PHS 463

MATERIAL SCIENCE

Source:

Wikipedia

G. Gottstein. Physcal Foundations of Materials Science, Springer, 2004.

C. Kittel, Introduciton to Solid State Physics, Wiley, eighth edition, 2004.

T. H. Courtney, Mechanical Behavior of Materials, Mcgraw Hill, 2004

Online Lecture Notes and e-materials

1.0 Basic Elements of Elasticity

Deformation of materials: The change in shape and/or size of a continuum body after it undergoes a displacement. Deformations results from *stresses* within the continuum induced by external forces or due to changes in its temperature.

1.1. Stress and strain

1.1.1. Stress is a measure of the average amount of force exerted per unit area.

- A measure of the intensity of the total internal forces acting within a body across imaginary internal surfaces, as a reaction to external applied forces and body forces.

– It was introduced into the theory of elasticity by Cauchy around 1822.

– Stress is a concept that is based on the concept of continuum. In general, stress is expressed as:

$$\sigma = \frac{F}{A}$$

where F is the average stress, also called engineering or nominal stress, and A is the force acting over the area .

1.1.2. Stress cannot be measured directly but is usually inferred from measurements of strain and knowledge of elastic properties of the material.

– Strain is the geometrical measure of deformation representing the relative displacement between particles in the material body.

– Strain defines the amount of stretch or compression along a material line elements or fibers, (normal strain) and the amount of distortion associated with the sliding of plane layers over each other, (shear strain), within a deforming body

– Strain is a dimensionless quantity, which can be expressed as a decimal fraction, a percentage or in parts-per notation.

– The state of strain at a material point of a continuum body is defined as the totality of all the changes in length of material lines or fibers (normal strain) which pass through that point and also the totality of all the changes in the angle between pairs of lines initially perpendicular to each other (shear strain) radiating from this point.

– However, it is sufficient to know the normal and shear components of strain on a set of three mutually perpendicular directions.

- If there is an increase in length of the material line, the normal strain is called tensile strain, otherwise, if there is reduction or compression in the length of the material line, it is called compressive strain.

1.1.3. Engineering Stress

- Tensile stress, σ:

F_t

Area, A_o

F_t
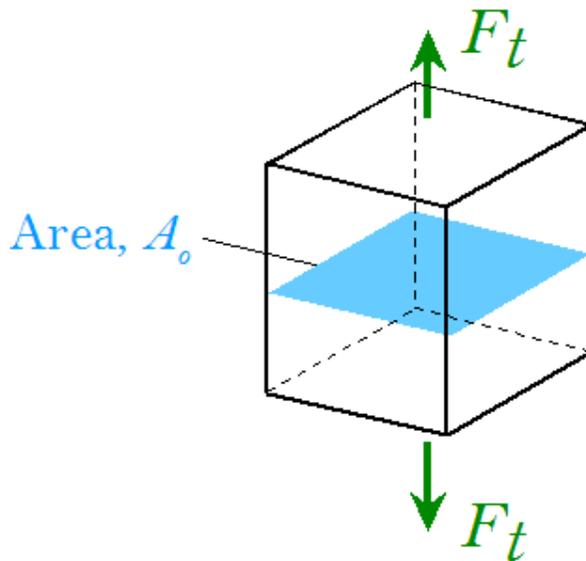
$$\sigma = \frac{F_t}{A_o} = \frac{\text{lb}_f}{\text{in}^2} \quad \text{or} \quad \frac{\text{N}}{\text{m}^2}$$

original area
before loading

- Shear stress, τ:

F_t F

Area, A_o

F_s

F_s

F F_t

$$\tau = \frac{F_s}{A_o}$$
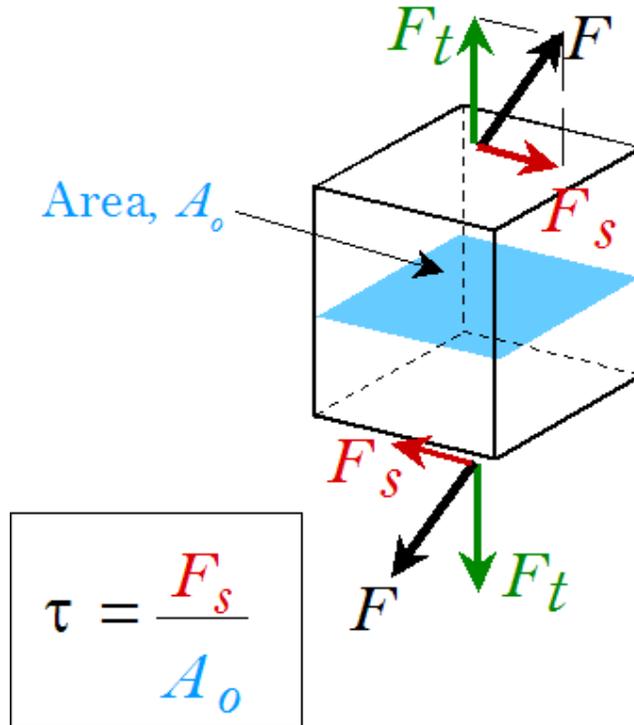
1.1.4. Engineering Strain

The engineering strain is the most common definition applied to materials used in mechanical and structural engineering, which are subjected to very small deformations. The engineering strain is expressed as the ratio of total deformation to the initial dimension of the material body in which the forces are being applied.

- **Tensile strain:**

$$\varepsilon = \frac{\delta}{L_o}$$

- **Lateral strain:**

$$\varepsilon_L = \frac{-\delta_L}{w_o}$$
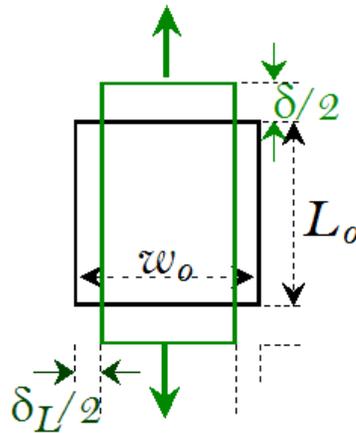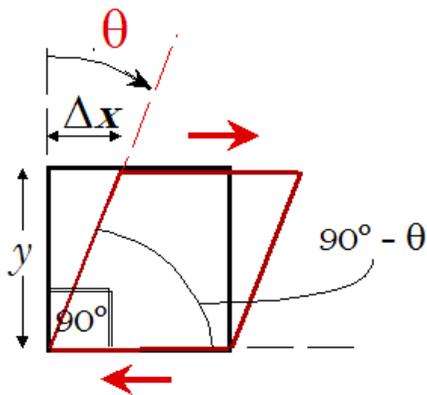
- **Shear strain:**

$$\gamma = \Delta x / y = \tan\theta$$

1.1.5. Hooke's law

- The normal stress σ is proportional to the strain ε. The proportionality constant is Young's modulus E :

$$\frac{F_t}{A} = E\frac{\Delta l}{l_0}$$

- The shear stress τ is proportional to the shear strain γ . The proportionality constant is the shear modulus G:

$$\frac{F_t}{A} = G\frac{\Delta x}{y}$$

2. The Flow curve

2.1. Linear Elastic curve

$\sigma = E\varepsilon$

Elastic means reversible

2.2 Plastic deformation

Stress-strain curve:

A: Elastic limit

B: Upper Yield Stress

C: Lower Yield Stress

D: Ultimate Yield Stress

E: Break Stress
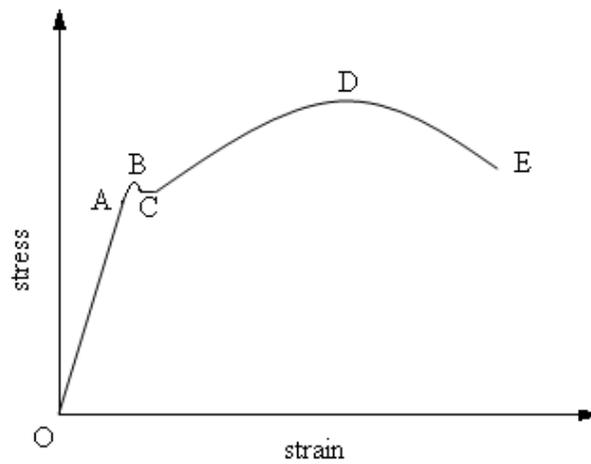
Plastic (Permanent) Deformation

(at lower temperatures, i.e. $T < T_{melt}/3$)

• Simple tension test:

engineering stress, $\square$

Elastic+Plastic
at larger stress

Elastic
initially

permanent (plastic)
after load is removed

ε_p

engineering strain, $\square$
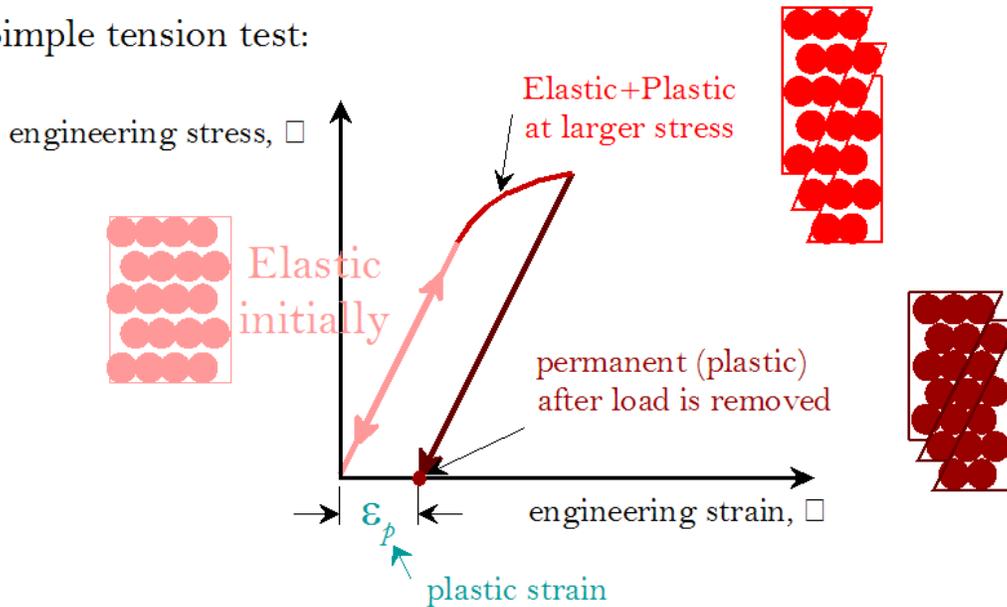
plastic strain

Ductility: Plastic strain at failure.

- brittle materials $\square$ $f \leq 0.1\%$ (ceramics, hard materials)

- ductile materials $\square f \approx 10\%$ (metals and commertial alloys)

- superplastic materials $\square$ $f \approx 1000\%$ (special fine graind materials)

True stress and strain

Note: S.A. changes when sample stretched

- True stress $\qquad \sigma_T = F/A_i$ $\qquad \boxed{\sigma_T = \sigma(1+\varepsilon)}$
- True strain $\qquad \varepsilon_T = \ln(\ell_i/\ell_o)$ $\qquad \boxed{\varepsilon_T = \ln(1+\varepsilon)}$
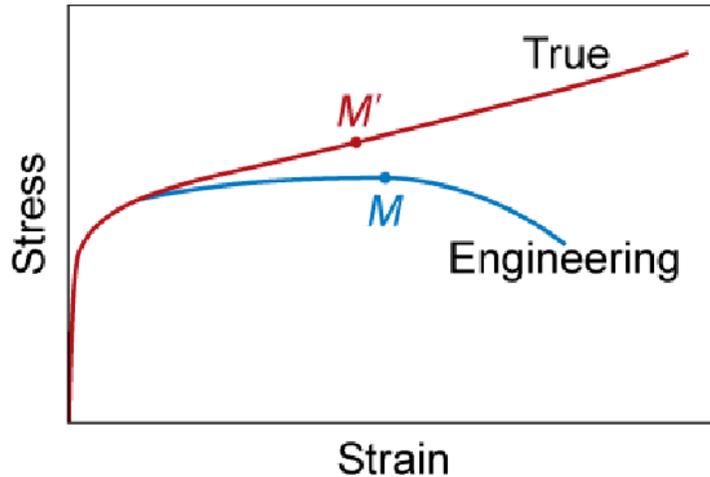
The true stress-strain curve can be calculated from the engineering stress-strain curve.

