Technique). In this process, trenches are formed and silicon dioxide is deposed inside. The LOCOS technology can't be used in this way, because of the change of the volume during the thermal oxidation, which would induce too much stress in the trenches.



LOCOS process steps:
I. Preparation of silicon substrate
II. CVD deposition of SiO_2, pad/buffer oxide
III. CVD deposition of Si_3N_4, nitride mask
IV. Etching of nitride layer and silicon oxide layer
V. Thermal growth of silicon oxide
VI. Further growth of thermal silicon oxide
VII. Removal of nitride mask
LOCOS process materials:
1) Si, silicon substrate
2) SiO_2, pad/buffer oxide, chemical vapor deposition silicon oxide
3) Si_3N_4, nitride mask
4) SiO_2, isolation oxide, thermal oxide



Fully recessed LOCOS structure process steps:
I. Preparation of silicon substrate
II. CVD deposition of SiO_2, pad/buffer oxide
III. CVD deposition of Si_3N_4, nitride mask
IV. Etching of nitride layer and silicon oxide layer
V. Silicon etching
VI. Thermal growth of silicon oxide
VII. Further growth of thermal silicon oxide
VIII. Removal of nitride mask

Optical beam-induced currents

Optical Beam-Induced Current (**OBIC**) is a semiconductor analysis technique that employs a scanning laser beam to induce a current flow within a semiconductor sample which may be collected and analyzed to generate images that represent the sample's properties. It is a useful imaging technique for detecting or locating various defects or anomalies on a semiconductor sample. Conventional OBIC scans an ultrafast laser beam over the surface of the sample, exciting some electrons into the conduction band through what is known as 'single-photon absorption'. As its name implies, single-photon absorption involves just a single photon to excite the electron into conduction. This can only occur if that single photon carries enough energy to overcome the band gap of the semiconductor (1.2 eV for Si) and provide the electron with enough energy to make it jump into the conduction band.

Proximity communication

Proximity communication is a Sun microsystems technology of wireless chip-to-chip communications. Partly by Robert Drost and Ivan Sutherland. Research done as part of High Productivity Computing Systems DARPA project.

Proximity communication replaces wires by capacitive coupling, promises significant increase in communications speed between chips in an electronic system, among other benefits. Partially funded by a $50 million award from the Defense Advanced Research Projects Agency.

Comparing traditional area ball bonding, proximity communication has one order smaller scale, so it can be two order densier (in terms of connection number/pin number) than ball bonding. This technique require very good alignment between chips and very small gaps between tx and rx parts (2-3 micrometers), which can be destroyed by thermal expansion, vibration, dust, etc.

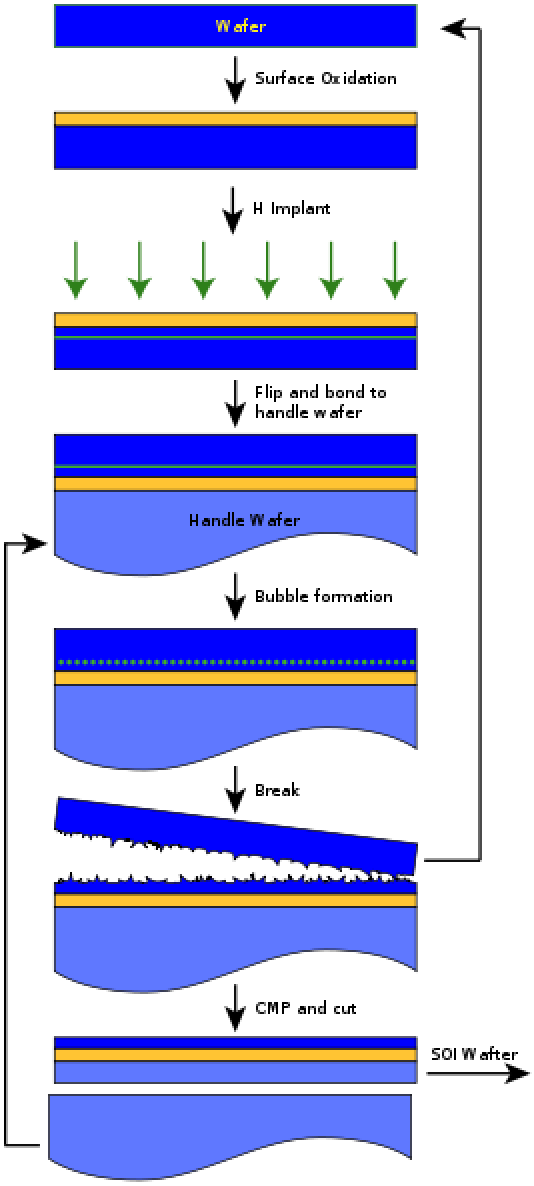
Chip transmitter consists (according to presentation slide) of big 32x32 array of very small Tx micropads, 4x4 array of bigger Rx micropads (four times bigger than tx micropad), and two linear arrays of 14 X vernier and 14 Y vernier.

Proximity communication can be used with 3D packing on chips in Multi-Chip Module, allowing to connect several MCM without sockets and wires.

Speed was up to 1.35 Gbps/channel in tests of 16 channel systems. BER < 10^{-12}. Static power is 3.6 mW/channel, dynamic power is 3.9 pJ/bit

Current status of the project is unknown.

Smart Cut



Smart Cut process

Smart Cut is a technological process that enables the transfer of very fine layers of crystalline material onto a mechanical support. The application of this technological procedure is used mainly in silicon-on-insulator (SOI). The role of SOI is to electronically insulate a fine layer of monocrystalline silicon from the rest of the silicon wafer; an ultra-thin silicon film is transferred to a mechanical support, thereby introducing an intermediate, insulating layer. Semiconductor manufacturers can then fabricate integrated circuits on the top layer of the SOI wafers using the same processes they would use on plain silicon wafers. The wafers are then cut up and the chips packaged for mounting on the cards that are integrated into electronic systems such as personal computers. The Smart Cut process was developed and is patented by SOITEC corporation from France.

Chapter-13

Semiconductor Technology Material

Aluminium nitride

Aluminium nitride (AlN) is a nitride of aluminium. Its wurtzite phase (w-AlN) is a wide band gap (6.2 eV) semiconductor material, giving it potential application for deep ultraviolet optoelectronics.

History

AlN was first synthesized in 1877, but it was not until the middle of the 1980s that its potential for application in microelectronics was realized due to its relative high thermal conductivity for an electrical insulating ceramic (70–210 $W \cdot m^{-1} \cdot K^{-1}$ for polycrystalline material, and as high as 285 $W \cdot m^{-1} \cdot K^{-1}$ for single crystals).

Stability and chemical properties

Aluminium nitride is stable at high temperatures in inert atmospheres and melts at 2800 °C. In a vacuum, AlN decomposes at ~1800 °C. In the air, surface oxidation occurs above 700°C, and even at room temperature, surface oxide layers of 5-10 nm have been detected. This oxide layer protects the material up to 1370°C. Above this temperature bulk oxidation occurs. Aluminium nitride is stable in hydrogen and carbon dioxide atmospheres up to 980°C.

The material dissolves slowly in mineral acids through grain boundary attack, and in strong alkalies through attack on the aluminium nitride grains. The material hydrolyzes slowly in water. Aluminium nitride is resistant to attack from most molten salts, including chlorides and cryolite.

Manufacture

AlN is synthesized by the carbothermal reduction of alumina or by direct nitridation of aluminium. The use of sintering aids and hot pressing is required to produce a dense technical grade material.

Applications

Metallization methods are available to allow AlN to be used in electronics applications similar to those of alumina and beryllium oxide.

Currently there is much research into developing light-emitting diodes to operate in the ultraviolet using the gallium nitride based semiconductors and, using the alloy aluminum gallium nitride, wavelengths as short as 250 nm have been achieved. In May 2006, an inefficient AlN LED emission at 210 nm has been reported.

Among the applications of AlN are

- opto-electronics,
- dielectric layers in optical storage media,
- electronic substrates, chip carriers where high thermal conductivity is essential,
- military applications,
- as a crucible to grow crystals of gallium arsenide,
- steel and semiconductor manufacturing.

Epitaxially grown thin film crystalline aluminium nitride is also used for surface acoustic wave sensors (SAW's) deposited on silicon wafers because of the AlN's piezoelectric properties. One application is an RF filter used in mobile phones called a thin film bulk acoustic resonator (FBAR). This is a MEMS device that uses aluminium nitride sandwiched between two metal layers.

Aluminium phosphide

Aluminium phosphide is the chemical compound with the empirical formula AlP. This colourless solid is generally sold as a grey-green-yellow powder due to the presence of impurities arising from hydrolysis and oxidation. This material is a wide band gap semiconductor and is used as a fumigant.

Structure, synthesis, and chemical properties

AlP crystallizes in the cubic zinc blende lattice, wherein all atoms have tetrahedral coordination. Related materials crystallize similarly, including GaAs. At pressures of 14-17 GPa, AlP transforms into a rocksalt phase .

Crude aluminium phosphide can be prepared in the laboratory by igniting a mixture of red phosphorus and powdered aluminium.

Aluminium phosphide reacts with water or acids to release phosphine.

$AlP + 3\ H_2O \rightarrow Al(OH)_3 + PH_3$
$AlP + 3\ H^+ \rightarrow Al^{3+} + PH_3$

Physical properties

Aluminium phosphide has a hardness of 5.5 on the Mohs scale .

Pesticide

AlP is used as a rodenticide, insecticide, and fumigant for stored cereal grains. It is used to kill small verminous mammals such as moles, rabbits, and rodents. The tablets or pellets typically also contain other chemicals that evolve ammonia which helps to reduce the potential for spontaneous ignition or explosion of the phosphine gas.

As a rodenticide, aluminium phosphide pellets are provided as a mixture with food for consumption by the rodents. The acid in the digestive system of the rodent reacts with the phosphide to generate the toxic phosphine gas. Other pesticides similar to aluminium phosphide are zinc phosphide and calcium phosphide.

As a rodenticide, aluminium phosphide can be encountered under various brand names, e.g. **Celphos**, **Fumitoxin**, **Phostoxin**, and **Quick Phos**.

