Solid and Liquid Phases

Mark Boone

SOLID AND LIQUID PHASES

SOLID AND LIQUID PHASES

Mark Boone

BIBLIOTEX
Digital Library

Solid and Liquid Phases
by Mark Boone

www.bibliotex.com

Ebook ISBN: 9781984666024

BIBLIOTEX
Digital Library

Published by:
Bibliotex
Canada
Website: www.bibliotex.com

Contents

1

Phases Equilibrium: Solid and Liquid

Major Concept Area

Equilibrium. The concept of dynamic equilibrium is an extremely important one in chemistry. We then apply energy and entropy ideas to the physical equilibria established among the phases of one or more pure substances under varying conditions of temperature and pressure.

The concept of dynamic equilibrium is then extended to a variety of situations involving a balance of opposing rates across a boundary between two phases. In all cases we will find that the rate of passage from one phase to another depends on the surface area of contact between the two phases, and the concentration of substance in the phase of origin. This idea will help us to understand equilibria involving 2 (or 3) phases of a single pure substance.

Such equilibria occur all around us in nature and are important in weather and the deterioration of infrastructure. We will then apply the idea to equilibria in which a substance distributes itself between 2 other substances that are in contact with each other; and to equilibria involving the dissolution of molecular and ionic solids and gases in liquids.

Solubility of solids and gases is extremely important. For example, fluoridation of water is done to convert the principal structural component of teeth, apatite ($Ca_5(OH)(PO_4)_3$), to fluoroapatite ($Ca_5(F)(PO_4)_3$), which is both less soluble and more resistant to chemical attack by food components (particularly acids!) than is apatite. The dissolution of gases such as CO_2 and O_2 is essential to the aquatic life cycle.

Specific Concepts

Disorder increases in spontaneous processes: Dynamic equilibrium is a balance of opposing processes occuring at equal rates.

In a closed container, a dynamic equilibrium involving two phases of a pure substance can be established at certain pressures and temperatures; the rate of passage from phase 1 to phase 2 equals the rate of passage from phase 2 to phase 1.

When salts dissolve, a dynamic equilibrium is established between undissolved solid and ions in solution; the rate of dissolution equals the rate of crystallization.

When gases dissolve, a dynamic equilibrium is established between gas above the solution and dissolved gas; the rate of dissolution equals the rate of escape.

Equilibrium Phase Diagrams of Two-component Systems

To construct an equilibrium phase diagram of a binary system, it is a necessary and sufficient condition that the boundaries of one-phase regions be known. In other words, the equilibrium diagram is a plot of solubility relations between components of the system.

It shows the number and composition of phases present in any system under equilibrium conditions at any given temperature. Construction of the diagram is often based on solubility limits determined by thermal analysis – i.e., using cooling curves. Changes in volume, electrical conductivity, crystal structure and dimensions can also be used in constructing phase diagrams.

The solubility of two-component (or binary) systems can range from essential insolubility to complete solubility in both liquid and solid states, as mentioned above. Water and oil, for example, are substantially insoluble in each other while water and alcohol are completely intersoluble. Let us visualize an experiment on the water-ether system in which a series of mixtures of water and ether in various proportions is placed in test tubes.

After shaking the test tubes vigorously and allowing the mixtures to settle, we find present in them only one phase of a few per cent of ether in water or water in ether, whereas for fairly large percentages of either one in the other there are two phases. These two phases separate into layers, the upper layer being ether saturated with water and the lower layers being water saturated with ether.

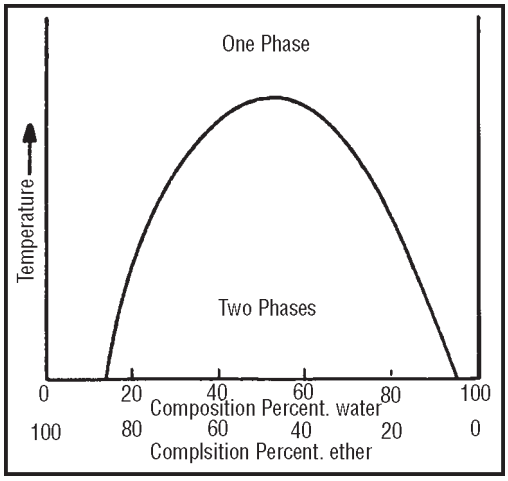


Fig. Schematic Representation of the Solubilities of Ether and Water in Each other.

After sufficiently increasing the temperature, we find, regardless of the proportions of ether and water, that the two phases become one. If we plot solubility limit with temperature as ordinate and composition as abscissa, we have an isobaric [constant pressure (atmospheric in this case)] phase diagram. This system exhibits a solubility gap.

Cooling Curves

Equilibrium phase diagrams, also called constitutional diagrams, are usually constructed from the data obtained from cooling curves. These are secured by plotting the measured temperatures at equal time intervals during the cooling period of a melt to a solid. The main types of cooling curves.

The cooling curve shown in figure is that of a pure molten metal or a pure compound.

The cooling proceeds uniformly along curve AB at a decreasing rate until point B is reached – when the first crystals being to form. As freezing continues the latent heat

of fusion is liberated in such amounts that the temperature remains constant from B to C until the whole mass has entirely solidified.

Period BC is known as the "horizontal thermal arrest". Further cooling from point C will cause the temperature to drop along curve CD. On examining the phase rule equation for a condensed system, it is seen that this one-component system has one degree of freedom (univariant) in regions AB and CD. There is only one phase present – either liquid in the region AB or solid in the region CD:

$F = C - P + 1 = 1 - 1 + 1 = 1$

Therefore the temperature can be varied independently without the disappearance or appearance of a phase. In region BC, however, there are two phases present – liquid and solid – and the number of the degrees of freedom will be:

$F = C - P + 1 = 1 - 2 + 1 = 0$

Therefore the system is nonvariant. Consequently the temperature must be constant as long as two phases are present, but when all the liquid has solidified to a solid only one phase will be present and the system will again become univariant.

Curve shows a cooling curve for a binary system consisting of two metals forming a solid solution. Section AB of the curve is similar in character to that, but during the freezing period the temperature does not remain constant but drops along line BC until the whole mass as completely solidified. The application of the phase rule for section BC shows that the system is univariant; therefore the temperature will vary independently.

$F = C - P + 1 + 2 - 2 + 1 = 1$

The slope of the cooling curve, however, will change due to the evolution of latent heat of crystallization. From point C on there will be only one solid phase and the temperature falls along line CD.

The melting or freezing range of alloys is due to the changes in the composition of the solid and liquid phase and naturally results in variable freezing or melting points.

A cooling curve of a binary system whose two components are completely soluble in the liquid state but entirely insoluble in the solid state. The liquid cools along line AB until temperature B is reached where a component that is in excess will crystallize and the temperature will fall along line BC, having a slope different than that of line AB.

At point C the liquid composition has been reached at which the two components crystallize simultaneously from the solution and the temperature remains constant until all the liquid solidifies. This is known as the eutectic reaction.

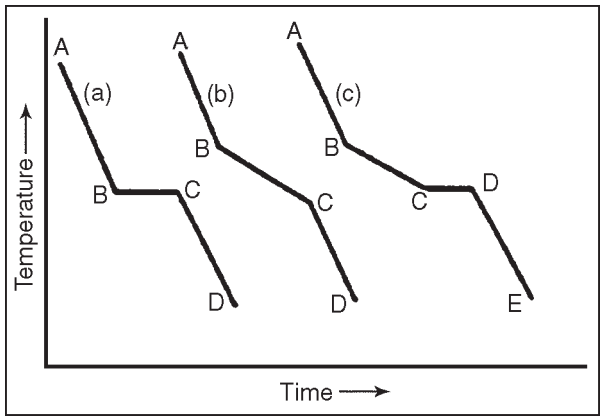


Fig. Cooling Curves: (a) Pure Compound; (b) Binary Solid Solution; (c) Binary eutectic system.

In addition to these types of cooling curves further modifications can make possible somewhat differently shaped curves.

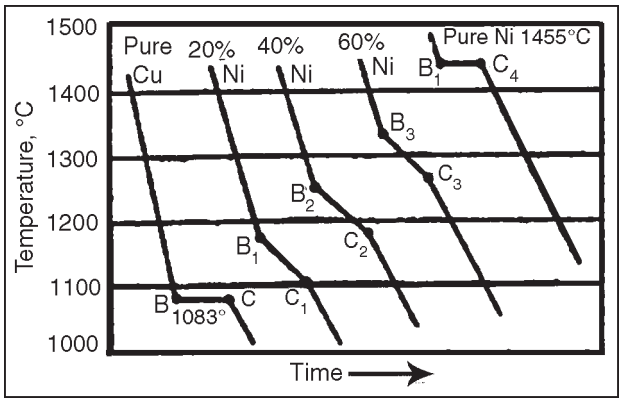
Solid Solution Equilibrium Diagrams

A plot of an equilibrium diagram for a solid solution can be made from a series of cooling curves obtained for different alloy compositions, as illustrated by the copper-nickel alloys.

The temperatures corresponding to the upper points on the cooling curves (B, B1, B2,...), when plotted against suitable alloy compositions, give a curve called the "liquidus line". A plot of temperatures corresponding to lower points (C, C1, C2,...) against alloy compositions gives a curve called the "solidus line".

The solidus line represents the melting points of the different solid solutions whereas the liquid line is the freezing point curve.

Points B and B4 correspond to the melting points of pure components. The region between the liquidus and solidus is a two-phase region in which solid crystals of homogeneous solid solution are in equilibrium with a liquid of suitable composition. Above the liquidus line there will be one liquid phase whereas below the solidus line there will be only one solid phase.



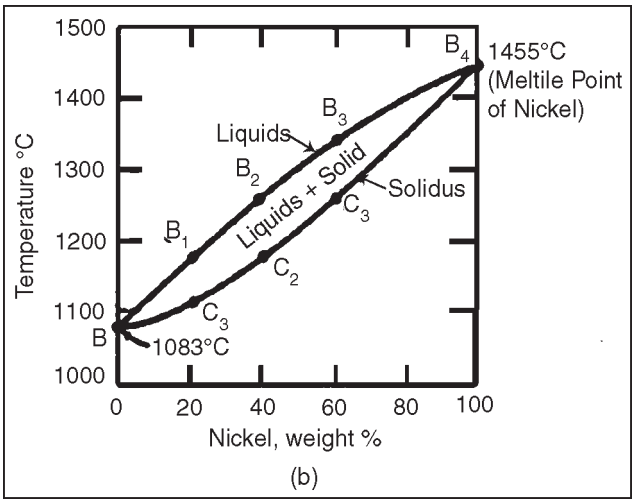


Fig. Plotting Equilibrium Diagrams from Cooling Curves for Cu-Ni Solid Solution alloys. (a) Cooling curves; (b) equilibrium diagram.

In figure we find that addition of Cu to Ni decreases its melting point. This is a commonly observed phenomenon. For example, we can depress the freezing point of water by adding common salt. At the same time, however, we find addition of Ni to Cu increases the melting point of Cu. This seems unusual and an explanation is in order.

We must first recognize that vapour pressures of a solid and a coexisting liquid are equal at the melting point. The explanation of change of melting point lies in the vapour pressures of the liquid and solid phases. If solute atoms are soluble in the liquid but insoluble in the solid, it is clear that in the liquid the attraction of solvent for solute atoms must be greater than that of solvent atoms for themselves – otherwise the two kinds of atoms would not mix.

This greater attraction means that fewer solvent atoms escape per second at the surface of the liquid – thus the vapour pressure of the liquid phase is lowered, but the vapour

pressure of the solid phase is unchanged. The melting point may then either rise or fall depending on the relative degree of lowering in each phase. It should be emphasized that the above reasoning and results hold equally well for other phase changes, including polymorphic changes, as for melting and freezing.

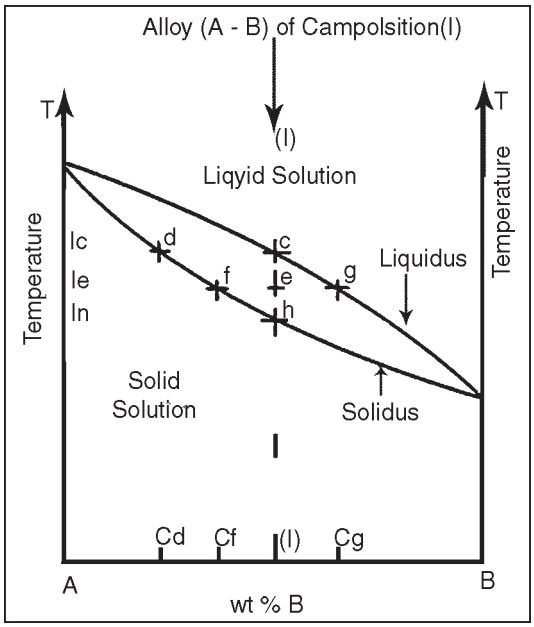


Fig. Behaviour of an Alloy of Composition.

The complete phase diagram makes it possible to predict the state of any alloy in the system at any temperature included in the diagram, if sufficient time is allowed for the system to reach equilibrium. In normal operation, however, it is fairly rare that a system reaches (or sometimes even approaches) true equilibrium, especially at lower temperatures. The equilibrium diagram is, even then, still of considerable importance since it can serve as a guide in predicting the behaviour of the system under nonequilibrium conditions. Consider the behaviour of an alloy of composition

with the elements A and B, initially completely liquid, which is cooled to room temperature. As temperature decreases, there is no discernible change until the liquidus is reached (i.e., Tc). At any lower temperature, solid must appear. The composition of the solid is given by the solidus line (point d) at that temperature. Thus the first solid is much richer in A than the liquid, which is still essentially the composition of alloy I, although it is slightly richer in B than before.

The liquid now has a slightly lower melting point. As more heat is removed additional solid is formed with a composition following the solidus, whereas the liquid becomes increasingly richer in B with composition following the liquidus. A decrease in temperature to Te causes the melt to precipitate solid of f composition which both encases the existing crystals and forms new, separate crystals. For equilibrium at this temperature diffusion must take place between the d cores and the f encasements. In addition, for the entire solid to be of f composition, some B atoms from the liquid must diffuse into the A-rich centre of the crystals since diffusion between d and f compositions can only result in a composition intermediate to d and f, not in composition f. A schematic representation of these solid and liquid combinations.

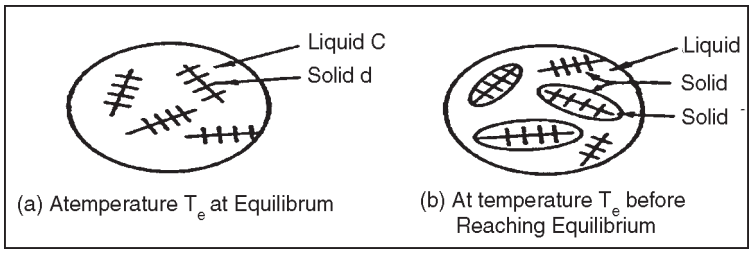


Fig. Schematic Representation of the Physical Condition of Alloy I Sketched.

If diffusion keeps pace with crystal growth (i.e., if true equilibrium is constantly maintained), melt composition

moves downward along the liquidus and solid composition moves downward along the solidus. Under the microscope the separating solid appears the same as a pure metal.

This continues until temperature Th is reached. At that temperature, under equilibrium conditions, solidification is complete and on further cooling there are no further discernible changes in the solid. If the temperature of the alloy is raised from room temperature whiled constantly maintaining equilibrium to the point where the alloy is completely molten, the behaviour is exactly the reverse of the cooling behaviour.

Phase Rule And Equilibrium

The phase rule, also known as the Gibbs phase rule, relates the number of components and the number of degrees of freedom in a system at equilibrium by the formula

$$F = C - P + 2$$

where F equals the number of degrees of freedom or the number of independent variables, C equals the number of components in a system in equilibrium and P equals the number of phases. The digit 2 stands for the two variables, temperature and pressure.

The number of degrees of freedom of a system is the number of variables that may be changed independently without causing the appearance of a new phase or disappearance of an existing phase. The number of chemical constituents that must be specified in order to describe the composition of each phase present. For example, in the reaction involving the decomposition of calcium carbonate on heating, there are three phases – two solid phases and one gaseous phase.

$$CaCO_3\ (s) \leftrightarrows CaO\ (s) + CO_2\ (g)$$

There are also three different chemical constituents, but the number of components is only two because any two constituents completely define the system in equilibrium. Any third constituent may be determined if the concentration of the other two is known.

Substituting into the phase rule we can see that the system is univariant, since $F = C - P + 2 = 2 - 3 + 2 = 1$. Therefore only one variable, either temperature or pressure, can be changed independently. (The number of components is not always easy to determine at first glance, and it may require careful examination of the physical conditions of the system at equilibrium.)

The phase rule applies to dynamic and reversible processes where a system is heterogeneous and in equilibrium and where the only external variables are temperature, pressure and concentration. For one-component systems the maximum number of variables to be considered is two – pressure and temperature.

Such systems can easily be represented graphically by ordinary rectangular coordinates. For two-component (or binary) systems the maximum number of variables is three – pressure, temperature and concentration.

Only one concentration is required to define the composition since the second component is found by subtracting from unity. A graphical representation of such a system requires a three-dimensional diagram. This, however, is not well suited to illustration and consequently separate two-coordinate diagrams, such as pressure vs temperature, pressure vs composition and temperature vs composition, are mostly used. Solid/liquid systems are usually investigated

at constant pressure, and thus only two variables need to be considered – the vapour pressure for such systems can be neglected.

This is called a condensed system and finds considerable application in studying phase equilibria in various engineering materials. A condensed system will be represented by the following modified phase rule equation:

$$F = C - P + 1$$

where all symbols are the same as before, but (because of a constant pressure) the digit 2 is replaced by the digit 1, which stands for temperature as variable. The graphical representation of a solid/liquid binary system can be simplified by representing it on ordinary rectangular coordinates: temperature vs concentration or composition.

H vs T Phase Diagram

With the aid of a suitable calorimeter and energy reservoir, it is possible to measure the heat required to melt and evaporate a pure substance like ice. The experimental data obtainable for a mole of ice.

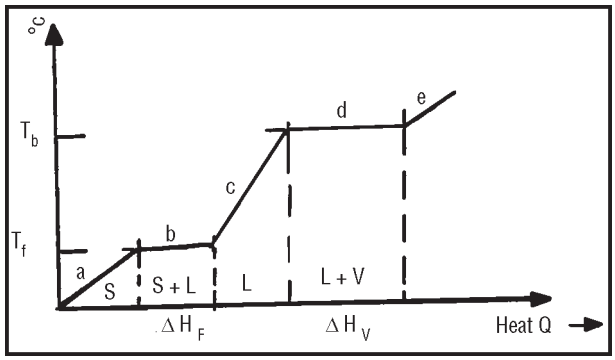


Fig. H vs T Diagram for Pure H_2O. (Not to Scale.)

As heat is added to the solid, the temperature rises along line "a" until the temperature of fusion (T_f) is reached. The amount of heat absorbed per mole during melting is

represented by the length of line "b", or ΔH_F. The amount of heat absorbed per mole during evaporation at the boiling point is represented by line "d". The reciprocal of the slope of line "a", (dH/dT), is the heat required to change the temperature of one mole of substance (at constant pressure) by 1CF. (dH/dT) is the molar heat capacity of a material, referred to as "Cp". As the reciprocal of line "a" is Cp (solid), the reciprocals of lines "c" and "e" are C_p (liquid) and C_p (vapour) respectively.

From a thermodynamic standpoint, it is important to realise that figure illustrates the energy changes that occur in the system during heating. Actual quantitative measurements show that 5.98 kJ of heat are absorbed at the melting point (latent heat of fusion) and 40.5 kJ per mole (latent heat of evaporation) at the boiling point.

The latent heats of fusion and evaporation are unique characteristics of all pure substances. Substances like Fe, Co, Ti and others, which are allotropic (exhibit different structures at different temperatures), also exhibit latent heats of transformation as they change from one solid state crystal modification to another.

Energy Changes

When heat is added from the surroundings to a material system (as described above), the energy of the system changes. Likewise, if work is done on the surroundings by the material system, its energy changes. The difference in energy (Δ E) that the system experiences must be the difference between the heat absorbed (Q) by the system and the work (W) done on the surroundings. The energy change may therefore be written as:

$\Delta E = Q - W$

If heat is liberated by the system, the sign of Q is negative and work done is positive. Q and W depend on the direction of change, but ΔE does not. The above relation is one way of representing the First Law of Thermodynamics which states that the energy of a system and its surroundings is always conserved while a change in energy of the system takes place. The energy change, ΔE, for a process is independent of the path taken in going from the initial to the final state.

In the laboratory most reactions and phase changes are studied at constant pressure.

The work is then done solely by the pressure (P), acting through the volume change, ΔV.

$$W = PDV \text{ and } DP = 0$$

Hence:

$$Q = \Delta E + P\Delta V$$

Since the heat content of a system, or the enthalpy H, is defined by:

$$H = E + PV$$
$$DH = DE + PDV$$

so that:

$$DH = Q - W + PDV$$

or

$$DH = Q$$

Reactions in which ΔH is negative are called exothermic since they liberate heat, whereas endothermic reactions absorb heat. Fusion is an endothermic process, but the reverse reaction, crystallization, is an exothermic one.

Entropy and Free Energy

When a gas condenses to form a liquid and a liquid freezes to form a crystalline solid, the degree of internal order

increases. Likewise, atomic vibrations decrease to zero when a perfect crystal is cooled to 0°K. Since the term entropy, designated by S, is considered a measure of the degree of disorder of a system, a perfect crystal at 0°K has zero entropy.

The product of the absolute temperature, T, and the change in entropy, Δ S, is called the entropy factor, TΔ S. This product has the same units (Joules/mole) as the change in enthalpy, Δ H, of a system. At constant pressure, P, the two energy changes are related to one another by the Gibbs free energy relation:

$$D\ F = D\ H - TD\ S$$

where

$$F = H - TS$$

The natural tendency exhibited by all materials systems is to change from one of higher to one of lower free energy. Materials systems also tend to assume a state of greater disorder whereby the entropy factor TΔ S is increased. The free energy change, Δ F, expresses the balance between the two opposing tendencies, the change in heat content (Δ H) and the change in the entropy factor (TΔ S). If a system at constant pressure is in an equilibrium state, such as ice and water at 0°C, for example, at atmospheric pressure it cannot reach a lower energy state. At equilibrium in the ice-water system, the opposing tendencies, Δ H and TΔ S, equal one another so that Δ F = 0. At the fusion temperature, T_F:

$$\Delta S_F = \frac{\Delta H_F}{T_F}$$

Similarly, at the boiling point:

$$\Delta S_V = \frac{\Delta H_V}{T_V}$$

Thus melting or evaporation only proceed if energy is supplied to the system from the surroundings. The entropy of a pure substance at constant pressure increases with temperature according to the expression:

$$\Delta S = \frac{C_P \Delta T}{T} (\text{Since} : \Delta H = C_P \Delta T)$$

where Cp is the heat capacity at constant pressure, Δ Cp, Δ H, T and Δ T are all measurable quantities from which Δ S and Δ F can be calculated.

F vs T

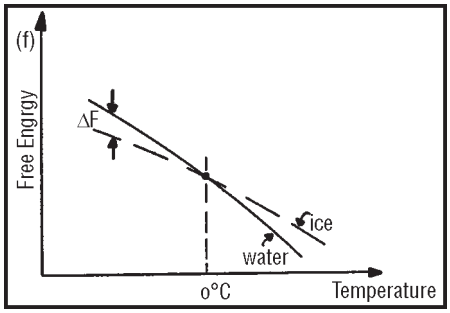


Fig. Free Energy is a Function of Temperature for Ice and Water.

Any system can change spontaneously if the accompanying free energy change is negative. This may be shown graphically by making use of F vs T curves.

The general decrease in free energy of all the phases with increasing temperature is the result of the increasing dominance of the temperature-entropy term. The increasingly negative slope for phases which are stable at increasingly higher temperatures is the result of the greater entropy of these phases.

Solid-Vapour Equilibrium

Just as liquid and vapour can coexist in equilibrium under certain temperature-pressure conditions, so too can solid

and vapour coexist, without any liquid present. This situation can be directly observed by placing some solid iodine (I2) in a closed Erlenmeyer flask. Purple iodine vapour is visible in the space above the solid when the flask is held against a white background.

$$s \Leftrightarrow v$$

The double arrow means that conversion of solid to vapour (sublimation) and vapour to solid (deposition) occur at equal rates, so that there is no macroscopic change in the system.

The temperature dependence of the pressure of vapour in equilibrium with solid is similar to that of the liquid vapour pressure. It is governed by a form of the Clausius-Clapeyron equation in which the enthalpy and entropy of vaporization are replaced by the corresponding quantities for the sublimation process, ΔH^o_{sub} and ΔS^o_{sub}:

$\ln P_{vap} = -\Delta H^o_{sub}/RT + \Delta S^o_{sub}/R$

As shown in the figure, the straight line for the solid has a more negative slope and a more positive intercept than the liquid line. This is true because

$$\Delta H^o_{sub} > \Delta H^o_{vap}$$
$$\Delta S^o_{sub} > \Delta S^o_{vap}$$

The increase in potential energy for conversion of solid to vapour must exceed that for conversion of liquid to vapour, and similarly for the increase in disorder, because the solid is the phase of lowest PE and disorder.

The inequalities that the two lines cross at some temperature, Tt. At T > Tt, the vapour pressure of the solid exceeds that of the liquid; at T < Tt, the vapour pressure of the solid is less than that of the liquid.

Ultimately we would like to develop a graphical presentation of the system that shows directly the vapour pressure of the system as a function of temperature. To do this, we translate the graph of ln P_{vap} versus 1/T to one in which P_{vap} is plotted directly against temperature. Each possible state of the system is represented by a (P,T) point on this plot. To perform the translation, we make use of 3 facts:

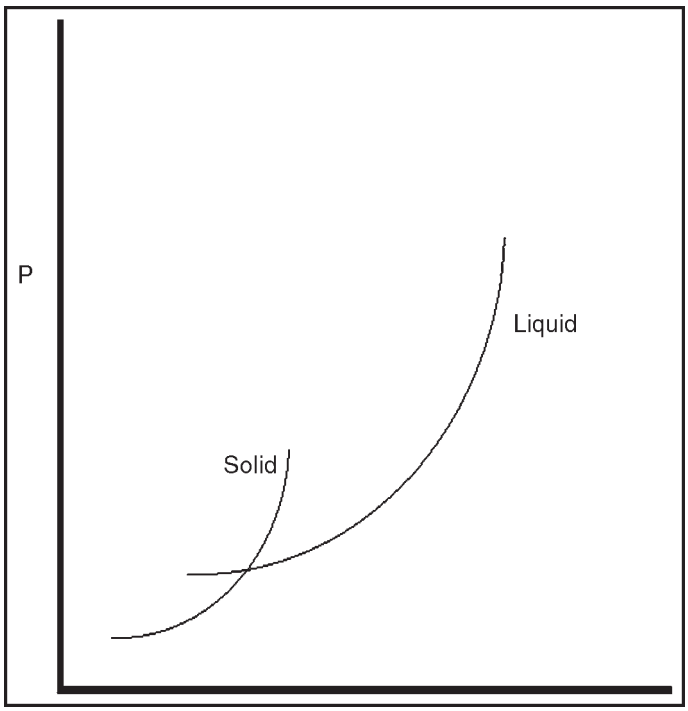


Fig. Vapour Pressure Temperature Curves for Solid and Liquid

ln P_{vap} for the solid is less than ln Pvap for the liquid at high 1/T (low T). Therefore P_{vap} for the solid is less than P_{vap} for the liquid at low T.

The Clausius-Clapeyron equation for the solid has a steeper slope than that for the liquid. Therefore the vapour pressure of the solid rises more rapidly with T than does that for the liquid.

As we have already seen, plots of P_{vap} versus T curve exponentially upward.

The curve labelled "l-v" represents all P-T combinations where liquid and vapour can coexist in equilibrium. The curve labelled "s-v" represents all P-T combinations at which solid and vapour can coexist in equilibrium. Because the point at which these two curves intersect lies on both curves simultaneously, all three phases must coexist together here. This is called the triple point, with temperature Tt.

Solid-Liquid Equilibrium

Completion of our graphical presentation of the states of the system requires only the addition of a curve of all P,T combinations at which solid and liquid coexist in equilibrium. One point of this curve — the triple point — is already plotted. For most substances, the "s-l" line slopes sharply up from the triple point, with a positive slope.

The steep slope of this line reflects the fact that solid and liquid are both condensed phases, with properties that depend very little on pressure. A big change in pressure results in only a small change in the temperature of the equilibrium system.

The intersection of the horizontal dotted line at P = 1 atm with the s-l curve is called the normal melting point of the substance. This is the temperature at which solid and liquid exist at equilibrium under 1 atm pressure. For most substances, the vapour pressure at the triple point is less than 1 atm. It follows that the melting temperature, Tm, is somewhat greater than the triple point temperature, Tt (although the difference is usually slight). The intersection

with the l-v curve is the normal boiling point of the substance. The temperature at which the vapour pressure of the liquid reaches 1 atm.

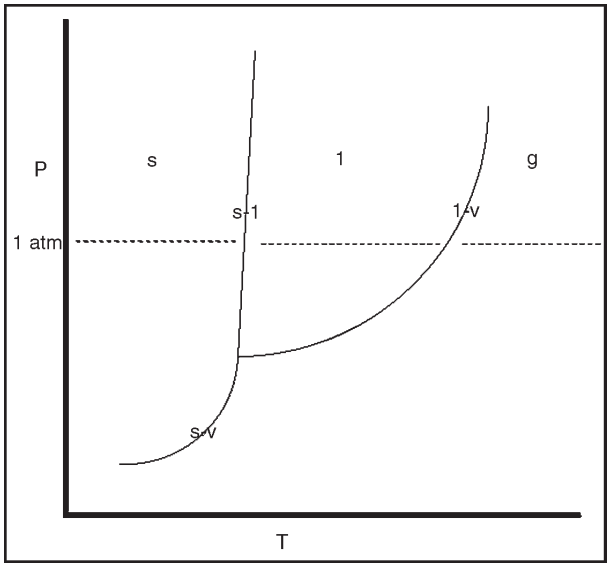


Fig. Phase Diagram for a Pure Substance

Liquid-Vapour Equilibrium

Our observation following injection of water into the box was that the pressure rose gradually from zero to 23.8 torr at a temperature of 25°C, where it levelled off and remained constant indefinitely. The pressure resulted from formation of water vapour by evaporation of some of the liquid water.

We address two interesting questions:

1) Why does evaporation occur?
2) Why does the pressure rise to 23.8 torr, then stop changing?

The drive to minimum energy favors the liquid phase. However, the drive to increased disorder favors the gas phase. Two natural tendencies try to drive the process in equation in opposite directions. The result is a compromise in which

some water is present as liquid and some as vapour. This is our macroscopic interpretation of the vapour pressure phenomenon. We now seek an understanding at the molecular level.

An important result from our models of the gas, liquid, and solid phases can be stated succinctly as follows: at a given temperature, all molecules have the same average kinetic energy, whether they are present in the solid, liquid, or gas phase. Further, the distribution of kinetic energies follows the Maxwell-Boltzmann distribution, regardless of phase. The general form of the Maxwell-Boltzmann distribution. In the ensuing discussion, we will apply this plot to the liquid phase of a pure substance.

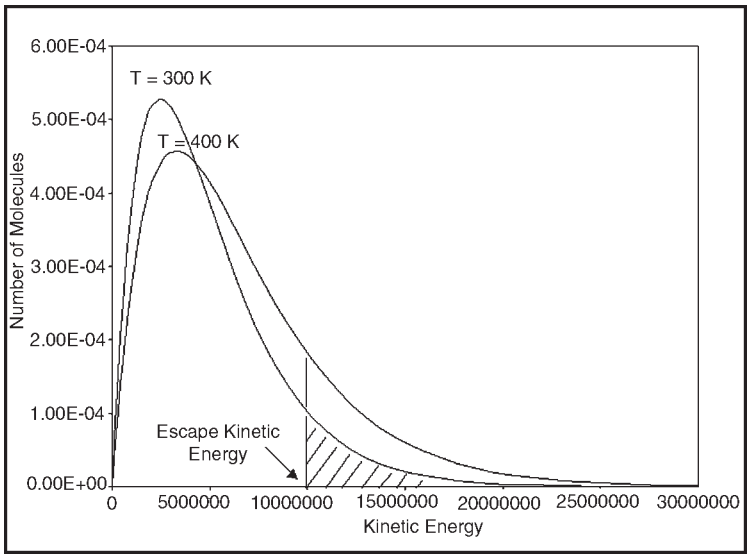


Fig. Maxwell-Boltzmann Distribution

As the curve indicates, at any moment there are some molecules moving very slowly, a large number moving with intermediate (near-average) kinetic energies, and a few molecules with very high kinetic energy. We focus now on the molecules near the surface of the liquid, because it is

these that have a chance to escape from the liquid into the space above. Molecules near the surface possess a range of kinetic energies, like those in the bulk. A molecule near the surface and moving toward the surface will escape the liquid if its kinetic energy is sufficient to overcome the attractive forces of nearby molecules in the liquid. In other words, the kinetic energy must be at least equal in magnitude to the depth of the liquid potential well. We will call this minimum required kinetic energy the escape kinetic energy. The number of liquid molecules having at least this kinetic energy is proportional to the area under the curve to the right of a vertical line passing through the escape KE.