$F=\sigma t \Box q$

$$\frac{dF}{d\varepsilon_t} = q.\frac{d\sigma_t}{d\varepsilon_t} = \sigma \frac{dq}{d\varepsilon_t}$$

The stability of plastic deformation is determined by physical hardening (slope, $.\frac{d\sigma_t}{d\varepsilon_t}$)

and geometrical softening ($\frac{dq}{d\varepsilon}$). This is used to explain necking.
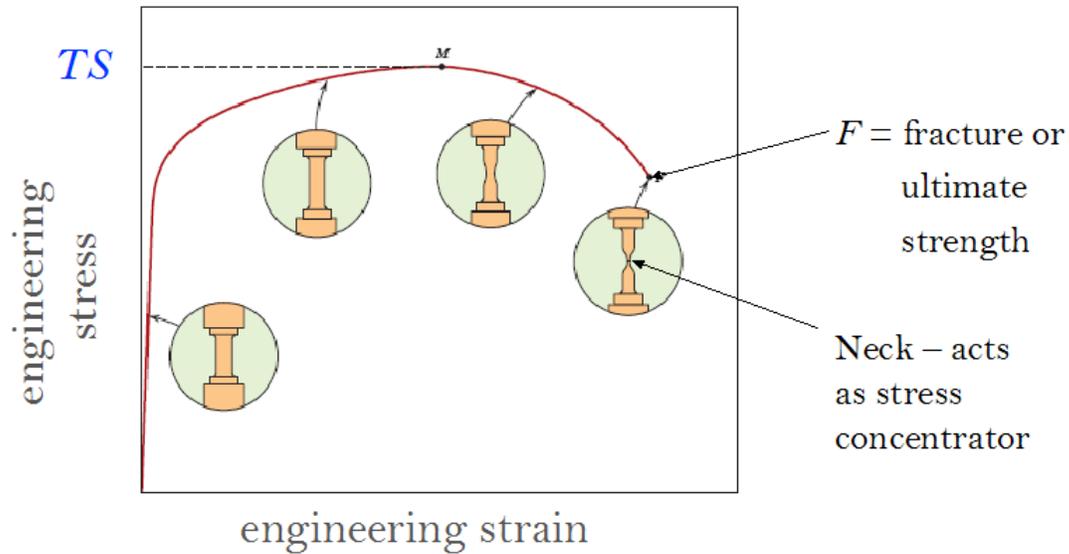
3.0 Defects in Metal Structures

Imperfections

Imperfections in the lattice structures of metals known as defects have a considerable effect on the corrosion resistance properties.

Types of defects within the grains:

3.1 Point defects: single atom defect
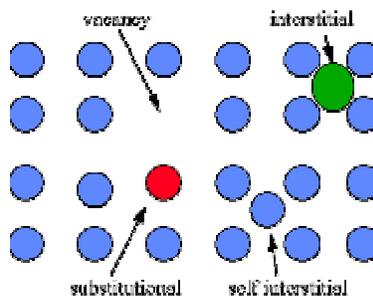
- vacancies - a missing atom
- substitutional defects a foreign atom
- interstitial defects a squeezed atom

3.2 Line defects: -planes of atoms not perfectly fitted into the lattice (dislocations)
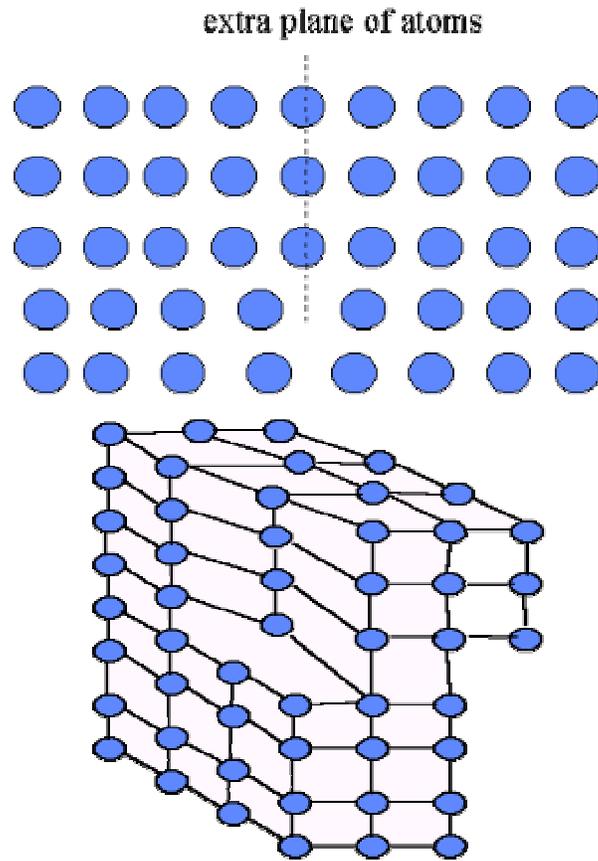
edge dislocations screw dislocation

3.3 Volume defects

- voids: holes in the materials
- cracks introduced during processing
- inclusions particles of foreign matter embedded in the solid.

Volume defects play an important role in corrosion mechanisms.

4.0 Griffith's Criterion

Fracture mechanics was developed during World War I by English aeronautical engineer, A. A. Griffith, to explain the failure of brittle materials.[1] Griffith's work was motivated by two contradictory facts:

- The stress needed to fracture bulk glass is around 100 MPa (15,000 psi).
- The theoretical stress needed for breaking atomic bonds is approximately 10,000 MPa (1,500,000 psi).

A theory was needed to reconcile these conflicting observations. Also, experiments on glass fibers that Griffith himself conducted suggested that the fracture stress increases as the fiber diameter decreases. Hence the uniaxial tensile strength, which had been used extensively to predict material failure before Griffith, could not be a specimen-independent material property. Griffith suggested that the low fracture strength observed in experiments, as well as the size-dependence of strength, was due to the presence of microscopic flaws in the bulk material.

To verify the flaw hypothesis, Griffith introduced an artificial flaw in his experimental specimens. The artificial flaw was in the form of a surface crack which was much larger than other flaws in a specimen. The experiments showed that the product of the square root of the flaw length (a) and the stress at fracture (σ_f) was nearly constant, which is expressed by the equation:

$$\sigma_f \sqrt{a} \approx C$$

An explanation of this relation in terms of linear elasticity theory is problematic. Linear elasticity theory predicts that stress (and hence the strain) at the tip of a sharp flaw in a linear elastic material is infinite. To avoid that problem, Griffith developed a thermodynamic approach to explain the relation that he observed.

The growth of a crack requires the creation of two new surfaces and hence an increase in the surface energy. Griffith found an expression for the constant C in terms of the surface energy of the crack by solving the elasticity problem of a finite crack in an elastic plate. Briefly, the approach was:

- Compute the potential energy stored in a perfect specimen under an uniaxial tensile load.
- Fix the boundary so that the applied load does no work and then introduce a crack into the specimen. The crack relaxes the stress and hence reduces the elastic energy near the crack faces. On the other hand, the crack increases the total surface energy of the specimen.
- Compute the change in the free energy (surface energy – elastic energy) as a function of the crack length. Failure occurs when the free energy attains a peak value at a critical crack length, beyond which the free energy decreases by increasing the crack length, i.e. by causing fracture. Using this procedure, Griffith found that

$$C = \sqrt{\frac{2E\gamma}{\pi}}$$

where E is the Young's modulus of the material and γ is the surface energy density of the material. Assuming $E = 62$ GPa and $\gamma = 1$ J/m^2 gives excellent agreement of Griffith's predicted fracture stress with experimental results for glass.