Evidently poisonous, aluminium phosphide has been used for suicide. Fumigation has also caused unintentional deaths, such as examples in Saudi Arabia and the United States. Known as "rice tablet" in Iran, for its use to preserve rice, there have been frequent incidents of accidental or intentional death. There is a campaign by Iranian Forensic Medicine Organization to stop its use as a pesticide.

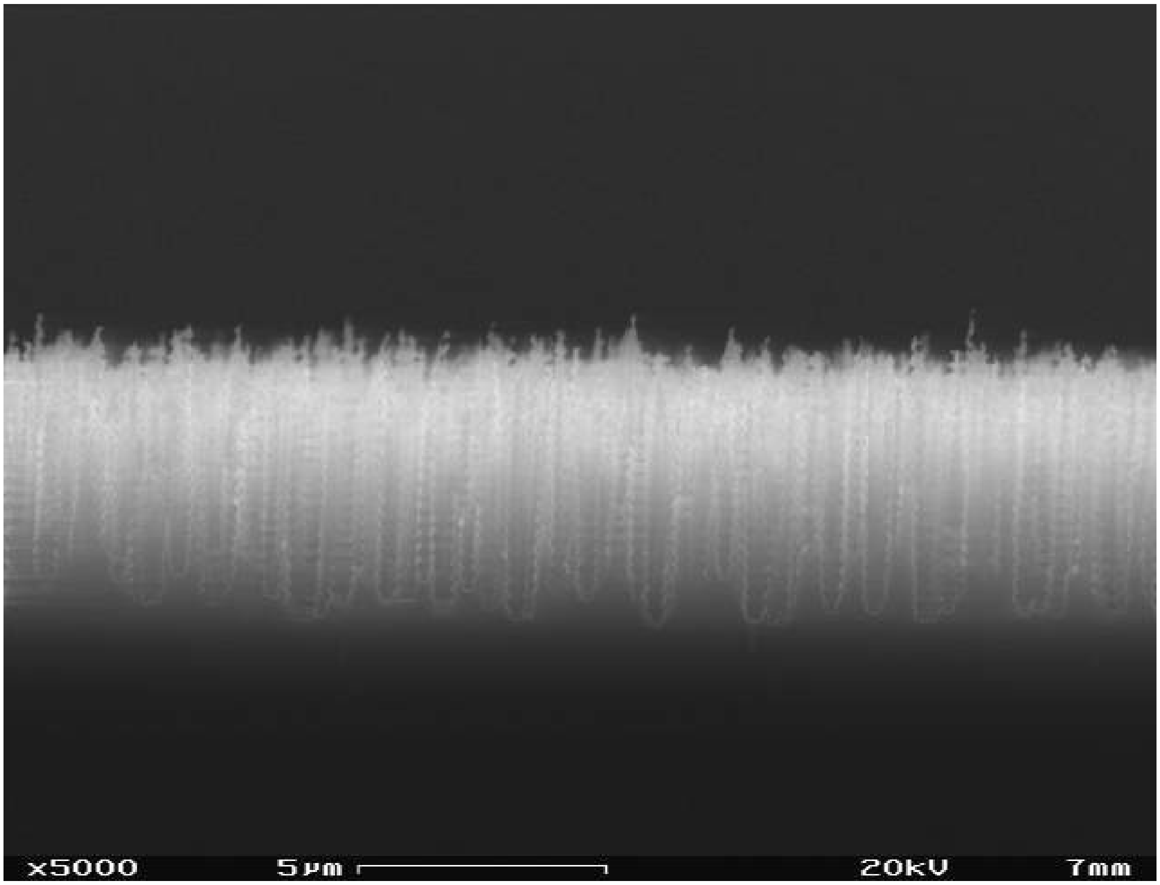
Semiconductor applications

Industrially, AlP is a semiconductor material that is usually alloyed with other binary materials for applications in devices such as light-emitting diodes (e.g. aluminium gallium indium phosphide).

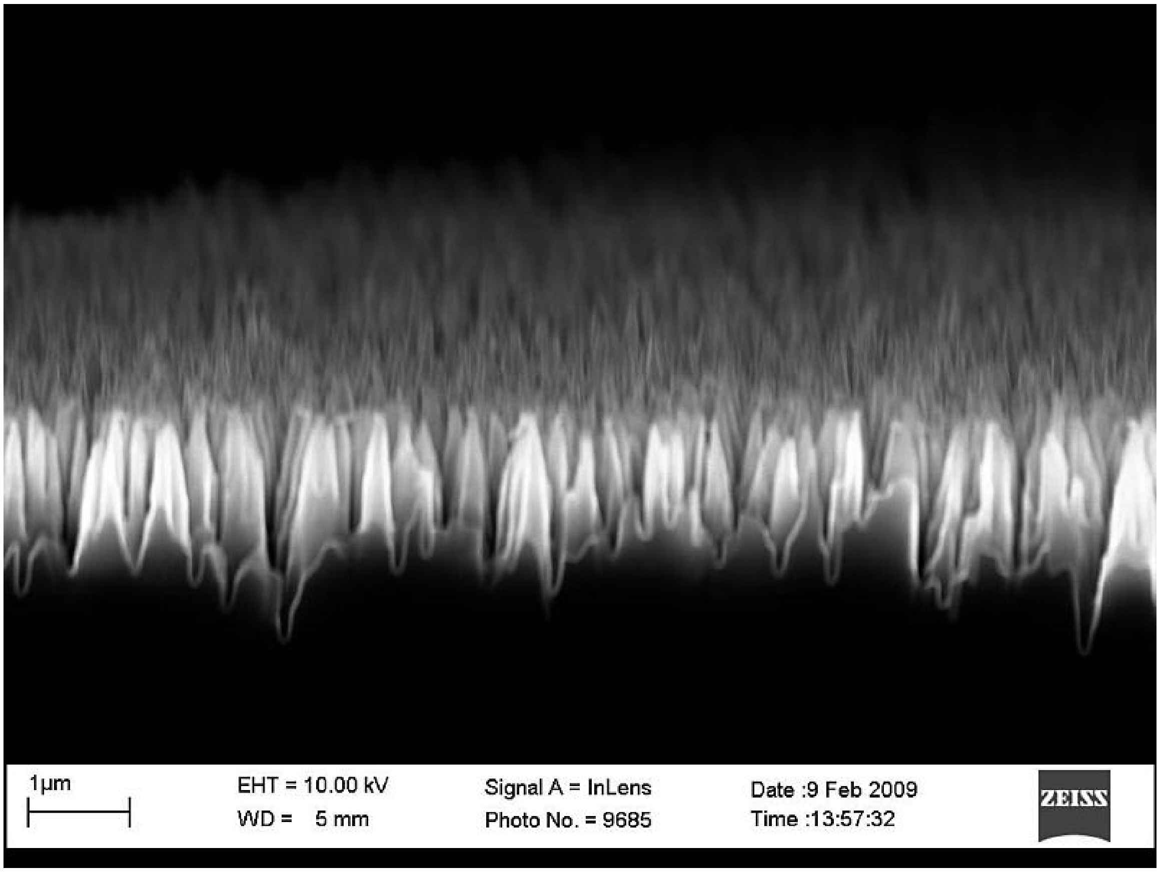
Black silicon

Black silicon is a semiconductor material, a surface modification of silicon with very low reflectivity and correspondingly high absorption of visible (and infrared) light. The modification was discovered in the 1980s as an unwanted side effect of reactive ion etching (RIE). Another method for forming a similar structure was developed in Eric Mazur's laboratory at Harvard University (1998).

Properties



Scanning electron micrograph of black silicon, produced by RIE (ASE process)



SEM micrograph of black silicon formed by cryogenic RIE. Notice the smooth, sloped surfaces, unlike the undulated sidewalls obtained with the Bosch process RIE.

Black silicon is a needle-shaped surface structure where needles are made of single-crystalline silicon and have a height above 10 microns and diameter <1 micron. Its main feature is an increased absorption of incident light – the high reflectivity of the silicon, which is usually 20–30% for quasi-normal incidence, is reduced to about 5%. This is due to the formation of a so-called effective medium by the needles. Within this medium, there is no sharp interface, but a continuous change of the refractive index that reduces Fresnel reflection.

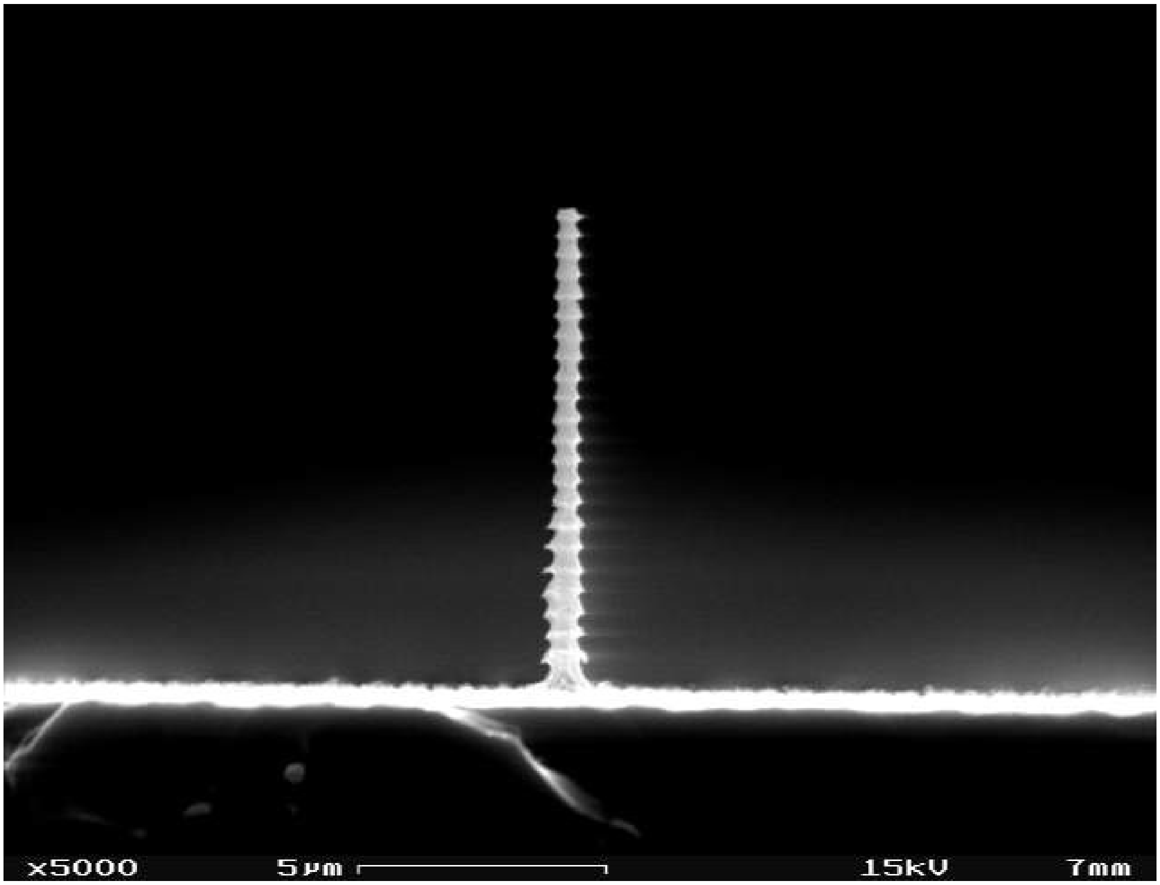
Applications

The unusual optical characteristics, combined with the semiconducting properties of silicon make this material interesting for sensor applications. The potential applications include:

- Image sensors with increased sensitivity
- Thermal imaging cameras
- Photodetector with high efficiency through increased absorption.
- Mechanical contacts and interfaces
- Terahertz applications.