When liquid water is injected into an evacuated box, it spreads out on the bottom and molecules begin to evaporate at a rate depending on two quantities. First, the number of molecules that evaporate per unit time is proportional to the number having at least the escape kinetic energy. Second, the evaporation rate is proportional to the number of molecules at the surface — i.e., to the exposed liquid surface area.

In equation form,

- RE = kE(T)*A
- RE = evaporation rate, with units of amount per time
- A = liquid surface area
- kE(T) = rate constant for evaporation, with units of amount per unit area per unit time.

(The constant, kE, is called a rate constant. It is a proportionality constant relating the evaporation rate to macroscopic variables on which rate depends. In this case, surface area is the only variable of importance, as long as

temperature is constant. The size of the rate constant depends upon the fraction of molecules with KE greater than the escape KE. Because this depends on temperature, kE is temperature-dependent.

This is indicated in the equation.) Evaporation leads to a buildup of molecules in the vapour phase over time. These exert pressure on the container walls, which increases directly as the number of molecules. Further, they collide occasionally with the liquid surface, where they may once again be captured by the intermolecular forces of the closely packed molecules at the liquid surface — in other words, they condense.

The rate of condensation is proportional to the exposed surface area of liquid and to the number of molecules in the vapour phase.

This number is directly proportional to the pressure exerted by the vapour. The rate of condensation, in equation form, is,

$RC = kC*A*P$

RC = Condensation rate, with units of amount per time

kC = Rate constant for condensation, a proportionality constant that is independent of T, with units amount per time per area per unit pressure.

As the number of molecules in the vapour phase continues to increase as a result of a steady rate of evaporation, a point is eventually reached at which

$$RE = RC$$

Although both evaporation and condensation still occur, they occur at equal rates. There will be no further change in the number of molecules in the vapour phase, hence no

further change in pressure. We describe this situation as a state of phase equilibrium.

$$l \Leftrightarrow v$$

The double arrow indicates that the two processes, evaporation and condensation, occur simultaneously and at equal rates. Phase equilibrium is dynamic, rather than static, because although it appears at the macroscopic level that nothing is happening (vapour pressure is constant), there is a lot of action at the molecular level.

This can be easily substantiated by starting with a system consisting of H_2O in the liquid phase and D_2O (deuterium oxide) in the vapour phase. After some time has elapsed, both forms of water will be found (by, for example, mass spectrometry) in both phases, indicating that evaporation and condensation are continuously occurring.

The pressure exerted by the vapour at equilibrium is called the equilibrium vapour pressure of the liquid, symbolized Pvap. Equations can be used to obtain an expression for Pvap in terms of the constants, kE and kC. Since at equilibrium, RE and RC are equal, we write

$$kE(T)A = kCAPvap$$

Solving for Pvap and cancelling the surface area terms in numerator and denominator gives equation

$$Pvap = kE(T)/kC$$

This equation shows that for a particular liquid at a fixed temperature, Pvap is constant, in agreement with experiment. The vapour pressure of a pure liquid provides the first example of an equilibrium constant, denoted Keq. For a liquid in equilibrium with vapour, Pvap = Keq. Keq is a ratio of rate constants.

Each liquid has a characteristic vapour pressure at a given T whose value depends primarily on kE. The greater the intermolecular forces, the greater the escape KE, and the smaller the value of kE. Small kE means smaller Pvap. Since the constant kC does not depend on the potential well depth, it is almost the same for all liquids. Equilibrium vapour pressures for several common laboratory liquids at 25oC are given in Table.

Table. Equilibrium Pvap at 25oC

Substance	Pvap (torr)
Water (H2O)	24
Ethanol (H2O)	65
Chloroform (H2O)	215
Diethylether (H2O)	545

Qualitative Aspects of Entropy

We have also seen that a major driving force for physical and chemical systems is the tendency toward minimum potential energy. This leaves us with a question: why do solids melt to form liquids, and why do these vaporize to gases? If the solid is the lowest-energy phase, and if all systems seek lowest energy, it seems that all substances should be solids at all temperatures and pressures. But this is clearly wrong. To understand the way pure substances behave requires an additional concept.

An ideal gas is confined in bulb A, connected via a closed stopcock to evacuated bulb B. What will happen when the stopcock is opened? Experience tells us that the gas will expand to occupy bulb B until the pressure in both bulbs is uniform. Why does this happen? An initial response might be that by expanding, the gas decreases in potential energy.

But closer examination reveals that this is not true. First, if the gas expands at constant temperature (which we can arrange), the average molecular kinetic energy remains the same.

Thus $\Delta KE = 0$. Second, the gas is ideal, so there are no forces between molecules. There is therefore no potential energy involved, and $\Delta PE = 0$. (If the gas is real rather than ideal, there are weak attractive forces; expansion results in an increase in distance of separation of the molecules and a correponding increase in PE. For a real gas, ΔPE is small but > 0.) Thus the drive to minimum energy cannot be responsible for the tendency of a gas to expand. The only change on expansion is that each molecule has a larger volume available to it. The location of a molecule is less certain than it was before expansion. Expansion has the effect of increasing the disorder, or randomness, of the gas molecules.

What is the chance that the molecules in bulb B will spontaneously move through the stopcock into bulb A to give the state at the left of the figure? Such a process would lead to a decrease in the disorder of the system. But such a process has never been observed. Thus experience tells us that it is very improbable that the reverse process will occur. Expansion is spontaneous, meaning it occurs without outside assistance; compression is not spontaneous.

There is a natural direction for this process that leads to an increase in the randomness, or disorder, in the system. The amount of disorder in a system is measured by a quantity called entropy, symbolized S. The higher the disorder in a system, the higher the entropy. The change in entropy in a process is the difference between the entropies of the final

and initial states, and is symbolized ΔS. It seems to be a law of nature that entropy increases (ΔS > 0) in any spontaneous process.

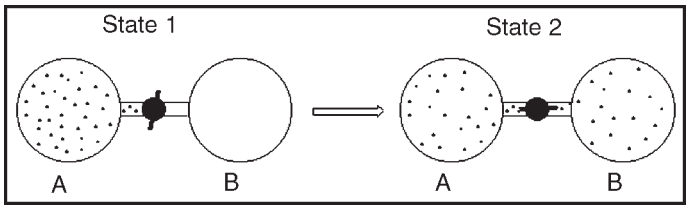


Fig. Ideal Gas is Confined in Bulb

A valid question at this point is this: "Why is there a tendency toward increased disorder?" A simple macroscopic example, based on the puzzle toy called Rubik's cube, will shed some light on this question. The toy consist of a cube with each face a different colour.

Each face can be independently rotated in either direction (clockwise or counterclockwise); rotation causes scrambling of the colours. We now consider what happens when we begin with the cube in the most ordered state — each face showing only one colour — and make a few random 90o face rotations. Begin with a 90^o clockwise rotation of the top face of the cube. This produces a situation in which four of the faces show 2 colours.

If we agree that any configuration of the cube in which 2 or more faces show more than one colour is a disordered state, then our single rotation has produced a disordered state. Now give the cube to a blindfolded person. Inform the person that by making a single 90^o rotation of one face, it is possible for him/her to return the cube to the ordered state.

Tell the person to make a single random move. What is the probability that this move will produce the ordered state?

There are 12 possible 90o moves (6 faces, 2 directions per face). Of these, 1 will produce the ordered state, but 11 will produce other disordered states.

The probability that the person's move will produce the ordered state is 1 in 12. The probability that a random move will produce another disordered state is 11 times larger than the probability of producing the ordered state.

The odds are very unfavorable, even when the cube is only one move away from the ordered state. If the cube were 6 moves away from the ordered state, the odds against restoring the ordered state by a sequence of 6 random moves would be 2,985,983 to 1. Disordered states are more probable than ordered ones.

Collections of molecules, like Rubik's cube, tend toward the most probable situation, because molecular motions are completely random. This requires only that some molecules in bulb A be moving toward the right.

On the other hand, once the gas has expanded, it can spontaneously contract only if, completely by chance, all of the 10^{23} molecules in bulb B happen to move toward the stopcock at the same time. The probability of this happening is essentially zero. The drive to maximum entropy is a consequence of the fact that a collection of particles, each moving randomly, is governed by the laws of statistics and probability.

We have indicated that the process is accompanied by increases in enthalpy and disorder. It is common knowledge that for any pure substance, the gas phase is favored at relatively high temperature and the solid phase at relatively low temperature.

This simple, familiar fact allows us to conclude that law 2 is more important than law 1 at high T; and 1 is more important than 2 at low T. Although this statement has been based on a single simple process, it is generally true, for all systems.

The question posed at the beginning of the section can now be addressed. Although the solid is the phase of minimum energy, the gas is the phase of maximum disorder. At low T, the drive to minimum energy predominates, and the solid is favored. At high T, the drive to maximum S predominates, and the gas is favored.

Temperature Dependence of Vapour Pressure

When the temperature of a liquid-vapour equilibrium system is increased, the equilibrium vapour pressure increases. There are two reasons for this.The pressure due to the molecules already in the gas phase increases by the ideal gas law. This makes a small contribution to the vapour pressure increase. More molecules enter the gas phase. This is by far the larger contribution.

We now attempt to justify these statements at the molecular level.

An increase in temperature means that the kinetic energy of the molecules increases in both phases. Molecules at the liquid surface jiggle more vigorously, and a greater number of them have sufficient KE to escape the potential well of the liquid and enter the vapour phase. The increase in the number of molecules in the vapour phase results in an increased vapour pressure.

As the curve shifts right with increasing T, the area under the curve to the right of the escape KE increases

exponentially. Pvap therefore increases exponentially with T. Mathematically, the T-dependence of Pvap can be understood in terms of equation.

The minimum KE requirement for escape from the liquid is imbedded in the constant kE, which depends markedly on T. kC has almost no temperature dependence because no minimum KE is required for a molecule to fall into a potential well. The temperature dependence of Pvap therefore parallels that of kE , which directly reflects the shift in the Maxwell-Boltzmann distribution.

It can be shown, both experimentally and theoretically, that vapour pressure varies with temperature according to equation:

$P_{vap} = e^{-\Delta Ho/RTe\Delta So/R}$

This is usually written in logarithmic form as the Clausius-Clapeyron equation:

$\ln P_{vap} = -\Delta H^{o}_{vap}/RT + \Delta S^{o}_{vap}/R$

ΔH^{o}_{vap} and ΔS^{o}_{vap} are the standard molar enthalpy and entropy of vaporization, respectively. A plot of $\ln P_{vap}$ vs. 1/T should yield the enthalpy of vaporization, ΔH^{o}_{vap}, from the slope, and the entropy of vaporization, ΔS^{o}_{vap}, from the intercept.

This is a convenient experimental approach to the measurement of these quantities. The form of the Clausius-Clapeyron equation is general in physical science; all molecular processes exhibit temperature dependence of this form.

The vapour pressure is plotted against temperature for several common laboratory liquids. The magnitude of P_{vap} for a liquid at a particular temperature and the rate of change

of P_{vap} with temperature depend on the magnitude of the intermolecular forces in the liquid phase.

The stronger the forces, the lower P_{vap} at a given T and the steeper the rise in the curve. Of the liquids plotted water clearly has the largest intermolecular forces, due to strong hydrogen bonding. Having seen how P_{vap} depends on T, we consider a problem involving conceptual aspects of l-v phase equilibrium.

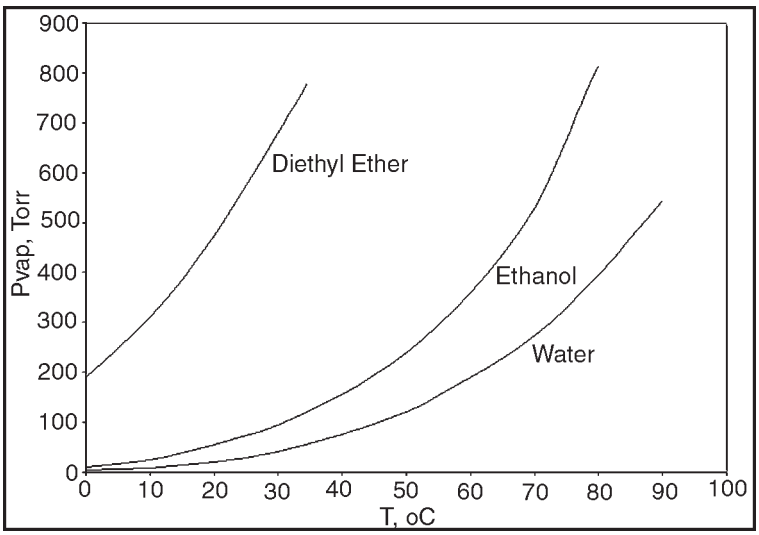


Fig. Vapour Pressure Versus Temperature

2

Structures and Types of Solids

In the broadest sense, solids are divided into two categories – crystalline and amorphous solids. Crystalline solids have regular arrangements or patterns of lattice structures of the components (atoms, ions, or molecules); amorphous solids have their components arranged in a rather disorderly manner. Common glass is an example of amorphous solid.

Many characteristic shapes of crystalline solids, such as the diamond structure and the rock salt structure are due to the regular arrangement of the particles called lattice in the crystals. The smallest repeating unit containing a minimum number of lattice points that may represent the entire crystal is called a unit cell. The entire crystal can be generated by repeating its unit cell in all three dimensions.

Types of Crystalline Solids

There are many types of crystalline solids. They are classified by the types and lattice patterns of components

that make up the substance and how these components (atoms, ions, or molecules) are bonded together. The following classifications for crystalline solids may also apply to amorphous solids.

- Molecular solids - solids that contain molecules: ice, dry ice, sugar, etc.
- Ionic solids - solids that contain ions: all salts;
- Atomic solids – metallic, covalent network, and Group 8A solids.

Molecular solids contain molecules that are held together by weak intermolecular forces such as dispersion forces, permanent dipole-dipole attractions, and hydrogen bonding. For examples, in ice, water molecules are held together by hydrogen bonds; in dry ice (solid CO_2) and iodine molecules are held together by London dispersion forces. These intermolecular forces are relatively very weak forces and molecular solids are characterised by low melting points. They are also non-conductors.

All ionic compounds form ionic crystals they contain ions held firmly together by strong ionic bonds. Ionic solids are characterised by high melting points. For example, NaCl melts at about 801 ^{o}C. Ionic solids cannot conduct electricity because the ions are not free to flow. However, when melted or dissolved in water, the ions are free to flow and they form excellent conductors.

Atomic solids are divided into three subgroups: metallic solids, covalent network solids, and solids formed by the noble gas (Group 8A) elements. All metals form metallic solids, which are characterised by high thermal and electrical conductivity, malleability, and ductility. In metallic solids

atoms are held together by a “sea” of delocalised valence electrons, which make metals an excellent thermal and electrical conductors. The “sea-of-electrons” model also explains why metals are malleable and ductile. The layers of atoms that form metallic crystals are held firmly by the “sea” of valence electrons, which act like a glue. These layers of atoms can be forced to slide or glide over each other without breaking. The interaction of delocalised electrons with light also accounts for the lustrous appearance of metals.

Metals have variable hardness depending on the number of valence electrons holding the atoms together. For example, Group 1A metals are generally soft because they have the least number of valence electrons. Most transition metals are relatively hard because they have more electrons in the ns and (n-1)d subshells, which make up the "sea of electrons”.

Covalent network solids such as diamond and silica, contain atoms that are covalently bonded to each other in a network pattern. They are very hard and have extremely high melting points. Diamond is the hardest known natural substance with melting point of approximately 3600 ^{o}C. They are non-conductors.

Graphite is a semi-covalent network solid, which is made of layers of carbon atoms; within each layer carbon atoms are covalently bonded, but the forces holding these layers of carbon atoms are weak van der Waals forces. These weak forces allow the carbon layers to slide over each other and can flake off. Thus graphite is much softer than diamond although both are composed of carbon atoms. Graphite is used as pencil lead and lubricant, whereas diamond is used as glass cutting utensil and in rock drill bits. Unlike diamond,

graphite can conduct electricity because it contains delocalised π-electron system.

In the Group 8A solids, noble gas atoms are held together by weak dispersion forces. As a result, this type of atomic solids has very low melting points. Unlike metallic solids, this type of solids is a non-conductor because their valence electrons are not delocalised.

Structure and Bonding in Metals

Metallic crystals can be pictured as containing spherical atoms packed together in regular lattice and are held together by a "sea" of delocalised valence electrons. Metals crystallise either into hexagonal closest packing, cubic closest-packing (or face-centered cubic), body-centered cubic, or simple cubic structures. The simple cubic structure is the least common in metals.

Cubic Closest-Packed Arrangement and Face-Centered Cubic (FCC) Structures

In the closest-packing arrangement, each sphere has 12 nearest neighbours. Layers of spherical atoms are stacked in such a way that each atom of one layer is nested in the depression of the two adjacent layers.

When spheres in the second layer of a close-packed arrangement are nested on top of the first layer, two types of "holes" are formed in the second layer. These are the tetrahedral holes - those directly on top of spheres in the first layer, and octahedral holes - those directly on top of "holes" in the first layer. Sphere of the third layer may be nested on top of the tetrahedral holes, which will result in an ABABAB…pattern of arrangement of spheres, or on top

of the octahedral holes, which yields an ABCABC.. pattern.

The closest-packed arrangement with the ABCABC.. pattern forms the face-centered cubic (fcc) or cubic close-packed structure. While the ABABAB.. pattern is called the hexagonal closest-packed (hcp) structure. In both types of arrangements, the maximum space occupied by spheres is about 74 per cent (26% is void space).

Body-Centered Cubic (BCC)

In the body-centered cubic structure, each sphere has the coordination number 8 (or eight nearest neighbours). The body-centered cubic structure has a higher percentage of occupied space than the simple cubic structure. The maximum occupied space in a body-centered cubic arrangement is about 69 per cent. Many metals crystallise into the body-centered cubic structure, such as chromium, manganese, and iron.

Simple Cubic (SC)

In the simple cubic structure spheres (atoms or ions) are arranged in layers, one placed directly on top, below, or next to the other. Each sphere has 6 nearest neighbours, which is also called the coordination number. In the simple cubic structure, the maximum space occupied by spheres is only 52.36 per cent of the available space (47.64% is void space).

Types of Unit Cells in Metals

The structures of crystalline solids are determined by the manner atoms are arranged and the resulting lattice patterns of the unit cells. All cubic unit cells have a lattice point at each corner of the cube, which are the only lattice points in

the simple cubic unit cell. In the body-centered cubic (bcc) unit cell an additional lattice point occurs at the centre of the cube; while the face-centered cubic (fcc) unit cell has a lattice point at the centre of each face of the cube in addition to the eight corner positions.

Bonding in Solids

According to molecular orbital theory, a covalent molecule or a metal-ligand complex forms its molecular orbitals by combining the atomic orbitals of the constituent atoms. Similarly, we can treat a solid particle as a *giant molecule* consisting of an extremely large number of molecular orbitals. We can treat molecular orbitals of a polyatomic molecule from two extreme points of view, *i.e.*, the localised and delocalised bond models. Both models are useful in the discussion of solid structures. With the localised bond approach, we can imagine a situation where one of the bonded atoms draws the shared electrons completely to itself. This will result in a completely ionic bond, and the resulting giant molecule will be in reality an *ionic solid*, where the three-dimensional structure is based on the electrostatic interaction (bonding) between the constituent ions. On the other hand, if the bond is completely covalent, we get a three-dimensional covalent solid such as diamond, where the structure consists of C-C bonds. At the other extreme, we can consider the giant molecule from the viewpoint of the delocalised bond model. In this multi-centered molecular orbital approach, the atomic nuclei of the solid are viewed as embedded in a sea of electrons. This situation is characteristic of the solids of the metallic elements, and this type of delocalised bonding has

acquired the name *metallic bonding*. In some cases, in addition to the ionic, covalent and metallic bonds, other intermolecualar interactions, such as van der Waals forces and hydrogen bonding, contribute to the development of the three-dimensional structure.

Atomic Packing

We suggested that a solid can be usefully considered as a giant molecule. As regards the shape of this molecule, however, we were silent. In this section we wish to turn to this question: How are the constituent atoms of a solid organised spatially in the solid phase? To begin, we consider the number of ways spheres can be densely packed into space. Then we relax the constraint on dense packing; this enables us to treat the crystal structures of ionic solids. It is helpful to organise our thinking around two-dimensional layers of spheres in contact. Then we build three dimensional structures by depositing these layers one on top the other. We can begin the layer by layer stacking by allowing the spheres of the second layer to fall in the cavities in the first layer. We then have two ways to add a third layer. In one scenario, we can use an ABAB procedure, where the third layer spheres are vertically colinear with first layer spheres. In the other scenario, which is based on an ABCABC pattern, the third layer spheres are vertically colinear with the holes in the first layer. These two structures of ABAB and ABCABC are called *polytypes* - they are identical in two dimensions while they are mismatched in the third. For each of these polytypes, the coordination number is 12.

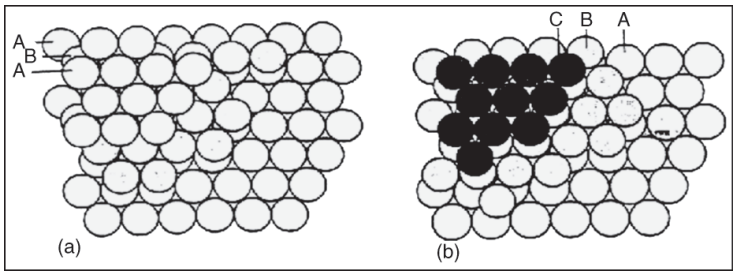


Fig. Layer by Layer Build-up of Three-dimensional Structures.

Crystal Structures

Metallic Elements and Alloys

That the bonding in solids of the metallic elements is best described in terms of delocalised bonds. As a result of the non-directional nature of this kind of bonding, the atoms of metals in the solid state can acquire the maximum coordination numbers allowed by radius ratio considerations. Therefore, the structures of the metallic elements tend to be close-packed. This close packing accounts for the typically high densities of metals. What factors determine whether a given metal prefers the FCC or HCP structure? A useful rule of thumb is that a hexagonal structure is preferred if the number of valence electrons/valence orbitals is low, while cubic structures are preferred when the number is high. (Why?)

Besides HCP and FCC, other structures are also known. Thus, with increase in temperature, it is not uncommon for metals to undergo a transition from close-packed to less close-packed structures. The onset of structural transition can be rationalised by recognizing that greater atomic movements are more compatible with a more loose crystal structure. At relatively low temperatures, the tendency of a given metal to form a close-packed structure may be related to electronic structure.

If a metal is viewed in terms of positive charges immersed in an ocean of electrons, then the higher the electron density in this sea, the higher the atomic packing. Accordingly, the alkali metals which have relatively low valence electrons will have a greater tendency to adopt the bcc structure. The temperature marking the transition from close-packed to non-close-packed structures may be above room temperature (*e.g.*, Ca, Ti, Mn), or it may be below room temperature (*e.g.*, Li, Na).

When two or more metals are mixed in the molten state, the resulting solid product is termed an *alloy*. At one extreme, an alloy may represent a *solid solution*, where the constituent atoms are randomly intermixed. At the other extreme, the atoms in the mixture combine to give compounds with specific compositions and structure. The simplest alloy is a binary mixture where the majority atoms represent the solvent and the minority atoms the solute. In *substitutional* solid solutions, the solute atoms are located at sites normally occupied by the solvent atoms. In the case of interstitial solid solutions, the solute atoms insert themselves in the spaces between the solvent atoms.

The following set of conditions must be met in order for a substitutional solid solution to form:

- The solute atom must have a size comparable with that of the solvent atom (± 15%),
- The different metals should have similar crystal structures, and
- The metals must have comparable electronegativities so that the possible formation of a compound is minimized.

In the case of interstitial solid solutions, the solute atoms must have sizes compatible with the interstitial sites in the solvent lattice. Examples of intermetallic compounds: β-brass (CuZn), $MgZn_2$, Cu_3Au, and Na_5Zn_{21}.

Metal Hydroxo and Hydrated Compounds

In this section we consider the following types of compounds:

- *Metal hydroxides*, M(OH)n,
- *Metal oxyhydroxides*, MO(OH),
- *Basic salts*, and
- *Hydrated salts*.

A metal oxyhydroxide may be viewed as a compound intermediate between the corresponding oxide and hydroxide, whereas a basic salt lies in between the salt and the hydroxide. We are concerned here with *hydroxy halides* and *hydroxy-oxysalts*.

Metal Hydroxides and Oxyhydroxides

A major factor that determines the structures of metal hydroxides and oxyhydroxides is the manner in which the OH^- ion behaves in the vicinity of the constituent cations. Three main types of behaviour may be identified, with the OH^- ion behaving as: (a) a spherically symmetric anion, (b) a cylindrically symmetric (*i.e.*, dipolar) species, and (c) a quadrupolar species.

As the cation decreases in size and increases in charge, c-type behaviour becomes more favourable in comparison with b-type behaviour. In its quadrupolar form, the OH^- ion can undergo hydrogen bonding; the negative and positive regions of different OH^- ions interact. As a result of the above

considerations, we may divide metal hydroxides into two main groups, *i.e.*, those which have no hydrogen bonds (a- and b-type behaviours), and those with hydrogen bonds (c-type behaviour). Table below presents structural information for selected metal hydroxides and oxyhydroxides.

Table. The Crystal Structures of Metal Hydroxides and Oxyhydroxides

Type of Structure	Formula Type	M CN:O CN	Name of Structure	Examples
Infinite complexes	MOH	6:6	Sodium chloride	KOH, RbOH
	$M(OH)_3$	9:3	UCl_3	M = La, Y, Pr, Nd, Sm, Gd, Tb, Dy, Er, Yb, Am
	$M(OH)_3$	6:4?	$Al(OH)_3$	a-$Al(OH)_3$ (bayerite), g-$Al(OH)_3$ (gibbsite)
	MO.OH	7:4:3*	YO.OH	M = Y, 4f metal
	MO.OH	?	InO.OH	b-CrO.OH, e-FeO.OH
	MO.OH	?	a-MO.OH	a-AlO.OH (diaspore) a-FeO.OH (goethite), a-MnO.OH, a-ScO.OH, a-GaO.OH, a-VO.OH
	MO.OH	?	g-MO.OH	g-AlO.OH (boehmite) g-FeO.OH (lepidocrocite), g-MnO.OH manganite), g-ScO.OH
	MO.OH	?	a-CrO.OH	a-CrO.OH, a-CoO.OH
Layer structures	MOH			LiOH, NaOH
	$M(OH)_2$		Cadmium iodide	M = Mg, Ca, Mn, Fe, Co, Ni, Cd

Table. OH:M Ratios in Hydroxy-oxysalts

OH: M	Salt	OH: M	Salt
1/5	$Ca_5(OH)(PO_4)_3$	6/5	$Zn_5(OH)_6(CO_3)_2$
1/2	$Cu_2(OH)PO_4$	3/2	$Cu_2(OH)_3NO_3$
1/1	$Cu(OH)IO_3$	2/1	$Th(OH)_2SO_4$

Hydroxysalts

The metals which tend to form basic salts include Be, Mg, Al, several of the A subgroup transition metals (*e.g.*, Ti, Zr), 3d metals (*e.g.*, Fe, Co, Ni), 4f and 5f metals (*e.g.*, Ce, Th, U), a majority of the B subgroup elements (*e.g.*, In, Sn, Pb, Bi). In nature, some

important ore minerals occur as basic salts, *e.g.*, brochantite, $Cu_4(OH)_6SO_4$, and atacamite, $Cu_2(OH)_3Cl$. Corrosion products may also occur as basic salts, *e.g.*, hydrozincite, $Zn_5(OH)_6(CO_3)_2$, forms when metallic zinc is exposed to moist air. Other examples of basic salts include white lead, $Pb_3(OH)_2(CO_3)_2$, used as a pigment, and $Mg_2(OH)_3Cl_4 \times H_2O$, a reaction product in the hydration of Sorels cement. There is a wide variation in the OH:M ratios encountered in hydroxysalts for a series of hydroxy-oxysalts.

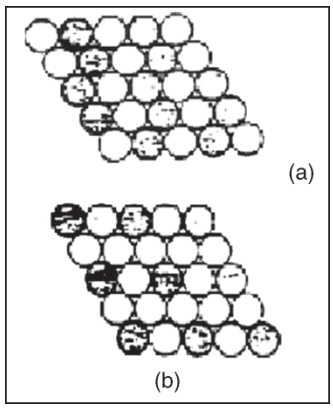


Fig. Structures of Hydroxyhalides: Closed-packed Spheres of (a) X(OH), (b) $X(OH)_3$.

Hydroxyhalides

The structures of hydroxyfluorides tend to follow those of oxyhdroxides or oxides. Thus, for example, ZnF(OH) has the diaspore structure, CdF(OH), the $CaCl_2$ structure, HgF(OH), a rutile-type structure, and $InF_2(OH)$, a ReO_3^-type structure.

The most common hydroxychlorides and hydroxybromides are those of the types, MXOH and $M_2X(OH)_3$, where M = Mg, Mn, Fe, Co, Ni, Cu(II). In the MXOH structures, X and OH ions form a close-packed layer. The metal atoms reside in 25% of the octahedral sites. In the case of the $M_2X(OH)_3$

structures, X and OH ions form a close-packed layer. The metal atoms reside in 50% of the octahedral sites.

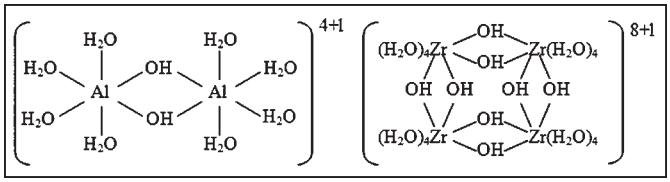


Fig. Structures Based on Finite M-OH Complexes: (a) $[Al_2(OH)_2(H_2O)_8]^{4+}$, (b) $[Zr_4(OH)_8(H_2O)_{16}]^{8+}$

Hydroxy-oxysalts

It is helpful to think of the structures of hydroxy-oxysalts in terms of the nature of the M-OH complex. Thus the M-OH complex or subunit may be (a) finite, (b) infinite one-dimensional (chain), (c) infinite two-dimensional (sheet), and (d) infinite three-dimensional. In the case of the first three situations, the oxyanions play the important role of joining the subunits into three-dimensional structures.

An example of a structure based on a finite M-OH subunit is the basic aluminum sulfate with the empirical formula $Al_2O_3.2SO_3.11H_2O$. The structural formula can be expressed as $[Al_2(OH)_2(H_2O)_8](SO_4)_2.2H_2O$. There are two OH bridges linking two Al^{3+} centers, *i.e.*, a dimeric complex. The finite M-OH subunit can also be in the form of a ring, as exemplified by the structure of $[Zr_4(OH)_8(H_2O)_{16}]^{8+}$.

Structures based on M-OH chains are characteristic of many basic salts of stoichiometry OH:M = 1. The M-OH chains are linked via oxyanions to give three-dimensional structures. Figure below illustrates the $Cu(OH)IO_3$ structure.

A number of structures based on two-dimensional M-OH subunits can be viewed as derivatives of the CdI_2 structure. First, we take the CdI_2 structure and replace Cd by M and I_2 by

$(OH)_2$. Then we replace 25% of the OH groups with oxygen atoms of oxyanions, such that the plane of the XO_n^{z-} group is perpendicular to that of the layer.