Irwin's modification

Plastic zone

Plastic zone

Crack

Crack

Plane Strain

Plane Stress
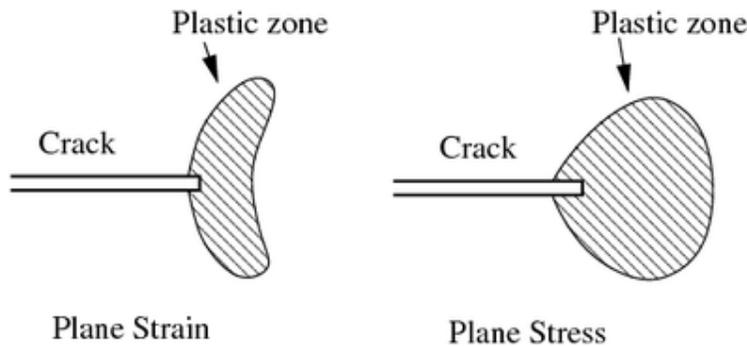
The plastic zone around a crack tip in a ductile material.

Griffith's work was largely ignored by the engineering community until the early 1950s. The reasons for this appear to be (a) in the actual structural materials the level of energy needed to cause fracture is orders of magnitude higher than the corresponding surface energy, and (b) in structural materials there are always some inelastic deformations around the crack front that

would make the assumption of linear elastic medium with infinite stresses at the crack tip highly unrealistic (F. Erdogan (2000)).

Griffith's theory provides excellent agreement with experimental data for brittle materials such as glass. For ductile materials such as steel, though the relation $\sigma_y\sqrt{a} = C$ still holds, the surface energy (γ) predicted by Griffith's theory is usually unrealistically high. A group working under G. R. Irwin[3] at the U.S. Naval Research Laboratory (NRL) during World War II realized that plasticity must play a significant role in the fracture of ductile materials.

In ductile materials (and even in materials that appear to be brittle[4]), a plastic zone develops at the tip of the crack. As the applied load increases, the plastic zone increases in size until the crack grows and the material behind the crack tip unloads. The plastic loading and unloading cycle near the crack tip leads to the dissipation of energy as heat. Hence, a dissipative term has to be added to the energy balance relation devised by Griffith for brittle materials. In physical terms, additional energy is needed for crack growth in ductile materials when compared to brittle materials.

Irwin's strategy was to partition the energy into two parts:

- the stored elastic strain energy which is released as a crack grows. This is the thermodynamic driving force for fracture.
- the dissipated energy which includes plastic dissipation and the surface energy (and any other dissipative forces that may be at work). The dissipated energy provides the thermodynamic resistance to fracture. Then the total energy dissipated is

$G = 2\gamma + G_p$

where γ is the surface energy and G_p is the plastic dissipation (and dissipation from other sources) per unit area of crack growth. The modified version of Griffith's energy criterion can then be written as

$$\sigma_f\sqrt{a} = \sqrt{\frac{E\,G}{\pi}}.$$

For brittle materials such as glass, the surface energy term dominates and $G \approx 2\gamma = 2 \ J/m^2$. For ductile materials such as steel, the plastic dissipation term dominates and $G \approx G_p = 1000 \ J/m^2$. For polymers close to the glass transition temperature, we have intermediate values of $G \approx 2 - 1000 \ J/m^2$.

5.0 Introduction to Diffusion

Mass transport in a gas or liquid generally involves the flow of fluid (e.g. convection currents) although atoms also diffuse. Solids on the other hand, can support shear stresses and hence do not flow except by diffusion involving the jumping of atoms on a fixed network of sites. Assume that such jumps can somehow be achieved in the solid state, with a frequency υ with each jump over a distance $_{\text{ʒ}}$.

$$\bar{x} = \lambda\sqrt{n}$$
$$= \lambda\sqrt{\upsilon t}$$
$$diffusion = dis\tan ce \propto \sqrt{t}$$

Concentration of solute, C, number m^{-3}

Each plane has $C_{\text{ʒ}}$ atoms m^{-2}

$$\delta C = \lambda \left\{ \frac{1}{\partial x} \frac{\partial C}{} \right\}$$

$Atomic flux, J, atoms(m^{-2}s^{-1})$

$$J_{L \to R} = \frac{1}{6} v C \lambda$$

$$J_{R \to L} = \frac{1}{6} v (C + \delta C) \lambda$$

$Net flux along x =$

$$J_{net} = -\frac{1}{6} v \delta C \lambda$$

$$= -\frac{1}{6} v \lambda^2 \left\{ \frac{\partial C}{\partial x} \right\}$$

$$= -D \left\{ \frac{\partial C}{\partial x} \right\}$$

This is Fick's first law where the constant of proportionality is called the diffusion coefficient in m^2 s^{-1}. Fick's first law applies to steady state flux in a uniform concentration gradient. Thus, our equation for the mean diffusion distance can now be expressed in terms of the diffusivity as **Fick's first law of Diffussion**

$$\bar{x} = \lambda \sqrt{t},$$

$$D = \frac{1}{6} v^2$$

$giving$

$$\bar{x} = \sqrt{6Dt} \approx \sqrt{Dt}$$

Non–Uniform Concentration Gradients Suppose that the concentration gradient is not uniform

$$Flux(in) = -D\left\{\frac{\partial C}{\partial x}\right\}_1$$

$$Flux(out) = -D\left\{\frac{\partial C}{\partial x}\right\}_2$$

$$= -D\left[\left\{\frac{\partial C}{\partial x}\right\}_1 + \partial x\left\{\frac{\partial^2 C}{\partial x^2}\right\}\right]$$

In the time interval ±t, the concentration changes ±C

$$\partial C \partial x = \left[Flux(in) - Flux(out)\right]\partial t$$

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2}$$

assuming that the diffusivity is independent of the concentration. This is **Fick's second law of diffusion**. This is amenable to numerical solutions for the general case but there are a couple of interesting analytical solutions for particular boundary conditions. For a case where a fixed quantity of solute is plated onto a semi–infinite bar

boundary conditions:

$$\int_0^\infty C\{x,t\}dx = B$$

$$C\{x,t = 0\} = 0$$

$$C\{x,t\} = \frac{B}{\sqrt{\pi Dt}}\exp\left\{\frac{-x^2}{4Dt}\right\}$$

Now imagine that we create the diffusion couple illustrated, by stacking an infinite set of thin sources on the end of one of the bars. Diffusion can thus be treated by taking a whole set of the exponential functions obtained above, each slightly displaced along the x axis, and asumming (integrating) up their individual effects. The integral is in fact the error function