- Solar Cells

Production by RIE



Scanning electron micrograph of a single "needle" of black silicon, produced by RIE (ASE process)

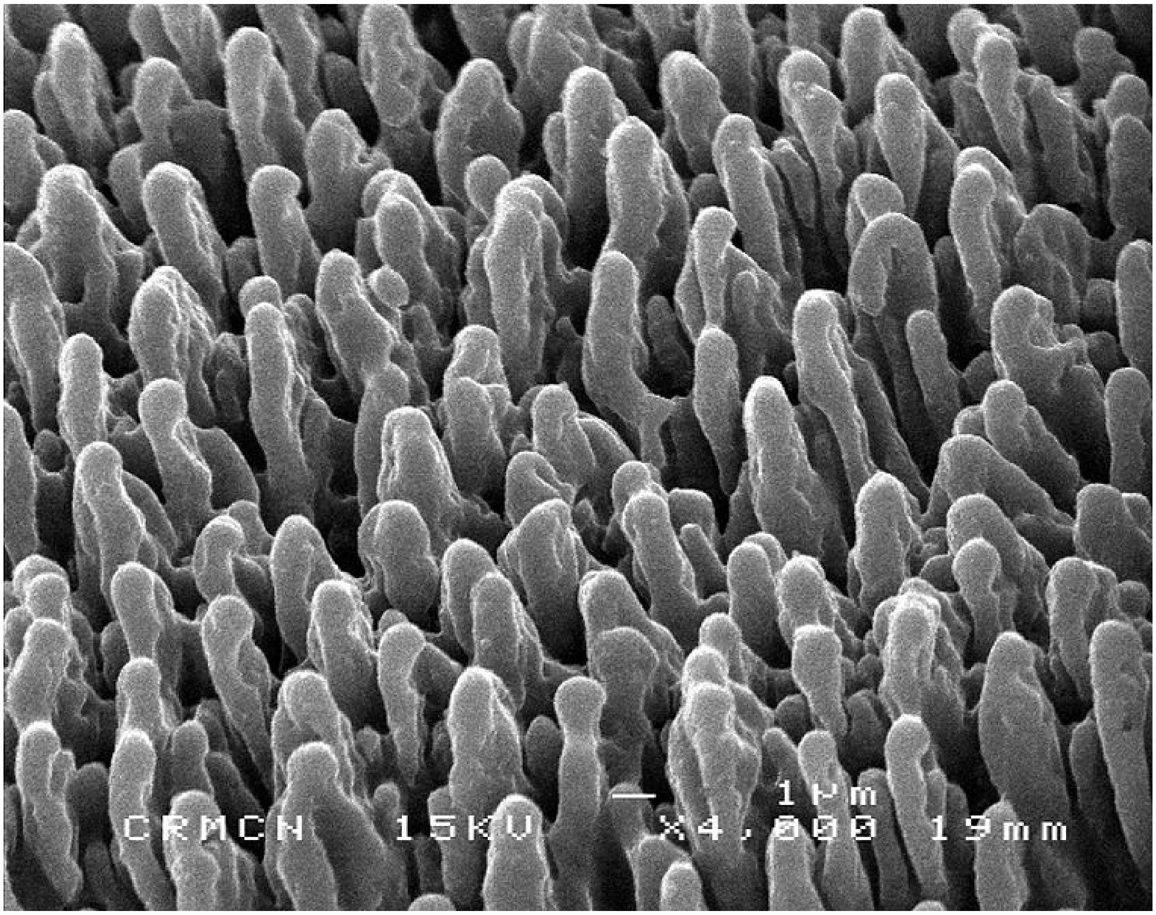
In semiconductor technology, RIE is a standard procedure for the producing trenches and holes with a depth of up to several hundred microns and very high aspect ratios. In Bosch process RIE, this is achieved by repeatedly switching between an etching and passivation. With cryogenic RIE, the low temperature and oxygen gas achieve this sidewall passivation by forming SiO_2, easily removed from the bottom by directional ions. Both RIE methods can produce black silicon, but the morphology of the resulting structure differs substantially. The switching between etching and passivation of the Bosch process creates undulated sidewalls, which are visible also on the black silicon formed this way.

During etching, however, small debris remain on the substrate; they mask the ion beam and produce structures that are not removed and in the following etching steps and result in tall silicon pillars. The process can be set so that a million needles are formed on an area of one square millimeter.

Production by Mazur's method

Material

In 1999, a group led by Eric Mazur and James Carey at the Harvard University developed a process in which black silicon was produced by irradiating silicon with femtosecond laser pulses. After irradiation in the presence of a gas containing sulfur hexafluoride and other dopants, the surface of silicon develops a self-organized microscopic structure of micrometer-sized cones. The resulting material has many remarkable properties, such as an enhanced absorption that extends to the infrared below the band gap of silicon, including the wavelengths for which unmodified silicon is transparent. This property is caused by sulfur atoms being forced to the silicon surface, creating a structure with a lower band gap and therefore the ability to absorb longer wavelengths.



Black silicon made without special gas ambient - laboratory LP3-CNRS

Similar surface modification can be achieved in vacuum using the same type of laser and laser processing conditions. In this case, the individual silicon micro-cones are lack of sharp tip but more in a penguin-like form. The reflectivity of such a micro-structured surface is very low, 3~14% in the spectral range 350–1150 nm. Such reduction in reflectivity is considered to be contributed by the geometry of these micro-cones, which increases the light internal reflections between themselves and hence the possibility of light absorption by the silicon is increased. The gain in absorption achieved by fs laser texturization is found to be superior to that achieved by using alkaline chemical etch method, which is a standard industrial approach for surface texturization of mono-crystalline silicon wafers in solar cell manufacturing. It is also found that such surface modification is independent to local crystalline orientation. Uniform texturisation effect can be achieved across the whole surface of a multi-crystalline silicon wafer. These findings make it an attractive potential alternative to chemical surface texturisation process in solar cell manufacturing.

Function

When the material is biased by a small electric voltage, absorbed photons are able to excite dozens of electrons. The sensitivity of black silicon detectors is 100–500 times higher than that of untreated silicon (conventional silicon), in both the visible and infrared spectra.

Uses and commercialization

The material has found commercial applications in a number of photodetectors for various imaging and night vision applications. Black silicon is currently being commercialized by SiOnyx, a Massachusetts-based venture-funded startup company which acquired licensing for the process from Harvard in 2006.

Black silicon also has potential application for high efficiency solar cells, which is being explored by Solasys, an EU Seventh Framework Programme funded demonstration project aiming at lowering manufacturing costs while increasing cell efficiency at the same time.

Boron nitride

Boron nitride

Identifiers

CAS number	10043-11-5 ✓
PubChem	66227
ChemSpider	59612 ✓
EC number	233-136-6
RTECS number	ED7800000

SMILES

InChI

Properties

Molecular formula	BN
Molar mass	24.818 g/mol
Appearance	white powder or translucent crystals (both c-BN and h-BN)
Density	3.4870 g/cm^3 (cBN)
Melting point	2973 °C (cBN) (sublimes)
Solubility in water	insoluble
Solubility	insoluble in acids
Electron mobility	200 cm^2/(V·s) (cBN)
Refractive index (n_D)	2.17 (cBN)

Structure

Crystal structure	hexagonal, sphalerite, wurtzite

Thermochemistry

Std enthalpy of	476.98 kJ mol^{-1}

formation $\Delta_f H^\ominus_{298}$	
Std enthalpy of combustion $\Delta_c H^\ominus_{298}$	-250.91 kJ mol^{-1}
Standard molar entropy $S^\ominus_{298}$	14.77 J mol^{-1} K^{-1}
Hazards	
R-phrases	R36, R37
S-phrases	S26, S36
Related compounds	
Other anions	BP, BAs B_4C, B_2O_3

Boron nitride is a chemical compound with chemical formula BN, consisting of equal numbers of boron and nitrogen atoms. BN is isoelectronic to a similarly structured carbon lattice and thus exists in various crystalline forms. The hexagonal form corresponding to graphite is the most stable and softest among BN polymorphs, and is therefore used as a lubricant and an additive to cosmetic products. The cubic (sphalerite structure) variety analogous to diamond is called c-BN. Its hardness is inferior only to diamond, but its thermal and chemical stability is superior. The rare wurtzite BN modification is similar to lonsdaleite and may even be harder than the cubic form.