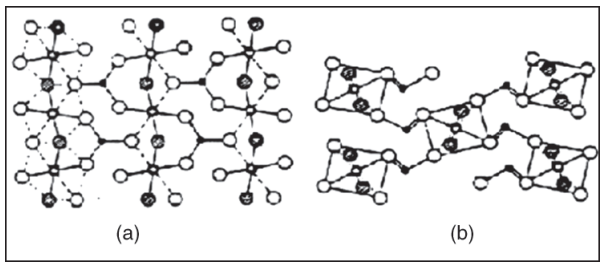


Fig. Structures of Hydroxy-oxy Salts Based on Two-dimensional M-OH Subunits. $Cu(OH)IO_3$: (a) Plan; (b) Elevation, with the Octahedral Chains Perpendicular to the Plane of the Paper.

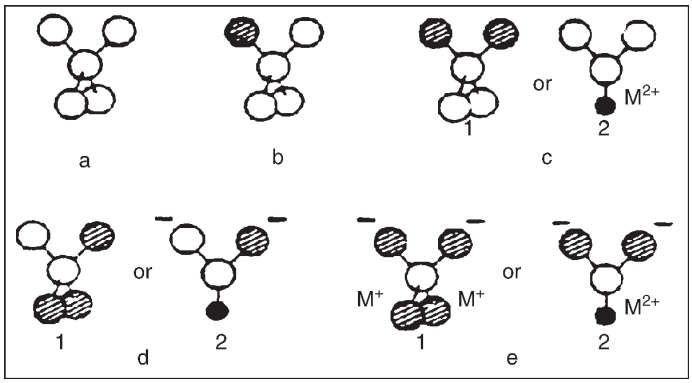


Fig. Interactions of Water Molecules in Crystals. The Larger Shaded Circles Represent M^+, OH^-, F^-, or Oxygen of oxy-ion.

Hydrated Salts

Here we shall be concerned with hydrates of *halides, oxysalt*s, and *hydroxides*. The oxygen atom in a water molecule possesses two lone pairs of electrons. These result in two regions of negative charge on one end of the molecule.

At the same time, due to the polarisation of the water molecule, the two hydrogen atoms are relatively positive, compared with the oxygen atom.

Thus the two hydrogen atoms represent two regions of positive charge.

As a result, the water molecule may be viewed as a *quadrupolar* species, with a tetrahedral arrangement of the two positive and two negative regions. The manner in which a water molecule is linked in a hydrate may be rationalised in terms of the interactions of anions, metal ions, or other water molecules with these charged regions.

A convenient classification, based on the coordination of the cation, for a hydrated salt $M_xA_y.wH_2O$, where A symbolises an anion such as Cl^-, SO_4^{2-}, and n is the coordination number of the metal ion.

In class AI, M is fully hydrated, and there is excess water ((w/xn) > 1). This is a relatively small group. A representative compound is $MgCl_2.12H_2O$. Here the cation and anion are each octahedrally surrounded by six water molecules.

Table. Classification of Hydrated Salts

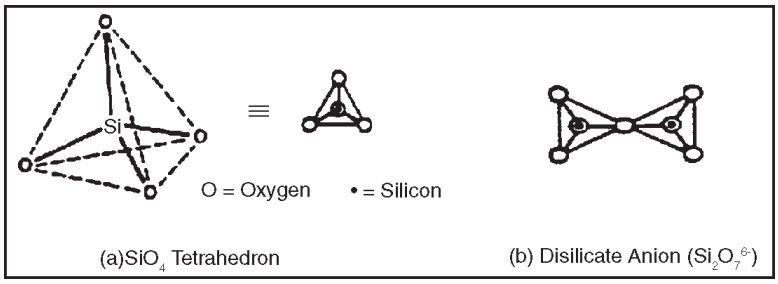
		M Fully Hydrated		M Incompletely Hydrated	
	w/xn	**Excess H_2O (I)**	**No Excess H_2O (II)**	**Excess H_2O (III)**	**Excess H_2O (IV)**
A	2	$MgCl_2\times12H_2O$			
	9/6	$FeBr_2\times9H_2O$			
	7/6	$NiSO_4\times7H_2O$			
	1	-	$Sm(BrO_3)_3\times9H_2O$	$CoCl_2\times6H_2O$	
	1	-	$AlCl_3\times6H_2O$	$NiCl_2\times6H_2O$	
	1	-	$BeSO_4\times4H_2O$	$FeCl_3\times6H_2O$	
B	5/6	$Na_2SO_4\times10H_2O$	$Na_2CO_3\times10H_2O$	$CuSO_4\times5H_2O$	
	3/4				$CaCO_3\times6H_2O$
	1/2		$LiClO_4\times3H_2O$	$FeF_3\times3H_2O$	$CuSO_4\times3H_2O$
C	1/3				$CoCl_2\times2H_2O$
					$NiCl_2\times2H_2O$
	1/4				$CaSO_4\times2H_2O$
	1/6				$CuSO_4\times H_2O$

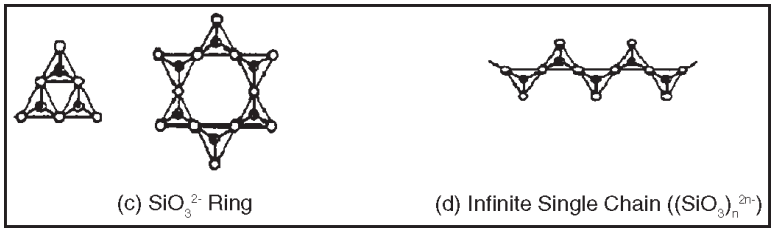
Metal Silicates

Silicon and oxygen make up about 75% of the earth's crust, where they exist primarily as silicate minerals. The basic building block in silicates is the SiO_4 tetrahedron, it is convenient to represent this as a triangle, which symbolises a top-down view of the tetrahedron.

The complex structures of silicates can be resolved in terms of the manner in which the SiO_4 tetrahedra are linked. Examples of the possible linkages are respectively for the disilicate anion ($Si_2O_7^{6-}$), a ring (SiO_3^{2-}), an infinite single chain $(SiO_3)_n^{2n-}$, an infinite double chain or band, $(Si_4O_{11})_n^{6n-}$, and a sheet or layer, $(Si_2O_5)_n^{2n-}$.

The charge balance in the linked structures can be obtained by ascribing unit negative charge (–1) to a terminal O atom and zero charge to a shared O atom. Traditional nomenclature: the label *ortho* is used to signify that a species is completely hydrated, while the label *meta* indicates the elimination of a mole of water from the ortho species. A selection of silicate compounds, with corresponding SiO_4 tetrahedral linkages.





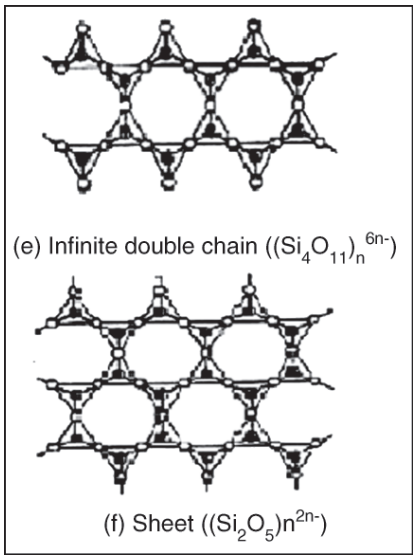


Fig. SiO_4 Tetrahedral Linkages.

Table: Silicate Compounds

Building Block	**Compounds**
Discrete orthosilicate (SiO_4^{4-}) anions	Mg_2SiO_4, $(Mg, Fe)_2SiO_4$ (olivine) Na_4SiO_4, K_4SiO_4, Be_2SiO_4 (phenacite) Zn_2SiO_4 (willemite), $ZrSiO_4$ (zircon) $M(II)_3M(III)_2(SiO_4)_3$, M(II) = Ca^{2+}, Mg^{2+}, Fe^{2+}, M(III) = Al^{3+}, Cr^{3+}, Fe^{3+} (garnets)
Discrete disilicate ($Si_2O_7^{6-}$) anions	$Sc_2Si_2O_7$ (thortveitite)
Trisilicate rings ($Si_3O_9^{6-}$)	$BaTiSi_3O_9$ (benitoite), $Na_2ZrSi_3O_9 \times H_2O$ (catapleite)
Hexasilicate rings ($Si_6O_{18}^{12-}$)	$Cu_6Si_6O_{18} \times 6H_2O$ (dioptase), $Be_3Al_2Si_6O_{18} \times 6H_2O$ (beryl)
Single Chains (Pyroxenes; $(SiO_3)_n^{2n-}$)	$MgSiO_3$ (enstatite), $CaMg(SiO_3)_2$ (diopside), $LiAl(SiO_3)_2$ (spodumene)
Double Chains (Amphiboles; $(Si_4O_{11})_n^{6n-}$)	$Na_2Fe_5(OH)_2(Si_4O_{11})_2$ (crocidolite, blue asbestos) $(Mg,Fe)_7(OH)_2(Si_4O_{11})_2$ (amosite)
Sheets ($(Si_2O_5)_n^{2n-}$)	$Al_2(OH)_4Si_2O_5$(kaolinite) $Al_2(OH)_2Si_4O_{10}$ (pyrophyllite) $Mg_3(OH)_4Si_2O_5$ (serpentine) $Mg_3(OH)_2Si_4O_{10}$ (talc)
Al-substituted sheets ($(AlSi_3O_{10})^{5-}$)	$KAl_2(OH)_2Si_3AlO_{10}$ (muscovite, white mica) $KMg_3(OH)_2Si_3AlO_{10}$ (phlogopite, Mg-mica) $K(Mg,Fe)_3(OH)_2Si_3AlO_{10}$ (biotite, black mica)
Framework	SiO_2 (quartz; tridymite, cristobalite, coesite, stishovite) $MAl_{2-x}Si_{2+x}O_8$ (feldspars) $M_{x/n}^{n+}[Al_xSi_yO_{2x+2y}]^{x-} \times zH_2O$ (zeolites)

It should be noted that silicate structures can also be considered in terms of a close-packed array of O^{2-} ions, where tetrahedral holes are occupied by Si^{4+} ions, and octahedral and tetrahedral holes by metal ions.

Metal Sulfides

The electronegativity of sulfur is less than that of oxygen. As a result,M-S bonds tend to be less polar than M-O bonds. Thus, whereas the crystal structures of metal oxides tend to be similar to those of the corresponding fluorides, metal sulfides prefer the structures of the more covalent chlorides, bromides, and iodides. For example, layer structures are unusual among oxides and fluorides, but are frequently observed among sulfides and the heavier halides. Many of the sulfides of interest in aqueous processing have structures which are more complex. However, the structures of these more complex sulfides can often be viewed as derivatives of the simpler structures.

Table. Crystal Structures of Metal Sulfides

Type of Structure		M CN:S CN	Name of Structure	Examples
Infinite	6:6	Sodium chloride	CaS, MnS, PbS, LaS	
3-dimensional	6:6	Nickel arsenide	FeS, CoS, NiS, VS, TiS	
complexes	6:6	Pyrite	FeS_2, CoS_2, NiS_2, MnS_2, OsS_2, RuS_2	
	4:4	Sphalerite	BeS, ZnS, CdS, HgS	
	4:4	Wurtzite	ZnS, CdS, MnS	
	4:4	Cooperite	PtS	
Layer structures	6:3	Cadmium iodide	TiS_2, ZrS_2, SnS_2, PtS_2, TaS_2, HfS_2	
	6:3	Cadmium chloride	TaS_2	
	6:3	Molybdenum sulfide	MoS_2, WS_2	
Chain structures			Sb_2S_3, Bi_2S_3, HgS	

Table. Metal Sulfides

Copper Sulfides	CuS (covellite), digenite ($Cu_{1.8}S$), djurleite ($Cu_{1.96}S$), Cu2S (chalcocite), $Cu_{1+x}S$ (blaubleinder covellite), Cu_3AsS_4 (enargite),Cu_5FeS_4 (bornite), $CuFeS_2$ (chalcopyrite), $CuFe_2S_3$ (cubanite), Cu_3FeS_4 (idaite)

Iron Sulfides	Fe_9S_{10} (Im pyrrhotite (5C)), $Fe_{11}S_{12}$ (Im pyrrhotite (6C)), $Fe_{10}S_{11}$ (Im pyrrhotite (11C)), FeS_2 (pyrite), FeS_2 (marcasite), $Fe_{1+x}S$ (mackinawite), Fe_9S_{10} ® $Fe_{11}S_{12}$ (Non-integral pyrrhotite (nC)), FeS (troilite), Fe_7S_8 (monoclinic pyrrhotite (4C)), Fe_3S_4 (greigite), FeAsS (arsenopyrite), $FeNi_2S_4$ (violarite), $(Fe, Ni)_9S_{11}$ (smythite)
Nickel Sulfides	NiS_2 (vaesite), NiS (millerite), Ni_3S_2 (heazlewoodite), a-Ni_7S_6, b-Ni_7S_6 (godlevskite), Ni_3S_4 (polydymite), $(Ni,Fe)_9S_8$ (pentlandite)
Cobalt Sulfides	CoS (jaipurite ?), Co_3S_4 (linnaeite), CoS_2 (cattierite), Co_9S_8 (cobalt pentlandite), (Co,Fe)AsS (cobaltite)
Silver Sulfides	a-Ag_2S (argentite), b-Ag_2S (acanthite)
Arsenic Sulfides	b-AsS (realgar), a-AsS (a-arsenic sulfide), As_2S_3 (orpiment)
Zinc Sulfides	b-ZnS (sphalerite), a-ZnS (wurtzite)

3

Solid Crystallography

Crystal Structure

In mineralogy and crystallography, a crystal structure is a unique arrangement of atoms in a crystal. A crystal structure is composed of a motif, a set of atoms arranged in a particular way, and a lattice. Motifs are located upon the points of a lattice, which is an array of points repeating periodically in three dimensions.

The points can be thought of as forming identical tiny boxes, called unit cells, that fill the space of the lattice. The lengths of the edges of a unit cell and the angles between them are called the lattice parameters. The symmetry properties of the crystal are embodied in its space group. A crystal's structure and symmetry play a role in determining many of its properties, such as cleavage, electronic band structure, and optical properties.

Solids

Amorphous solids are homogeneous and isotropic because there is no long range order or periodicity in their internal atomic arrangement. By contrast, the crystalline state is characterised by a regular arrangement of atoms over large distances. Crystals are therefore anisotropic – their properties vary with direction. For example, the interatomic spacing varies with orientation within the crystal, as does the elastic response to an applied stress. Engineering materials are usually aggregates of many crystals of varying sizes and shapes; these polycrystalline materials have properties which depend on the nature of the individual crystals, but also on aggregate properties such as the size and shape distributions of the crystals, and the orientation relationships between the individual crystals. The randomness in the orientation of the crystals is a measure of texture, which has to be controlled in the manufacture of transformer steels, uranium fuel rods and beverage cans. The crystallography of interfaces connecting adjacent crystals can determine the deformation behaviour of the polycrystalline aggregate; it can also influence the toughness through its effect on the degree of segregation of impurities to such interfaces.

Metals and ceramics are the basic materials of nuclear reactors. The regular arrangement of the atoms in such solids is termed a crystal lattice. Pure solids (whether elemental or binary) need not retain the same crystal structure from low temperature to the melting point; iron and uranium exhibit three crystal structures, each confined to a definite temperature interval, but uranium dioxide has only one. The change from one crystal structure to another is called a phase

transformation. Such changes are the solid-solid equivalents of melting and vaporization.

The crystal structure can exert a profound effect on material's behaviour in the environment of a nuclear reactor, as seen from the following examples:

- Void swelling can cause unacceptable dimensional changes in structures of fast reactors or fusion reactors (but not generally in light water reactors, except at higher doses and intermediate temperatures in austenitic stainless steels). The crystal structure of the austenitic form of steel is susceptible to this radiation effect but the different crystal structure of the ferritic form is immune until very high doses (>H"50 displacements per atom, dpa).
- Under fast-neutron irradiation, Zircaloy cladding of LWRs undergoes growth in the axial direction. This effect is due to the anisotropy of the crystal structure of zirconium.
- The relation between crystal anisotropy and the directional dependence of macroscopic properties in zirconium was covered in Sect I.2.
- The utility of uranium metal as a nuclear fuel is diminished because of the volume change accompanying a solid phase transformation that occurs at relatively low temperature.Imperfections in the crystal lattice are even more important than the details of the atomic structure.
- Missing atoms in the atomic lattice called vacancies, or extra atoms in nonregular positions called interstitials, are responsible for the mobility of atoms in the solid

- Many of the mechanical properties of metals are governed by a linear defect in the atomic structure called a dislocation
- All structural alloys and ceramic fuels consist of a large number of small crystals, or grains, joined at their peripheries. The discontinuities in the crystal structure at these internal surfaces are called grain boundaries. Their presence influences a variety of mechanical and transport properties, including cracking in steel and fission gas movement in fuel.

For elemental solids, only 14 distinct structures, or spatial arrangements of atoms, are possible. Each crystal structure is characterized by a unit cell, which is a small group of atoms that contains the unique features of the particular lattice type. A macroscopic crystal is composed of large numbers of unit cells stacked together.

The elemental and binary solids important in LWRs are restricted to only a few of the 14 possible lattice types, mainly those with the highest degrees of symmetry and the tightest packing of atoms. With one exception, these unit cells are parallelepipeds (all angles 90o) with side lengths denoted by a, b, and c.

The exception is a structure whose unit cell is a hexagonal prism with only two characteristic dimensions. Only the parallelepiped with all sides of equal length (*i.e.*, a cube) is isotropic. All the others are anisotropic, meaning that properties differ in the principal coordinate directions.

The Lattice

Crystals have translational symmetry: it is possible to identify a regular set of points, known as the lattice points,

each of which has an identical environment. The set of these lattice points constitutes a three dimensional lattice. A unit cell may be defined within this lattice as a space–filling parallelepiped with origin at a lattice point, and with its edges defined by three non-coplanar basis vectors a1, a2 and a3, each of which represents translations between two lattice points. The entire lattice can then be generated by stacking unit cells in three dimensions. Any vector representing a translation between lattice points is called a lattice vector.

The unit cell defined above has lattice points located at its corners. Since these are shared with seven other unit cells, and since each cell has eight corners, there is only one lattice point per unit cell. Such a unit cell is primitive and has the lattice symbol P. Non–primitive unit cells can have two or more lattice points, in which case, the additional lattice points will be located at positions other than the corners of the cell. A cell with lattice points located at the centres of all its faces has the lattice symbol F; such a cell would contain four lattice points. Not all the faces of the cell need to have face–centering lattice points; when a cell containing two lattice points has the additional lattice point located at the centre of the face defined by a2 and a3, the lattice symbol is A and the cell is said to be A-centred. B-centred and C-centred cells have the additional lattice point located on the face defined by a3 & a1 or a1 & a2 respectively.

A unit cell with two lattice points can alternatively have the additional lattice point at the body-centre of the cell, in which case the lattice symbol is I. The lattice symbol R is for a trigonal cell; the cell is usually defined such that it contains three lattice points. The basis vectors a1, a2 and a3 define

the the unit cell; their magnitudes a1, a2 and a3 respectively, are the lattice parameters of the unit cell. The angles a1^a2, a2^a3 and a3 ^a1 are conventionally labelled g,aand b respectively. Note that our initial choice of the basis vectors was arbitrary since there are an infinite number of lattice vectors which could have been used in defining the unit cell. The preferred choice includes small basis vectors which are as equal as possible, provided the shape of the cell reflects the essential symmetry of the lattice.

The Bravais Lattices

The number of ways in which points can be arranged regularly in three dimensions, such that the stacking of unit cells fills space, is not limitless; Bravais showed in 1848 that all possible arrangements can be represented by just fourteen lattices. The fourteen Bravais lattices can be categorised into seven crystal systems (cubic, tetragonal, orthorhombic, trigonal, hexagonal, monoclinic and triclinic); the cubic system contains for example, the cubic-P, cubic-F and cubic-I lattices. Each crystal system can be characterised uniquely by a set of defining symmetry elements, which any crystal within that system must possess as a minimum requirement.

The Bravais lattices are each as a projection along the a_3 axis. Projections like these are useful as simple representations of three–dimensional objects. The coordinate of any point with respect to the a_3 axis is represented as a fraction of a_3 along the point of interest; points located at 0a3 are unlabelled; translational symmetry requires that for each lattice point located at $0a_3$, there must exist another at $1a_3$.

Directions

Any vector u can be represented as a linear combination of the basis vectors ai of the unit cell (i = 1, 2, 3):

$$u = u_1a_1 + u_2a_2 + u_3a_3$$

and the scalar quantities u_1, u_2 and u_3 are the components of the vector u with respect to the basis vectors a_1, a_2 and a_3. Once the unit cell is defined, any direction u within the lattice can be identified uniquely by its components [u1 u2 u3], and the components are called the Miller indices of that direction and are by convention enclosed in square brackets.

Planes

If a plane intersects the a1, a2 and a3 axes at distances x1, x2 and x3 respectively, relative to the origin, then the Miller indices of that plane are given by (h1 h2 h3) where:

$$h_1 = \phi\frac{a_1}{x_1}, \quad h_2 = \phi\frac{a_2}{x_2}, \quad h_3 = \phi\frac{a_3}{x_3},$$

ϕ is a scalar which clears the numbers *hi* off fractions or common factors. Note that *xi* are negative when measured in the "*ai* directions. The intercept of the plane with an axis may occur at 1, in which case the plane is parallel to that axis and the corresponding Miller index will be zero. Miller indices for planes are by convention written using round brackets: (h1 h2 h3) with braces being used to indicate planes of the same form: {h1 h2 h3}.

The Reciprocal Lattice

The reciprocal lattice is a special co–ordinate system. For a lattice represented by basis vectors a1, a2 and a3, the corresponding reciprocal basis vectors are written a_1^*, a_2^* and a_3^*, such that:

$$a_1^* = \frac{(a_2 \wedge a_3)}{(a_1.a_2 \wedge a_3)}$$

$$a_2^* = \frac{(a_3 \wedge a_1)}{(a_1.a_2 \wedge a_3)}$$

$$a_3^* = \frac{(a_1 \wedge a_2)}{(a_1.a_2 \wedge a_3)}$$

In equation 2a, the term $(a_1.a_2 \wedge a_3)$ represents the volume of the unit cell formed by ai, while the magnitude of the vector $(a_2 \wedge$ a3) represents the area of the (1 0 0)A plane. Since $(a_2 \wedge a_3)$ points along the normal to the (1 0 0)A plane, it follows that a_1^* also points along the normal to (1 0 0)A and that its magnitude $||$ is the reciprocal of the spacing of the (1 0 0)A planes. The components of any vector referred to the reciprocal basis represent the Miller indices of a plane whose normal is along that vector, with the spacing of the plane given by the inverse of the magnitude of that vector. For example, the reciprocal lattice vector u* = (1 2 3) is normal to planes with Miller indices (1 2 3) and interplanar spacing $1/|u^*|$.

$a_i \cdot a_j^* = 1$ when i = j, and $a_i \cdot a_j^* = 0$ when i = *j*

or in other words, $a_i \cdot a_j^* = \delta_{ij}$

δ_{ij} is the Kronecker delta, which has a value of unity when i = j and is zero when i = *j*

Electrons in Metals

Metals are electrical conductors because some of the electrons are able to move freely, even though they move in a periodic array of positive metal ions. Electrons are waves, characterised by a wave number *k*,

$$k = \pm \frac{2\pi}{\lambda}$$

where λ is the wavelength. The lattice does not affect the motion of these electrons except at critical values of k, where the lattice planes reflect the electrons†. For a square lattice, these critical values of *k* are given by,

$$k = \frac{n\pi}{a \sin \theta}$$

where a is the lattice parameter and μ is the direction of motion. The component of k along the *x* axis is $kx = k \sin \mu$ and since μ = 0, it follows that reflection occurs when

$$kx = \pm \frac{1/4}{a}$$

and similarly when

$$ky = \pm \frac{1/4}{a},$$

This can be represented in *k*–space, which is reciprocal space with a reciprocal lattice parameter of magnitude 2p/a. Electrons are reflected in the lattice whenever k falls on any point on the square ABCD.

This square is known as the first Brillouin zone boundary. Electron energy contours are frequently plotted in *k*–space. Inside a Brillouin zone, the energy of electrons can increase 'smoothly' with *k*; there is then a discontinuity (an energy gap) at the zone boundary.

Unit Cell

The crystal structure of a material or the arrangement of atoms in a crystal structure can be described in terms of its unit cell. The unit cell is a tiny box containing one or more motifs, a spatial arrangement of atoms. The units cells stacked in three-dimensional space describes the bulk arrangement of atoms of the crystal. The unit cell is given by its lattice parameters, the length of the cell edges and the

angles between them, while the positions of the atoms inside the unit cell are described by the set of atomic positions (*xi*, *yi*, *zi*) measured from a lattice point.

Although there are an infinite number of ways to specify a unit cell, for each crystal structure there is a conventional unit cell, which is chosen to display the full symmetry of the crystal. However, the conventional unit cell is not always the smallest possible choice.

A primitive unit cell of a particular crystal structure is the smallest possible volume one can construct with the arrangement of atoms in the crystal such that, when stacked, completely fills the space. This primitive unit cell will not always display all the symmetries inherent in the crystal. A Wigner-Seitz cell is a particular kind of primitive cell which has the same symmetry as the lattice. In a unit cell each atom has an identical environment when stacked in 3 dimensional space. In a primitive cell, each atom may not have the same environment.

Symmetry

Although the properties of a crystal can be anisotropic, there may be different directions along which they are identical. These directions are said to be equivalent and the crystal is said to possess symmetry. That a particular edge of a cube cannot be distinguished from any other is a measure of its symmetry; an orthorhombic parallelepiped has lower symmetry, since its edges can be distinguished by length. Some symmetry operations are illustrated in Fig. 6; in essence, they transform a spatial arrangement into another which is indistinguishable from the original. The rotation of a cubic lattice through 90° about an axis along the edge of

the unit cell is an example of a symmetry operation, since the lattice points of the final and original lattice coincide in space and cannot consequently be distinguished. We have already encountered translational symmetry when defining the lattice; since the environment of each lattice point is identical, translation between lattice points has the effect of shifting the origin.

Symmetry Operations

An object possesses an n-fold axis of rotational symmetry if it coincides with itself upon rotation about the axis through an angle $360°\pm/n$. The possible angles of rotation, which are consistent with the translational symmetry of the lattice, are 360°, 180°, 120°, 90°and 60° for values of n equal to 1, 2, 3, 4 and 6 respectively. A five–fold axis of rotation does not preserve the translational symmetry of the lattice and is forbidden. A one-fold axis of rotation is called a monad and the terms diad, triad, tetrad and hexad correspond to $n = 2$, 3, 4 and 6 respectively.

All of the Bravais lattices have a centre of symmetry. An observer at the centre of symmetry sees no difference in arrangement between the directions $[u_1\ u_2\ u_3]$ and $[-u_1\ -u_2\ -u_3]$. The centre of symmetry is such that inversion through that point produces an identical arrangement but in the opposite sense. A rotoinversion axis of symmetry rotates a point through a specified angle and then inverts it through the centre of symmetry such that the arrangements before and after this combined operation are in coincidence. For example, a three-fold inversion axis involves a rotation through 120° combined with an inversion, the axis being labelled $\bar{3}$.

A rotation operation can also be combined with a translation parallel to that axis to generate a screw axis of symmetry. The magnitude of the translation is a fraction of the lattice repeat distance along the axis concerned. A 31 screw axis would rotate a point through 120° and translate it through a distance t/3, where t is the magnitude of the shortest lattice vector along the axis. A 32 operation involves a rotation through 120° followed by a translation through 2t/3 along the axis. For a right–handed screw axis, the sense of rotation is anticlockwise when the translation is along the positive direction of the axis. A plane of mirror symmetry implies arrangements which are mirror images. Our left and right hands are mirror images. The operation of a 2 axis produces a result which is equivalent to a reflection through a mirror plane normal to that axis. The operation of a glide plane combines a reflection with a translation parallel to the plane, through a distance which is half the lattice repeat in the direction concerned. The translation may be parallel to a unit cell edge, in which case the glide is axial; the term diagonal glide refers to translation along a face or body diagonal of the unit cell. In the latter case, the translation is through a distance which is half the length of the diagonal concerned, except for diamond glide, where it is a quarter of the diagonal length.

Classification of Crystals by Symmetry

The defining property of a crystal is its inherent symmetry, by which we mean that under certain operations the crystal remains unchanged. For example, rotating the crystal 180 degrees about a certain axis may result in an atomic configuration which is identical to the original configuration.

The crystal is then said to have a twofold rotational symmetry about this axis. In addition to rotational symmetries like this, a crystal may have symmetries in the form of mirror planes and translational symmetries, and also the so-called compound symmetries which are a combination of translation and rotation/mirror symmetries. A full classification of a crystal is achieved when all of these inherent symmetries of the crystal are identified.

Crystal systems

The crystal systems are a grouping of crystal structures according to the axial system used to describe their lattice. Each crystal system consists of a set of three axes in a particular geometrical arrangement. There are seven unique crystal systems. The simplest and most symmetric, the cubic (or isometric) system, has the symmetry of a cube, that is, it exhibits four threefold rotational axes oriented at 109.5 degrees (the tetrahedral angle) with respect to each other. These threefold axes lie along the body diagonals of the cube. This definition of a cubic is correct, although many textbooks incorrectly state that a cube is defined by three mutually perpendicular axes of equal length – if this were true there would be far more than 14 Bravais lattices. The other six systems, in order of decreasing symmetry, are hexagonal, tetragonal, rhombohedral (also known as trigonal), orthorhombic, monoclinic and triclinic. Some crystallographers consider the hexagonal crystal system not to be its own crystal system, but instead a part of the trigonal crystal system. The crystal system and Bravais lattice of a crystal describe the (purely) translational symmetry of the crystal.

Point and Space Groups

The crystallographic point group or crystal class is the mathematical group comprising the symmetry operations that leave at least one point unmoved and that leave the appearance of the crystal structure unchanged. These symmetry operations can include reflection, which reflects the structure across a reflection plane, rotation, which rotates the structure a specified portion of a circle about a rotation axis, inversion which changes the sign of the coordinate of each point with respect to a center of symmetry or inversion point and improper rotation, which consists of a rotation about an axis followed by an inversion. Rotation axes (proper and improper), reflection planes, and centers of symmetry are collectively called symmetry elements. There are 32 possible crystal classes. Each one can be classified into one of the seven crystal systems.

The space group of the crystal structure is composed of the translational symmetry operations in addition to the operations of the point group. These include pure translations which move a point along a vector, screw axes, which rotate a point around an axis while translating parallel to the axis, and glide planes, which reflect a point through a plane while translating it parallel to the plane. There are 230 distinct space groups.

Physical properties

Defects or Impurities in Crystals

Real crystals feature defects or irregularities in the ideal arrangements described above and it is these defects that critically determine many of the electrical and mechanical

properties of real materials. When one atom substitutes for one of the principal atomic components within the crystal structure, alteration in the electrical and thermal properties of the material may ensue. Impurities may also manifest as spin impurities in certain materials. Research on magnetic impurities demonstrates that substantial alteration of certain properties such as specific heat may be affected by small concentrations of an impurity, as for example impurities in semiconducting ferromagnetic alloys may lead to different properties as first predicted in the late 1960s. Dislocations in the crystal lattice allow shear at lower stress than that needed for a perfect crystal structure.

Crystal Symmetry and Physical Properties

Twenty of the 32 crystal classes are so-called piezoelectric, and crystals belonging to one of these classes (point groups) display piezoelectricity. All 21 piezoelectric classes lack a center of symmetry. Any material develops a dielectric polarization when an electric field is applied, but a substance which has such a natural charge separation even in the absence of a field is called a polar material. Whether or not a material is polar is determined solely by its crystal structure. Only 10 of the 32 point groups are polar. All polar crystals are pyroelectric, so the 10 polar crystal classes are sometimes referred to as the pyroelectric classes.

There are a few crystal structures, notably the perovskite structure, which exhibit ferroelectric behaviour. This is analogous to ferromagnetism, in that, in the absence of an electric field during production, the ferroelectric crystal does not exhibit a polarisation. Upon the application of an electric field of sufficient magnitude, the crystal becomes

permanently polarised. This polarisation can be reversed by a sufficiently large counter-charge, in the same way that a ferromagnet can be reversed. However, it is important to note that, although they are called ferroelectrics, the effect is due to the crystal structure, not the presence of a ferrous metal. the angle between the normals to the two intersecting faces is called interfacial angle. Incommensurate crystals have period-varying translational symmetry. The period between nodes of symmetry is constant in most crystals. The distance between nodes in an incommensurate crystal is dependent on the number of nodes between it and the base node.