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\{-u^2\} du$$

so the solution to the diffusion equation is

boundary conditions:

$$C\{x = 0, t\} = C_8$$

$$C\{x, t = 0\} = C_0$$

$$C\{x, t\} = C_8 - (C_8 - C_0) erf\left\{\frac{x}{2\sqrt{Dt}}\right\}$$

This solution can be used in many circumstances where the surface concentration is maintained constant, for example in the carburisation or decarburisation processes (the concentration profiles would be the same as in Fig. 4, but with only one half of the couple illustrated). The solutions described here apply also to the diffusion of heat. Mechanism of Diffusion. Atoms in the solid–state migrate by jumping into vacancies. The vacancies may be interstitial or in substitutional sites. There is nevertheless, a barrier to the motion of the atoms because the motion is associated with a transient distortion of the lattice. Assuming that the atom attempts jumps at a frequency $^o0$, the frequency of successful jumps is given by:

$$v = v_0 \exp\left\{-\frac{|G^*}{KT}\right\}$$

$$\equiv v_0 \exp\left\{\frac{S^*}{k}\right\} \times \exp\left\{-\frac{H^*}{kT}\right\}$$

where k and T are the Boltzmann constant and the absolute temperature respectively, and H* and S* the activation enthalpy and activation entropy respectively. Since

$$D \propto v$$

and

$$D = D_0 \exp\left(-\frac{H^*}{kT}\right)$$

A plot of the logarithm of D versus 1/T should therefore give a straight line, the slope of which is $-H^*/k$. Note that H^ϵ is frequently called the activation energy for diffusion and is often designated Q.

The activation enthalpy of diffusion can be separated into two components, one the enthalpy of migration (due to distortions) and the enthalpy of formation of a vacancy in an adjacent site. After all, for the atom to jump it is necessary to have a vacant site; the equilibrium concentration of vacancies can be very small in solids. Since there are many more interstitial vacancies, and since most interstitial sites are vacant interstitial atoms diffuse far more rapidly than substitutional solutes.

5.1 Kirkendall Effect

Diffusion is at first sight difficult to appreciate for the solid state. A number of mechanisms have been proposed historically. This includes a variety of ring mechanisms where atoms simply swap positions, but controversy remained because the strain energies associated with such swaps made the theories uncertain. One possibility is that diffusion occurs by atoms jumping into vacancies. But the equilibrium concentration of vacancies is typically 10^{-6}, which is very small. The theory was therefore not generally accepted until an elegant experiment by Smigelskas and Kirkendall. The experiment applies to solids as well as cible liquids. Consider a couple made from A and B. If the diffusion fluxes of the two elements are different ($|JA| > |JB|$) then there will be a net flow of matter past the inert markers, causing the couple to shift bodily relative to the markers. This can only happen if diffusion is by a vacancy mechanism. An observer located at the markers will see not only a change in concentration due to intrinsic diffusion, but also because of the Kirkendall flow of matter past the markers. The net effect is described by the usual Fick's laws, but with an interdiffusion coefficient D which is a weighted average of the two intrinsic diffusion coefficients:

$$\overline{D} = X_B D_A + X_A D_B$$

where X represents a mole fraction. It is the interdiffusion coefficient that is measured in most experiments.

Diffusion occurs in response to a concentration gradient expressed as the change in concentration due to a change in position, $\partial C/\partial x$. The local rule for movement or flux J is given by Fick's 1st law of diffusion:

$$J = -\chi \frac{\partial C}{\partial x}$$

in which the flux J [cm^{-2} s^{-1}] is proportional to the diffusivity χ [cm^2/s] and the negative gradient of concentration, $\partial C/\partial x$ [cm^{-3} cm^{-1}] or [cm^{-4}]. The negative sign indicates that J is positive when movement is down the gradient, i.e., the negative sign cancels the negative gradient along the direction of positive flux.
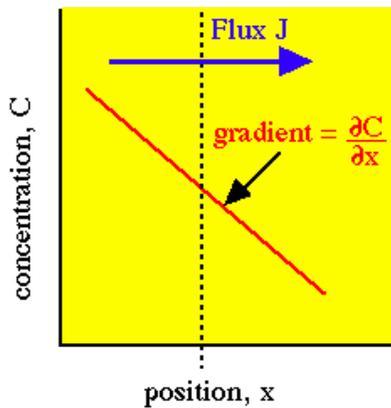
The flux J is driven by the negative gradient $\partial C/\partial x$ in the direction of increasing x. For light, the diffusivity is proportional to the diffusion length D [cm] and the speed of light c:

$$\chi = cD$$

where D = 1/(3 μ_s(1-g)). The units of velocity [cm/s] times the units of length [cm] yield the units of diffusivity [cm^2/s].

For optical diffusion, Fick's 1st law is expressed as the energy flux J [W cm^{-2}] proportional to the diffusion constant D [cm] and the negative fluence gradient dF/dx:

$$J = -D\frac{\partial F}{\partial x}$$

which was obtained by substituting cD for ϰ and substituting F/c for C. The factors c and 1/c cancel to yield the above equation.

6.0 Corrossion

Corrosion can be categorized in some common types

- uniform corrosion
- pitting corrosion
- galvanic corrosion
- crevice corrosion
- concentration cell corrosion
- Microbial corrosion

Uniform or General Corrosion

The metal loss is uniform from the surface and it is usually combined with high-velocity fluid erosion, with or without abrasives.

Pitting Corrosion

Certain conditions, such as low concentrations of oxygen or high concentrations of species such as chloride which compete as anions, can interfere with a given alloy's ability to re-form a passivating film. In the worst case, almost all of the surface will remain protected, but tiny local fluctuations will degrade the oxide film in a few critical points. Corrosion at these points will be greatly amplified, and can cause *corrosion pits* of several types, depending upon conditions. While the corrosion pits only nucleate under fairly extreme circumstances, they can continue to grow even when conditions return to normal, since the interior of a pit is naturally deprived of oxygen and locally the pH decreases to very low values and the corrosion rate increases due to an auto-catalytic process. In extreme cases, the sharp tips of extremely long and narrow corrosion pits can cause stress concentration to the point that

otherwise tough alloys can shatter; a thin film pierced by an invisibly small hole can hide a thumb sized pit from view. These problems are especially dangerous because they are difficult to detect before a part or structure fails. Pitting remains among the most common and damaging forms of corrosion in passivated alloys, but it can be prevented by control of the alloy's environment.