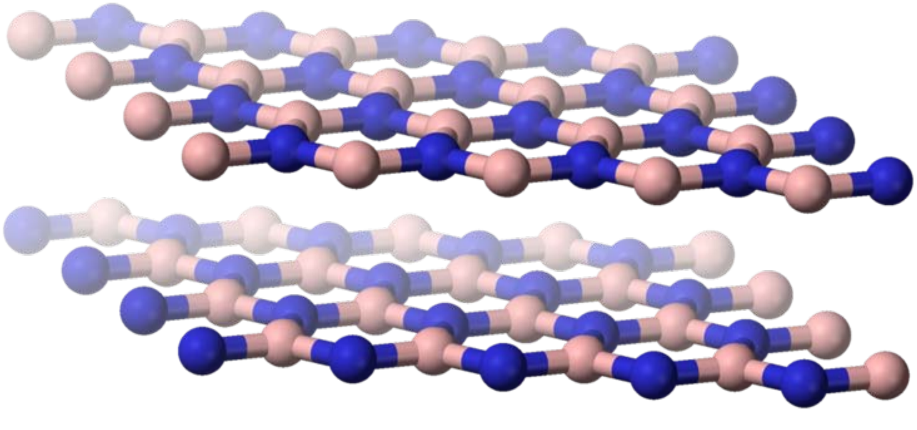
Boron nitride is not found in nature and is therefore produced synthetically from boric acid or boron trioxide. The initial product is amorphous BN powder, which is converted to crystalline h-BN by heating in nitrogen flow at temperatures above 1500 °C. c-BN is made by annealing h-BN powder at higher temperatures, under pressures above 5 GPa. Contrary to diamond, larger c-BN pellets can be produced by fusing (sintering) relatively cheap c-BN powders. As a result, c-BN is widely used in mechanical applications.

Because of excellent thermal and chemical stability, boron nitride ceramics are traditionally used as parts of high-temperature equipment. Boron nitride has a great potential in nanotechnology. Nanotubes of BN can be produced that have a structure similar to that of carbon nanotubes, i.e. graphene (or BN) sheets rolled on themselves, however the properties are very different: whereas carbon nanotubes can be metallic or semiconducting depending on the rolling direction and radius, a BN nanotube is an electrical insulator with a wide bandgap of ~5.5 eV (same as in diamond), which is almost independent of tube chirality and morphology. Similar to other BN forms, BN nanotubes are more thermally and chemically stable than carbon nanotubes which favors them for some applications.

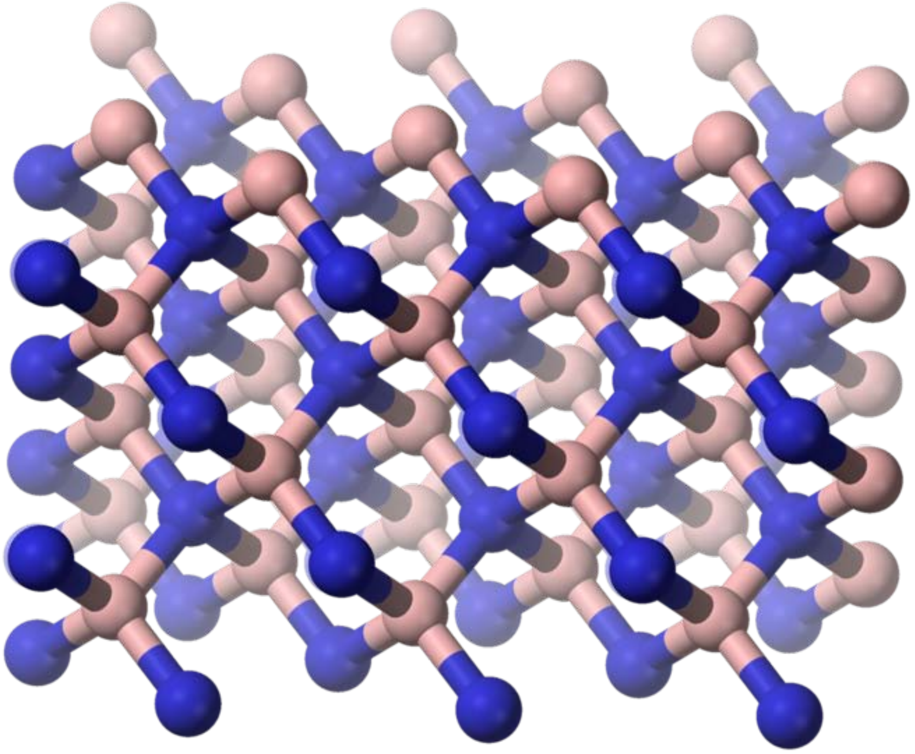
Structure

Boron nitride has been produced in an amorphous (a-BN) and crystalline forms. The most stable crystalline form is the hexagonal one, also called h-BN, α-BN, or g-BN (graphitic BN). It has a layered structure similar to graphite. Within each layer, boron and nitrogen atoms are bound by strong covalent bonds, whereas the layers are held together by weak van der Waals forces. The interlayer "registry" of these sheets differs, however, from the pattern seen for graphite, because the atoms are eclipsed, with boron atoms lying over and above nitrogen atoms. This registry reflects the polarity of the B-N bonds. Still, h-BN and graphite are very close neighbors and even the BC_6N hybrids have been synthesized where carbon substitutes for some B and N atoms.

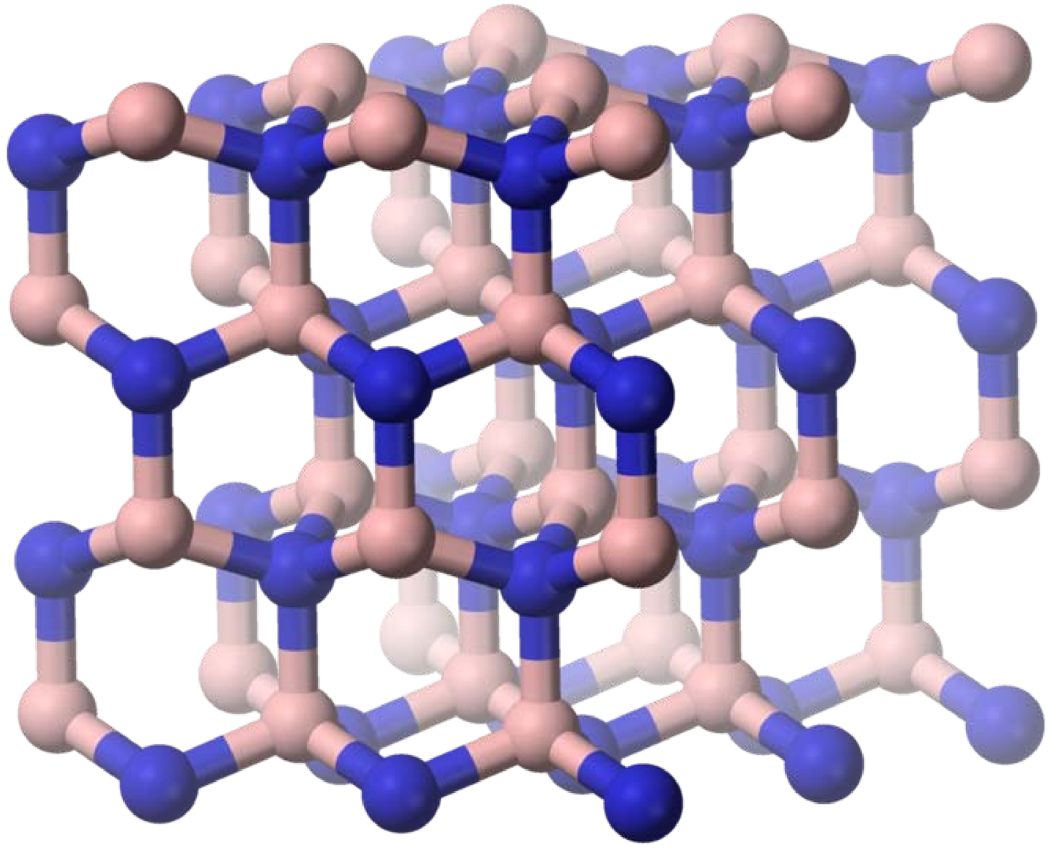
As diamond is less stable than graphite, cubic BN is less stable than h-BN, but the conversion rate between those forms is negligible at room temperature. The cubic form has the sphalerite crystal structure, same as diamond structure, and is also called β-BN or c-BN. The wurtzite BN form (w-BN) has similar structure as lonsdaleite, rare hexagonal polymorph of carbon. In both c-BN and w-BN boron and nitrogen atoms are grouped into tetrahedra, but the angles between neighboring tetrahedra are different.



α-BN, hexagonal



β-BN, sphalerite structure



BN, wurtzite structure

Properties

Physical

Properties of amorphous and crystalline BN, graphite and diamond.
Some properties of h-BN and graphite differ within the basal planes (‖) and perpendicular to them (╧)

Material	a-BN	h-BN	c-BN	w-BN	graphite	diamond
Density (g/cm^3)	2.28	~2.1	3.45	3.49	~2.1	3.515
Mohs hardness		1-2	~10	~10	1-2	10
Knoop hardness (GPa)	10		45	34		100
Bulk modulus (GPa)	100	36.5	400	400	34	440

Thermal conductivity (W/cm K)	0.03	6∥; 0.3⊥	7.4		2-20∥; 0.02-0.8⊥	6-20
Thermal expansion (10^{-6}/°C)		-2.7∥; 38⊥	1.2	2.7	-1.5∥; 25⊥	0.8
Bandgap (eV)	5.05	5.2	6.4	4.5-5.5	0	5.5
Refractive index	1.7	1.8	2.1	2.05		2.4
Magnetic susceptibility (µemu/g)		-0.48∥; -17.3⊥			-0.2..-2.7∥; -20..-28⊥	-1.6

Sources: amorphous BN, crystalline BN, graphite, diamond.

The partly ionic structure of BN layers in h-BN reduces covalency and electrical conductivity, whereas the interlayer interaction increases resulting in higher hardness of h-BN relative to graphite. The reduced electron-delocalization in hexagonal-BN is also indicated by its absence of color and a large band gap. Very different bonding - strong covalent within the basal planes (planes where boron and nitrogen atoms are covalently bonded) and weak between them - causes high anisotropy of most properties of h-BN.

For example, the hardness, electrical and thermal conductivity are much higher within the planes than perpendicular to them. On the contrary, the properties of c-BN and w-BN are more homogeneous.