Interstices

The atoms inside a unit cell do not fill all space. The empty space represents the interstices. It is often the case that these interstices can accommodate small impurity atoms. As an example, we shall consider the crystal structure of iron which at ambient temperature has the cubic–I lattice with an atom of iron at each lattice point. There are two kinds of interstitial sites capable of accommodating small atoms such as carbon or nitrogen. These are the tetrahedral and octahedral sites as illustrated in.

Crystallography and Crystal Defects

A plane of atoms can glide rigidly over its neighbour in process described as slip with the slip system defined by the plane and the direction of slip. This kind of deformation requires enormous stresses, far greater than those required to actually deform a crystal. This is because almost all crystals contain defects known as dislocations. A dislocation enables the planes to glide in a piecewise manner rather than the

rigid displacement of the entire plane. This greatly reduces the stress required to cause slip.

A good analogy to illustrate the role of a dislocation is to imagine the force required to pull an entire carpet along the floor. On the other hand, if a bump is introduced into the carpet, the force needed to move the bump along is much smaller. Crystals may also contain point defects, which are imperfections of occupation. A vacancy is when an atom is missing from a site which should be occupied. An interstitial occurs when an atom is forced into a space within the crystal structure, where atoms are not normally located. Polycrystalline materials contain many crystals; another common term for crystals in such materials is grains. Atoms in the grain boundary between crystals must in general be displaced from positions they would occupy in the undisturbed crystal. Therefore, grain boundaries are defects.

Model Crystal Structures

Bonding in Solids

According to molecular orbital theory, a covalent molecule or a metal-ligand complex forms its molecular orbitals by combining the atomic orbitals of the constituent atoms. Similarly, we can treat a solid particle as a *giant molecule* consisting of an extremely large number of molecular orbitals. We can treat molecular orbitals of a polyatomic molecule from two extreme points of view, *i.e.*, the localised and delocalised bond models. Both models are useful in the discussion of solid structures. With the localised bond approach, we can imagine a situation where one of the bonded atoms draws the shared electrons completely to itself. This will result in a completely ionic bond, and the resulting giant molecule will

be in reality an *ionic solid*, where the three-dimensional structure is based on the electrostatic interaction (bonding) between the constituent ions.

On the other hand, if the bond is completely covalent, we get a three-dimensional covalent solid such as diamond, where the structure consists of C-C bonds. At the other extreme, we can consider the giant molecule from the viewpoint of the delocalised bond model. In this multi-centered molecular orbital approach, the atomic nuclei of the solid are viewed as embedded in a sea of electrons. This situation is characteristic of the solids of the metallic elements, and this type of delocalised bonding has acquired the name *metallic bonding*. In some cases, in addition to the ionic, covalent and metallic bonds, other intermolecualar interactions, such as van der Waals forces and hydrogen bonding, contribute to the development of the three-dimensional structure.

Atomic Packing

We suggested that a solid can be usefully considered as a giant molecule. As regards the shape of this molecule, however, we were silent. In this section we wish to turn to this question: How are the constituent atoms of a solid organised spatially in the solid phase? To begin, we consider the number of ways spheres can be densely packed into space. Then we relax the constraint on dense packing; this enables us to treat the crystal structures of ionic solids. It is helpful to organise our thinking around two-dimensional layers of spheres in contact. Then we build three dimensional structures by depositing these layers one on top the other. We can begin the layer by layer stacking by allowing the spheres of the second layer to fall in

the cavities in the first layer. We then have two ways to add a third layer. In one scenario, we can use an ABAB procedure, where the third layer spheres are vertically colinear with first layer spheres. In the other scenario, which is based on an ABCABC pattern, the third layer spheres are vertically colinear with the holes in the first layer.

These two structures of ABAB and ABCABC are called *polytypes* - they are identical in two dimensions while they are mismatched in the third. For each of these polytypes, the coordination number is 12.

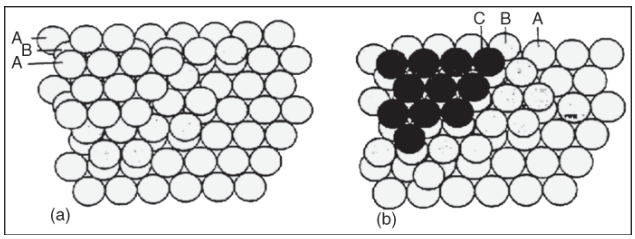


Fig. Layer by Layer Build-up of Three-dimensional Structures.

The Unit Cell

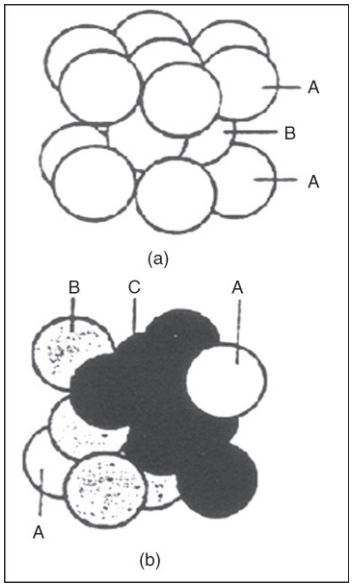


Fig. Close-packed Crystal Structures:
(a) Hexagonal Close-Packed (hcp), (b) Face-Centered Cubic (FCC).

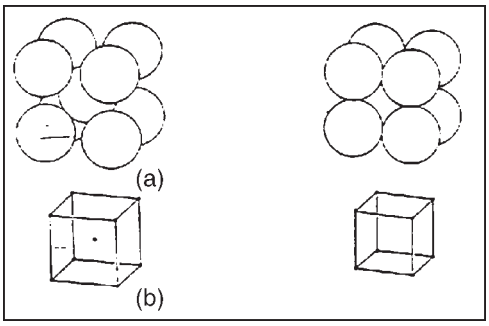


Fig. Less Close-Packed Crystal Structures:
(a) Body-Centered Cubic (BCC), (b) Primitive Cubic (Cubic-P).

The structure of a crystalline solid, it is helpful to identify the smallest three-dimensional unit, *i.e.*, *the unit cell*, that represents the repeat unit for the two close-packed structures discussed above. To help in visualizing the unit cell, it is convenient to represent each atom position as a point. The resulting system of points is termed the *lattice*, and by connecting the lattice points with straight lines, we can construct the *unit cell*. That the ABAB arrangement results in a hexagonal unit cell; the atoms are described as *Hexagonally Close-Packed (HCP)*. On the other hand, the ABCABC arrangement gives a face-centered cubic unit cell, the atoms being called *Face-Centered Cubic (FCC)*.

While the unit cells below involve close-packed equal-sized spheres, we can imagine other sphere packings which are not close-packed. A unit cell involving a coordination number of eight. This represents a *Body-Centered Cubic (BCC)* structure and it is characterised by a system of cube-center plus cube-corner lattice points. Another non-close-packed structure, the *primitive cubic (cubic-P*. This structure is derived from the bcc structure by removing the cube-center lattice point.

Typology of Crystal Structures

When we survey the myriad of solids available, whether the various materials are classified as metallic elements, ionic solids, or covalent solids, we find that the various crystal structures can be considered in terms of a small set of model crystal structures.

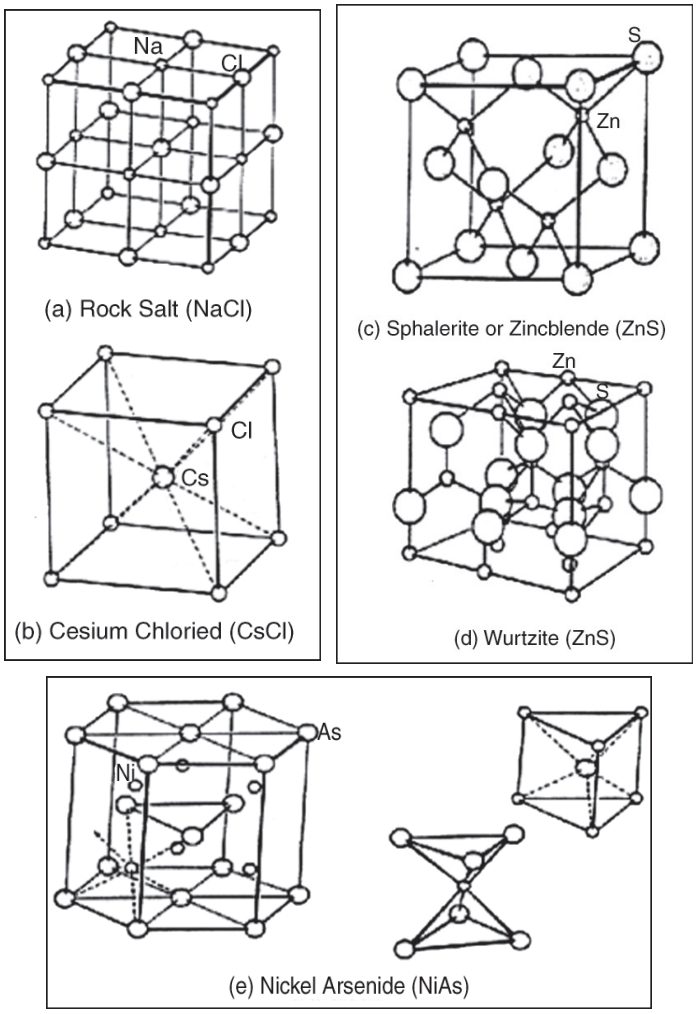


Fig. Model Crystal Structures for MX (1:1) Compounds.

In these structures the coordination number ranges from twelve to two. Metals are generally associated with large coordination numbers (8–12), while molecular solids usually

have low coordination numbers. For ionic solids, a coordination of 6 is typically encountered. Four of the model structures have already been discussed above, *i.e.*, the Face Centered Cubic (FCC), the Hexagonal Close-Packed (HCP), the body centered cubic (bcc), and simple or primitive cubic (cubic-P).

Model crystal structures for compounds with 1:1 stoichiometry (*i.e.*, chemical formula MX) are presented in Figure. In the *rock-salt* (NaCl) structure, chloride anions are arranged in an fcc pattern, and the sodium cations reside in the octahedral holes. The resulting coordination is described as (6,6) since each ion has six counterions as nearest neighbours. The *cesium chloride* (CsCl) structure, which is cubic, is derivative of the ideal BCC structure, with chloride anions at the corners and Cs cations at the cube center. The coordination is (8,8). The *sphalerite or zincblende* (ZnS) structure may be viewed in terms of an fcc array of sulfide anions with zinc cations in one set of the tetrahedral holes. The coordination is (4,4). In the case of the *wurtzite* (ZnS) structure, the sulfur anions are arranged in an HCP pattern and the zinc cations reside in one set of the tetrahedral holes. The *nickel arsenide* (NiAs) structure is founded on an HCP pattern of As atoms, with nickel atoms filling the octahedral holes. The Ni atoms are arranged into a trigonal prism, within which the an atom resides.

Model structures for compounds with 1:2 (or 2:1) stoichiometry, *i.e.*, compounds of the type MX_2 (or M_2X). The *rutile* (titanium dioxide, TiO_2) structure is based on an HCP lattice of anions in which only 50% of the octahedral holes are occupied by the cations. The coordination number of Ti

is six, while that of oxygen is three. In the *fluorite* (CaF_2) structure, the Ca_2^+ ions are arranged as an FCC array; the fluoride ions occupy both sets of tetrahedral holes, giving (8,4) coordination.

For compounds with the formula M_2X, the cation and anion positions are reversed, and the resulting structure is called *antifluorite*. In *b-cristobalite* (SiO_2), each Si atom is surrounded by four O atoms, while each O atom lies between two Si atoms.

This structure may be derived from the sphalerite structure by replacing each zinc and sulfur atom with a silicon atom and inserting oxygen atoms between the silicon atoms. In the *cadmium iodide* (CdI_2) structure, a layer of Cd atoms is sandwiched in between two layers of iodine atoms. Adjacent sandwiches are linked only via van der Waals forces.

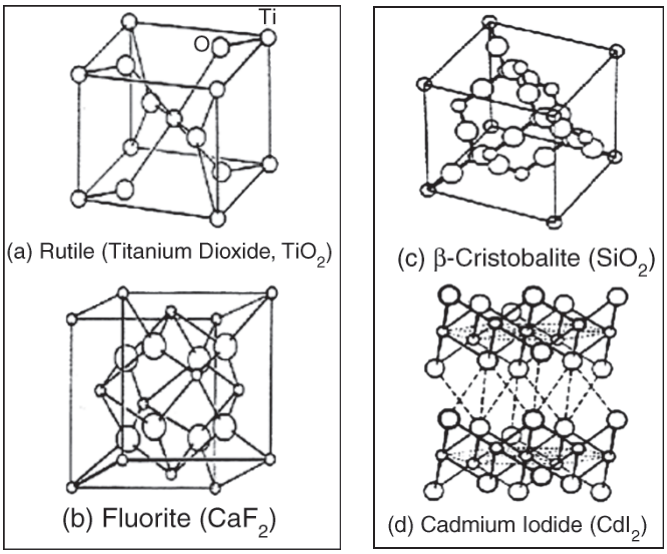


Fig. Model Crystal Structures for MX_2 (M_2X) compounds.

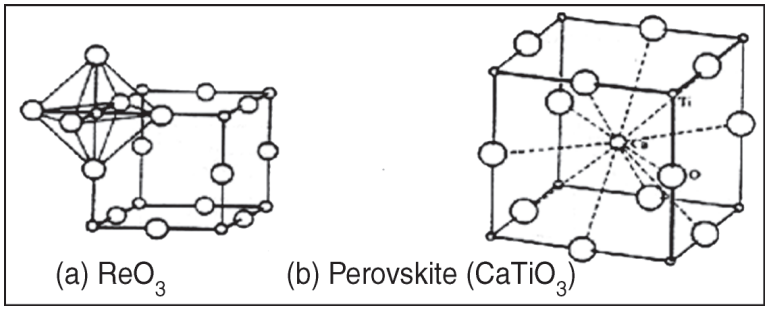


Fig. Model Crystal Structures for ABX_3 Compounds: (a) ReO_3, (b) Perovskite ($CaTiO_3$).

The model structure for many ABX_3 compounds is the *perovskite* ($CaTiO_3$) structure. This structure, in turn, may be derived from the ReO_3 structure. The ReO_3 structure is cubic; Re atoms occupy cube-corners, while O atoms are located at the edges of the unit cube. When the structure is extended beyond the unit cell, it can be seen that it is characterised by octahedra which are linked at the cube-corners. An important feature of the ReO_3 structure is the presence of a large twelve-coordinate central hole. By inserting a large anion A into this hole, one obtains the perovskite structure. In the perovskite formula of ABX_3, X is typically O^{2-} or F^-. In order to maintain electroneutrality, the total charge carried by the A and B atoms must add up to six.

Many metal compounds of general formula AB_2O_4, have crystal structures that are modeled after that of *spinel* ($MgAl_2O_4$). In this structure, O^{2-} ions are organised into an fcc pattern, and 1/8 of the tetrahedral holes are occupied by A ions, while 1/2 of the octahedral holes contain B ions. The identity of the octahedral ions may be highlighted by using square brackets, as in $A[B_2]O_4$. In some cases, an *inverse spinel* $A[A,B]O_4$ may form.

4

Solid State Phase

Real crystals are never perfect: they always contain a considerable density of defects and imperfections that affect their physical, chemical, mechanical and electronic properties. The existence of defects also plays an important role in various technological processes and phenomena such as annealing, precipitation, diffusion, sintering, oxidation and others.

It should be noted that defects do not necessarily have adverse effects on the properties of materials. There are many situations in which a judicious control of the types and amounts of imperfections can bring about specific characteristics desired in a system. This can be achieved by proper processing techniques. In fact, "defect engineering" is emerging as an important activity.

All defects and imperfections can be conveniently considered under four main divisions: point defects, line

defects or dislocations, planar defects or interfacial or grain boundary defects, and volume defects. We can also add here macroscopic or bulk defects such as pores, cracks and foreign inclusions that are introduced during production and processing of the solid state.

Point defects are inherent to the equilibrium state and thus determined by temperature, pressure and composition of a given system. The presence and concentration of other defects, however, depend on the way the solid was originally formed and subsequently processed.

Briefly consider the effects of imperfections or crystal defects on a few important properties of solids. The electrical behaviour of semiconductors, for example, is largely controlled by crystal imperfections. The conductivity of silicon can thus be altered in type (n or p) and by over eight orders of magnitude through the addition of minute amounts of electrically active dopant elements. In this case, each atom of dopant, substitutionally incorporated, represents a point defect in the silicon lattice.

The fact that such small amounts of impurity atoms can significantly alter the electrical properties of semiconductors is responsible for the development of the transistor and has opened up the entire field of solid state device technology.

Practically none of the semiconducting properties that led to these engineering accomplishments are found in a "perfect" crystal. They are properties peculiar to the defective solid state. The existence of dislocations (line defects) in crystals provides a mechanism by which permanent change of shape or mechanical deformation can occur.

A crystalline solid free of dislocations is brittle and practically useless as an engineering material. While the existence of dislocations in crystals insures ductility (ability to deform), the theoretical strength of crystalline solids is drastically reduced by their presence.

We should recognize that dislocations play a central role in the determination of such important properties as strength and ductility. In fact, virtually all mechanical properties of crystalline solids are to a significant extent controlled by the behaviour of line imperfections.

The ability of a ferromagnetic material (such as iron, nickel or iron oxide) to be magnetized and demagnetized depends in large part on the presence of two-dimensional imperfections known as Bloch walls. These interfaces are boundaries between two regions of the crystal which have a different magnetic state.

As magnetization occurs, these defects migrate and by their motion provide the material with a net magnetic moment. Without the existence of Bloch walls all ferromagnetic materials would be permanent magnets. In fact, electromagnets would not exist if it were not for this type of defect.

The presence of surface defects such as cracks causes brittle materials like glass to break at small applied stresses. This fact is familiar to anyone who has broken a glass tube by first filing a small notch (or crack) into the surface.

Removal of cracks from the surface of glass either by etching in hydrofluoric acid or by flame polishing almost always raises the fracture strength.

For example, glass in the absence of any surface cracks has a fracture strength of ~10^{10} Newton/m^2 (as opposed to real glass which has a fracture strength of ~10^7 Newton/m^2).

Formation of Point Defects

An incontrovertible law of nature states: "Nothing is perfect". This law applies to humans as well as to the inorganic world of crystalline solids and can be formulated as the 2nd law of thermodynamics:

$F = H - TS$

where F is the free energy of a given system, H is the heat content or enthalpy and TS is the entropy, or disorder, term. If a reaction takes place at a temperature T, we find the change in F (Δ F) related to a change in H (Δ H), the heat content, and possibly also a change in TS (TΔ S).

Such is the case when defects are formed in a perfect solid: The energy distribution in a solid (Maxwell-Boltzmann) suggests that a number of individual atoms may acquire enough thermal energy to be displaced from the equilibrium lattice site into an interstitial position.

This process of point defect formation requires energy and leads to lattice strain which constitutes, as discussed earlier, an increase in the heat content of the system (Δ H is positive and increases linearly with the number of defects formed). The departure from perfection by the generation of defects leads to disorder (Δ S is positive).

The magnitude of disorder generated (Δ S) is very large during the initial step from perfection to slight disarray, but the increase in disorder (with a given number of defects generated) decreases as the overall disorder increases.

Correspondingly the term TΔ S drops rapidly at the beginning and then flattens out.

The net result, free energy, exhibits a minimum for a certain number of defects in the solid [equilibrium defect density = f (temperature)]; the Fminimum suggests also that the transition from perfection to equilibrium defect structure is spontaneous: it occurs naturally.

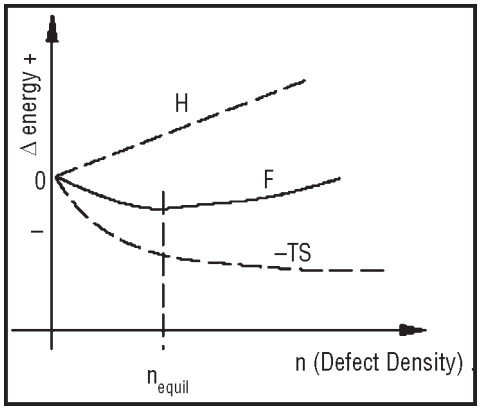


Fig. Thermodynamics of Point Defect Formation.

While the detailed mechanisms for the formation of atomic vacancies in solids are still the subject of extensive research, the associated equilibrium energetics are clear: calculations of the thermal energy of atoms in a lattice show that the average vibrational energy of lattice atoms is much less than 1 eV (the approximate energy change associated with vacancy formation, i.e., the least amount of energy required to form a vacancy) at room temperature.

Therefore a lattice atom will only acquire the energy ΔH_d, the energy required to form the defect, upon the occurrence of a large energy fluctuation. Since the relative probability of an atom having an energy ΔH_d or more in excess of the ground state energy is,

$$e^{-\Delta H_d / kT},$$

the probability that an atomic site is vacant varies in the same way. In a (molar) crystal containing N atomic sites, the number n_d of vacant sites is, therefore,

$$n_d = ANe^{-\Delta H_d / kT}$$

n_d = The number of defects (in equilibrium at T)

N =The total number of atomic sites per mole

DH_d =The energy necessary to form the defect

T = The absolute temperature (K)

k = The Boltzmann constant

A = proportionality constant

Point Defects in "Pure" Metallic Systems

Point defects in "pure" crystalline metals are defects of atomic dimensions, such as impurity atoms, the absence of a matrix atom and/or the presence of a matrix atom in the wrong place.

Some of these point defects. An impurity atom that:

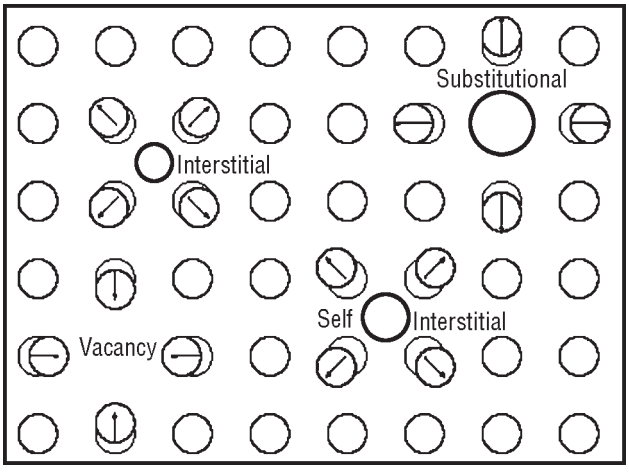


Fig. Point Defects in Crystalline Solids.

Interstitial site depends largely on the size of the atom relative to the size of the site. Small atoms are usually interstitial impurities, while larger atoms are usually substitutional impurities.

A vacancy is an atom site, normally occupied in the perfect crystal, from which an atom is missing. Often the term "vacancy" is used to denote a so-called Schottky defect, which is formed when an atom or an ion leaves a normal lattice site and repositions itself in a lattice site on the surface of the crystal.

This may be the result of atomic rearrangement in an existing crystal at a high temperature when atomic mobility is high because of increased thermal vibrations.

A vacancy may also originate in the process of crystallization as a result of local disturbances during the growth of new atomic planes on the crystal surface.

Vacancies are point defects of a size nearly equal to the size of the original (occupied) site; the energy of the formation of a vacancy is relatively low - usually less than 1 eV.

The number of vacancies at equilibrium at each temperature in a crystal can be determined from equation, in which ΔH_d is the energy necessary to take an atom from a regular site of the crystal and place it on the surface for a Schottky-type defect.

When a solid is heated a new higher equilibrium concentration of vacancies is established, usually first at crystal surfaces and then in the vicinity of dislocations and grain boundaries which provide sites for the atoms which have left their normal lattice site.

Vacancies gradually spread throughout the crystal (from the surfaces into the bulk). On cooling the vacancy concentration is lowered by "diffusion of vacancies" to grain boundaries or dislocations, which act as sinks.

In both cases, the new equilibrium vacancy concentration is established only after a finite amount of time. The rate at which vacancies move from point to point in the lattice decreases exponentially with decreasing temperature. Thus, on very rapid cooling (quenching) from a high temperature near the melting point most of the vacancies do not have time to diffuse to sinks and are said to be "frozen in". This gives a considerably greater ("non-equilibrium") concentration of vacancies in quenched specimens than that indicated by the thermal equilibrium value.

The concentration of vacant lattice sites in pure materials is very small at low temperatures - about one vacancy every 108 atom sites - and increases with increasing temperature to about one vacancy every 103 sites at the melting temperature.

Vacancies are important because they control the rate of matrix (or substitutional) atom diffusion - i.e., atoms are able to move around in a crystalline solid primarily because of the presence of vacancies.

The mechanism by which they move is the same as that associated with moving a car in a filled parking lot to the exit. Self-interstitials are generally not encountered in close-packed metallic systems, but may be introduced by irradiation.

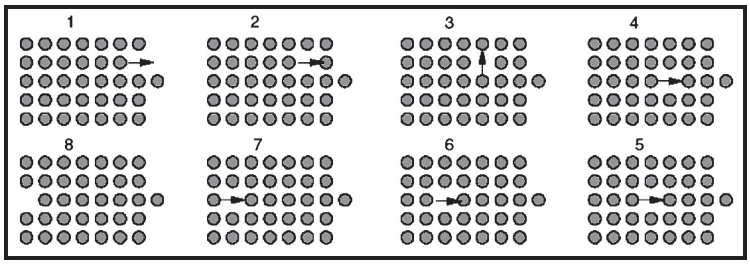


Fig. Dynamics of Vacancy Movements in a Close Packed Solid.

For example, high-energy neutrons from atomic fission can knock metal atoms from their regular sites into interstitial sites, creating vacancy-interstitial pairs.

Point Defects in Ionic Solids

Point defects in ionic structures differ from those found in pure elements because of the charge neutrality requirement. For example, in a pure monovalent ionic material a cation vacancy must have associated with it either a cation interstitial or an anion vacancy to maintain charge neutrality.

Similar requirements hold for anion vacancies. A vacancy pair defect (migration of a cation and an anion to the surface) is usually called a Schottky imperfection, and a vacancy-interstitial pair defect is referred to as a Frenkel imperfection (an anion or cation has left its lattice position, which becomes a vacancy, and has moved to an interstitial position).

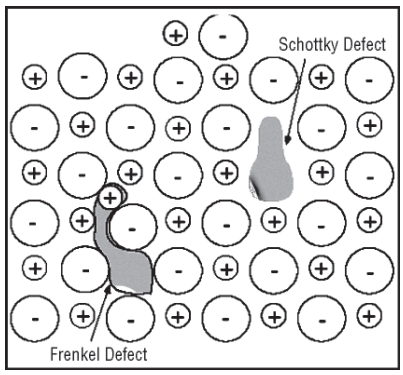


Fig : Point Defects in Ionic Solids.

Self-interstitials are much more common in ionic structures than in pure elements because many ionic compounds have relatively large interstitial sites available. That is, there are often interstitial sites in the unit cell that have nearly the same surroundings as normal atom sites. (For example, in BeO the Be atoms fill only one-half the

available tetrahedral sites, leaving four possible cation interstitial sites per unit cell. Thus a Be atom could go from a regular lattice site to an almost equivalent interstitial site with little distortion of the lattice.)

Foreign atoms in ionic crystals produce defects that also must maintain charge neutrality. For example, in NaCl a monovalent cation, such as lithium, may simply replace one of the sodium ions as a substitutional impurity. But a divalent cation, such as calcium, replacing a sodium ion must be accompanied by either a cation vacancy or an anion interstitial if charge neutrality is to be maintained. Correspondingly, monovalent impurity cations in a divalent structure (e.g., Na in MgO) must be accompanied by an appropriate number of cation interstitials or anion vacancies.

Point Defects in Covalently Bonded Solids

Substitutional impurities in covalently bonded materials can create a unique imperfection in the electronic structure if the impurity atom is from a group in the periodic table other than the matrix atoms. For example, you already considered Group V and Group III elements in a Group IV matrix, such as As or B in Si.

When foreign atoms are incorporated into a crystal structure, whether in substitutional or interstitial sites, we say that the resulting phase is a solid solution of the matrix material (solvent) and the foreign atoms (solute). The term "solid solution", however, is not restricted to the low solute contents of doped semiconductor systems; there are many solid solutions, such as metallic alloys, that comprise a wide composition range.

States of Matter

Condensed-matter physics is the fundamental science of solids and liquids, states of matter in which the constituent atoms are sufficiently close together that each atom interacts simultaneously with many neighbours. It also deals with states intermediate between solid and liquid (*e.g.*, liquid crystals, glasses, and gels), with dense gases and plasmas, and with special quantum states (superfluids) that exist only at low temperatures. All these states constitute what are called the *condensed states* of matter.

Condensed-matter physics is important for two reasons. The first is that it provides the quantum-mechanical foundation of the classical sciences of mechanics, hydrodynamics, thermodynamics, electronics, optics, metallurgy, and solid-state chemistry. The second is the massive contributions that it provides to high technology. It has been the source of such extraordinary technological innovations as the transistor, superconducting magnets, solid-state lasers, and highly sensitive detectors of radiant energy.

It thereby directly affects the technologies by which people communicate, compute, and use energy and has had a profound impact on non-nuclear military technology. At the fundamental level, research in condensed-matter physics is driven by the desire to understand both the manner in which the building blocks of condensed matter—electrons and nuclei, atoms and molecules—combine coherently in enormous numbers ($\sim 10^{24}/cm^3$) to form the world that is visible to the naked eye, and much of the world that is not, and the properties of the systems thus formed. It is in the

fact that condensed-matter physics is the physics of systems with an enormous number of degrees of freedom that the intellectual challenges that it presents are found. A high degree of creativity is required to find conceptually, mathematically, and experimentally tractable ways of extracting the essential features of such systems, where exact treatment is an impossible task. Condensed-matter physics is intellectually stimulating also because of the discoveries of fundamentally new phenomena and states of matter, the development of new concepts, and the opening up of new subfields that have occurred continuously throughout its 60-year history. It is the field in which advances in quantum and other theories most directly confront experiment and has repeatedly served as a source or testing ground for new conceptual ways of viewing complex systems. In fact, condensed-matter physics is unique among the various subfields of physics in the frequency with which it feeds its fundamental ideas into other areas of science. Thus, advances in such sub-areas of condensed-matter physics as many-body problems, critical phenomena, broken symmetry, and defects have had a major impact on nuclear physics, elementary-particle physics, astrophysics, molecular physics, and chemistry. These advances continue and offer the promise of equally fundamental discoveries in the next decade.

At the same time, condensed-matter physics excites interest because of the well-founded expectations for applications of discoveries in it. Of all the branches of physics, condensed matter has the greatest impact on our daily lives through the technological developments to which it gives rise.

Such familiar devices as the transistor, which has led to the miniaturisation of a variety of electronic appliances; the semiconductor chip, which has made possible all the myriad aspects of the computer; magnetic tapes used in recording of all kinds; plastics for everything from kitchen utensils to automobile bodies; catalytic converters to reduce automobile emissions; composite materials used in fan jets and modern tennis rackets; and NMR tomography are but a few of the practical consequences of research in condensed-matter physics. A whole new technology, optical communications, is being developed at this time from research in condensed-matter physics, optics, and the chemistry of optical fibres.

These examples serve to illustrate the intimate connection between fundamental science and the development of basic new technology in condensed-matter physics. In both universities and industry they are carried out by people with the same research training, who use the same physics concepts and the same advanced instrumentation. Because fundamental science in condensed-matter physics is so deeply involved with technological innovation, it has a strong natural bond with industry. This is the main reason why condensed matter physics has been so successful in leading industrial innovation.

Indeed, the full extent to which the consequences of research in condensed matter physics play a role in the quality of our everyday lives, and in meeting national needs, is far greater than any such listing can indicate. In order to show this explicitly we have constructed the matrix displayed in Table, the first column of which lists the subareas of condensed-matter physics, and the first row the major areas

of human and technological activity that are of national interest. The elements of the matrix are filled in with a solid circle, indicating a critical connection between the corresponding subarea of condensed-matter physics and the area of application; a half-filled circle, indicating an important or emerging connection; an open circle, denoting the possibility of a connection; or a blank, implying that the connection is not known.

Artificially Structured Materials

One area of condensed-matter physics that has progressed remarkably in the past decade is that of artificially structured materials—materials that have been structured either during or after growth to have dimensions or properties that do not occur naturally.