Galvanic Corrosion

Galvanic corrosion occurs when two different metals and/or alloys have electrical contact with each other and are immersed in an electrolyte. This effect is a galvanic couple where the more active metal corrodes at an accelerated rate and the more noble metal corrodes at a retarded rate. When immersed, neither metal would normally corrode as quickly without the electrically conductive connection. What type of metal(s) to use is readily determined by following the galvanic series. For example, zinc is often used as a sacrificial anode for steel structures. Galvanic corrosion is of major interest to the marine industry and also anywhere water (via impurities such as salt) contacts pipes or metal structures. Factors such as relative size of anode, types of metal, and operating conditions (temperature, humidity, salinity, etc.) affect galvanic corrosion. The surface area ratio of the anode and cathode directly affects the corrosion rates of the materials. Galvanic corrosion is often utilized in sacrificial anodes.

Crevice Corrosion

Crevice corrosion is a localized form of corrosion occurring in spaces to which the access of the working fluid from the environment is limited and a concentration cell, areas with different oxygen concentration, will take place with consequent high corrosion rate . These spaces are generally called crevices. Examples of crevices are gaps and contact areas between parts, under gaskets or seals, inside cracks and seams, spaces filled with deposits and under sludge piles. Crevice Corrosion occurs at places with gaskets, bolts and lap joints where crevice exists and usually creates pits similar to pitting corrosion.

Concentration Cell Corrosion

Occurs where the surface is exposed to an electrolytic environment where the concentration of the corrosive fluid or the dissolved oxygen varies. It is usually combined with stagnant fluid or in areas with low fluid velocity.

Microbial Corrosion

Microbial corrosion, or bacterial corrosion, is a corrosion caused or promoted by microorganisms, usually chemoautotrophs. It can apply to both metals and non-metallic materials, in both the presence and lack of oxygen. Sulfate-reducing bacteria are common in lack of oxygen; they produce hydrogen sulfide, causing sulfide stress cracking. In presence of oxygen, some bacteria directly oxidize iron to iron oxides and hydroxides, other bacteria oxidize sulfur and produce sulfuric acid causing biogenic sulfide corrosion. Concentration cells can form in the deposits of corrosion products, causing and enhancing galvanic corrosion.

Accelerated Low Water Corrosion (ALWC) is a particularly aggressive form of MIC that affects steel piles in seawater near the low water tide mark. It is characterised by an orange sludge, which smells of hydrogen sulphide when treated with acid. Corrosion rates can be very high and design corrosion allowances can soon be exceeded leading to premature failure of the steel pile.[1] Piles that have been coating and have cathodic protection installed at the time of construction are not susceptible to ALWC. For unprotected piles, sacrificial anodes can be installed local to the affected areas to inhibit the corrosion or a complete retrofitted sacrificial anode system can be installed. Affected areas can also be treated electrochemically by using an electrode to first produce chlorine to kill the bacteria, and then to produced a calcareous deposit, which will help shield the metal from further attack.

6.1 Corrosion Removal

Often it is possible to chemically remove the products of corrosion to give a clean surface, but one that may exhibit artifacts of corrosion such as pitting. For example phosphoric acid in the form of naval jelly is often applied to ferrous tools or surfaces to remove rust. Corrosion removal should not be confused with Electropolishing which removes some layers of the underlying metal to make a smooth surface. For example phosphoric acid (again) may be

used to electropolish copper but it does this by removing copper, not the products of copper corrosion.

6.2 Resistance to corrosion

Some metals are more intrinsically resistant to corrosion than others, either due to the fundamental nature of the electrochemical processes involved or due to the details of how reaction products form. For some examples, see galvanic series. If a more susceptible material is used, many techniques can be applied during an item's manufacture and use to protect its materials from damage.

7.0 Electrical Properties

7.1 Ohm's Law

When an electric potential V is applied across a material, a current of magnitude I flows. In most metals, at low values of V, the current is proportional to V, according to Ohm's law:

$$I = V/R$$

where R is the electrical resistance. R depends on the intrinsic *resistivity* ρ of the material and on the geometry (length l and area A through which the current passes).

$$R = \rho l/A$$

7.2 Electrical Conductivity

The *electrical conductivity* is the inverse of the resistivity: $\sigma = 1/\rho$.

The electric field in the material is $E=V/l$, Ohm's law can then be expressed in terms of the current density $j = I/A$ as:

$$j = \sigma E$$

The conductivity is one of the properties of materials that varies most widely, from 10^7 (Ω-m) typical of metals to 10^{-20} (Ω-m) for good electrical insulators. Semiconductors have conductivities in the range 10^{-6} to 10^4 (Ω-m).

7.3 Electronic and Ionic Conduction

In metals, the current is carried by electrons, and hence the name *electronic conduction*. In ionic crystals, the charge carriers are ions, thus the name *ionic conduction* (see Sect. 19.15).

7.4 Energy Band Structures in Solids

When atoms come together to form a solid, their valence electrons interact due to Coulomb forces, and they also feel the electric field produced by their own nucleus and that of the other atoms. In addition, two specific quantum mechanical effects happen. First, by Heisenberg's uncertainty principle, constraining the electrons to a small volume raises their energy, this is called *promotion.* The second effect, due to the Pauli exclusion principle, limits the number of electrons that can have the same property (which include the energy). As a result of all these effects, the valence electrons of atoms form wide valence bands when they form a solid. The bands are separated by gaps, where electrons cannot exist. The precise location of the bands and band gaps depends on the type of atom (e.g., Si vs. Al), the distance between atoms in the solid, and the atomic arrangement (e.g., carbon vs. diamond).

In semiconductors and insulators, the valence band is filled, and no more electrons can be added, following Pauli's principle. Electrical conduction requires that electrons be able to gain energy in an electric field; this is not possible in these materials because that would imply that the electrons are promoted into the forbidden band gap.

In metals, the electrons occupy states up to the *Fermi level*. Conduction occurs by promoting electrons into the *conduction band*, that starts at the Fermi level, separated by the valence band by an infinitesimal amount.

7.5 Conduction in Terms of Band and Atomic Bonding Models

Conduction in metals is by electrons in the conduction band. Conduction in insulators is by electrons in the conduction band and by holes in the valence band. *Holes* are vacant states in the valence band that are created when an electron is removed.

In metals there are empty states just above the Fermi levels, where electrons can be promoted. The promotion energy is negligibly small so that at any temperature electrons can be found in the conduction band. The number of electrons participating in electrical conduction is extremely small.

In insulators, there is an energy gap between the valence and conduction bands, so energy is needed to promote an electron to the conduction band. This energy may come from heat, or from energetic radiation, like light of sufficiently small wavelength.

A working definition for the difference between semiconductors and insulators is that in semiconductors, electrons can reach the conduction band at ordinary temperatures, where in insulators they cannot. The probability that an electron reaches the conduction band is about $\exp(-E_g/2kT)$ where E_g is the band gap and kT has the usual meaning. If this probability is, say, $< 10^{-24}$ one would not find a single electron in the conduction band in a solid of 1 cubic centimeter. This requires $E_g/2kT > 55$. At room temperature, $2kT = 0.05$ eV; thus $E_g > 2.8$ eV can be used as the condition for an insulator.