Those materials are extremely hard, with the hardness of c-BN being slightly smaller and w-BN even higher than that of diamond. Because of much better stability to heat and metals, c-BN surpasses diamond in mechanical applications. The thermal conductivity of BN is among the highest of all electric insulators (see table).

Boron nitride can be doped p-type with Be and n-type with boron, sulfur, silicon or if co-doped with carbon and nitrogen. Both hexagonal and cubic BN are wide-gap semiconductors with a band gap energy corresponding to the UV region. If voltage is applied to h-BN or c-BN, then it emits UV light in the range 215-250 nm and therefore can potentially be used as light emitting diodes (LEDs) or lasers.

Little is known on melting behavior of boron nitride. It sublimates at 2973 °C at normal pressure releasing nitrogen gas and boron, but melts at elevated pressure.

Thermal stability

Hexagonal and cubic (and probably w-BN) BN show remarkable chemical and thermal stabilities. For example, h-BN is stable to decomposition in temperatures up to 1000 °C in air, 1400 °C in vacuum, and 2800 °C in an inert atmosphere. The reactivity of h-BN and c-BN is relatively similar, and the data for c-BN are summarized in the table below.

Reactivity of c-BN with solids

Solid	**Ambient**	**Action**	**Threshold T (°C)**
Mo	10^{-2} Pa vacuum	reaction	1360
Ni	10^{-2} Pa vacuum	wetting	1360
Fe, Ni, Co	argon	react	1400–1500
Al	10^{-2} Pa vacuum	wetting and reaction	1050
Si	10^{-3} Pa vacuum	wetting	1500
Cu, Ag, Au, Ga, In, Ge, Sn	10^{-3} Pa vacuum	no wetting	1100
B		no wetting	2200
$Al_2O_3 + B_2O_3$	10^{-2} Pa vacuum	no reaction	1360

Thermal stability of c-BN can be summarized as follows:

- In air or oxygen: B_2O_3 protective layer prevents further oxidation to ~1300 °C; no conversion to hexagonal form at 1400 °C.
- In nitrogen: some conversion to h-BN at 1525 °C after 12 h.
- In vacuum (10^{-5} Pa): conversion to h-BN at 1550 - 1600 °C.

Chemical stability

Boron nitride is insoluble in usual acids, but is soluble in alkaline molten salts and nitrides, such as LiOH, KOH, NaOH-Na_2CO_3, $NaNO_3$, Li_3N, Mg_3N_2, Sr_3N_2, Ba_3N_2 or Li_3BN_2, which are therefore used to etch BN.

Thermal conductivity

The theoretical thermal conductivity of hexagonal Boron nitride nanoribbons (BNNRs) can approach 1700-2000 W/(m·K), which has the same order of magnitude as the experimental measured value for graphene, and can be comparable to the theoretical calculations for graphene nanoribbons. Moreover, the thermal transport in the BNNRs is anisotropic. The thermal conductivity of zigzag-edged BNNRs is about 20% larger than that of armchair-edged nanoribbons at room temperature.

Synthesis

Boron nitride has not been found in nature and therefore is produced synthetically. The most common raw materials for BN synthesis, boric acid and boron trioxide are produced on industrial scales by treating minerals borax and colemanite with sulfuric acid or hydrochloric acid:

$$Na_2B_4O_7\cdot 10H_2O \text{ (borax)} + H_2SO_4 \rightarrow 4\ H_3BO_3 \text{ (boric acid)} + Na_2SO_4 + 5\ H_2O$$

Boron trioxide is obtained by heating boric acid.

Preparation and reactivity of hexagonal BN

Hexagonal boron nitride is obtained by the reacting boron trioxide (B_2O_3) or boric acid ($B(OH)_3$) with ammonia (NH_3) or urea ($CO(NH_2)_2$) in nitrogen atmosphere:

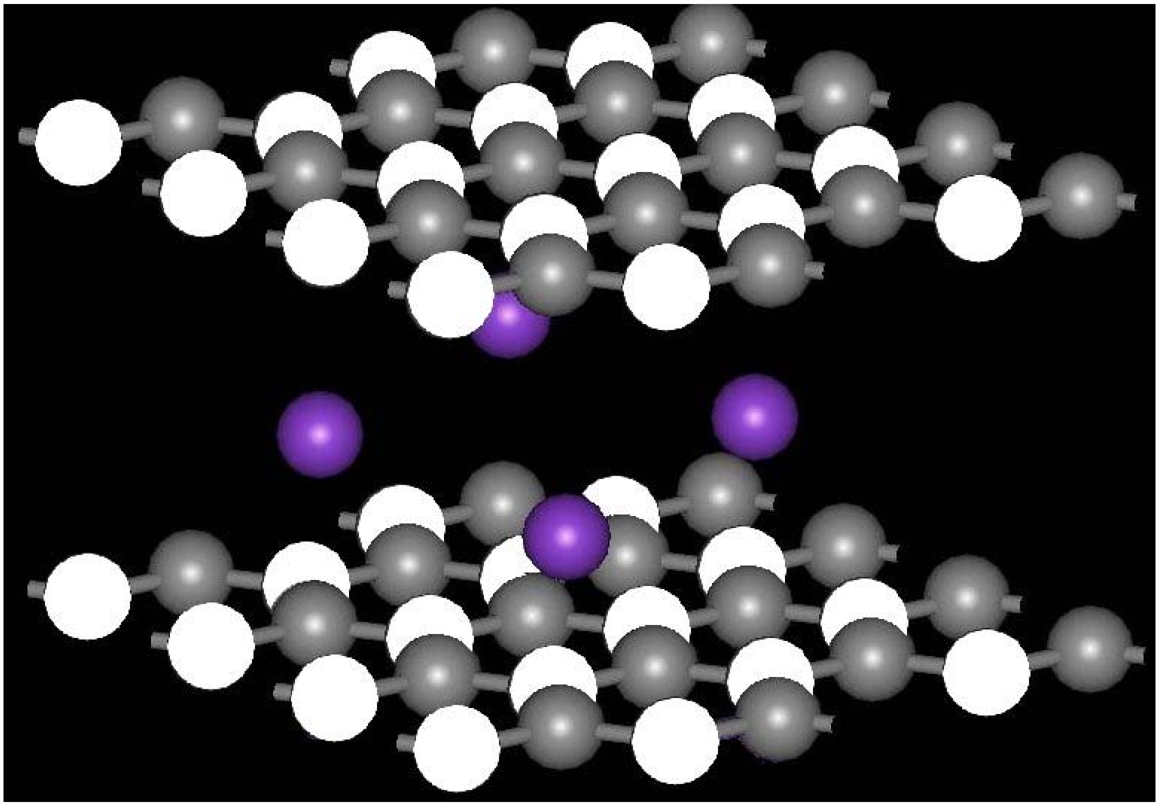
$$B_2O_3 + 2\ NH_3 \rightarrow 2\ BN + 3\ H_2O\ (T = 900\ °C)$$
$$B(OH)_3 + 3\ NH_3 \rightarrow BN + 2\ NH_3 + 3\ H_2O\ (T = 900\ °C)$$
$$B_2O_3 + CO(NH_2)_2 \rightarrow 2\ BN + CO_2 + 2\ H_2O\ (T > 1000\ °C)$$
$$B_2O_3 + 3\ CaB_6 + 10\ N_2 \rightarrow 20\ BN + 3\ CaO\ (T > 1500\ °C)$$

The resulting disordered (amorphous) boron nitride contains 92-95% BN and 5-8% B_2O_3. The remaining B_2O_3 can be evaporated in a second step at temperatures >1500 °C in order to achieve BN concentration >98%. Such annealing also crystallizes BN, the size of the crystallites increasing with the annealing temperature.

h-BN parts can be fabricated inexpensively by hot-pressing with subsequent machining. The parts are made from boron nitride powders adding boron oxide for better compressibility. Thin films of boron nitride can be obtained by chemical vapor deposition from boron trichloride and nitrogen precursors. Combustion of boron powder in nitrogen plasma at 5500 °C yields ultrafine boron nitride used for lubricants and toners.

Boron nitride reacts with iodine fluoride in trichlorofluoromethane at -30 °C to produce an extremely sensitive contact explosive, NI_3, in low yield.

Intercalation of hexagonal BN



Structure of hexagonal boron nitride intercalated with potassium (B_4N_4K)

Similar to graphite, various molecules, such as NH_3 or alkali metals, can be intercalated into hexagonal boron nitride, that is inserted between its layers. Both experiment and theory suggest the intercalation is much more difficult for BN than for graphite.

Preparation of cubic BN

Synthesis of c-BN uses same methods as that of diamond: Cubic boron nitride is produced by treating hexagonal boron nitride at high pressure and temperature, much as synthetic diamond is produced from graphite. Direct conversion of hexagonal boron nitride to the cubic form has been observed at pressures between 5 and 18 GPa and temperatures between 1730 and 3230 °C, that is similar parameters as for direct graphite-diamond conversion. The addition of a small amount of boron oxide can lower the required pressure to 4-7 GPa and temperature to 1500 °C. As in diamond synthesis, to further reduce the conversion pressures and temperatures, a catalyst is added, such as lithium, potassium, or magnesium, their nitrides, their fluoronitrides, water with ammonium compounds, or hydrazine. Other industrial synthesis methods, again borrowed from diamond growth, use crystal growth in a temperature gradient, or explosive shock wave. The shock wave method is used to produce material called heterodiamond, a superhard compound of boron, carbon, and nitrogen.

Low-pressure deposition of thin films of cubic boron nitride is possible. As in diamond growth, the major problem is to suppress the growth of hexagonal phases (h-BN or graphite, respectively). Whereas in diamond growth this is achieved by adding hydrogen gas, boron trifluoride is used for c-BN. Ion beam deposition, plasma-enhanced chemical vapor deposition, pulsed laser deposition, reactive sputtering, and other physical vapor deposition methods are used as well.

Preparation of wurtzite BN

Wurtzite BN can be obtained via static high-pressure or dynamic shock methods. The limits of its stability are not well defined. Both c-BN and w-BN are formed by compressing h-BN, but formation of w-BN occurs at much lower temperatures close to 1700 °C.