The most important techniques for the creation of such materials are molecular-beam epitaxy (MBE), the molecule-by-molecule deposition of material of the desired composition from a molecular beam, and metallo-organic chemical vapor deposition (MOCVD). These are prime examples of technological breakthroughs, used primarily to make semiconductor lasers and other devices, feeding back to fundamental physics. One can fabricate artificial periodic superlattices consisting of alternating layers of different semiconductors, different metals, or semiconductors and metals, and one can also create artificial, purely two-dimensional electron gases. The latter have unique and important properties, *e.g.*, extremely high electron mobilities, which cannot be provided by metal-oxide-semiconductor (MOS) inversion layers. The new physical phenomena to which the resulting structures have given rise include the

quantised Hall effect and the fractionally quantised Hall effect. It has also been possible to grow metallic superlattices in which the electronic mean free path is appreciably longer than the period of the superlattices (the sum of the thicknesses of the two alternating metal layers). It is found that it is possible to induce new lattice structures rather easily in such superlattices. Metal/insulator superlattices are ideal systems for the study of dimensional effects in metals, *e.g.*, the crossover from two- to three-dimensional superconductivity in Nb/Ge superlattices as the Ge thickness is decreased.

The Quantised Hall Effect

Modern technology has made possible unique, purely two-dimensional electron gases (in the sense that only one quantum state is excited in the direction perpendicular to the plane of the gas, so that electronic motion in it is strictly confined to that plane). These systems show exciting properties and are a new laboratory for the study of fundamental physics. The most remarkable property of such systems is undoubtedly the quantised Hall effect.

At low temperature and high perpendicular magnetic field, the electron states are split into so-called Landau or cyclotron energy levels. It is found that when the Fermi level is between two such levels one sees an almost perfectly flat plateau or constant value of the Hall conductance, the conductance perpendicular to the electric and magnetic fields, as well as zero parallel conductance.

These plateaus are found to be quantised in units of e^2/h = 1/25,812.8 ohm^{-1}. The precision of this result, at least one part in a hundred million, has led to improvement in the

measurement of this fundamental constant and to a new portable resistance standard. More recently, quantisation of the Hall conductance in simple fractions like 1/3, 2/5, and 2/7 of e^2/h has been seen, and an explanation of this effect has been proposed, and widely accepted, that involves a completely new and unexpected ordered state of matter. In this state one proposes that a new type of elementary excitation with fractional electronic charge plays a major role.

Condensed-matter Physics in the Next Decade

The list of outstanding achievements during the past decade given in the previous section demonstrates the vitality of condensed-matter physics and is a strong indication of progress to be made in the coming decades. Many of the developments provide new experimental or theoretical tools for studying physical problems that are as yet unsolved. As examples, we cite the applications of vacuum tunneling microscopy and of intense synchrotron radiation sources, as well as the various renormalisation-group methods of theoretical analysis. There are newly discovered materials or phenomena that are only partially understood and that therefore open up new fields of enquiry. Examples here include the heavy-fermion superconductors, modulated semiconductor structures, the quantised and fractionally quantised Hall effect, and the random magnetic-field problem.

In the remainder of this section we describe some of the areas of condensed-matter physics that appear to us to provide particularly exciting research opportunities in the next decade. In compiling such a list, we are, however, cognizant of the fact that each of the preceding two or three decades has seen the discovery of physical phenomena or

methods that could not have been predicted at the beginning of that decade. For example, several of the outstanding discoveries listed in the previous section, such as the new phases of ^{3}He and the quantised Hall effect, would have been absent from a list of opportunities drawn up in 1970. It is virtually certain that the coming decade will have its share of such unexpected discoveries.

A great deal of effort is expected to be devoted to the determination of the structures and excitations of surfaces of crystalline solids, both clean and covered with adsorbed layers, and of the interfaces between two different solids and between solids and liquids. These investigations will be carried out by the use of such instruments and techniques as the recently developed scanning vacuum tunneling microscope, Rutherford ion backscattering, grazing-incidence x-ray scattering, low-energy electron diffraction, electron energy loss spectroscopy, and atomic diffraction. The results of these determinations are crucial to a detailed understanding of various surface and interface excitations, electronic, vibrational, and magnetic. We can hope that eventually we may not only determine surface structures, but control them by deliberate doping or other physical and chemical techniques. The great goal of understanding catalysis must always remain in our minds, but few experiments yet undertaken approach the real problems of catalysis in practical systems as opposed to atomically clean models.

The determination of the structure, ground-state properties, and elementary excitations in glasses, amorphous materials, and other disordered systems both magnetic and

non-magnetic will be an active area of research, because of both the interesting physical properties displayed by such systems and the technological importance of many of them. There is certain to be a vigourous effort in developing new types of disordered materials for electronic and optical applications.

The interplay of this field with computer science and even more distant fields such as neurobiology is a fascinating new development that is sure to be a major area of activity in the next decade. Here, for the first time, physics is pushing at the theoretical limits of computational complexity, and hence it is not even clear that meaningful simulation of the structure of glass, or even the folding of a random polymer or protein molecule, for example, can be carried out in principle without recourse to completely new ideas in the computer field and new and complex theoretical concepts. We may also assume that new physical and chemical insights will be necessary in this area, especially as we approach problems involving even more complicated materials such as polymer glasses and gels.

Artificially structured materials can be ordered in their structure, as in the case of artificial superlattices, or they can be ordered on a microscopic scale and yet be disordered overall, as in the cases of systems composed of small particles of one material embedded in a matrix of a second. Such materials possess transport, elastic, and optical properties that can differ considerably from those of their constituents in the bulk state, and, perhaps more importantly, these properties can be tuned in desirable ways by varying the constituents and their thicknesses in the case of superlattices

or by altering the constituents, their size distributions, and their relative concentrations in the case of mixed media. The length scales involved in artificially structured materials, however, are so small that many of the conventional methods of solid-state physics are no longer applicable in determining their physical properties. In this area interface states, ballistic transport, Kapitza resistance, quantum-well effects, noise, and electromigration and thermomigration are all topics of fundamental interest and will offer research opportunities for the future.

In the area of phase transitions, we understand equilibrium phenomena in principle very well, but the kinetics of phase transitions remains an important field of materials science, especially because of its practical value, as for example under the exotic conditions of laser pulse annealing. Kinetic questions may even underlie the problems of phase transitions in the early universe. The renormalisation group will continue to expand beyond its classic uses into more or less exotic domains, as it has already enlightened studies of chaos and disorder.

The last decade has seen the creation of new organic and polymeric materials with striking physical properties, among them metallic electrical conductivity and even superconductivity. Although a theoretical explanation of *some* of the properties of *some* of these materials is beginning to emerge, much remains to be understood, and opportunities for good theoretical work in this field will continue to exist into the foreseeable future.

The underlying physics of exotic new crystalline materials, such as heavy fermion conductors, charge-density wave

materials, and high T_c and $H_c{}^2$ superconductors such as the Chevrel phases, is expected to be elucidated in the coming decade.

It has been shown repeatedly that as materials are subjected to more extreme physical conditions they display new physical properties. Thus, within the past decade the combination of small size, low temperature, and high magnetic fields made possible the discovery of the fractionally quantised Hall effect; the ability to reach ultralow temperatures made possible the discovery of the superfluid phases of ^{3}He; the use of very high pressures enabled the rare gas xenon to be solidified into a metal; the use of very low pressures (ultrahigh vacuum) has made possible the first efforts to determine structures of surfaces uncontaminated by unwanted foreign atoms or oxide layers.

It is expected that the study of condensed matter at ever lower and higher temperatures, at higher magnetic and electric fields, at higher and lower pressures, and at much higher purities will continue into the next decade, with the range of properties of known materials being broadened thereby.

In the field of non-linear dynamics, instabilities, and chaos, many questions have been raised by the work of the past decade, and efforts to answer them will surely parallel the discovery of new phenomena in the next decade. We have already mentioned some of these, but a particularly important one is how to treat turbulence and instabilities in real systems with many degrees of freedom when we know that the dynamics of systems with only a few degrees of freedom is already incredibly complicated.

The effect of sequences of instabilities on heat, mass, and momentum transport in fluids have not been adequately studied and could have practical applications, *e.g.*, in the understanding of lubrication. There is as yet no really fundamental understanding of the mechanisms that control solidification patterns in, for example, snowflakes, quenched alloys, or directionally solidified eutectic mixtures.

An important problem for the future is the computer simulation of the onset and growth of turbulence in hydrodynamic systems. Fundamental questions that remain to be answered include: Is chaos a meaningful concept in quantum mechanics? How does one characterise and classify the steady states that are obtained under constant external conditions away from thermodynamic equilibrium? While our ability to calculate electronic structures in relatively simple situations with relatively weak interelectronic interactions, such as semimetals and semiconductors, has increased enormously, we still have great difficulties with a variety of types of strongly interacting systems. Theories of magnetic metals and the various types of metal-insulator transitions remain controversial. Even more primitive is our understanding of metals where the electron-phonon interaction is strong, such as the charge-density wave materials as well as the strong-coupled superconductors. Many as yet unexplained experimental data exist. The same difficulties probably extend to many surface electronic states as well as to the newer organic and other linear compounds. The mixed valence problem promises to be an even more difficult one to understand. The methods of pseudo potential and density functional theory work well for chemically simple

systems, but one hopes to extend electronic calculations at the same level of rigour and accuracy to chemically sophisticated systems such as organic compounds. Equally sophisticated molecular orbital methods may be within our grasp, but so far their exploitation has rested with more chemically oriented theorists.

From liquid crystals one should expect that more complex mesophases will begin to be of importance in condensed-matter theory and experiment. A start has been made with lyotropic liquid crystals and membrane phase transitions, for example, as well as with the physics of biologically interesting molecules such as proteins and nucleic acid chains. The next decade may be the decade in which the physics of random polymers becomes as interesting as more conventional disordered systems are now.

Applications of femtosecond laser spectroscopy to studies of condensed matter will increase in number and scope over the coming decade. Structural changes, such as melting and structural phase transitions, can now be studied in a time-resolved fashion on the femtosecond scale. The response of crystalline solids to short electrical pulses can now also be studied by femtosecond spectroscopy. Efforts to create even shorter pulses will continue, in parallel with efforts to extend the wavelengths of such short pulses towards the ultraviolet and infrared regions. Free-electron lasers show great promise as high-power tunable sources. The first free-electron laser operated in the near infrared, and recently lasing action has been achieved in the blue region of the visible spectrum. Two that produce both far-infrared and infrared radiation are just coming into operation. The availability of such

instruments will open the door to the experimental study of a wide variety of non-linear optical phenomena in condensed matter. These include the generation of magnetic and vibrational excitations with wave vectors at arbitrary points of the first Brillouin zone of the corresponding crystals; driving a displacive ferroelectric crystal above its nominal transition temperature to induce the ferroelectric phase transition at this higher temperature; studying the pinning of charge-density waves in one-dimensional metals; and investigating the expected large non-linear response of two-dimensional plasmons at semiconductor heterojunction interfaces.

It is expected that epithermal neutrons produced by pulsed neutron sources will be used to study excitations in condensed matter whose energies exceed thermal energies and extend up into the electron-volt range. In addition, they will include time-dependent effects such as high-frequency vibrations in solids, particularly those containing hydrogen atoms. It should also prove possible to measure in this way the momentum distribution of light atoms in their ground state. Of particular interest in this regard is the superfluid ^{4}He, for which a zeromomentum condensate fraction is presumed to exist but whose magnitude remains uncertain.

We believe that the next decade will see the increased use of insertion devices (undulators and wigglers) to increase the brightness of synchrotron radiation sources. This development will make it possible, for example, to study the properties of defects in crystals in the 10^{-5}-10^{-6} concentration range, in contrast with the 10^{-3}-10^{-4} concentration range that can be studied at present. The

higher resolution in energy and momentum expected from the use of insertion devices will also make it possible to study small samples of materials that are difficult to prepare in the form of large crystals. Inelastic x-ray scattering from the bulk and from surfaces should become a reality, as well as the ability to determine experimentally electronic structures of solids, particularly of systems with small Brillouin zones, such as artificially structured materials and semiconductors with reconstructed surfaces.

Line Defects

Line imperfections, or dislocations, in crystalline solids are defects that produce lattice distortions centered about a line. A dislocation is simply the edge of an extra inserted fractional plane of atoms. Normally the symbol ⊥ is used to represent a positive dislocation (extra fractional plane) is used to represent a negative dislocation (missing fractional plane).

The importance of dislocations is readily demonstrated in the deformation of crystalline materials. The plane in which a dislocation moves through the lattice is called a slip plane. With an applied shear stress the dislocation moves, atomic row by atomic row, and one part of the crystal is displaced relative to the other.

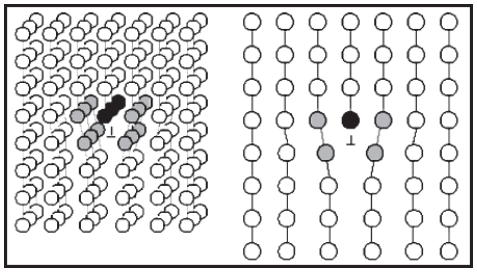


Fig. Schematic Presentation of a Dislocation; The Last Row of Atoms (Dark) in the Inserted Fractional Plane.

When the dislocation has passed through the crystal, the portion of the crystal above the slip plane has shifted one atomic distance relative to the portion below the slip plane. In other words, the motion of the dislocation has caused the crystal to change its shape - to be permanently deformed.

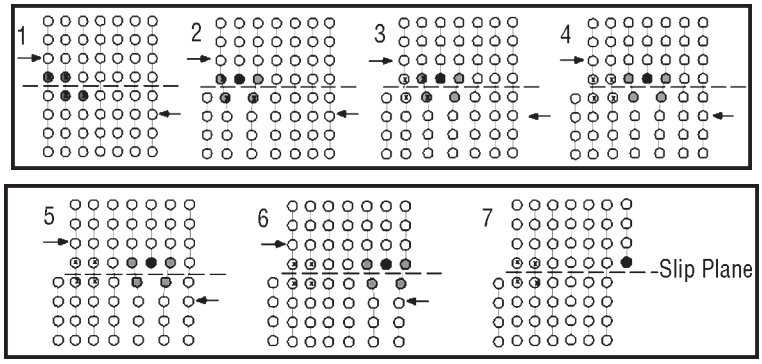


Fig. Plastic Deformation of Crystalline Solid by Slip Associated with Stress Induced Motion of Dislocation.

Please note: on either side of the dislocation the crystal lattice is essentially perfect, but in the immediate vicinity of the dislocation the lattice is severely distorted.

For a positive edge dislocation, the presence of the extra half plane causes the atoms above the slip plane to be put in compression, while those below the slip plane are put in tension.

Consequently, the edge dislocation will have a stress field around it that is compressive above the slip plane and tensile below the slip plane.

Plastic Deformation By Slip

When single crystals of metal (or semiconductor) are pulled in tension, they will begin to deform (elongate) plastically at relatively low stress levels, and “blocks” of the crystals slide over one another because of dislocation motion.

Simultaneously, so-called slip lines appear on their surface.

It is found that deformation by slip occurs most easily on planes with high atomic density and with large interplanar spacing, while the direction of slip is in all instances an atomically "close-packed direction".

For FCC structures we therefore observe as the primary slip system {111} planes in <110> direction, while in BCC structures the primary slip occurs on {110} planes in <111> directions.

(It should be noted that an alternate deformation mechanism is "deformation twinning", presently not to be considered.)

Dislocation Climb

Climb is the name given to the motion of dislocations when the extra "half" plane is extended farther into a crystal or partially withdrawn from it. Clearly, the climb process is not a motion of the plane, but rather its growth or shrinking as a result of the addition of atoms or "vacancies" respectively from the environment of the dislocation.

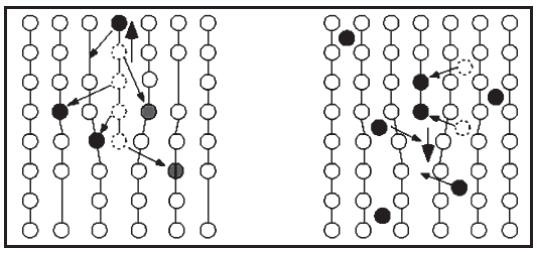


Fig. Dislocation Climb by (a) Loss of Atoms to Surrounding Vacancies and (b) Incorporation of Interstitial Atoms.

Multiplication of Dislocations

Since during slip each dislocation leaves the matrix, macroscopic deformation could not take place given normal

dislocation densities in the range of 10^6-10^8/cm^3. Examination of the deformed crystals indicates that multiplication of dislocations takes place during deformation.

Dislocation Interactions

The relative ease with which dislocations move across a solid matrix can be attributed to the severe displacements of atoms in the core of dislocations. If these local stresses are reduced, the mobility of dislocations - and thus the ease of slip - is reduced. It is found that impurities in the vicinity of dislocation cores tend to reduce the local distortion energy of the dislocations and thus stabilize the system against slip.

In many systems impurities are intentionally added (e.g., solid solution hardening) to increase the strength of materials. Similarly, micro-precipitates tend to impede dislocation motion (e.g., precipitation hardening).

Interfacial Imperfections

The several different types of interfacial, or planar imperfections, in solids can be grouped into the following categories:

- Interfaces between solids and gases, which are called free surfaces;
- Interfaces between regions where there is a change in the electronic structure, but no change in the periodicity of atom arrangement, known as domain boundaries;
- Interfaces between two crystals or grains of the same phase where there is an orientation difference in the atom arrangement across the interface; these interfaces are called grain boundaries;

- Interfaces between different phases, called phase boundaries, where there is generally a change of chemical composition and atom arrangement across the interface.

Grain boundaries are peculiar to crystalline solids, while free surfaces, domain boundaries and phase boundaries are found in both crystalline and amorphous solids.

Free Surfaces

Because of their finite size, all solid materials have free surfaces. The arrangement of atoms at a free surface differs slightly from the interior structure because the surface atoms do not have neighboring atoms on one side. Usually the atoms near the surface have the same crystal structure but a slightly larger lattice parameter than the interior atoms.

Perhaps the most important aspect of free surfaces is the surface energy associated with surfaces of any solid. The source of this surface energy may be seen by considering the surroundings of atoms on the surface and in the interior of a solid. To bring an atom from the interior to the surface, we must either break or distort some bonds - thereby increasing the energy.

The surface energy is defined as the increase in energy per unit area of new surface formed. In crystalline solids, the surface energy depends on the crystallographic orientation of the surface - those surfaces that are planes of densest atomic packing are also the planes of lowest surface energy.

This is because atoms on these surfaces have fewer of their bonds broken or, equivalently, have a larger number of nearest neighbors within the plane of the surface. Typical

values of surface energies of solids range from about 10^{-1} to 1 J/m^2. Generally, the stronger the bonding in the crystal, the higher the surface energy.

Surface energies can be reduced by the adsorption of foreign atoms or molecules from the surrounding atmosphere. For example, in mica the surface energy of freshly cleaved material in a vacuum is much higher than the surface energy of the same surface cleaved in air. In this instance, oxygen is adsorbed from the air to partially satisfy the broken bonds at the surface.

Impurity atom adsorption makes it almost impossible to maintain atomically clean surfaces. As a result, surface properties such as electron emission, rates of evaporation and rates of chemical reactions are extremely dependent on the presence of any adsorbed impurities. These properties will be different if the measurements are made under conditions giving different surface adsorption.

Grain Boundaries

Grain boundaries separate regions of different crystallographic orientation. The simplest form of a grain boundary is an interface composed of a parallel array of edge dislocations.

This particular type of boundary is called a tilt boundary because the misorientation is in the form of a simple tilt about an axis, parallel to the dislocations. Tilt boundaries are referred to as low-angle boundaries because the angle of misorientation is generally less than 10°. When a grain boundary has a misorientation greater than 10 or 15, it is no longer practical to think of the boundary as being made up of dislocations because the spacing of the dislocations

would be so small that they would lose their individual identity. The grain boundary represents a region a few atomic diameters wide where there is a transition in atomic periodicity between adjacent crystals or grains.

Grain boundaries have an interfacial energy because of the disruption in atomic periodicity in the vicinity of the boundary and the broken bonds that exist across the interface. The interfacial energy of grain boundaries is generally less than that of a free surface because the atoms in a grain boundary are surrounded on all sides by other atoms and have only a few broken or distorted bonds.

Solids with grain boundaries are referred to as polycrystalline, since the structure is composed of many crystals - each with a different crystallographic orientation. In the case of iron the grain boundary structure can be revealed by preferential chemical attack (etching) at the grain boundaries, while the grain structure in polyethylene is revealed by the use of polarized light. The grain structure is usually specified by giving average grain diameter or by using a scheme developed by the American Society for Testing and Materials (ASTM). In the ASTM procedure the grain size is specified by a "grain size number" (n) where

$$N = 2^{n-1}$$

with N equal to the number of grains per square inch when the sample is viewed at 100X magnification. For example, at a magnification of X = 100, a material with grain size number 8 will show 128 grains per inch2 - this material in effect has (at X = 1) 1.28×10^6 grains per square inch. If the grains are approximately square in cross section, this corresponds to an average grain dimension of 8.8×10^{-4} in.

In polycrystalline samples the individual grains usually have a random crystallographic orientation with respect to one another, and the grain structure is referred to as randomly oriented.

In some instances, however, the grains all have the same orientation to within a few degrees. In this instance the material is said to have a preferred orientation or texture.

Phase Boundaries

A phase is defined as a homogeneous, physically distinct and mechanically separable portion of the material with a given chemical composition and structure. Phases may be substitutional or interstitial solid solutions, ordered alloys or compounds, amorphous substances or even pure elements; a crystalline phase in the solid state may be either polycrystalline or exist as a single crystal.

Solids composed of more than one element may - and often do - consist of a number of phases. For example, a dentist's drill, something painfully familiar to all of us, consists of a mixture of small single crystals of tungsten carbide surrounded by a matrix of cobalt.

Here the cobalt forms a continuous phase. Polyphase materials such as the dentist's drill are generally referred to as composite materials. Composite materials have great importance in the engineering world because they have many attractive properties that set them apart from single-phase materials.

For example, the dentist's drill has good abrasive characteristics (due to the hard carbide particles) and good toughness and impact resistance (due to the continuous cobalt matrix). Neither the tungsten carbide nor the cobalt

has both abrasion resistance and impact resistance, yet the proper combination of the two phases yields a composite structure with the desired properties.

The nature of the interface separating various phases is very much like a grain boundary. Boundaries between two phases of different chemical composition and different crystal structure are similar to grain boundaries, while boundaries between different phases with similar crystal structures and crystallographic orientations may be analogous to low-angle grain boundariess in both energy and structure. The concept of a solid consisting of a continuous phase and a discontinuous phase (or phases) leads to a simple classification of the various types of composite materials. Table gives this classification, which is based on the structure (whether amorphous or crystalline) of the continuous and discontinuous phases.

5

Liquid Chromatography Stationary Phases

Traditionally the stationary phase used in LC has been silica gel which separates solutes largely on the basis of polarity, although, due to its unique structure, silica gel also exhibits strong exclusion characteristics. The bonded phases were introduced to provide a material that would separate solutes by dispersive interactions and also to provide some semie polar stationary phases. The bonded phases were also based on silica gel. More recently, polymeric stationary phases were introduced to provide materials that were insoluble in water and that were stable at extremes of pH.

Silica Gel

The Preparation of Irregular Silica Gel

Silica gel is manufactured by releasing silicic acid from a strong solution of sodium silicate by hydrochloric acid.

(Sodium silicate is prepared by heating sand at a high temperature in contact with caustic soda or sodium carbonate).

Initially, silicic acid is released,

$$Na_2SiO_3 + H_2O + 2HCl = Si(OH)_4 + 2NaCl$$

and then the free acid quickly starts to condense with itself with the elimination of water to form dimers, trimers and eventually polymeric silicic acid. The polymer grows, initially forming polymer aggregates and then polymer spheres, a few Angstrom in diameter. These polymeric spheres are called primary silica particles. These primary particles continue to grow until, at a particular size, the surface silanol groups on adjacent primary polymer particles, condense with the elimination of water. This condensation causes the primary particles to adhere to one another and at this stage the solution begins to gel. During this process, the primary particles of silica gel will have diameters ranging from a few Angstrom to many thousands of Angstrom depending on the conditions of formation.

The size of the primary particles depend, among other factors, on the conditions of synthesis, (e.g., reaction temperature, the pH of the mixture at the time of gelling and even the subsequent treatment of the gel, including the conditions of washing). The product at this stage is called a 'hydrogel'.

If a neutralized silicate solution is allowed to age with gentle stirring to prevent gel formation, Ostwald Ripening occurs, causing the larger primary particles to grow in size at the expense of the formation of smaller particles. In general, the aging and ripening of silica solutions will increase the size of

the larger particles which, in turn, will decrease the surface area of the silica and increase its porosity.

It is apparent that the formation of the primary particles of condensed silica, and their subsequent fusion by gelling, that confers on silica gel its high porosity and high surface area, which are so important in its use as a stationary phase in LC. Furthermore, as has already been stated, it is the condensation of the surface silanol groups of the primary particles, that causes their adhesion and the onset of gel formation and, consequently, the mechanical strength of the gel.

After the gel has first set, it is a very soft and is usually transferred to vats or trays where it is then allowed to stand for a number of days. During this period, condensation between the primary particles continues to take place and the gel shrinks and exudes salinated water. This shrinking process accompanied by saline elimination is called sinerisis and eventually a firm, almost rigid gel is produced which is called the hydrogel.

The compact hydrogel is then well washed under controlled conditions to eliminate the last of the sodium chloride and then heated for a few hours at 120°C. The resulting product is a hard amorphous mass called the xerogel. The xerogel, ground and graded is the material that is used for packing LC columns and manufacturing bonded phases. The product, prepared in this way, is called irregular silica gel, to differentiate it from spherical silica gel which is prepared by employing an entirely different synthetic procedure. Irregular silica gel is also the basic material from which bonded phases can be prepared.

The Preparation of Spherical Silica Gel

Spherical particles of silica can be prepared by spraying a neutralized silicate solution (the colloidal silica sol.) into fine droplets before gelling has taken place and subsequently drying the droplets in a stream of hot air. It has also been shown possible to disperse a silica sol in the form of an emulsion in a suitable organic solvent where the droplets gel in spherical form. Unfortunately, details for the preparation of spherical silica have tended to be kept very confidential for commercial reasons and so information is a little sparse.

One of the first methods reported was that of Le Page et al. A stable silica sol. (generated at low pH so that gelling does not take place) is passed through a non aqueous solvent in such a manner as to produce droplets. These droplets rapidly solidify and are then filtered off, dried, and heated to 400°C to 800°C to form a rigid xerogel. The structure of the resulting silica is strongly affected by the alkali metal content and the calcining temperature. The higher the temperature employed the lower the surface area and pore volume of the resulting silica. An alternative method devised and patented by Unger called the polyethoxy silane procedure involved a two stage process.

Firstly tetraethoxy-silane is partially hydrolyzed to polyethoxysiloxane, a viscous liquid which is then emulsified in an ethanol water mixture by vigorous stirring. The stirring produces spheres of polyethoxysiloxane which by hydrolytic condensation, initiated by a catalyst, are changed to silica hydrogel.

The hydrogel spheres are then washed and converted to the xerogel by heating. Spherical silica gel is readily available commercially, both in the form of silica and as different types of bonded phase. Today the majority of silica-based LC column packings are spherical although with modern methods of grinding, the so called irregular particles are more rounded and the difference between the performance of spherical and irregular particles is less clearly defined.

The Structure of Silica Gel

The matrix of the primary silica gel particle consists of a core of silicon atoms joined together with oxygen atoms by siloxane bonds (silicon-oxygen-silicon bonds). On the surface of each primary particle some residual, uncondensed hydroxyl groups from the original polymeric silicic acid remain. These residual hydroxyl groups confer upon silica gel its polar properties. These hydroxyl groups react with the silane reagents to form bonded phases. The silica surface is quite complex and contains more than one type of hydroxyl group, strongly bound or 'chemically' adsorbed water and loosely bound or 'physically adsorbed' water. There are three types of hydroxyl group.

The first is a single hydroxyl group attached to a silicon atom which has three siloxane bonds joining it to the gel matrix. The second is one of two hydroxyl groups attached to the same silicon atom which, in turn, is joined to the matrix by only two siloxane bonds. These twin hydroxyl groups are called Geminal hydroxyl groups. The third is one of three hydroxyl groups attached to a silicon atom which is now only joined to the silica matrix by only a single siloxane bond. An example of each type of hydroxyl bond is shown in Figure.

Employing NMR techniques Sindorf and Maciel, has shown that the single hydroxyl group is likely to be the most prolific. The next most common is the geminal hydroxyl groups followed by the tertiary hydroxyl group.

The silica surface, however, has additional complexities. Water can be hydrogen bonded to the hydroxyl groups and multi-layers of water physically adsorbed on top of these. Water can be hydrogen bonded to the silica gel surface in a number of different ways which are depicted in figure.

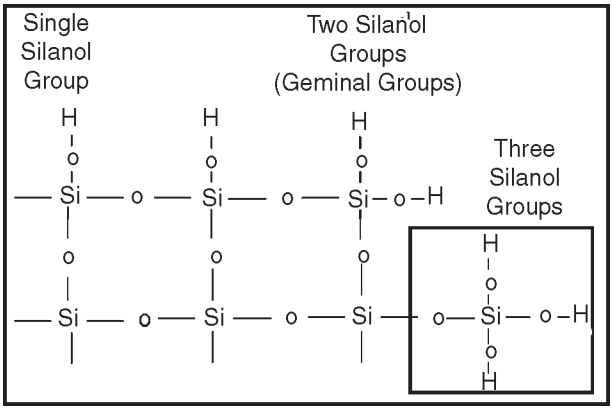


Fig. Different Forms of Hydroxyl Group that can Occur on the Surface of Silica Gel.

None of the above structures has been confirmed in an unambiguous manner but all are reasonably possible. The centre and right hand side structures contain a type of double hydrogen bond and would have high energies of formation and, thus, more stable than the simple hydrogen bond depicted on the left. The right hand structure would be particularly stable as it constitutes a four membered hydrogen bonded ring similar to that which might be expected to form in the strong association of water with itself.

Silica gel adsorbs relatively large quantities of water which was explained on the basis of multi-layer adsorption. This

concept was supported by Vleeskens and experimentally validated by gravimetric measurements. An example of one type of multi-layer adsorption is shown in figure.

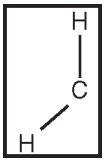


Fig. Different Ways in Which Water May be Hydrogen Bonded to Silica Gel Hydroxyl Groups

The multi-layer adsorption depicted in figure is much over simplified, as adsorption could also take place the surface of siloxane bonds as well.

Probably the most informative experimental procedure to help in the elucidation of the structure of the silica surface is thermogravimetric analysis (TGA).

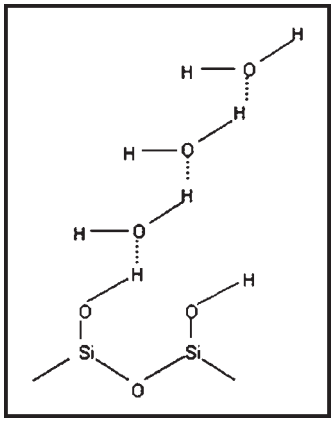


Fig. Multi-Layers of Physically Adsorbed Water

Formally, the silica was heated for known times at known temperatures and the loss in weight noted. With the advent of the programmed TGA apparatus this procedure has been simplified a great deal and the basic thermogram can now provide considerable information on the nature of the surface hydroxyl groups and adsorbed water.

The Thermogravimetric Analysis of Silica Gel

The sample is suspended from the arm of a continuously recording micro balance in a temperature controlled furnace and is heated from a defined starting temperature to a specified final temperature at a designated heating rate usually given as temperature change per unit time. The temperature and the sample weight are continuously digitized and the data stored.

The results can then be printed out or presented as an appropriate graph relating sample mass to temperature. To help identify the desorption of different species, derivative curves can also be produced. The results obtained from a sample of Matrex 20 LC silica gel taken directly from the Perkin Elmer TGA instrument is shown in figure.

The derivative curve is seen to have three distinctly different desorption processes. The first takes place from about 30°C to 130°C; the second between about 200°C and 450°C and the third between about 400°C and 900°C. The three different desorption processes are distinct and unambiguous and are similar to those previously identified by Scott and Traiman (22). The total loss from the sample was about 5%w/w but it would appear from the TGA curve that the desorption may not be entirely complete at the temperature of 900°C. The curve relating mass of water lost (obtained by subtracting each data point from the initial total mass of sample) against temperature in a TGA analysis, was a type of desorption isotherm that could be described by assuming three distinct and separate desorbable species on the silica surface, each evolving water over a specific temperature range.