Besides having relatively small E_g, semiconductors have covalent bond, whereas insulators usually are partially ionic bonded.

7.6 Electron Mobility

Electrons are accelerated in an electric field E, in the opposite direction to the field because of their negative charge. The force acting on the electron is $-eE$, where e is the electric charge. This force produces a constant acceleration so that, in the absence of obstacles (in vacuum, like inside a TV tube) the electron speeds up continuously in an electric field. In a solid, the situation is different. The electrons scatter by collisions with atoms and vacancies that change drastically their direction of motion. Thus electrons move randomly but with a net drift in the direction opposite to the electric field. The drift velocity is *constant*, equal to the electric field times a constant called the *mobility μ*,

$$v_d = -\mu_e E$$

which means that there is a friction force proportional to velocity. This friction translates into energy that goes into the lattice as heat. This is the way that electric heaters work.

The electrical conductivity is:

$$\sigma = n \, |e| \, \mu_e$$

where n is the concentration of electrons (n is used to indicate that the carriers of electricity are negative particles).

7.7 Electrical Resistivity of Metals

The resistivity then depends on collisions. Quantum mechanics tells us that electrons behave like waves. One of the effects of this is that electrons do not scatter from a perfect lattice. They scatter by defects, which can be:

- o atoms displaced by lattice vibrations
- o vacancies and interstitials
- o dislocations, grain boundaries
- o impurities

One can express the total resistivity ρ_{tot} by the Matthiessen rule, as a sum of resistivities due to thermal vibrations, impurities and dislocations. Fig. 19.8 illustrates how the resistivity increases with temperature, with deformation, and with alloying.

7.8 Electrical Characteristics of Commercial Alloys

The best material for electrical conduction (lower resistivity) is silver. Since it is very expensive, copper is preferred, at an only modest increase in ρ. To achieve low ρ it is necessary to remove gases occluded in the metal during fabrication. Copper is soft so, for applications where mechanical strength is important, the alloy CuBe is used, which has a nearly as good ρ. When weight is important one uses Al, which is half as good as Cu. Al is also more resistant to corrosion.

When high resistivity materials are needed, like in electrical heaters, especially those that operate at high temperature, nichrome (NiCr) or graphite are used.

7.9 Intrinsic Semiconduction

Semiconductors can be *intrinsic* or *extrinsic*. Intrinsic means that electrical conductivity does not depend on impurities, thus intrinsic means pure. In extrinsic semiconductors the conductivity depends on the concentration of impurities.

Conduction is by electrons and holes. In an electric field, electrons and holes move in opposite direction because they have opposite charges. The conductivity of an intrinsic semiconductor is:

$$\sigma = n \, |e| \, \mu_e + p \, |e| \, \mu_h$$

where p is the hole concentration and μ_h the hole mobility. One finds that electrons move much faster than holes:

$$\mu_e > \mu_h$$

In an intrinsic semiconductor, a hole is produced by the promotion of each electron to the conduction band. Thus:

$$n = p$$

Thus, $\sigma = 2 \, n \, |e| \, (\mu_e + \mu_h)$ (only for intrinsic semiconductors).

7.10 Extrinsic Semiconduction

Unlike intrinsic semiconductors, an extrinsic semiconductor may have different concentrations of holes and electrons. It is called *p-type* if $p>n$ and *n-type* if $n>p$. They are made by *doping*, the addition of a very small concentration of impurity atoms. Two common methods of doping are diffusion and ion implantation.

Excess electron carriers are produced by substitutional impurities that have more valence electron per atom than the semiconductor matrix. For instance phosphorous, with 5 valence electrons, is an electron *donor* in Si since only 4 electrons are used to bond to the Si lattice when it substitutes for a Si atom. Thus, elements in columns V and VI of the periodic table are donors for semiconductors in the IV column, Si and Ge. The energy level of the donor state is close to the conduction band, so that the electron is promoted (ionized) easily at room

temperature, leaving a hole (the ionized donor) behind. Since this hole is unlike a hole in the matrix, it does not move easily by capturing electrons from adjacent atoms. This means that the conduction occurs mainly by the donated electrons (thus *n-type*).

Excess holes are produced by substitutional impurities that have fewer valence electrons per atom than the matrix. This is the case of elements of group II and III in column IV semiconductors, like B in Si. The bond with the neighbors is incomplete and so they can capture or accept electrons from adjacent silicon atoms. They are called *acceptors*. The energy level of the acceptor is close to the valence band, so that an electron may easily hop from the valence band to complete the bond leaving a hole behind. This means that conduction occurs mainly by the holes (thus *p-type*).

7.11 The Temperature Variation of Conductivity and Carrier Concentration

Temperature causes electrons to be promoted to the conduction band and from donor levels, or holes to acceptor levels. The dependence of conductivity on temperature is like other thermally activated processes:

$\sigma = A \exp(-E_g/2kT)$

where *A* is a constant (the mobility varies much more slowly with temperature). Plotting *ln* σ vs. $1/T$ produces a straight line of slope $E_g/2k$ from which the band gap energy can be determined. Extrinsic semiconductors have, in addition to this dependence, one due to the thermal promotion of electrons from donor levels or holes from acceptor levels. The dependence on temperature is also exponential but it eventually saturates at high temperatures where all the donors are emptied or all the acceptors are filled.

This means that at low temperatures, extrinsic semiconductors have larger conductivity than intrinsic semiconductors. At high temperatures, both the impurity levels and valence electrons are ionized, but since the impurities are very low in number and they are exhausted, eventually the behavior is dominated by the intrinsic type of conductivity.

7.12 Semiconductor Devices

A semiconductor *diode* is made by the intimate junction of a *p*-type and an *n*-type semiconductor (an *n-p junction*). Unlike a metal, the intensity of the electrical current that

passes through the material depends on the polarity of the applied voltage. If the positive side of a battery is connected to the *p*-side, a situation called *forward bias*, a large amount of current can flow since holes and electrons are pushed into the junction region, where they recombine (annihilate). If the polarity of the voltage is flipped, the diode operates under *reverse bias*. Holes and electrons are removed from the region of the junction, which therefore becomes depleted of carriers and behaves like an insulator. For this reason, the current is very small under reverse bias. The asymmetric current-voltage characteristics of diodes (Fig. 19.20) is used to convert alternating current into direct current. This is called rectification.