Production statistics

Whereas the production and consumption figures for the raw materials used for BN synthesis, namely boric acid and boron trioxide, are well known, the corresponding numbers for the boron nitride are not listed in statistical reports. An estimate for the 1999 world production is 300 to 350 metric tons. The major producers and consumers of BN are located in the United States, Japan, China and Germany. In 2000, prices varied from about \$75/kg to \$120/kg for standard industrial-quality h-BN and were about up to \$200–\$400/kg for high purity BN grades.

Applications

Hexagonal BN

Ceramic BN crucible

Hexagonal BN is the most widely used polymorph. It is a good lubricant at both low and high temperatures (up to 900 °C, even in an oxidizing atmosphere). h-BN lubricant is particularly useful when the electrical conductivity or chemical reactivity of graphite (alternative lubricant) would be problematic. Another advantage of h-BN over graphite is that its lubricity does not require water or gas molecules trapped between the layers. Therefore, h-BN lubricants can be used even in vacuum, e.g. in space applications. The lubricating properties of fine-grained h-BN are used in cosmetics, paints, dental cements, and pencil leads.

Hexagonal BN was first used in cosmetics around 1940 in Japan. However, because of its high price, h-BN was soon abandoned for this application. Its use was revitalized in the late 1990s with the optimization h-BN production processes, and currently h-BN is used by nearly all leading producers of cosmetic products for foundations, make-up, eye shadows, blushers, kohl pencils, lipsticks and other skincare products.

Because of its excellent thermal and chemical stability, boron nitride ceramics are traditionally used as parts of high-temperature equipment. h-BN can be included in ceramics, alloys, resins, plastics, rubbers, and other materials, giving them self-lubricating properties. Such materials are suitable for construction of e.g. bearings and in steelmaking. Plastics filled with BN have less thermal expansion, higher thermal conductivity and electrical resistivity. Due to its excellent dielectric and thermal properties, BN is used in electronics e.g. as a substrate for semiconductors, microwave-transparent windows, structural material for seals.

Hexagonal BN is used in xerographic process and laser printers as a charge leakage barrier layer of the photo drum. In the automotive industry, h-BN mixed with a binder (boron oxide) is used for sealing oxygen sensors, which provide feedback for adjusting fuel flow. The binder utilizes the unique temperature stability and insulating properties of h-BN.

Parts can be made of h-BN by hot pressing. Union Carbide Corporation produces three grades of BN. HBN, with boron oxide binder, usable to 550-850 °C in oxidizing atmosphere and up to 1600 ° in vacuum, but due to the boron oxide content is sensitive to water. HBR uses calcium borate binder and is usable to 1600 °C. HBC grade uses no binder and can be used to 3000 °C.

Cubic boron nitride

Cubic boron nitride (CBN or c-BN) is widely used as an abrasive. Its usefulness arises from its insolubility in iron, nickel, and related alloys at high temperatures, whereas diamond is soluble in these metals to give carbides. Polycrystalline c-BN (PCBN) abrasives are therefore used for machining steel, whereas diamond abrasives are preferred for aluminum alloys, ceramics, and stone. When in contact with oxygen at high temperatures, BN forms a passivation layer of boron oxide. Boron nitride binds well with metals, due to formation of interlayers of metal borides or nitrides. Materials with cubic boron nitride crystals are often used in the tool bits of cutting tools. For grinding applications, softer binders, e.g. resin, porous ceramics, and soft metals, are used. Ceramic binders can be used as well. Commercial products are known under names "Borazon" (by Diamond Innovations), and "Elbor" or "Cubonite" (by Russian vendors). Similar to diamond, the combination in c-BN of highest thermal conductivity and electrical resistivity is ideal for heat spreaders. Contrary to diamond, large c-BN pellets can be produced in a simple process (called sintering) of annealing c-BN powders in nitrogen flow at temperatures slightly below the BN decomposition temperature. This ability of c-BN and h-BN powders to fuse allows cheap production of large BN parts.

As cubic boron nitride consists of light atoms and is very robust chemically and mechanically, it is one of the popular materials for X-ray membranes: low mass results in small X-ray absorption, and good mechanical properties allow usage of thin membranes, thus further reducing the absorption.

Amorphous boron nitride

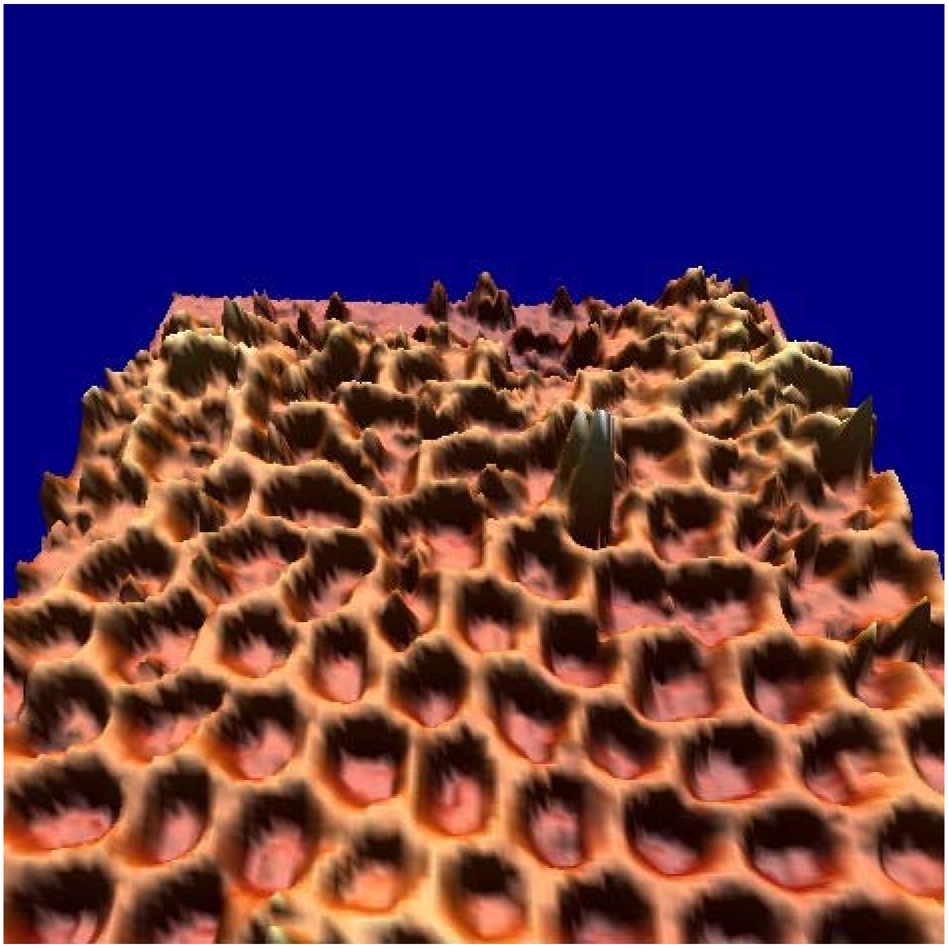
Layers of amorphous boron nitride (a-BN) are used in some semiconductor devices, e.g. MISFETs. They can be prepared by chemical decomposition of trichloroborazine with caesium, or by thermal chemical vapor deposition methods. Thermal CVD can be also used for deposition of h-BN layers, or at high temperatures, c-BN.

Other BN forms

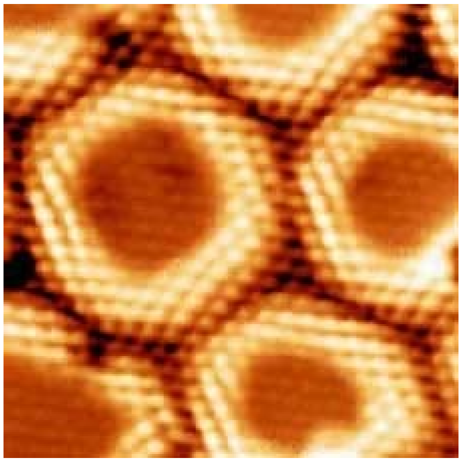
Boron nitride fibers

Hexagonal BN can be prepared in the form of fibers, structurally similar to carbon fibers, by thermal decomposition of extruded borazine ($B_3N_3H_6$) fibers with addition of boron oxide in a nitrogen atmosphere at 1800 °C. An alternative method is thermal decomposition of cellulose fibers impregnated with boric acid or ammonium tetraborate in an atmosphere of ammonia and nitrogen above 1000 °C. Boron nitride fibers are used as reinforcement in composite materials, with the matrix materials ranging from organic resins to ceramics to metals.

Boron nitride nanomesh



Perspective view of nanomesh (structure ends at the back of the figure)



BN nanomesh observed by scanning tunneling microscope. The center of each ring corresponds to the center of the pores

Boron nitride nanomesh is a inorganic nanostructured two-dimensional material. It consists of a single BN layer, which forms by self-assembly a highly regular mesh after high-temperature exposure of a clean rhodium or ruthenium surface to borazine under ultra-high vacuum. The nanomesh looks like an assembly of hexagonal pores. The distance between 2 pore centers is 3.2 nm and the pore diameter is ~2 nm.

The boron nitride nanomesh is not only stable to decomposition under vacuum, air and some liquids, but also up to temperatures of 800 °C. In addition, it shows the extraordinary ability to trap molecules and metallic clusters which have similar sizes to the nanomesh pores, forming a well-ordered array. These characteristics promise interesting applications of the nanomesh in areas like nanocatalysis, surface functionalisation, spintronics, quantum computing and data storage media like hard drives.

Boron nitride nanotubes

Boron nitride nanotubes, were theoretically predicted in 1994 and experimentally discovered in 1995. They can be imagined as a rolled up sheet of boron nitride. Structurally, it is a close analog of the carbon nanotube, namely a long cylinder with diameter of several to hundred nanometers and length of many micrometers, except carbon atoms are alternately substituted by nitrogen and boron atoms. However, the properties of BN nanotubes are very different: whereas carbon nanotubes can be metallic or semiconducting depending on the rolling direction and radius, a BN nanotube is an electrical insulator with a bandgap of ~5.5 eV, basically independent of tube chirality and morphology. In addition, a layered BN structure is much more thermally and chemically stable than a graphitic carbon structure.