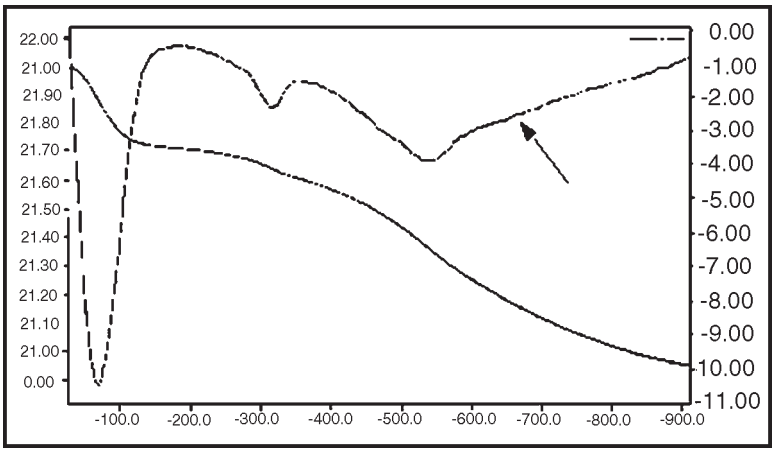


Fig. Thermogram of Silica Gel

The first loss of water between 50°C and 150°C can be attributed to physically adsorbed water. The second source that is lost between 150°C and 400°C appears to be strongly held (probably hydrogen bonded) water.

The third and last loss appears to come from the condensation of the silanol groups to siloxane bonds with the elimination of water. The amount of water lost during the three desorption processes shown as desorption curves in figure.

Bonded Phases

Bonded phases are formed by reacting the surface hydroxyl groups with an appropriate reagent to chemical link an organic moiety to the silica surface.

The nature of the organic moiety will determine the type of interaction that will take place between the solute and the surface. If the moiety is a hydrocarbon chain, then the interaction will be dispersive, if, for example, it contains a cyano group, it will be polar and if it carries a group that can dissociate into ions, the interactions will be ionic.

The Synthesis of Bonded Phases

The particle size will be determined by the nature of the separation and will vary with the complexity of the mixture and the instrument characteristics. For example a difficult separation will require high efficiencies and thus a small particle diameter (e.g. 3 mM). If however, the available column pressure is limited then the particle diameter made need to be relatively large (e.g. 10 mM) to provide adequate flow rate.

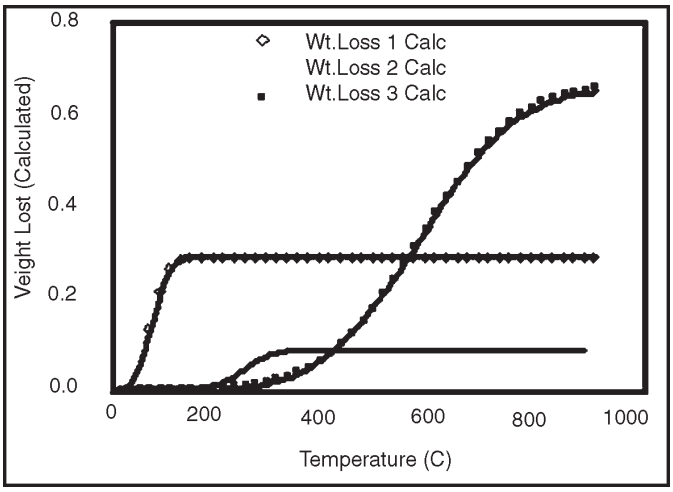


Fig. Graph of Calculated Weight Loss/Temperature for Each Desorption Stage

The choice of the pore size and surface area of the silica is more complex. The surface area of a silica tends to vary inversely as the pore size, so the larger the pore size the smaller the surface area. Thus, when one is defined the other, to some extent, is also specified.

The most efficient bonded phase has the maximum surface coverage. It is understood, that due to stearic hindrance from the bonded moiety itself, only a proportion of the silanol groups can be bonded and there is little that can be done to avoid this problem. However, there are other reasons for

incomplete silanization of the silica. Incomplete silanization can result from the reagent molecule being excluded from the smaller pores of the silica.

Exclusion can be a particular problem when bonding relatively large molecular weight materials such as long chain hydrocarbons onto the silica surface. It is therefore, important to choose a silica gel that has a relatively large pore size (e.g., a mean pore diameter of 150Å) which may limit the surface area to between 150 and 250 sq.m per gram and thus, reduce the retentive capacity of the stationary phase. Nevertheless, this will ensure that the vast majority of the pores will be accessible to the silanizing reagent.

Care must be taken to ensure that the pore size is not chosen to be so large that the silica gel is mechanically weak and collapses under pressure during packing or even during normal use. An alternative approach is to subject the silica to hydrothermal treatment.

Hydrothermal treatment removes silica from the surface of large pores and deposits it on the smaller pores, which eventually become blocked and thus, impermeable. This would remove the small pores, and consequently possible sites for future erosion and make the large pores even larger and more accessible to the silanizing reagents.

Bonded Phase Synthesis by Reaction in a Solvent

The solvents normally used in bonded phase synthesis are aromatic hydrocarbons e.g., toluene that boils at 110°C or mixed xylenes that boil 138-140°C. The procedure varies a little depending on the size of the batch and the type of

silanizing reagent. An example of a laboratory scale synthesis using a chlorosilane is as follows.

10 g of the chosen silica is dried at 250°C for about 2 hours and dispersed in a flask containing 100 ml of toluene dried over sodium. If a mono-chlorosilane reagent is used a small trace of water in the toluene can be tolerated and can be eventually eliminated by the use of excess chlorosilane reagent. Under such circumstances the toluene need not be dried over sodium. If, a dichlorosilane is used however, for example in the initial step in the synthesis of an oligomeric phase, the presence of water may cause linear polymerization.

Consequently, stringent precautions must be taken to eliminate all traces of water. A slight excess of the chlorosilane is then added to the silica dispersion together with 5 ml of pyridine. The pyridine acts as scavenger for the hydrochloric acid released during the reaction.

The mixture is refluxed for about 5 hours and the product is then filtered on a sintered glass filter, washed sequentially with toluene, tetrahydrofuran (THF), methanol, methanol water (50:50 v/v) and finally with methanol and dried under suction.

The bonded phase now needs end-capping; that is, any unreacted silanol groups are treated with a small molecular weight silanizing reagent to react with those hydroxyl groups that were stearically unavailable to the larger reagent due to exclusion.

To end-cap the product, the bonded phase is refluxed for two hours in a mixture of 100 ml of toluene and 25 ml of hexamethyldisilazane. The product is again filtered free of the reaction liquid mixture and washed sequentially with,

toluene tetrahydrofuran, methanol, methanol water (50:50 v/v) and finally with methanol and then dried under vacuum.

It should be pointed out that end-capping cannot eliminate the hydroxyl groups that are stearically hindered by the bonded moiety, or at best, only a small proportion of them will be removed. It can, however, react with any readily available hydroxyl groups, particularly those contained in pores that the original reagent could not enter but to which the smaller capping reagent has access.

Capping can also eliminate any hydroxyl groups attached to the bonded moiety resulting from the presence of dichloro or tricoloro impurities in the silanizing reagent.

After capping, the carbon content of the product can then be determined to estimate the extent of reaction.

The carbon content of a bonded phase is often used to determine the efficacy of bonding, but its value must be used in conjunction with a knowledge of the surface area of the native silica in order to arrive at a meaningful conclusion.

The amount of material bonded to the silica will depend, not only on the efficiency of the reaction, but also on the number of silanol groups that were available with which it could react; ipso facto it will also depend on the surface area of the parent silica.

The carbon content of the bonded phase is usually determined by micro-analysis and the result expressed as %w/w of the combined bonded organic material and the silica gel.

Consider a bonded phase where the carbon content is (y)%w/w coated with a hydrocarbon moiety having (n) carbon atoms per aliphatic chain (e.g., for the dimethyl octyl brush

phase, n=10). Then, the concentration of aliphatic chains in mols. per gram of silica (m')will be,

$$m' = \frac{y}{100}\frac{1}{12n}$$

Consequently, if the surface area of the silica is (A) m^2g^{-1}, the number of mols of aliphatic chains (m) in micromols per square meter will be given by,

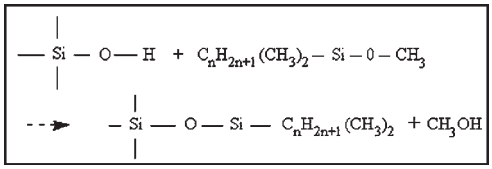
$$m = \frac{y}{100}\frac{1}{12n}\frac{10^6}{A}$$

or,

$$m = \frac{833y}{nA} \quad (1)$$

The amount of material bonded per square meter, which indicates the efficacy of the synthetic process, will be directly proportional to the carbon content of the product and inversely proportional to the surface area of the original silica gel.

The method of synthesis is very similar for the alkoxysilane reagents. The same solvents can be used but, as no hydrochloric acid is generated, there is no need to have pyridine present as a scavenger. The most reactive alkoxy reagents are the methoxy and ethoxysilanes and their reaction with a hydroxyl group is accompanied by the release of methanol or ethanol.



The reaction is best carried out in a distillation flask and the methanol removed as it is formed, the heating rate is adjusted to ensure that the aromatic solvent is not removed as well. The reaction is allowed to proceed for about 5 hours

and the product then filtered through a sintered glass filter and washed with the same sequence of solvents as those used in the chlorosilane synthesis.

The product is then refluxed with the THF/water mixture, filtered and again washed with the appropriate solvents and dried. The final capping process is also the same as that employed in the method using the chlorosilanes reagents, utilizing hexamethyldisilazane as the capping reagent.

The alkoxy-silanes are almost as readily available as the chlorosilanes and are easier and more pleasant to handle. They are, however, just as hygroscopic as the chlorosilanes and must be kept under the same anhydrous conditions. Approximately the same yields can be expected and the same degree of bonding.

The Fluidized Bed Method for Bonded Phase Synthesis

The concept of synthesizing bonded phases by means of a fluidized bed reactor was first suggested by Unger and was later implemented by Simpson and Khong.

The term fluidization describes a contact process in which particulate matter is transformed into a fluid-like state by a stream of gas or liquid. This state is achieved by the upward passage of a fluid through the bed of particles to a velocity at which the drag force acting on the surface of the particles is equal to the gravitational force downwards. At this point the particles move apart and become suspended by the fluid flow and the bed is said to be fluidized.

The fluidized bed apparatus can be operated in such a way that it allowed far more complex multi-stage syntheses to be carried out rapidly and with relative ease; e.g., the

synthesis of oligomeric phases. The fluidized bed apparatus is shown in figure. The fluidized bed is enclosed in a tube about 25 cm long, 4.5 cm in diameter situated in a heating jacket. The temperature of the bed is monitored by three thermocouples placed in the centre and at either end of the bed. A condenser is situated at the top of the bed which returns unreacted silanizing reagent to the vapour generator. The vapour generator consists of a simple boiling flask that can be provided with a nitrogen stream if the silanizing reagent does not boil at reasonable temperatures. The silanizing reagent vapour passes from the vaporizer to a preheater and then to the base of the fluidizer.

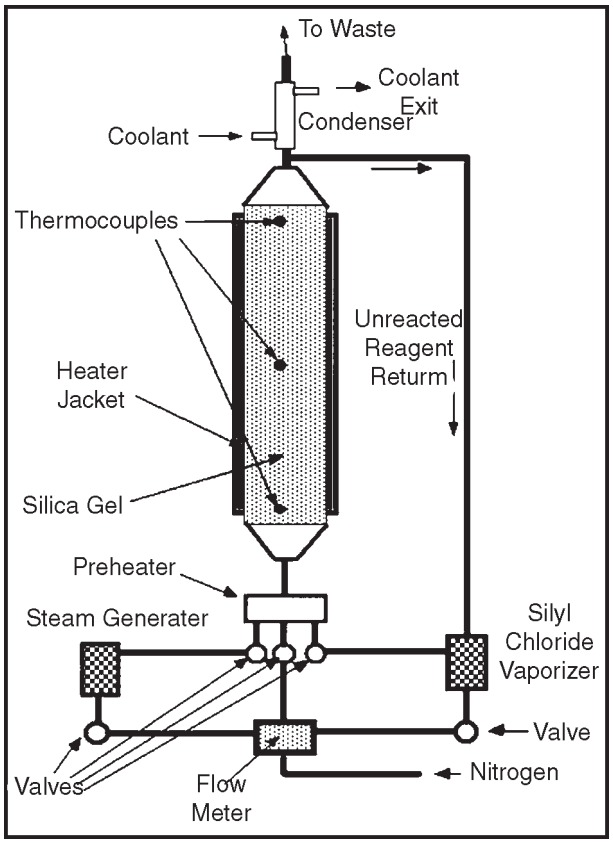


Fig. The Fluidized Bed Apparatus for the Synthesis of Bonded Phases

A steam generator, similar in form to the silanizing reagent vaporizer, is also attached to the preheater and consequently, can supply steam to the fluidized bed if, and when required. This means that the fluidized bed can be used to hydrothermally treat the silica prior to silanization. The nitrogen stream is continuously monitored by a flow-meter.

All vapour and nitrogen streams are adequately fitted with valves to provide the maximum operating flexibility. The fluidizer is normally loaded with about 25 g of silica (which provides a fluidized bed about 5 cm long) and heated to 200°C for about 3 hr. in a fluidizing stream of preheated nitrogen. The silanizing reagent is then heated to its boiling point in the vaporizer or to a temperature where it has a significant vapour pressure and the reagent blown through the fluidized bed by a stream of nitrogen. Excess, unreacted reagent is condensed and passed directly back to the vaporizer.

The reaction is allowed to proceed for about 6 hours and then the reagent vapour flow stopped and replaced by a stream of pure nitrogen. When reagent is no longer returning to the vaporizer, the bed is allowed to cool to room temperature with the nitrogen still flowing. When cool, the nitrogen flow is stopped, and the bonded phase can be removed from the fluidizer. Alternatively, the silanizing reagent can be replaced with a capping reagent and the material capped by a similar procedure and the product then removed.

The fluidized bed synthesis gives a more reproducible product and, at the same time, eliminates many of the tedious operations that are involved in the alternative method, such

as solvent removal and recovery, washing procedures and other manipulations. It also offers the possibility of carrying out very complicated syntheses involving multiple steps by a relatively simple procedure. However, the technique is more complicated, requires more elaborate apparatus and needs to be operated by an experienced technician.

Choosing a Bonded Phase

Different organic moieties can be bonded to all types of silica gel particles including the very popular spherical variety. One of the most important features to consider when choosing a bonded phase is the reproducibility of the product. The packing must obviously achieve the separation that is required, but unfortunately any bonded phase has a limited lifetime that may range from a few hours, if operated at extremes of pH, to several months if operated under mild conditions.

It follows, that all columns will eventually need replacing and if a specific analytical procedure has been established on a particular column, then the replacement must have as near identical chromatographic properties as possible. The reproducibility of the column is particularly important in forensic analysis or for analyses that are carried out to ensure adherence to regulatory standards as in environmental and pollution studies. The following tests are recommended as minimal for a new column and the results should be compared with the data obtained from the previous column as received.

- The column permeability should be measured i.e., the pressure required to produce a given flow rate e.g., a flow rate of 1 ml per minute.

- The column dead volume should be measured by determining the retention volume of an unretained solute
- The column efficiency should be measured for a set of standard solutes. If possible, the solutes should be chosen, from those likely to be present in the samples to be analyzed. Solutes eluting at (k') values of 2, 5 and 10 would be appropriate.
- The corrected retention volumes of a series of solutes spanning a (k') range of 1 to 20 should be determined and their retention ratios calculated.

All the measurements should fall within 5% of the specified values.

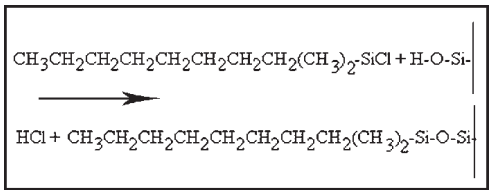
However, these criteria may not be sufficiently stringent for some forensic purposes and, consequently, the tests given above should be considered as minimum requirements for litigation purposes..

Types of Bonded Phase

There are three basic types of bonded phases, the brush phases, the bulk phases and the oligomeric phases. The different phases are produced by the use of the mono, di- and tri- substituted silanes respectively in the bonding process, e.g. the monochloro, dichloro and tricoloro silanes.

The monochlorosilanes, for instance octyldimethyl-chlorosilane, react with the hydroxyl groups on the silica surface to produce dimethyloctylsilyl chains attached to the silica.

Brush Phases



The alkyl chains are thought to stand out from the surface like bristles of a brush, hence the term brush phase. The extent to which the silanyl groups are reacted is a subject of debate at this time. It is thought that the two methyl groups next to the silicon atom hinder the reaction of adjacent hydroxyl groups with the reagent and thus there will be a considerable amount of unreacted hydroxyl groups remaining even after capping.

In the extreme, it has been suggested that there is a hydroxyl group situated between each bonded chain. There is certainly evidence of some polar interactions with reverse phases which if completely covered with hydrocarbon chains should only exhibit dispersive interactions.

However, reverse phases are predominantly dispersive in character and it would appear that if there are any hydroxyl groups still present on the surface it is likely that they would be relatively few in number compared with the bonded moieties.

Oligomeric Phases

The di-substituted silanes such as methyloctyldichlorosilanes can be used to synthesis the oligomeric bonded phases which is a far more complicated procedure. The silica is first reacted with the methyloctyldichlorosilane to link methyloctylchlorosilyl groups to the surface.

The product is then treated with water, which hydrolyses the methyloctylchlorosilyl groups to methyloctylhydroxysilyl groups with the elimination of hydrochloric acid. The "hydroxy" product is then reacted with more methyloctyldichlorosilane, attaching another methyloctylchlorosilyl group to the previous groups. This process of alternately treating the product with the silane reagent and then water can be repeated until, in the original synthesis, eight or ten oligomers are linked to each other and attached to each stearically available hydroxyl group on the surface.

However, it should be noted that the hydroxylation of the bonded moiety is not the only possible reaction that can take place with the water. There will be situations where it will be stearically possible for a water molecule to react with two adjacent chains and thus produce some cross-linking. However, a small amount of cross-linking would indeed strengthen the bonded system and could be advantageous.

The product is finally treated with trimethylchlorosilane or some other capping reagent to eliminate the last hydroxyl groups formed at the end of the oligomer. The oligomers are layered over the surface making the product extremely stable exhibiting almost no polar characteristics whatsoever. However, due to the complexity of the synthesis [which needs to be carried out in a fluidized bed for efficient reaction] oligomeric phases are expensive to manufacture and, consequently, are not often used and, at this time, are not commercially available.

Nevertheless, this type of procedure, if used to synthesize a reversed phase, would produce material that would probably be completely dispersive in character with no measurable polarity.

Bulk Phases

If the silica surface is saturated with water and octyltrichlorosilane is used as the reagent, reaction occurs with both the hydroxyls of the silica surface and the adsorbed water causing a cross-linking reaction, and an octylsilyl polymer can be built up on the surface. The same procedure can be used as that in the synthesis of oligomeric phases and the material can be alternatively treated with water and the trichlorosilane reagent. Layers of bonded phase are built up on the surface but, in this case, due to the tricoloro function of the reagent, extensive cross-linking occurs.

As a result of the polymerization process, the stationary phase has a chemically cross-linked, multi-layer character and, consequently, is termed a "bulk" phase.

The "bulk" phases are almost as popular as the "brush" phases as they tend to have a higher carbon content (more organic material bonded to the surface) and thus provide a little greater retention and selectivity. "Bulk" phases have about the same stability to aqueous solvents and pH as the "brush" phases. Using appropriate organic chlorosilanes, polar or polarizable groups such as nitriles or aromatic rings can be bonded to the silica to provide stationary phases covering a wide range of polarities. Bonded ion exchange materials have also been synthesized, although they are not as stable to salt solutions and extremes of pH as the ion exchange resins.

Interactions Between 'Brush' and 'Bulk' Reverse Phases and Aqueous Solvents

The interactions between aqueous solvents and brush reverse phases differ very significantly from those with a bulk

reverse phase at very low concentrations of solvent. This difference has been investigated by a number of workers and the basic difference between the two types of phase are shown in the curves relating retention volume of methanol to the concentration of methanol in the mobile phase in figure.

The phases shown are the RP-18 brush, reverse phase manufactured by E. M. Laboratories, which had a C_{18} (dimethyloctadecyl) chain and ODS-3 a bulk reverse phase which had a C_{18} (octadecyl) chain and was manufactured by Whatman Inc. The curves relating retention volume with solvent composition for the two phases show very different behaviour patterns.

The ODS-3 bulk reverse phase behaves in the expected manner, as the concentration of methanol increases the retention volume of the ethanol decreases smoothly and continuously up to a concentration of 10%w/v of methanol.

The brush phase, however, behaves in a very unexpected fashion. The retention volume of ethanol at first increases as the solvent concentration increases, to a maximum at a concentration of about 3%w/v of methanol.

At methanol concentrations above 3%w/v the retention values begin to fall and eventually follow a curve parallel to that for the bulk phase but somewhat higher. Lochmüller and Wilder carried out some similar experiments and obtained the same results while Gilpin and Squires carried out some thermodynamic measurements on the two systems and showed an anomalous behaviour by the brush type phase at very low concentrations of solvent.

It was suggested that the behaviour of the brush phase could be explained as a result of the free chains interacting

with themselves in preference to interacting with the aqueous solvent. In effect, this was the same phenomena of immiscibility that occurs between water and a liquid hydrocarbon.

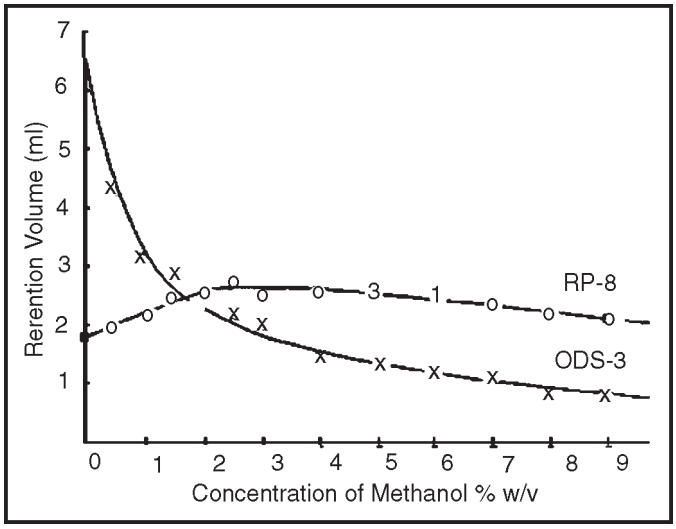


Fig. A Graph of the Corrected Retention Volume of Ethanol against the Concentration of Methanol in the Mobile Phase for a Bulk and Brush Reverse Phase

The dispersive forces between the hydrocarbon chains themselves are greater than the forces between the hydrocarbon chains and the aqueous solvent.

As a result the chains interact with themselves and collapse on the surface of the silica. When the methanol content of the mixture is increased, the solvent becomes more dispersive in nature and the hydrocarbon chains can then interact with the solvent and no longer exist in a collapsed state. The situation is depicted in figure.

The relationship between the retention of ethanol on the brush phase with water/methanol mixtures, as shown in figure, can now be explained. In pure water the hydrocarbon chains of the brush phase are collapsed on the surface and

thus, the effective surface area of the stationary phase is much reduced. Consequently, the retention volume of the solute, being proportional to the available surface area, is also reduced. As methanol is added to the solvent mixture, the solvent becomes more dispersive and the hydrocarbon chains can begin to interact with it and, as a consequence, begin to unfold.

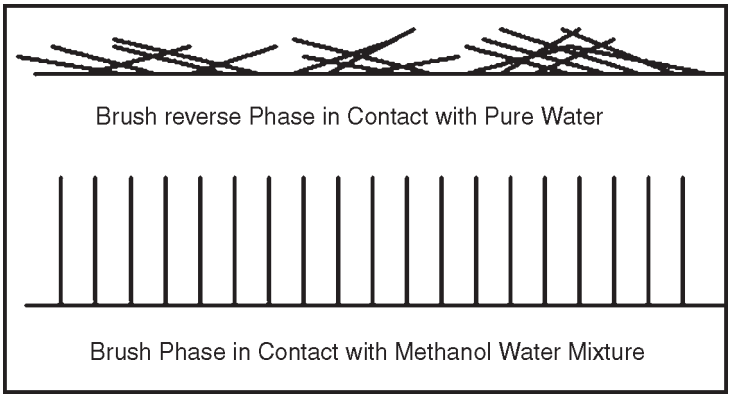


Fig. Effect of Solvent Composition on the Orientation of the Brush Phase.

The liberation of the chains from the surface results in an increase in the effective surface area of the stationary phase and the retention of the solute also starts to increase. This process continues until there is sufficient methanol in the solvent for the hydrocarbon chains to totally interact with the solvent and be completely released from the surface. At this concentration (about 3%w/v of methanol) the retention volume of the ethanol reaches a maximum. Subsequent increase in methanol concentration merely increases the interactions of the ethanol with the mobile phase and, by adsorption of the solvent onto the surface of the reverse phase, reduces the interactive forces with the reverse phase. Consequently, the retention volume steadily decreases in the expected manner.

Having explained the behaviour of the brush phase it is now interesting to consider the behaviour of the bulk phase. There appears to be no interaction between the hydrocarbon chains themselves and no change in surface area at high water concentrations.

In fact, the retention of ethanol falls steadily as the methanol content increases in the expected manner. The explanation given for this is that the cross-linking that takes place when the bulk phase is synthesized, keeps the polymeric chain system rigid and does not allow the individual hydrocarbon chains to collapse on the surface. As a consequence, the surface area is not reduced and the retention behaviour is normal.

Another aspect of the behaviour of the bulk phase at low solvent concentrations is that it does, in fact, confirm the cross-linked nature of the bulk phase.

It would also appear that for certain solutes, the best reverse phase for operation with aqueous mixtures containing very little solvent might be a bulk reverse phase. The retention mechanism on brush type phases under these conditions might be anomalous.

The Retention Properties of Bulk and Brush Phases

The equation for the corrected retention volume of a solute ($V'_{r)}$ (see Plate Theory and Extensions) is as follows,

$$V'_r = KV_S \text{ or } V'_r = KA_S$$

where (K) is the distribution coefficient of the solute between the two

phases,

(V_S) is the effective volume of stationary phase in the column,

and (A_S) is the effective surface area of the packing.

The two equations are given to illustrate that the reverse phase system may be considered as a liquid/liquid or liquid/solid distribution system where,

$$V_S = d_f A_S$$

and (d_f) is the effective film thickness of the bonded material and

any solvent that may be adsorbed on its surface

It follows that the retention of the solute will depend only on the volume or surface area of the bonded material.

Thus, providing all the bonded phase is available for solute interaction, the retention volume will be proportional to the carbon content of the phase. Scott and Kucera examined a series of commercially available reverse phases and determined the carbon content of each phase and the retention volume of a series of solutes on columns packed with each adsorbent.

The retentive properties of the five reverse phase are shown in figure where the corrected retention volume (V'_r) of 2-ethyl anthraquinone is plotted against carbon content of the reverse phase.

It is seen, somewhat surprisingly, that there is a linear relationship between retention volume and carbon content of the brush phases (R2, R8, R18).

This relationship can only be expected to occur if all the stationary phase is available to the solute and the packing procedure is very reproducible so that each column contains the same amount of packing.

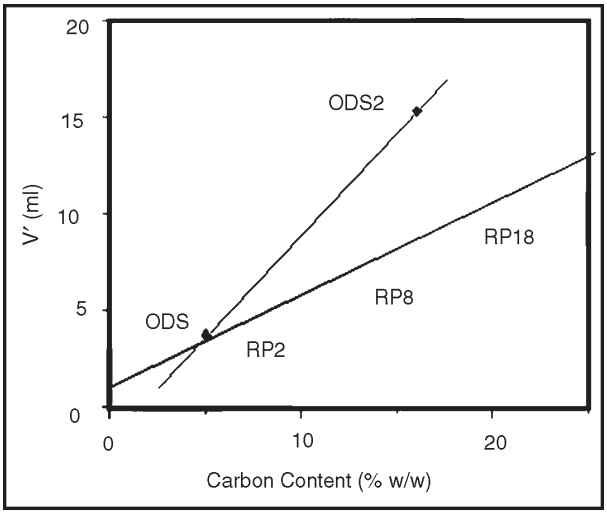


Fig. Graph of Retention Volume against Carbon Content (%w/w)

It should again be stressed that all three reverse phases were produced from base silicas of very different surface areas and, despite this, the linear relationship between carbon content and corrected retention volume remained.

This relationship may well depend, not only on the type of silica gel that is used but also on h bonding process and his relationship has not been proved generally. The relative resolution obtained from different reverse phases carrying diverse carbon contents and extreme chain lengths is shown by the chromatograms in figure. The higher retentive capacity of the bulk phase ODS_2 phase is again clearly demonstrated.

Short chain reverse phases reduce the extent to which proteins are denatured in the separation of substances of biological origin, it is seen by the chromatogram from the C2 reverse phase, that a serious price must be paid in loss of resolution if the nature of the separation demands the use of such material. However, the development of the polymer packings have, at last, partly solved this problem.

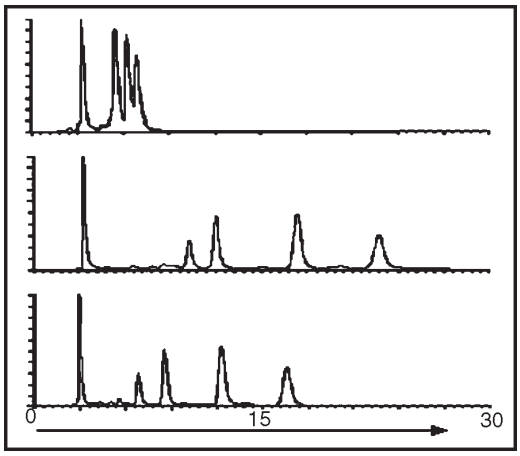


Fig. Chromatograms from Brush Phases Carrying Different Carbon Loads and Different Chain Lengths

In general, because the brush type phases can be synthesized in a more reproducible manner, particularly if carried out in a fluidized bed, the brush phases are generally recommended for the majority of applications. For high retentive capacity and for systems that will be operated with aqueous solvent mixtures having a very high water content, the bulk phases might be preferred. The partially reacted, low carbon content bulk phase also have special areas of application particularly in sample preparation.

Macroporous Polymers

Polymeric ion exchange materials were developed for chromatography in the early sixties resulting in the introduction of macro-porous polymers. The advantages of this material lay in the macro-porous nature of the resin packing, which consisted of resin particles a few microns in diameter, which, in turn, comprised of a fused mass of polymer micro-spheres a few Angstroms in diameter. The resin polymer micro-spheres play the same part as the silica gel primary particles, and confer on the polymer a relatively

high surface area together with a high porosity. The high surface area provided increased solute retention and selectivity together with a superior loading capacity and, consequently, a wide dynamic range of analysis.

The material consists of a highly cross-linked polystyrene resin with about a 50Å pore diameter. In the case of the ion exchange materials, inorganic groups of appropriate charge were chemically attached (e.g., by sulfonation). The initial resins manufactured by Rohm and Haas were called Amberlite but were largely employed in production processes and, consequently, had very large particle diameters.

Modern macroporous resin-type stationary phases have a range of particle diameters which can be as small as 2.5 or 3.0 microns. The more popular resin based packings are based on the co-polymerization of polystyrene and divinylbenzene. The degree of cross-linking determines the rigidity of the resin and the greater the cross-linking the harder the resin becomes. Ultimately, at extremely high cross-linking, the resin becomes brittle. To produce an ion exchange resin, the surface of the polystyrene-divinylbenzene copolymer is reacted with suitable reagents and covered with the required ionogenic interacting groups.

Most cross-linked polystyrene resins employed in LC are the macro-reticular type and can be produced with almost any desired pore size, ranging from 20Å to 5,000Å. Underivatized, they exhibit strong dispersive type interactions with solvents and solutes together with some induced polarizability arising from the aromatic nuclei in the polymer if strongly polar solutes are being separated. Consequently, the untreated resin has found some use as an alternative to the C8 and C18 reverse phase columns based on silica.