A *p-n-p junction transistor* contains two diodes back-to-back. The central region is very thin and is called the *base*. A small voltage applied to the base has a large effect on the current passing through the transistor, and this can be used to amplify electrical signals (Fig. 19.22). Another common device is the MOSFET transistor where a *gate* serves the function of the base in a junction transistor. Control of the current through the transistor is by means of the electric field induced by the gate, which is isolated electrically by an oxide layer.

7.13 Conduction in Ionic Materials

In ionic materials, the band gap is too large for thermal electron promotion. Cation vacancies allow ionic motion in the direction of an applied electric field, this is referred to as *ionic conduction.* High temperatures produce more vacancies and higher ionic conductivity.

At low temperatures, electrical conduction in insulators is usually along the surface, due to the deposition of moisture that contains impurity ions.

7.14 Electrical Properties of Polymers

Polymers are usually good insulators but can be made to conduct by doping. Teflon is an exceptionally good insulator.

7.15 Dielectric

A *dielectric* is an electrical insulator that can be made to exhibit an electric dipole structure (displace the negative and positive charge so that their center of gravity is different).

7.16 Capacitance

When two parallel plates of area *A,* separated by a small distance *l*, are charged by $+Q$, $-Q$, an electric field develops between the plates

$$E = D/\varepsilon\varepsilon_0$$

where $D = Q/A$. ε_0 is called the vacuum permittivity and ε the relative permittivity, or *dielectric constant* ($\varepsilon = 1$ for vacuum). In terms of the voltage between the plates, $V = E\,l$,

$$V = Dl/\varepsilon\varepsilon_0 = Q\,l/A\varepsilon\varepsilon_0 = Q\,/\,C$$

The constant $C = A\varepsilon\varepsilon_0/l$ is called the capacitance of the plates.

7.17 Field Vectors and Polarization

The dipole moment of a pair of positive and negative charges ($+q$ and $-q$) separated at a distance *d* is $p = qd$. If an electric field is applied, the dipole tends to align so that the positive charge points in the field direction. Dipoles between the plates of a capacitor will produce an electric field that opposes the applied field. For a given applied voltage *V*, there will be an increase in the charge in the plates by an amount Q' so that the total charge becomes $Q = Q' + Q_0$, where Q_0 is the charge of a vacuum capacitor with the same *V*. With $Q' = PA$, the charge density becomes $D = D_0\,E + P$, where the *polarization* $P = \varepsilon_0\,(\varepsilon-1)\,E$.

7.18 Types of Polarization

Three types of polarization can be caused by an electric field:

- *Electronic polarization*: the electrons in atoms are displaced relative to the nucleus.
- *Ionic polarization*: cations and anions in an ionic crystal are displaced with respect to each other.
- *Orientation polarization*: permanent dipoles (like H_2O) are aligned.

7.19 Frequency Dependence of the Dielectric Constant

Electrons have much smaller mass than ions, so they respond more rapidly to a changing electric field. For electric field that oscillates at very high frequencies (such as light) only electronic polarization can occur. At smaller frequencies, the relative displacement of positive and negative ions can occur. Orientation of permanent dipoles, which require the rotation of a molecule can occur only if the oscillation is relatively slow (MHz range or slower). The time needed by the specific polarization to occur is called the *relaxation time*.

7.20 Dielectric Strength

Very high electric fields ($>10^8$ V/m) can free electrons from atoms, and accelerate them to such high energies that they can, in turn, free other electrons, in an avalanche process (or electrical discharge). This is called *dielectric breakdown*, and the field necessary to start the is called the *dielectric strength* or breakdown strength.

7.21 Dielectric Materials

Capacitors require dielectrics of high ε that can function at high frequencies (small relaxation times). Many of the ceramics have these properties, like mica, glass, and porcelain). Polymers usually have lower ε.

7.22 Ferroelectricity

Ferroelectric materials are ceramics that exhibit permanent polarization in the absence of an electric field. This is due to the asymmetric location of positive and negative charges within the unit cell. Two possible arrangements of this asymmetry results in two distinct polarizations, which can be used to code "0" and "1" in ferroelectric memories. A typical ferroelectric is barium titanate, $BaTiO_3$, where the Ti^{4+} is in the center of the unit cell and four O^{2-} in the central plane can be displaced to one side or the other of this central ion (Fig. 19.33).

7.23 Piezoelectricity

In a *piezolectric* material, like quartz, an applied mechanical stress causes electric polarization by the relative displacement of anions and cations.

8.0 Phase Transformations

Phase transformations that involve a change in the microstructure can occur through:

- Diffusion

- Maintaining the type and number of phases (e.g., solidification of a pure metal, allotropic transformation, recrystallization, grain growth.
- Alteration of phase composition (e.g., eutectoid reactions, see 10.5)

- Diffusionless

- Production of metastable phases (e.g., martensitic transformation, see 10.5)

8.1 The Kinetics of Solid-State Reactions

Change in composition implies atomic rearrangement, which requires diffusion. Atoms are displaced by *random walk*. The displacement of a given atom, d, is not linear in time t (as would be for a straight trajectory) but is proportional to the square root of time, due to the tortuous path: $d = c(Dt)^{1/2}$ where c is a constant and D the diffusion constant. This time-dependence of the rate at which the reaction (phase transformation) occurs is what is meant by the term reaction *kinetics*.

D is called a constant because it does not depend on time, but it depends on temperature as we have seen in Ch. 5. Diffusion occurs *faster at high temperatures.*

Phase transformation requires two processes: *nucleation* and *growth.* **Nucleation** involves the formation of very small particles, or *nuclei* (e.g., grain boundaries, defects). This is similar to rain happening when water molecules condensed around dust particles. During **growth**, the nuclei grow in size at the expense of the surrounding material.

The kinetic behavior often has the S-shape form of Fig. 10.1, when plotting percent of material transformed vs. the logarithm of time. The nucleation phase is seen as an incubation period, where nothing seems to happen. Usually the transformation rate has the form $r = A\,e^{-}$

$^{Q/RT}$ (similar to the temperature dependence of the diffusion constant), in which case it is said to be *thermally activated.*

8.2 Multiphase Transformations

To describe phase transformations that occur during cooling, *equilibrium* phase diagrams are inadequate if the transformation rate is slow compared to the cooling rate. This is usually the case in practice, so that equilibrium microstructures are seldom obtained. This means that the transformations are delayed (e.g., case of supercooling), and *metastable* states are formed. We then need to know the effect of **time** on phase transformations.

Textbooks and Reading materials:

Wikipedia

G. Gottstein. Physcal Foundations of Materials Science, Springer, 2004.

C. Kittel, Introduciton to Solid State Physics, Wiley, eighth edition, 2004.

T. H. Courtney, Mechanical Behavior of Materials, Mcgraw Hill, 2004

Online Lecture Notes