All well-established techniques of carbon nanotube growth, such as arc-discharge, laser ablation and chemical vapor deposition, are used to synthesize BN nanotubes. BN nanotubes can also be produced by ball milling of amorphous boron, mixed with a catalyst - iron powder, under NH_3 atmosphere. Subsequent annealing at ~1100 °C in nitrogen flow transforms most of the product into BN.

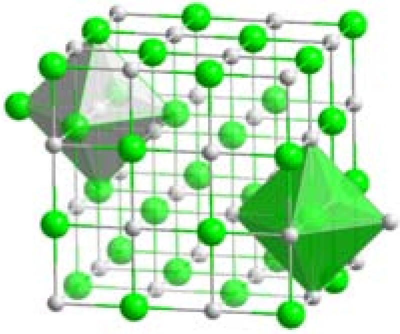
Electrical and field emission properties of the thus prepared nanotubes can be tuned by doping with gold atoms via sputtering of gold on the nanotubes. Doping rare-earth atoms of europium turns a BN nanotube into a phosphor material emitting visible light under electron excitation.

Like BN fibers, boron nitride nanotubes show promise for aerospace applications where integration of boron and in particular the light isotope of boron (^{10}B) into structural materials improves their radiation-shielding properties; the improvement is due to strong neutron absorption by ^{10}B. Such ^{10}BN materials are of particular theoretical value as composite structural materials in future manned interplanetary spacecraft, where absorption-shielding from cosmic ray spallation neutrons is expected to be a particular asset in light construction materials.

Composites containing BN

Addition of boron nitride to silicon nitride ceramics improves the thermal shock resistance of the resulting material. For the same purpose, BN is added also to silicon nitride-alumina and titanium nitride-alumina ceramics. Other materials being reinforced with BN are, e.g., alumina and zirconia, borosilicate glasses, glass ceramics, enamels, and composite ceramics with titanium boride-boron nitride and titanium boride-aluminium nitride-boron nitride and silicon carbide-boron nitride composition.

Cadmium oxide

Cadmium oxide	
	
	
IUPAC name Cadmium oxide	
Other names Cadmium(II) oxide, Cadmium monoxide	
Identifiers	
CAS number	1306-19-0 ✓
ChemSpider	14099 ✓
EC number	215-146-2
UN number	2570
RTECS number	EV1925000
Properties	
Molecular formula	CdO

Molar mass	128.41 g mol^{-1}
Appearance	colorless powder (alpha form) red-brown crystal (beta form)
Density	8.150 g/cm^3(crystalline), 6.95 g/cm^3 (amorphous) solid.
Melting point	900-1000 °C (decomposition of amorphous form)
Boiling point	1559 °C (sublimation)
Solubility in water	insoluble
Solubility in acid and alkaline	degrades
Electron mobility	531 cm^2/V s
Magnetic susceptibility	-3×10^{-5} cm^3/mol
Thermal conductivity	0.7 W/m-K
Structure	
Crystal structure	cubic, cF8
Space group	Fm3m, No. 225
Lattice constant	a = 4.6958 Å
Hazards	
MSDS	External MSDS
EU Index	048-002-00-0
EU classification	Carc. Cat. 2 Muta. Cat. 3 Repr. Cat. 3 Very toxic (**T+**) Dangerous for the environment (**N**)
R-phrases	R45, R26, R48/23/25, R62, R63, R68, R50/53
S-phrases	S53, S45, S60, S61
Flash point	Non-flammable
Related compounds	
Other anions	Cadmium sulfide Cadmium selenide Cadmium telluride
Other cations	Zinc oxide Mercury oxide

Cadmium oxide is an inorganic compound with the formula CdO. It is one of the main precursors to other cadmium compounds It crystallizes in a cubic rocksalt lattice like sodium chloride, with octahedral cation and anion centers. It occurs naturally as the rare mineral monteponite. Cadmium oxide can be found as a colorless amorphous powder or as brown or red crystals. Cadmium oxide is an n-type semiconductor with a band gap of 2.16 eV at room temperature.

Production and structure

Since cadmium compounds are often found in association with zinc ores, cadmium oxide is a common by-product of zinc refining. It is produced by burning elemental cadmium in air. Pyrolysis of other cadmium compounds, such as the nitrate or the carbonate, also affords this oxide. When pure, it is red but CdO is unusual in being available in many differing colours due to its tendency to form defect structures resulting from anion vacancies. Cadmium oxide is prepared commercially by oxidizing cadmium vapor in air.

Uses

CdO is used as a transparent conductive material, which was prepared as a transparent conducting film back in 1907. Cadmium oxide in the form of thin films has been used in applications such as photodiodes, phototransistors, photovoltaic cells, transparent electrodes, liquid crystal displays, IR detectors, and anti reflection coatings. CdO microparticles undergo bandgap excitation when exposed to UV-A light and is also selective in phenol photodegradation.

Cadmium oxide is used in cadmium plating baths, electrodes for storage batteries, cadmium salts, catalyst, ceramic glazes, phosphors, and nematocide. Major uses for cadmium oxide is as an ingredient for electroplating baths and in pigments. Most commercial electroplating of cadmium is done by electrodeposition from cyanide baths. These cyanide baths consist of cadmium oxide and sodium cyanide in water, which likely form cadmium cyanide and sodium hydroxide. A typical formula is 32 g/L cadmium oxide and 75 g/L sodium cyanide. The cadmium concentration may vary by as much as

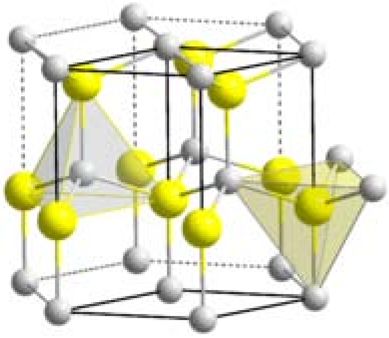
50%. Brighteners are usually added to the bath and the plating is done at room temperature with high purity cadmium anodes.

Reactivity

CdO is a basic oxide and is thus attacked by aqueous acids to give solutions of $[Cd(H_2O)_6]^{2+}$. Upon treatment with strong alkaline solutions, $[Cd(OH)_4]^{2-}$ forms. A thin coat of cadmium oxide forms on the surface of cadmium in moist air at room temperature. Cadmium will oxidize at room temperatures to form CdO. Cadmium vapor and steam will form CdO and hydrogen in a reversible reaction.

Cadmium selenide

Cadmium selenide



IUPAC name
Cadmium selenide
Other names
Cadmium(II) selenide,
Cadmoselite

Identifiers	
CAS number	1306-24-7 ✓
ChemSpider	14101 ✓
Properties	
Molecular formula	CdSe
Molar mass	191.37 g/mol
Appearance	Greenish-brown or dark red solid powder
Density	5.816 g/cm^3, solid
Melting point	1268 °C (1541 K)
Solubility in water	Insoluble
Band gap	1.74 eV (direct)
Refractive index (n_D)	2.5

Structure	
Crystal structure	Wurtzite
Space group	C_{6v}^{4}-$P6_3mc$
Coordination geometry	Tetrahedral
Hazards	
EU Index	048-001-00-5
EU classification	Harmful (**Xn**) Dangerous for the environment (**N**)
R-phrases	R20/21/22, R50/53
S-phrases	(S2), S60, S61
Related compounds	
Other anions	Cadmium oxide, Cadmium sulfide, Cadmium telluride
Other cations	Zinc selenide, Mercury(II) selenide

Cadmium selenide (CdSe) is a solid, binary compound of cadmium and selenium. Common names for this compound are **cadmium(II) selenide**, **cadmium selenide**, and **cadmoselite** (a very rare mineral).

Cadmium selenide is a semiconducting material, but has yet to find many applications in manufacturing. This material is transparent to infra-red (IR) light, and has seen limited use in windows for instruments utilizing IR light.

Much current research on cadmium selenide has focused on nanoparticles. Researchers are concentrating on developing controlled syntheses of CdSe nanoparticles. In addition to synthesis, scientists are working to understand the properties of cadmium selenide, as well as apply these materials in useful ways.

Structure

Three crystalline forms of CdSe are known: wurtzite (hexagonal), sphalerite (cubic) and rock-salt (cubic). The sphalerite CdSe structure is unstable and converts to the wurtzite form upon moderate heating. The transition starts at about 130 °C, and at 700 °C it completes within a day. The rock-salt structure is only observed under high pressure.

Production

The production of cadmium selenide has been carried out in two different ways. The preparation of bulk crystalline CdSe is done by the High-Pressure Vertical Bridgman method or High-Pressure Vertical Zone Melting.

Cadmium selenide may also be produced in the form of nanoparticles. Several methods for the production of CdSe nanoparticles have been developed: arrested precipitation in solution, synthesis in structured media, high temperature pyrolysis, sonochemical, and radiolytic methods are just a few.