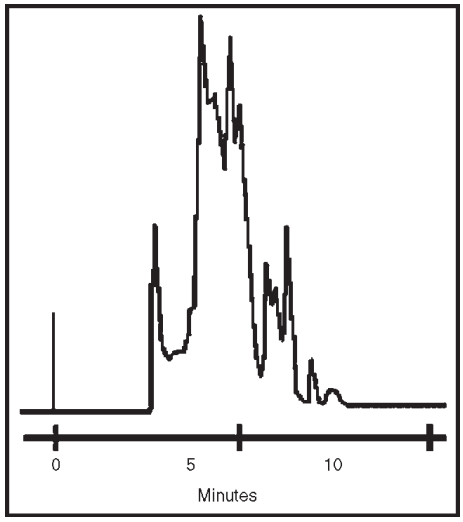


Fig. The Separation of a Crude Protein Extract by Exclusion on a Micro-Reticulated Resin Column

Due to their being stable to extremes of pH, their use for the separation of peptide and proteins at both high and low pH has been well established. An example of a macro-reticulated resin phase used as an exclusion medium in the separation of a crude protein extract is shown in figure. The column, 9TSKgel QC-PAK GFC 399GL0, was 15 cm long and 8 mm in diameter and the separation was carried out at a flow rate of 1 ml/min. The mobile phase consisted of 0.5M NaCl in 0.05M sodium phosphate buffer at pH 7.0. The sample was 5ml of a rat liver extract. An excellent separation based on molecular size is obtained from which the molecular weight range of the mixture and even that of the individual components could be estimated.

LC Mobile Phases

The choice of phase system can be very complex, particularly if multicomponent mixtures are to be separated. In the first instance the type of stationary phase needs to be chosen and this choice must be based on the interactive

character of the solutes to be separated. If the solutes are predominantly dispersive then the stationary phase must also be dispersive (a reversed phase) to promote dispersive interaction with the solutes and provide adequate retention and selectivity.

If the solutes are strongly polar then a polarizable stationary phase (one containing aromatic rings or cyano groups) would be appropriate to separate the solutes by polar and induced polar interactions. If the solutes are weakly polar then a strong polar stationary phase would be required (such as silica gel) to separate the solute by polar interactions.

The mobile phase must be chosen to complement the stationary phase so that the selected interactions are concentrated in the stationary phase. Thus, a reversed phase having strong dispersive interactions would be used with a strongly polar mobile phase (e.g., mixtures of methanol and water acetonitrile and water or tetrahydrofuran and water). In contrast, if the strongly polar silica gel is selected for the stationary phase then a strongly dispersive mobile phase would be appropriate (e.g., n-heptane, n-heptane/methylene chloride or n-heptane with a small quantity of n-propanol or ethanol). In general the mobile phase must be chosen so that the selected interactions strongly dominate in the stationary phase and are minimized in the mobile phase.

Solvent/Solute Interactions with the Silica Gel Surface

In all chromatography systems both the solvent and the solutes interact with the stationary phase. It follows that, when the silica surface is in contact with a solvent, the surface is covered with a layer of the solvent molecules.

If the mobile phase consists of a mixture of solvents the surface is partly covered by one solvent and partly with the other). Thus, any solute interacting with the stationary phase may well be presented with two, quite different types of surface with which to interact.

The probability that a solute molecule will interact with one particular type of surface will be statistically controlled by the proportion of the total surface area that is covered by that particular solvent.

Dispersive solvents appear to be adsorbed from a solvent mixture on the surface of silica gel according to the Langmuir adsorption isotherm. Examples of mono-layer adsorption isotherms obtained for benzene, chloroform and butyl chloride are shown in figure.

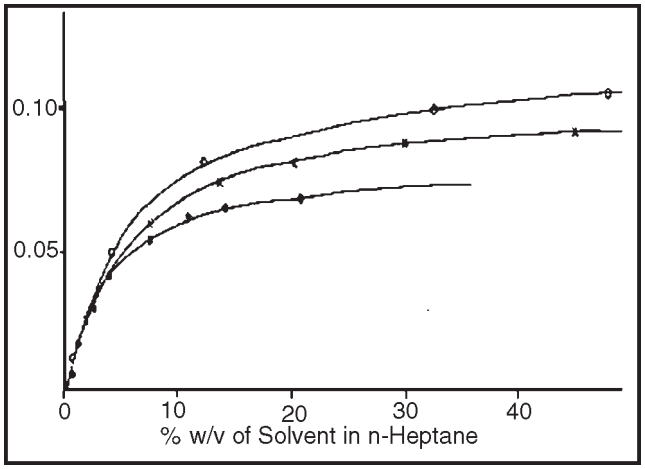


Fig. Langmuir Adsorption Isotherms for Benzene, Butyl Chloride and Chloroform

The adsorption isotherms of the more polar solvents, ethyl acetate, isopropanol and tetrahydrofuran from n-heptane solutions on to the silica gel surface did not fit the simple mono-layer adsorption equation but did fit the bi-layer adsorption isotherm which is a simple extension of the

monolayer formation process. The bi-layer adsorption isotherm for ethyl acetate on silica gel is shown in figure. The curve is theoretical and the points experimental.

The individual isotherms for the two adsorbed layers of ethyl acetate are included in figure.

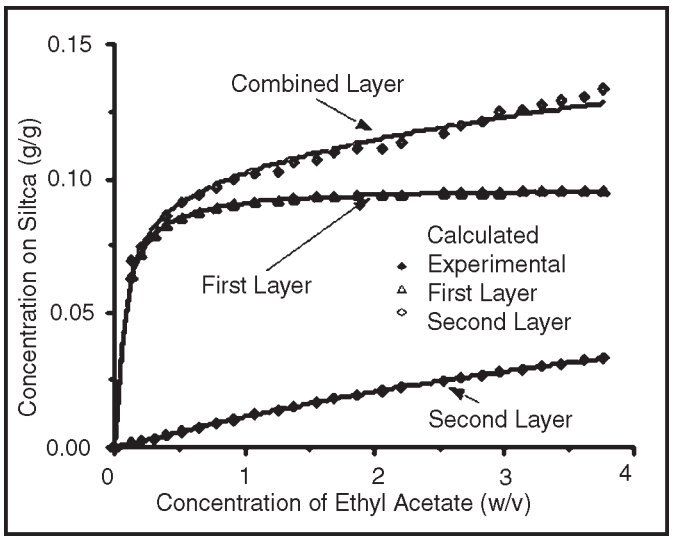


Fig. The Individual and Combined Adsorption Isotherms for Ethyl Acetate on Silica Gel

The two curves, although of the same form, are quite different in magnitude. The first layer is very strongly held to the surface and is complete when the concentration of ethyl acetate in the mobile phase is no more than 1%w/w.

As the concentration of ethyl acetate starts to rise above 1%w/w the second layer is only just being formed. The formation of the second layer of ethyl acetate is much slower and obviously the interactions between the solvent molecules with those already adsorbed on the surface are much weaker than their interaction with the silica gel silanol groups.

If it is assumed that the total area covered by the first layer of ethyl acetate will be very similar to the area covered by the second layer, then only about one third of the second

layer is complete at an ethyl acetate concentration of about 4%w/v. In contrast, the first layer is virtually complete at an ethyl acetate concentration of 1%w/v. It should be pointed out that solvent layer formation on the surface of the silica is not necessarily restricted to binary systems and multi-layers are quite feasible.

Solute Stationary Phase Interactions

There are basically two types of interaction that can take place between a solute and the silica gel surface. Firstly, the solute molecule can interact with the adsorbed solvent layer and rest on the top of it.

This type of interaction is called sorption interaction and occurs when the molecular forces between the solute and the silica are relatively weak compared with the forces between the solvent molecules and the silica.

The second type of interaction is where the solute molecules displace the solvent molecules from the surface and interact directly with the silica gel itself, for example, the silanol groups.

These two types of interaction are shown in figure. Displacement would occur if the solute was strongly polar such as an alcohol, which would interact more strongly with the polar silanol group than the dispersive chloroform layer. Sorption is depicted as a solute molecule on interacting with each solvent layer and can not interact strongly enough with the silica gel surface to displace the solvent..

Mobile phases consisting of mixtures of polar and dispersive solvents frequently produce surface bi-layers when used with silica gel as a stationary phase and therefore a far more complicated set of interactive possibilities exist. These

possibilities are depicted in figure. The surface offers the opportunity of a number of sorption and displacement processes that can take place between the solute and the stationary phase surface. There are three different surfaces on which a molecule can interact by sorption and three different surfaces from which molecules of solvent can be displaced and allow the solute molecule to penetrate. In any separation all the alternatives are possible but it is more likely that for one particular solute, one type of interaction will dominate. the various types of interaction are depicted in figure.

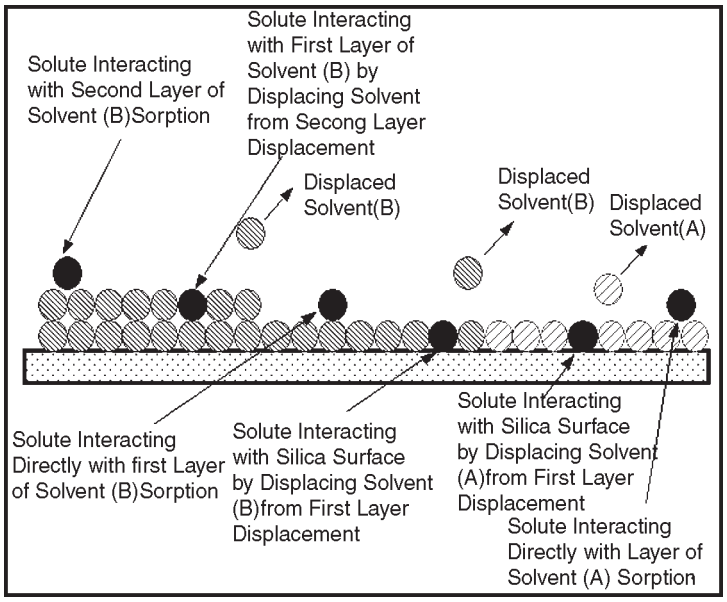


Fig. Different Types of Solute Interaction that can occur on Silica Surfaces Covered with a Solvent Bi-layer

Where there are multi-layers of solvent, the solvent that interacts directly with the silica surface is the most polar, and consequently constitutes the first layer. Depending on the concentration of the polar solvent the next layer may be a second layer of the same polar solvent as in the case of ethyl acetate.

If, however, the quantity of polar solvent is limited, then the second layer might consist of a less polar component of the solvent mixture. If a ternary mixture of solvents is used, the nature of the surface, and the solute interactions with the surface can become very complex indeed.

In general the stronger the polarity of the solute the more likely it is to interact with the surface by displacement even to the extent of displacing both layers of solvent (one of the alternative processes that is not depicted).

Solvent/Solute Interactions with the Reversed Phase Surface

Solvents interact with the surface of a reverse phase in a similar manner to the surface of silica gel. In figure, the adsorption isotherms for a series of aliphatic alcohols are shown. The effect of the carbon chain-length of the alcohol on the strength of the adsorption is clearly seen from the shape of the curves.

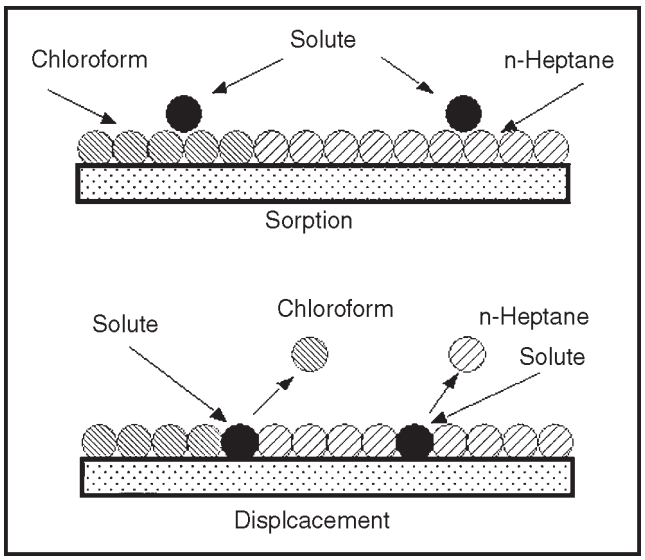


Fig. Diagram Depicting Sorption and Displacement Occurring on the Silica Surface.

The most strongly adsorbed alcohol, butanol, (the alcohol with the longest chain and, thus, the most dispersive) has only a four carbon chain and yet the surface is completely covered when the butanol concentration is only about 2%w/v.

Thus, any component of the mobile phase with a hydrocarbon chain length of four or more, will be rapidly adsorbed and modify the reverse phase surface extensively and, consequently, the magnitude of solute retention.

The curves shown in figure only cover a range of 0 to 0.05 $g.ml^{-1}$.

In order to show the shapes of the adsorption isotherms for the higher alcohols in proportion to those of the lower alcohols with reasonable clarity, the same curves are shown in figure for an alcohol concentration range of 0-100% (which is approximately 0-0.8g/ml). It should be noted that the mass adsorbed is expressed as $g.cm^{-2}$

The weak nature of the methanol adsorption, relative to the other alcohols is clear and it is seen that the surface of the reverse phase is being modified over one third of the methanol concentration range.

The reverse phase surface can be modified in a controlled manner, over the range of 0 to about 40 % methanol, but between methanol concentrations of 40% and 100 % the nature of the reverse phase surface remains sensibly constant and it is the solute interactions in the mobile phase that are progressively modified. Acetonitrile and tetrahydrofuran behave in a similar manner but there adsorption isotherms are closer in magnitude to those of ethanol than of methanol.

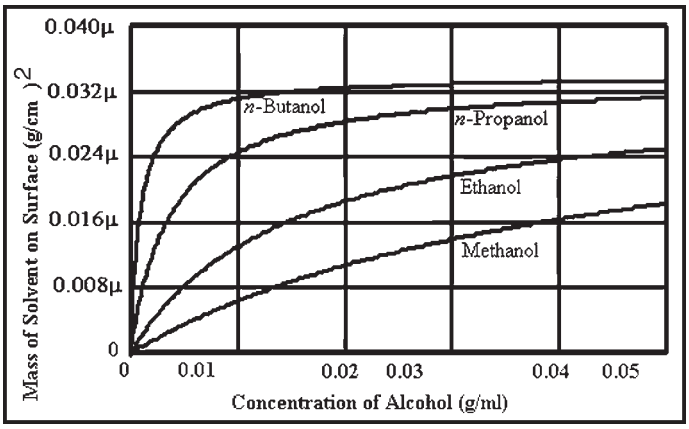


Fig. The Adsorption Isotherms of a Homologous Series of Aliphatic Alcohols over the Concentration Range of 0 to 0.0.5 $g.ml^{-1}$

The types of interactions that can take place between the solute and the reverse phase are similar to those that can take place between the solute and the silica gel surface. Solutes can interact by the sorption process, the displacement process or a combination of both. The same rules apply; if the solvent interacts more strongly with the surface than the solute then the solute interacts with the adsorbed layer of solvent by sorption.

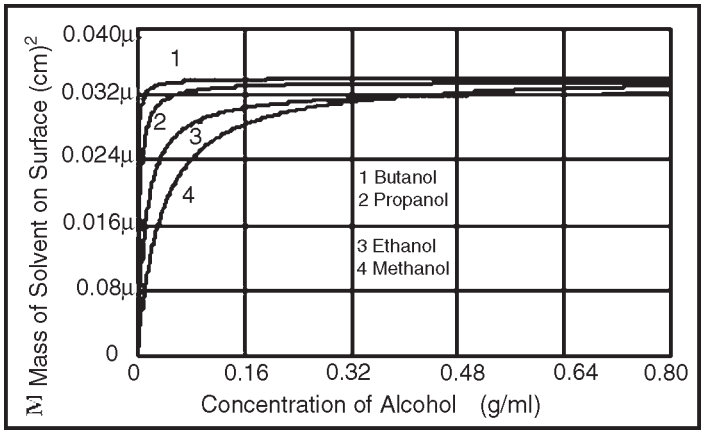


Fig. The Adsorption Isotherms of a Homologous Series of Aliphatic Alcohols

If, on the other hand, the solute interacts more strongly with the reverse phase than the layer of solvent molecules

then the solute will displace the solvent and interact directly with the surface by displacement. In, general, those solutes that elute early in the chromatogram will interact by sorption, those that elute late in the chromatogram will interact by displacement and at some intermediate point in the elution scale, solute stationary phase interactions will probably involve both sorption and displacement. Bi-layer adsorption is also possible with reverse phases but, at this time, experimental evidence of this does not appear to be available in the literature.

Molecular Interactions in the Mobile Phase

The 'elutive' capacity of the mobile phase (as opposed to the 'retentive' capacity of the stationary phase) depends on the strength of the different interactions that can take place in the mobile phase and the probability of a particular interaction occurring.

Purnel and Laub experimentally demonstrated in GC, that for a stationary phase consisting of a pair of non-associating liquids, the distribution coefficient of a solute was linearly related to the volume fraction of either liquid. This relationship indicated that the volume fraction of a solvent in a liquid mixture determined the probability of interaction.

Much the same as the partial pressure of gas determines the probability of collision. This relationship was challenged by a number of worker who also demonstrated that for certain solvent pairs this linear relationship broke down. However, it was also shown (35,36) that the nonlinear relationship occurred when there was strong association between the components of the mixture which resulted in a ternary

mixture containing the two components and the associate of the two components the concentration of which depended on the equilibrium constant and the experimental conditions.

Thus, for a mobile phase consisting of a binary mixture of solvents, as the retention volume will be inversely proportion to the elutive capacity of the mobile phase it will be also inversely proportional to the volume fraction of either component providing there is no strong association between the components.

This was experimentally demonstrated by Katz et al. who employed a liquid/liquid distribution system using water and a series of immiscible solvent mixtures as the two phases and measured the absolute distribution coefficient of a solute for different mixtures.

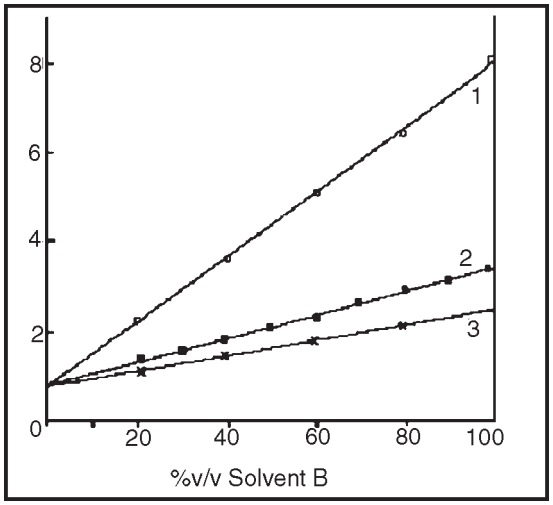


Fig. Graph of Distribution Coefficient against Solvent Composition.

The solute they used was n-pentanol and the immiscible solvent consisted of mixtures of n-heptane and chloroheptane, n-heptane and toluene and n-heptane and heptyl acetate. The two phase system was thermostatted at 25°C and, after equilibrium had been established, the

concentration of solute in the two phases was determined by GC analysis. The results they obtained are shown in figure. It is seen that linear relationships between solvent composition and distribution coefficient was obtained for all three solvent mixtures simulating the results that Purnell and Laub obtained in their gas chromatography experiments.

Aqueous Solvent Mixtures

When the relationship between the distribution coefficient of a solute and solvent composition, or the corrected retention volume and the solvent composition, was tested with aqueous solvent mixtures it was found that the relationship identified by Purnell and Laub and Katz et al failed.

It was suspected that the failure was due to the solvent strongly associating with the water and, in fact, an aqueous solution of methanol, for example, contained methanol, water and methanol associated with water.

The solvent mixture was thus, a ternary system and thus the linear relation ship between the volume fraction (before mixing) and retention would not be expected to hold.

The association of methanol and water was examined by Katz, Lochmüller and Scott using volume change on mixing and refractive index data and established that the methanol/water solvent system was indeed a complex ternary system.

They calculated both the association equilibrium constant and the distribution of the different components of a methanol water mixture form zero to 100% methanol. The curves they obtained are shown in figure.

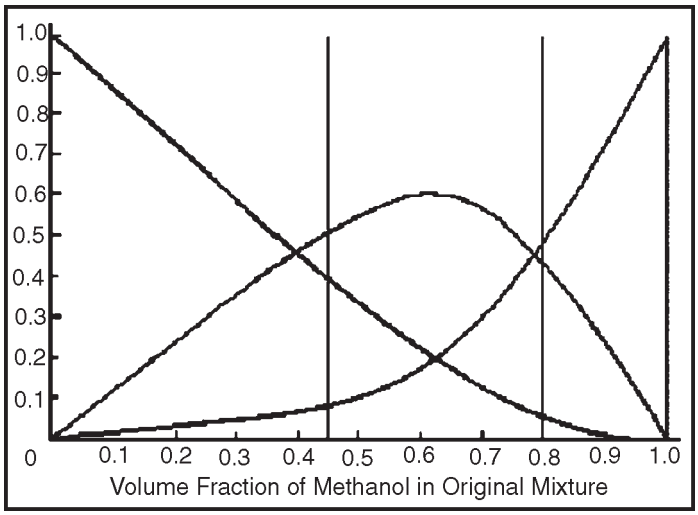


Fig. Diagram of the Ternary Solvent System for Methanol/Water Mixtures

It is seen from figure that there are three distinct ranges of methanol concentration where the solvent will behave very differently. From zero to 40%v/v of methanol in the original mixture, the solvent will largely behave as though it were a binary mixture of water and methanol associated with water. From 40%v/v to 80%v/v of methanol in the original mixture, the solvent will predominantly behave as though it were a ternary mixture of water, methanol and water associated with methanol. From 80%v/v to 100%v/v of methanol in the original mixture, the solvent will again behave as though it were again a binary mixture but this time a mixture of methanol and water associated with methanol.

The curves shown in figure explain some of the unique characteristics of mobile phases consisting of methanol water mixtures when used in reversed phase LC. From figure it is seen that when the original mixture contains 50%v/v of methanol there is little free methanol available in the mobile phase to elute the solutes as it is mostly associated with water. Subsequently, however, the amount of methanol

unassociated with water increases rapidly in the solvent mixture and this rapid increase must be accommodated by the use of a convex gradient profile when employing gradient elution.

The convex gradient will compensate for the strongly concave form of the unassociated methanol concentration profile shown in figure which will be the strongest eluting component of the mobile phase. The strong association of methanol with water could also account for the fact that proteins can tolerate a significant amount of methanol in the mobile phase before they become denatured. It is clear that this is because there is virtually no unassociated methanol present in the mixture which could cause protein denaturation since all the methanol is in a deactivated state by association with water.

Katz, Lochmüller and Scott also examined acetonitrile/ water, and tetrahydrofuran(THF)/water mixtures in the same way and showed that there was significant association between the water and both solvents but not to the same extent as methanol/water.

At the point of maximum association for methanol, the solvent mixture contained nearly sixty per cent of the methanol/water associate. In contrast the maximum amount of THF associate that was formed amounted to only about 17% and for acetonitrile the maximum amount of associate that was formed was as little as 8%.

It follows that acetonitrile water mixtures would be expected to behave more nearly as binary mixtures than methanol/ water or THF/water mixtures.

Katz et al. measured the distribution coefficient of benzene between n-hexadecane and some methanol/water mixtures.

Using the data from figure, they plotted the distribution coefficient of benzene against the volume fraction of methanol unassociated with water. The results they obtained are shown in figure.

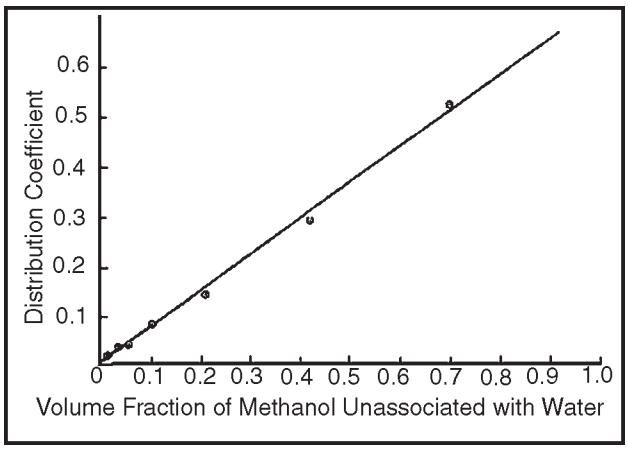


Fig. Graph of Distribution Coefficient of Benzene Between Methanol and Water Mixtures and n-Hexadecane against Volume Fraction of Methanol Unassociated with Water in the Aqueous Mixture

From the results in figure, benzene does not appear to interact with methanol associated with water or water itself but solely with methanol. The linear curve is obtained with zero intercept confirming the validity of this dependence. In fact, the methanol associated with water plays no significant part in competing for the benzene against the dispersive interactions of the n-hexadecane.

Chiral Stationary Phases

There are basically five general types of chiral stationary phase in common use in LC. The first is the protein based stationary phase. These stationary phases usually take the form of natural proteins bonded to a silica matrix.

As they are proteins, they contain a large number of chiral centers and are known to interact strongly with small

analytes exhibiting strong chiral selectivity. There are specific interactive sites that provide chiral selectivity, but there are many more sites that only contribute to general retention. These other sites can be deactivated by mobile phase additives (e.g. octylamine) which reduces the overall retention and increases the chiral selectivity.

The second type consists of relatively small molecular weight chiral substances bonded to silica 9 Pirkle. Each bonded group has a limited number of chiral centers available but, due to their small size, there can be a large number of groups bonded to the silica (as opposed to much larger complex chiral moieties).

It follows, that a relatively high probability is maintained of the solute interacting with a chiral centre. The advantage of the Pirkle chiral phases is that, as the overall interacting molecule is small, the solutes are not strongly retained and thus the chiral selectivity becomes the dominant factor.

The third type is based on polymers of cellulose and amylose which were developed by Okamato. These are derivatized to link appropriate interactive groups to the cellulose polymer which is then coated onto a silica support. The fourth type is based on the macrocyclic glycopeptides introduced by Armstrong. These are materials that also contain a large number of chiral centers, together with molecular cavities in which solute molecules can enter and interact with neighboring groups.

The spatial character of the solute will determine the degree of entry and consequently the proximity of interaction which, in turn, will determine the energy of interaction and the magnitude of the retention. Finally, the fifth group

contains the cyclodextrin based materials that control retention in a similar manner to that previously described for GC. In LC, the cyclodextrin stationary phases are bonded to a support such as silica and are prepared using similar techniques to those for making reverse phases. The more recent and most effective stationary phases are without doubt those based on the macrocyclic glycopeptides and the cyclodextrins.

Macrocyclic Glycopeptide Phases

The concept of using macrocyclic glycopeptides as chiral stationary phases was first introduced by Armstrong. One method of preparation is to covalently bond Vancomycin to the surface of silica gel particles.

Vancomycin contains 18 chiral centers surrounding three 'pockets' or 'cavities' which are bridged by five aromatic rings. Strong polar groups are proximate to the ring structures to offer strong polar interactions with the solutes. This type of stationary phase is stable in mobile phases containing 0–100% organic solvent. The proposed structure of Vancomycin is shown in figure.

A, B and C are inclusion cavities. Molecular weight 1449. Chiral centers 18. pK's 2.9, 7.2, 8.6, 9.6, 10.4, 11.7. Isoelectric point 7.2.

Vancomycin is very stable with a relatively high sample capacity, and, when covalently bonded silica gel has multiple linkages to the silica gel surface. It can be used as a reversed phase, with mobile phases having a high water content, or, alternatively, as polar stationary phase with a mobile phase of high solvent content (e.g., when used as a reversed phase,

strongly polar THF–water mixtures are very effective mobile phases. Conversely, when used as a polar stationary phase, n-hexane–ethanol mixtures are appropriate.

Vancomycin has a number of ionizing groups and thus can be used over a range of different pH values (pH 4.0 to 7.0) and exhibit a wide range of retention characteristics and chiral selectivities. Ammonium nitrate, triethylammonium acetate and sodium citrate buffers have all been used satisfactorily with this stationary phase.

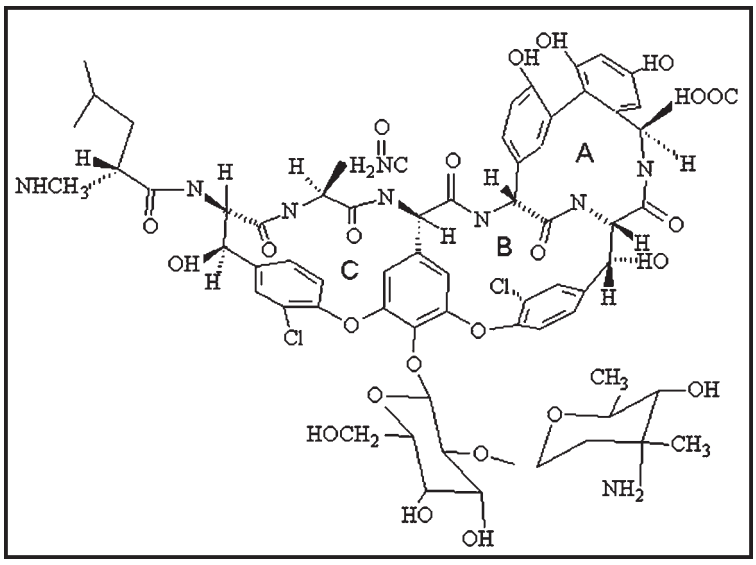


Fig. The Proposed Structure of Vancomycin

An example of the use of the stationary phase to separate the enantiomers of 3-methyl-5-phenylhydantoin is shown in figure.

t_o 1.85 min.	t_o 2.80 min.
k_1 0.78	k_1 1.57
k_2 1.32	k_2 2.13
K_1 1.69	K_2 1.35
R 2.18	R 3.0

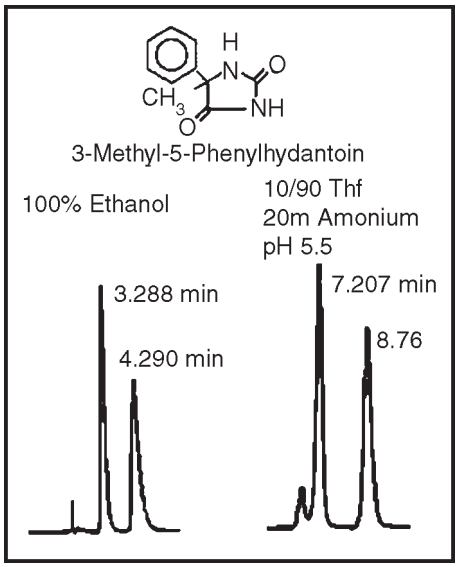


Fig. The Separation of the Enantiomers of 3-Methyl-5-Phenylhydantoin Using Polar and Dispersive Interactions

The separation was carried out under two conditions, the first used pure ethanol as the mobile phase, which is strongly dispersive, and in the second, the mobile phase that contains 90% of water. In the first case, the ethanol provides strong dispersive interactions in the mobile phase which would significantly exceed any dispersive interactions involved with the stationary phase. Consequently, the remaining dominant retentive forces will be polar or ionic.

In the second case, the mobile phase is predominantly water and thus provides strong polar interactions with the solute but weak dispersive interactions. It also follows, that the dispersive forces will dominate in the stationary phase. These two examples demonstrate the useful flexibility of Vancomycin. Another macrolytic glycopeptide used in chiral chromatography is the amphoteric glycopeptide Teicoplanin which is commercially available under the trade name of CHIROBIOTIC T.

A, B, C and D are inclusion cavities. Molecular weight 1885. Chiral centers 20, Sugar moieties 3, and R is CH_3-decanoic acid

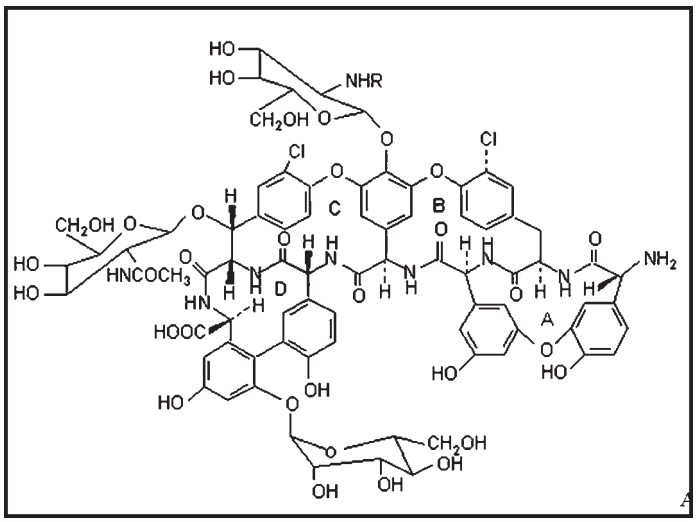


Fig. The Proposed Structure of Teicoplanin

This material can also be bonded to 5 mM silica gel particles by multiple covalent linkages. Teicoplanin contains 20 chiral centers surrounding four molecular 'pockets' or 'cavities'. Neighboring groups are strongly polar and aromatic rings provide ready polarizability.