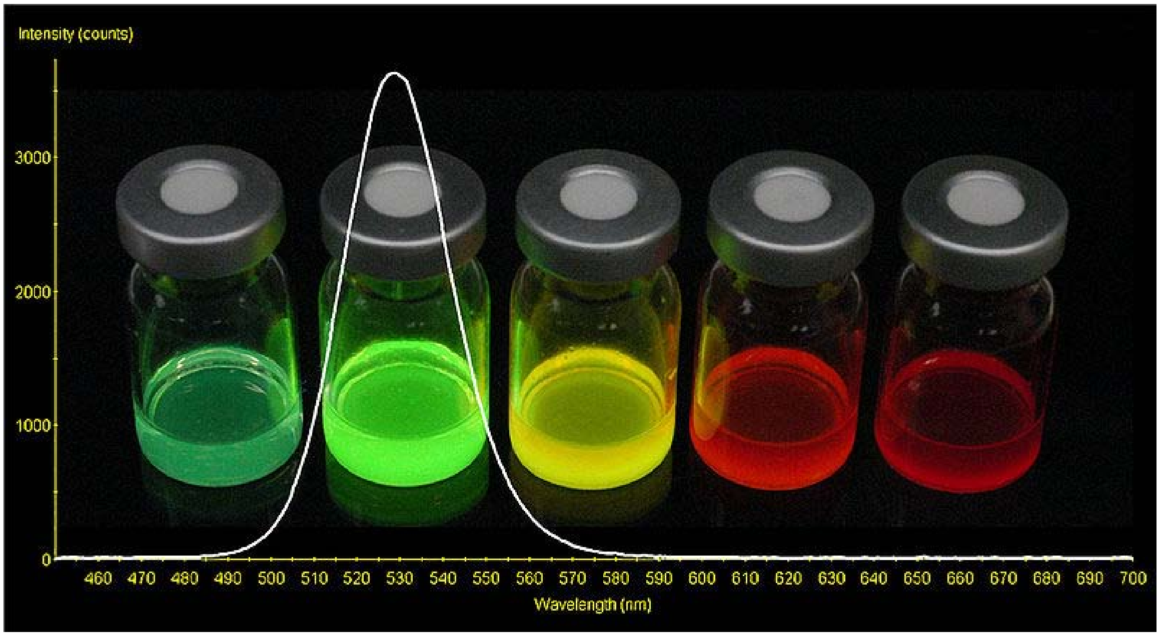
Production of cadmium selenide by arrested precipitation in solution is performed by introducing alkylcadmium and trioctylphosphine selenide (TOPSe) precursors into a heated solvent under controlled conditions.

$$Me_2Cd + TOPSe \rightarrow CdSe + (byproducts).$$

Synthesis in structured environments refers to the production of cadmium selenide in liquid crystal or surfactant solutions. The addition of surfactants to solutions often results in a phase change in the solution leading to a liquid crystallinity. A liquid crystal is similar to a solid crystal in that the solution has long range translational order. Examples of this ordering are layered alternating sheets of solution and surfactant, micelles, or even a hexagonal arrangement of rods.

High temperature pyrolysis synthesis is usually carried out using an aerosol containing a mixture of volatile cadmium and selenium precursors. The precursor aerosol is then carried through a furnace with an inert gas, such as hydrogen, nitrogen, or argon. In the furnace the precursors react to form CdSe as well as several by-products.

Applications



A photograph and representative spectrum of photoluminescence from colloidal CdSe quantum dots excited by UV light.

Cadmium selenide in its wurtzite crystal structure is an important II-VI semiconductor. As a semiconductor CdSe is an n-type semiconductor, which is difficult to dope p-type, however p-type doping has been achieved using nitrogen. CdSe is also being developed for use in opto-electronic devices, laser diodes, nanosensing, and biomedical imaging. They are also used being tested for use in high-efficiency solar cells

Most of the usefulness of CdSe stems from nanoparticles, that is particles with sizes below 100 nm. CdSe particles of this size exhibit a property known as quantum confinement. Quantum confinement results when the electrons in a material are confined to a very small volume. Quantum confinement is size dependent, meaning the properties of CdSe nanoparticles are tunable based on their size.

Since CdSe nanoparticles have a size dependent fluorescence spectrum, they are finding applications in optical devices such as laser diodes. Using these particles, engineers are able to manufacture laser diodes that cover a large part of the electromagnetic spectrum.

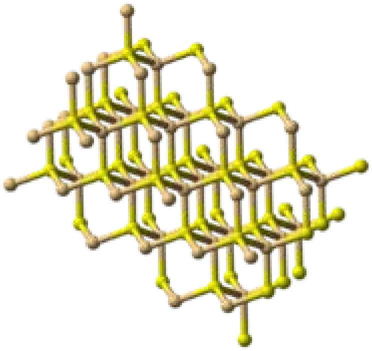
Along similar lines, doctors are developing these materials for use in biomedical imaging applications. Human tissue is permeable to far infra-red light. By injecting appropriately prepared CdSe nanoparticles into injured tissue, it may be possible to image the tissue in those injured areas.

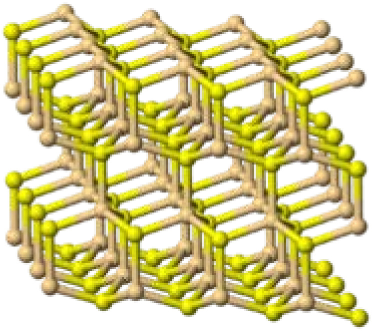
Safety information

Cadmium is a toxic heavy metal and appropriate precautions should be taken when handling it and its compounds. Selenides are toxic in large amounts. .

Cadmium sulfide

Cadmium sulfide





Other names
Cadmium(II) sulfide,
Greenockite
Hawleyite

Identifiers	
CAS number	1306-23-6 ✓
ChemSpider	7969586 ✓
UNII	057EZR4Z7Q ✓
EC number	215-147-8
UN number	2570
RTECS number	EV3150000
Properties	
Molecular formula	CdS
Molar mass	144.46 g/mol
Appearance	Yellow-orange solid.
Density	4.82 g/cm^3, solid.
Melting point	1750 °C at 100 bar (10 MPa)
Boiling point	980 °C *subl.*
Solubility in water	insoluble
Refractive index (n_D)	2.51
Structure	
Crystal structure	Hexagonal, Cubic
Hazards	
MSDS	ICSC 0404
EU Index	048-010-00-4
EU classification	Carc. Cat. 2 Muta. Cat. 3 Repr. Cat. 3 Toxic (**T**) Dangerous for the environment (**N**)
R-phrases	R45, R22, R48/23/25, R62, R63, R68, R50/53
S-phrases	S53, S45, S61
Flash point	Non-flammable
Related compounds	
Other anions	Cadmium oxide Cadmium selenide

Other cations	Zinc sulfide Mercury sulfide

Cadmium sulfide is the inorganic compound with the formula CdS. Cadmium sulfide is yellow solid. It occurs in nature with two different crystal structures as the rare minerals greenockite and hawleyite, but is more prevalent as an impurity substituent in the similarly structured zinc ores sphalerite and wurtzite, which are the major economic sources of cadmium. As a compound that is easy to isolate and purify, it is the principal source of cadmium for all commercial applications.

Production

Cadmium sulfide can be prepared by the precipitation from soluble cadmium(II) salts with sulfide ion and this has been used in the past for gravimetric analysis and qualitative inorganic analysis.

Pigment production usually involves the precipitation of CdS, the washing of the precipitate to remove soluble cadmium salts followed by calcination (roasting) to convert it to the hexagonal form followed by milling to produce a powder. When cadmium sulfide selenides are required the CdSe is co-precipitated with CdS and the cadmium sulfoselenide is created during the calcination step.

Industrially the production of thin films of CdS, required in e.g. photoresistors and chemical bath deposition (CBD), has been investigated using the hydrolysis of thiourea as the source of sulfide anions and an ammonium salt /ammonia buffer solution to control pH:

$$Cd^{2+} + 4\ NH_3 \rightarrow [Cd(NH_3)_4]^{2+}$$
$$(NH_2)_2CS + OH^- \rightarrow SH^- + H_2O + H_2CN_2$$
$$SH^- + Cd^{2+} \rightarrow CdS + H^+$$

Cadmium sulfide can be produced from volatile cadmium alkyls, an example is the reaction of dimethylcadmium with diethyl sulfide to produce a film of CdS using metalorganic vapour phase epitaxy techniques.

The preparative route and the subsequent treatment of the product, affects the polymorphic form that is produced. It has been asserted that chemical precipitation methods result in the cubic zincblende form

Routes to thin films of CdS

Thin films of CdS are components in some photoresistors and solar cells. Various methods have been used to deposit these thin films, for example (note: there is a large body of research in this area and only representative references are given):

- Chemical bath deposition, CBD

- Sol gel techniques
- MOCVD
- Sputtering
- Electrochemical deposition
- Spraying with precursor cadmium salt, sulfur compound and dopant
- Screen printing using a slurry containing dispersed CdS

Reactions

Cadmium sulfide is soluble in (actually degraded by) acids, and this conversion has been investigated as a method of extracting the pigment from waste polymers e.g. HDPE pipes:

$CdS + 2\ HCl \rightarrow CdCl_2 + H_2S$

When sulfide solutions containing dispersed CdS particles are irradiated with light hydrogen gas is generated:

$H_2S \rightarrow H_2 + S\ \Delta H_f = +9.4$ kcal/mol

The proposed mechanism involves the electron/hole pairs created when incident light is absorbed by the cadmium sulfide followed by these reacting with water and sulfide:

Production of an electron hole pair
$CdS + h\nu \rightarrow e^- + hole^+$
Reaction of electron
$2e^- + 2H_2O \rightarrow H_2 + 2OH^-$
Reaction of hole
$2hole^+ + S^{2-} \rightarrow S$

Structure and physical properties

Cadmium sulfide has, like zinc sulfide, two crystal forms; the more stable hexagonal wurtzite structure (found in the mineral Greenockite) and the cubic zinc blende structure (found in the mineral Hawleyite). In both of these forms the cadmium and sulfur atoms are four coordinate. There is also a high pressure form with the NaCl rock salt structure.

Cadmium sulfide is a direct band gap semiconductor (gap 2.42 eV). The magnitude of its band gap means that it appears coloured.
As well as this obvious property others properties result:

- the conductivity increases when irradiated with light (leading to uses as a photoresistor)
- when combined with a p-type semiconductor it forms the core component of a photovoltaic (solar) cell and a CdS/Cu_2S solar cell was one of the first efficient cells to be reported (1954)

- when doped with for example Cu^+ ("activator") and Al^{3+} ("coactivator") CdS luminesces under electron beam excitation (cathodoluminescence) and is used as phosphor
- both polymorphs are piezoelectric and the hexagonal is also pyroelectric
- electroluminescence
- CdS crystal can act as a solid state laser

Applications

CdS is mainly used as a pigment.

CdS and cadmium selenide are used in manufacturing of photoresistors (light dependent resistors) sensitive to visible and near infrared light.

Pigment

CdS is known as cadmium yellow (CI pigment yellow 37). By adding varying amounts of selenium as selenide it is possible to obtain a range of colors for example CI pigment orange 20 and CI pigment red 108.
Synthetic cadmium pigments based on cadmium sulfide are valued for their good thermal stability, light and weather fastness, chemical resistance and high opacity. The general commercial availability of cadmium sulfide from the 1840s lead to its adoption by artists notably Van Gogh, Monet (in his London series and other works) and Matisse (*Bathers by a river* 1916-1919). The presence of cadmium in paints has been used to detect forgeries in paintings alleged to have been produced prior to the 19th century. CdS is used as pigment in plastics.