The proposed structure of Teicoplanin is shown in figure. This stationary phase is claimed to be complementary to the Vancomycin phase and can be used with the same types of mobile phase, one often providing chiral selectivity, when the other does not.

Teicoplanin can be used in a reversed phase mode using strongly polar mixtures such as acetonitrile/aqueous buffer: 10/90 v/v, THF/aqueous buffer 10/90: v/v, and ethanol/ aqueous buffer: 20/80 v/v).

It can also be used as a polar stationary phase using n-hexane/ethanol mixtures as the mobile phase. In some cases

it is advisable to control the pH even when the solutes are not ionic, suitable buffers being ammonium nitrate and triethylamine acetate.

The separation of the Propranolol enantiomers on Teicoplanin is shown in figure.

The column used was 25 cm long, 4.6 mm I.D., packed with Chirobiotic T. The mobile phase was methanol containing acetic acid and triethylamine in the concentrations shown in figure. The column was operated at room temperature and at a flow rate of 2 ml/min. Teicoplanin is stable over a pH range of 3.8 to 6.5 although it can be used for limited periods of time outside this range. By suitable choice of mobile phase, Teicoplanin can be used in the reversed phase mode.

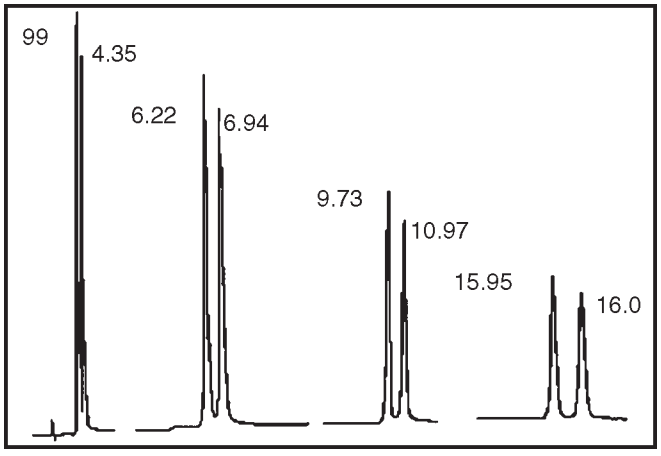


Fig. The Separation of the Enantiomers of Propranolol Employing Different Acid/Base Ratios

Another macrolytic glycopeptide that has been used as a chiral stationary phase, is the glycopeptide, Avoparcin. This stationary phase is available as CHIROBIOTIC A. Avoparcin is an antibiotic complex produced by Streptomysces candidus. In particular, it is used to prevent necrotic enteritus in chickens. The structure of avoparcin is shown in figure.

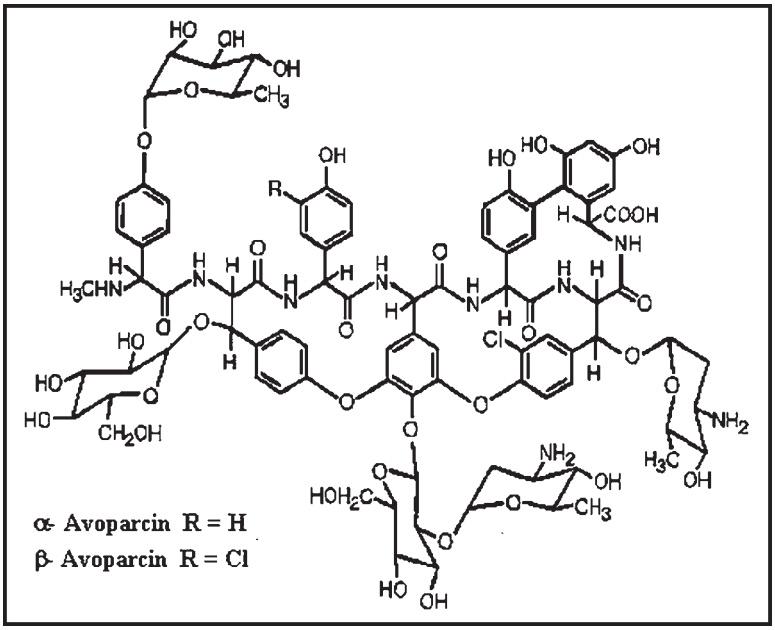


Fig. The Structure of Avoparcin

There are two forms of Avoparcin the unsubstituted b-Avoparcin structure and the chlorinated structure Avoparcin, the molecular weights being 1909 and 1944 respectively. The ratio of b-Avoparcin to b-Avoparcins is about 1:4. The aglycon portion of Avoparcin contains three connected semie rigid macrocyclic rings (one 12-membered, and two 16-membered) which form a pocket providing possible solute inclusion.

The glycopeptide contains seven aromatic rings with four phenol moieties, four carbohydrate chains, 16 hydroxyl groups, one carboxylic acid, two primary amines, one secondary amine, six amide linkages, two chlorine atoms for Avoparcin (only one for Avoparcin) and 32 stereogenic centers.

It is clear that, there is a wide diversity of interactive possibilities ranging from weak and strong dispersive interactions, to polar interactions that span from induced dipole interaction, through dipole–dipole interaction, to strong hydrogen bonding. In addition, at the right pK, basic and

acidic ionic interactions can also be invoked. More importantly, with 32 stereogenic centers the probability of interaction between chiral centers of solute and stationary phase is relatively high.

Cyclodextrin

The cyclodextrin based chiral stationary phases are some of the more popular materials used for contemporary chiral separations. One of their advantages lies in their use with all types of solvent. They can be used very effectively in the reversed phase mode and, as well as being usable as a normal phase. The cyclodextrins and their derivatives have been widely used for all types of chiral separations and can often be used for preparative separations.

Cyclodextrin-based phases are readily available, covalently bonded to spherical silica gel particles 5 mM in diameter. The cyclodextrins are produced by the partial degradation of starch followed by the enzymatic coupling of the glucose units into crystalline, homogeneous toroidal structures of different molecular size. The molecular structure of a, b, and g cyclodextrins are shown in figure.

The alpha-, beta- and gamma-cyclodextrins and have been shown to contain 6 (cyclohexamylose), 7 (cycloheptamylose) and 8 (cyclooctamylose) glucose units, respectively. These cyclic, chiral, torus shaped macromolecules contain the D(+)-glucose residues bonded through -(1-4) glycosidic linkages.

The mouth of the torus-shaped cyclodextrin molecule has a larger circumference than at the base and is linked to secondary hydroxyl groups of the C_2 and C_3 atoms of each glucose unit. The primary hydroxyl groups are located at the base of the torus on the C_6 atoms.

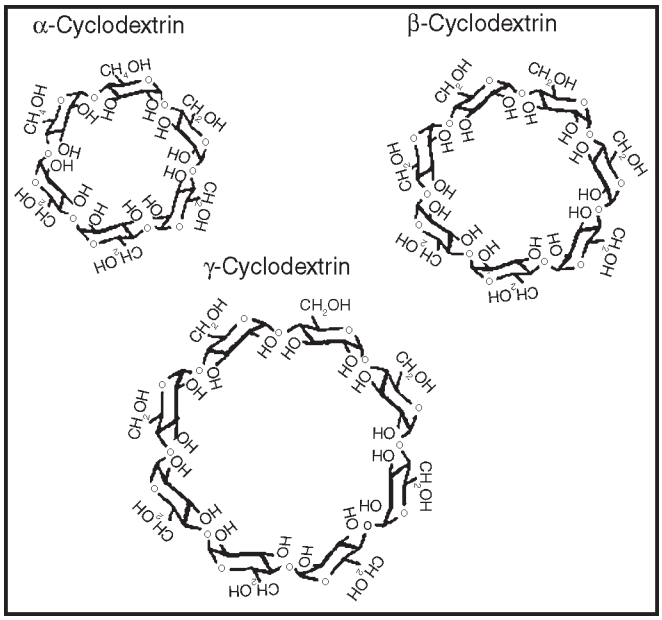


Fig. The Molecular Structure of a, b, and g Cyclodextrins

As these hydroxyl groups are free to rotate, they partially block the base aperture. The size of the cavity increases with increasing number of glucose units. The secondary hydroxyl groups can be reacted with appropriate reagents to introduce further interactive character to the cyclodextrin molecule. The very effective chiral characteristics of the cyclodextrins structures arise from the many chiral centers they contain, for example, b-cyclodextrin has 35 stereogenic centers.

When the a, b, or g cyclodextrins are derivatized, the hydroxyl group on the 2-position reacts first. However, the derivative is still size selective and interaction will be determined by the size and functional groups contained by the interacting molecule. Derivatizing the 6-hydroxyl position has little or no effect on chiral selectivity but does enhance the loading capacity of the stationary phase. This position is used for anchoring the cyclodextrins to silica gel in the preparation of LC stationary phases.

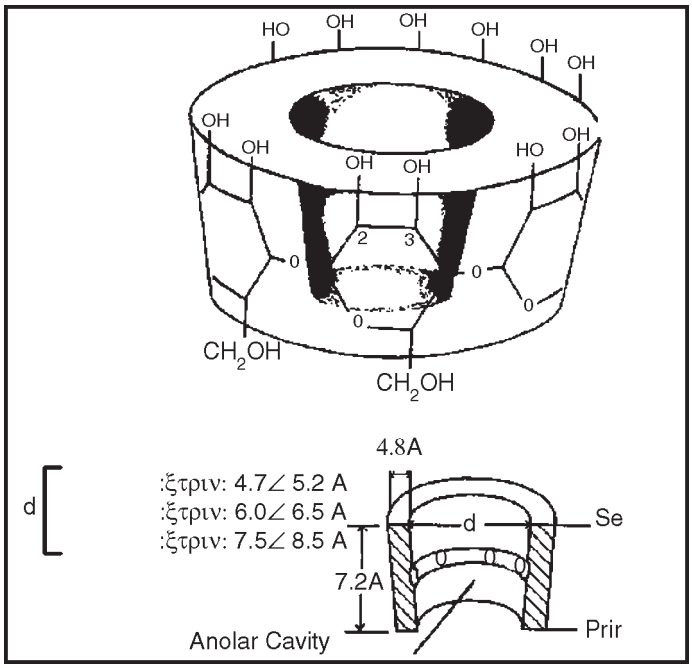


Fig. A Molecular Model of Cyclodextrin

The cyclodextrins have found a wide field of application in chiral chromatography and there arc many applications in the literature The following is a simple example of the use of this stationary phase.

In order to carry out in vivo pharmacological profiling of enantiomeric drugs, the direct analysis of biological fluids is required to reduce sample preparation time, and the chance of enantiomeric change. Unfortunately, many chiral stationary phases, such as the Pirkle types phases and the derivatized cellulose phases, demand the use of mobile phases that are incompatible with the biological fluids.

Haginaka and Wakai suggested that the silica should first be reacted with (3-glycidoxypropyl)trimethoxysilane, to cover a significant part of the silica surface with 'spacers', and then the remaining silanol groups reacted with cyclodextrin–carbamoylated triethoxysilane to attach the chiral agent.

Subsequent treatment, would convert the epoxy group to diols that would insulate the analytes from the silica surface. Stalcup and Williams analyzed a series of these type of materials and found that there was about 10 times as much spacer on the surface (ca. 2 mSmol/m^2) as there was derivatized cyclodextrin (ca. 0.2 mSmol/m^2). This type of chiral stationary proved to be very effective. The separation of the enantiomers of hexabarbital on this stationary phase by the direct injection of blood serum is shown in figure. Chromatogram A was obtained after 20 injections of serum and chromatogram B after 60 consecutive injections of blood serum. It is seen that here is very little column deterioration and that, although the tail of the major peak has become a little extended after 60 injections, the column could still be used very effectively for the analysis.

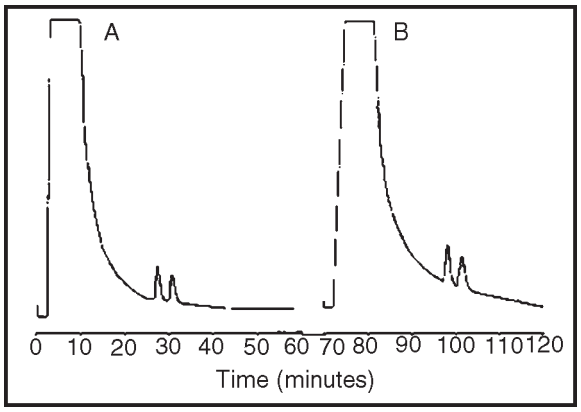


Fig. The Separation of the Hexobarbital Enantiomers Contained in Blood Serum by Direct Injection on a Cyclobond Phase

6

Specific Heats of Solids

Consider a simple solid containing N atoms. Now, atoms in solids cannot translate (unlike those in gases), but are free to vibrate about their equilibrium positions. Such vibrations are called lattice vibrations, and can be thought of as sound waves propagating through the crystal lattice. Each atom is specified by three independent position coordinates, and three conjugate momentum coordinates.

Let us only consider small amplitude vibrations. In this case, we can expand the potential energy of interaction between the atoms to give an expression which is quadratic in the atomic displacements from their equilibrium positions. It is always possible to perform a normal mode analysis of the oscillations. In effect, we can find $3N$ independent modes of oscillation of the solid.

Each mode has its own particular oscillation frequency, and its own particular pattern of atomic displacements. Any

general oscillation can be written as a linear combination of these normal modes.

Let q_i be the (appropriately normalized) amplitude of the ith normal mode, and p_i the momentum conjugate to this coordinate. In normal mode coordinates, the total energy of the lattice vibrations takes the particularly simple form

$$E = \frac{1}{2}\sum_{i=1}^{3N}(p_i^2 + \omega_i^2 q_i^2),$$

where ω_i is the (angular) oscillation frequency of the ith normal mode. It is clear that in normal mode coordinates, the linearized lattice vibrations are equivalent to $3N$ independent harmonic oscillators (of course, each oscillator corresponds to a different normal mode).

The typical value of ω_i is the (angular) frequency of a sound wave propagating through the lattice. Sound wave frequencies are far lower than the typical vibration frequencies of gaseous molecules.

In the latter case, the mass involved in the vibration is simply that of the molecule, whereas in the former case the mass involved is that of very many atoms (since lattice vibrations are non-localized).

The strength of interatomic bonds in gaseous molecules is similar to those in solids, so we can use the estimate $\omega \sim \sqrt{k/m}$ (k is the force constant which measures the strength of interatomic bonds, and m is the mass involved in the oscillation) as proof that the typical frequencies of lattice vibrations are very much less than the vibration frequencies of simple molecules.

It follows from $\Delta E = \hbar\omega$ that the quantum energy levels of lattice vibrations are far more closely spaced than the

vibrational energy levels of gaseous molecules. Thus, it is likely (and is, indeed, the case) that lattice vibrations are not frozen out at room temperature, but, instead, make their full classical contribution to the molar specific heat of the solid.

If the lattice vibrations behave classically then, according to the equipartition theorem, each normal mode of oscillation has an associated mean energy kT in equilibrium at temperature T [$(1/2)kT$ resides in the kinetic energy of the oscillation, and $(1/2)kT$ resides in the potential energy]. Thus, the mean internal energy per mole of the solid is

$$\bar{E} = 3NkT = 3\upsilon RT.$$

It follows that the molar heat capacity at constant volume is

$$c_V = \frac{1}{\upsilon}\left(\frac{\partial \bar{E}}{\partial T}\right)_V = 3R$$

for solids. This gives a value of 24.9 joules/mole/degree. In fact, at room temperature most solids (in particular, metals) have heat capacities which lie remarkably close to this value. This fact was discovered experimentally by Dulong and Petite at the beginning of the nineteenth century, and was used to make some of the first crude estimates of the molecular weights of solids (if we know the molar heat capacity of a substance then we can easily work out how much of it corresponds to one mole, and by weighing this amount, and then dividing the result by Avogadro's number, we can obtain an estimate of the molecular weight).

The experimental molar heat capacities Cp at constant pressure for various solids. The heat capacity at constant volume is somewhat less than the constant pressure value, but not by much, because solids are fairly incompressible.

It can be seen that Dulong and Petite's law (i.e., that all solids have a molar heat capacities close to 24.9 joules/mole/ degree) holds pretty well for metals. However, the law fails badly for diamond. This is not surprising.

As is well-known, diamond is an extremely hard substance, so its intermolecular bonds must be very strong, suggesting that the force constant k is large. Diamond is also a fairly low density substance, so the mass m involved in lattice vibrations is comparatively small. Both these facts suggest that the typical lattice vibration frequency of diamond ($\omega \sim \sqrt{k/m}$) is high. In fact, the spacing between the different vibration energy levels (which scales like $\hbar\omega$) is sufficiently large in diamond for the vibrational degrees of freedom to be largely frozen out at room temperature.

This accounts for the anomalously low heat capacity of diamond.

Table. Values of (Joules/Mole/Degree)
For Some Solids at T = 298° K. K. From Reif.

Solid	*Cp*	Solid	*Cp*
Copper	24.5	Aluminium	24.4
Silver	25.5	Tin (white)	26.4
Lead	26.4	Sulphur (rhombic)	22.4
Zinc	25.4	Carbon (diamond)	6.1

Dulong and Petite's law is essentially a high temperature limit. The molar heat capacity cannot remain a constant as the temperature approaches absolute zero, since, by Eq. this would imply S $\rightarrow \infty$, which violates the third law of thermodynamics. We can make a crude model of the behaviour of C_V at low temperatures by assuming that all the normal modes oscillate at the same frequency, ω, say.

According to Eq. the solid acts like a set of $3N$ independent oscillators which, making use of Einstein's

approximation, all vibrate at the same frequency. We can use the quantum mechanical result for a single oscillator to write the mean energy of the solid in the form

$$\bar{E} = 3N\hbar\omega\left(\frac{1}{2} + \frac{1}{\exp(\beta\hbar\omega) - 1}\right).$$

The molar heat capacity is defined

$$c_V = \frac{1}{v}\left(\frac{\partial \bar{E}}{\partial T}\right)_V = \frac{1}{v}\left(\frac{\partial \bar{E}}{\partial T}\right)_V \frac{\partial \beta}{\partial T} = -\frac{1}{vkT^2}\left(\frac{\partial \bar{E}}{\partial T}\right)_V,$$

giving

$$c_V = -\frac{3N_A\hbar\omega}{kT^2}\left[-\frac{\exp(\beta\hbar\omega)}{[\exp(\beta\hbar\omega) - 1]^2}\right],$$

which reduces to

$$c_V = 3R\left(\frac{\theta_E}{T}\right)^2 \frac{\exp(\theta_E/T)}{[\exp(\theta_E/T) - 1]^2}.$$

Here,

$$\theta_E = \frac{\hbar\omega}{k}$$

is called the Einstein temperature. If the temperature is sufficiently high that T>> θ_E then $kT >> \hbar\omega$, and the above expression reduces to C_V = 3R,, after expansion of the exponential functions. Thus, the law of Dulong and Petite is recovered for temperatures significantly in excess of the Einstein temperature. On the other hand, if the temperature is sufficiently low that T<< θ_E then the exponential factors in Eq. become very much larger than unity, giving

$$C_V \sim 3R\left(\frac{\theta_E}{T}\right)\exp(-\theta_E/T).$$

So, in this simple model the specific heat approaches zero exponentially as $T \to 0$.

In reality, the specific heats of solids do not approach zero quite as quickly as suggested by Einstein's model when $T \to 0$. The experimentally observed low temperature behaviour is more like $C_V \propto T^3$. The reason for this

discrepancy is the crude approximation that all normal modes have the same frequency. In fact, long wavelength modes have lower frequencies than short wavelength modes, so the former are much harder to freeze out than the latter (because the spacing between quantum energy levels, $\hbar\omega$, is smaller in the former case).

The molar heat capacity does not decrease with temperature as rapidly as suggested by Einstein's model because these long wavelength modes are able to make a significant contribution to the heat capacity even at very low temperatures.

A more realistic model of lattice vibrations was developed by the Dutch physicist Peter Debye in 1912. In the Debye model, the frequencies of the normal modes of vibration are estimated by treating the solid as an isotropic continuous medium. This approach is reasonable because the only modes which really matter at low temperatures are the long wavelength modes: i.e., those whose wavelengths greatly exceed the interatomic spacing. It is plausible that these modes are not particularly sensitive to the discrete nature of the solid: i.e., the fact that it is made up of atoms rather than being continuous.

Consider a sound wave propagating through an isotropic continuous medium. The disturbance varies with position vector r and time t like $\exp[-i\,(\mathrm{k.r} - \omega t)]$, where the wave-vector k and the frequency of oscillation ω satisfy the dispersion relation for sound waves in an isotropic medium:

$$\omega = kC_s.$$

Here, C_S is the speed of sound in the medium. Suppose, for the sake of argument, that the medium is periodic in the

x–, y–, and z–directions with periodicity lengths L_x, L_y, and L_z, respectively. In order to maintain periodicity we need

$$k_x(x + L_x) = k_x + 2\pi n_x,$$

where $n_x$is an integer. There are analogous constraints on k_y and k_z. It follows that in a periodic medium the components of the wave-vector are quantized, and can only take the values

$$k_x = \frac{2\pi}{L_x} n_x,$$
$$k_y = \frac{2\pi}{L_y} n_y,$$
$$k_z = \frac{2\pi}{L_z} n_z,$$

where n_x, n_y, and n_z are all integers. It is assumed that L_x, L_y, and L_z are macroscopic lengths, so the allowed values of the components of the wave-vector are very closely spaced. For given values of k_y and k_z, the number of allowed values of k_x which lie in the range k_x to $k_x + dk_x$ is given by

$$\Delta n_x = \frac{L_x}{2\pi} 2k_x.$$

It follows that the number of allowed values of k (i.e., the number of allowed modes) when k_x lies in the range k_z to $k_z + dk_z$, k_z, + k_y lies in the range k_y to $k_y + dk_y$, and k_z lies in the range k_z to $k_z + dk_z$, is

$$\rho d^3k = \left(\frac{L_x}{2\pi} dk_x\right)\left(\frac{L_y}{2\pi} dk_y\right)\left(\frac{L_z}{2\pi} dk_z\right) = \frac{V}{(2\pi)^3} dk_x dk_y dk_z,$$

where $V = L_x L_y L_z$ is the periodicity volume, and $d^3k \equiv dk_x dk_y dk_z$. The quantity ρ is called the density of modes. Note that this density is independent of k, and proportional to the periodicity volume.

Thus, the density of modes per unit volume is a constant independent of the magnitude or shape of the periodicity

volume. The density of modes per unit volume when the magnitude of k lies in the range k to $k + dk$ is given by multiplying the density of modes per unit volume by the "volume" in k-space of the spherical shell lying between radii k and $k + dk$. Thus,

$$\rho_k dk = \frac{4\pi k^2\, dk}{(2\pi)} = \frac{V^2}{2\pi^2} dk.$$

Consider an isotropic continuous medium of volume V. According to the above relation, the number of normal modes whose frequencies lie between ω and $\omega + d\omega$ (which is equivalent to the number of modes whose k values lie in the range ω/C_s to $\omega/C_s + d\omega/C_s$) is

$$\sigma_C(\omega)\, d\omega = 3\frac{k^2\, V}{2\pi^2} dk = 3\frac{V^2}{2\pi^2 {C_S}^3}\omega^2 d\omega.$$

The factor of 3 comes from the three possible polarizations of sound waves in solids. For every allowed wavenumber (or frequency) there are two independent torsional modes, where the displacement is perpendicular to the direction of propagation, and one longitudinal mode, where the displacement is parallel to the direction of propagation. Torsion waves are vaguely analogous to electromagnetic waves (these also have two independent polarizations). The longitudinal mode is very similar to the compressional sound wave in gases. Of course, torsion waves can not propagate in gases because gases have no resistance to deformation without change of volume.

The Debye approach consists in approximating the actual density of normal modes $\sigma(\omega)$ by the density in a continuous medium $\sigma_C(\omega)$, not only at low frequencies (long wavelengths) where these should be nearly the same, but also at higher frequencies where they may differ substantially.

Suppose that we are dealing with a solid consisting of N atoms. We know that there are only $3N$ independent normal modes. It follows that we must cut off the density of states above some critical frequency, ω_D say, otherwise we will have too many modes. Thus, in the Debye approximation the density of normal modes takes the form

$$\sigma_D(\omega) = \sigma_C(\omega) \text{ for } \omega < \omega_D$$
$$\sigma_D(\omega) = 0 \text{ for } \omega < \omega_D$$

Here, ω_D is the Debye frequency. This critical frequency is chosen such that the total number of normal modes is $3N$, so

$$\int_0^\infty \sigma_C(\omega)\, d\omega = \int_0^{\omega_D} \sigma_C(\omega)\, d\omega = 3N.$$

Substituting Eq. into the previous formula yields

$$\frac{3V}{2\pi^2 C_s^3}\int_0^{\omega_D} d\omega = \frac{3V}{2\pi^2 C_s^3}\omega_D^3 = 3N.$$

This implies that

$$\omega_D = C_s\left(6\pi^2\,\frac{N}{V}\right)^{1/3}.$$

Thus, the Debye frequency depends only on the sound velocity in the solid and the number of atoms per unit volume. The wavelength corresponding to the Debye frequency is $2\pi\, C_s/\omega_D$, which is clearly on the order of the interatomic spacing $a \sim (V/N)^{1/3}$.

It follows that the cut-off of normal modes whose frequencies exceed the Debye frequency is equivalent to a cut-off of normal modes whose wavelengths are less than the interatomic spacing. Of course, it makes physical sense that such modes should be absent.

Compares the actual density of normal modes in diamond with the density predicted by Debye theory. Not surprisingly,

there is not a particularly strong resemblance between these two curves, since Debye theory is highly idealized. Nevertheless, both curves exhibit sharp cut-offs at high frequencies, and coincide at low frequencies. Furthermore, the areas under both curves are the same. This is sufficient to allow Debye theory to correctly account for the temperature variation of the specific heat of solids at low temperatures.

We can use the quantum mechanical expression for the mean energy of a single oscillator, Eq. to calculate the mean energy of lattice vibrations in the Debye approximation. We obtain

$$\bar{E} = \int_0^\infty \sigma_D(\omega)\hbar\omega\left(\frac{1}{2} + \frac{1}{\exp(\beta\hbar\omega) - 1}\right)d\omega.$$

According to Eq. the molar heat capacity takes the form

$$C_V = \frac{1}{vkT^2}\int_0^\infty \sigma_D(\omega)\hbar\omega\left[\frac{\exp(\beta\hbar\omega)\hbar\omega}{[\exp(\beta\hbar\omega) - 1]^2}\right]d\omega.$$

Substituting in Eq. we find that

$$C_V = \frac{k}{v}\int_0^{\omega_D} \frac{\exp(\beta\hbar\omega)(\beta\hbar\omega)^2}{[\exp(\beta\hbar\omega) - 1]^2}\frac{3V}{2\pi^2 C_s^3}\omega^2 d\omega,$$

giving

$$C_V = \frac{3Vk}{2\pi^2 v(C_s\beta\hbar)^3}\int_0^{\beta\hbar\omega_D} \frac{\exp x}{(\exp x - 1)^2}x^4 dx,$$

in terms of the dimensionless variable $x = \beta\hbar\omega$. According to Eq. the volume can be written

$$V = 6\pi^2 N\left(\frac{Cs}{\omega_D}\right)^3,$$

so the heat capacity reduces to

$$C_V = 3Rf_D(\beta\hbar\omega_D) = 3Rf_D(\theta_D / T),$$

where the Debye function is defined

$$f_D(y) \equiv \frac{3}{y^3}\int_0^y \frac{\exp x}{(\exp x - 1)^2}x^4 dx.$$

We have also defined the Debye temperature θ_D as

$$k\theta_D = \hbar\omega_D.$$

Consider the asymptotic limit in which $T \gg \theta_D$. For small y, we can approximate exp x as 1 + x in the integrand of Eq. so that

$$f_D(y) \to \frac{3}{y^3}\int_0^y x^2 dx = 1.$$

Thus, if the temperature greatly exceeds the Debye temperature we recover the law of Dulong and Petite that $C_V = 3R$. Consider, now, the asymptotic limit in which $T \ll \theta_D$. For large y,

$$\int_0^y \frac{\exp x}{(\exp x - 1)^2} x^4 dx \to \int_0^\infty \frac{\exp x}{(\exp x - 1)^2} x^4 dx = \frac{4\pi^4}{15}.$$

Thus, in the low temperature limit

$$f_D(y) \to \frac{4\pi^4}{5}\frac{1}{y^3}.$$

This yields

$$c_V \to \frac{12\pi^4}{5} R\left(\frac{T}{\theta_D}\right)^3$$

in the limit $T \ll \theta_D$: i.e., c_V varies with temperature like T^3.

Table. Comparison of Debye Temperatures (In Degrees Kelvin)

Solid	θ_D *from low temp.*	θ_D *from sound speed*
NaCl	308	320
KCl	230	246
Ag	225	216
Zn	308	305

The fact that c_V goes like T^3 at low temperatures is quite well verified experimentally, although it is sometimes necessary to go to temperatures as low as 0.02 θ_D to obtain this asymptotic behaviour.

Theoretically, θ_D should be calculable from Eq. in terms of the sound speed in the solid and the molar volume. A comparison of Debye temperatures evaluated by this means

with temperatures obtained empirically by fitting the law to the low temperature variation of the heat capacity. It can be seen that there is fairly good agreement between the theoretical and empirical Debye temperatures. This suggests that the Debye theory affords a good, thought not perfect, representation of the behaviour of c_V in solids over the entire temperature range.

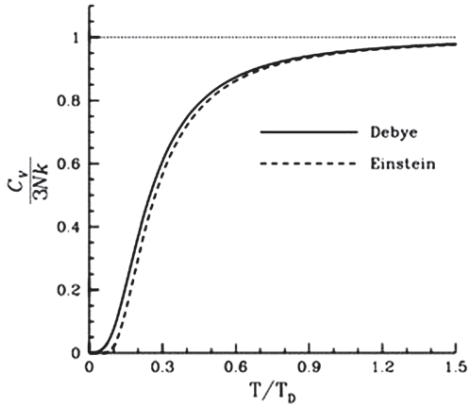


Fig. The Molar Heat Capacity of Various Solids.

Finally, the actual temperature variation of the molar heat capacities of various solids as well as that predicted by Debye's theory. The prediction of Einstein's theory is also show for the sake of comparison. Note that 24.9 joules/mole/degree is about 6 calories/gram-atom/degree (the latter are chemist's units).

The Maxwell distribution

Consider a molecule of mass m in a gas which is sufficiently dilute for the intermolecular forces to be negligible (i.e., an ideal gas). The energy of the molecule is written

$$\in = \frac{P^2}{2m} + \in^{\text{int}},$$

where p is its momentum vector, and $\in^{\text{int}}$ is its internal (i.e., non-translational) energy. The latter energy is due to

molecular rotation, vibration, etc. Translational degrees of freedom can be treated classically to an excellent approximation, whereas internal degrees of freedom usually require a quantum mechanical approach.

Classically, the probability of finding the molecule in a given internal state with a position vector in the range r to $r + dr$, and a momentum vector in the range P to $P + dP$, is proportional to the number of cells (of "volume" h_0) contained in the corresponding region of phase-space, weighted by the Boltzmann factor.

In fact, since classical phase-space is divided up into uniform cells, the number of cells is just proportional to the "volume" of the region under consideration. This "volume" is written $d^3r d^3p$. Thus, the probability of finding the molecule in a given internal state s is

$$P_s(r,p)d^3r\, d^3p \propto \exp(-\beta p^2/2m)\exp(\beta\in_s^{\text{int}})d^3r\, d^3p,$$

where Ps is a probability density defined in the usual manner. The probability $P(r, p)\, d^3r\, d^3p$ of finding the molecule in any internal state with position and momentum vectors in the specified range is obtained by summing the above expression over all possible internal states.

The sum over $\exp(-\beta\in_s^{\text{int}})$ just contributes a constant of proportionality (since the internal states do not depend on r or p), so

$$P(r,p)d^3r\, d^3p \propto \exp(-\beta p^2/2m)\, d^3r\, d^3p.$$

Of course, we can multiply this probability by the total number of molecules N in order to obtain the mean number of molecules with position and momentum vectors in the specified range.

Suppose that we now want to determine: $f(r, v)\, d^3r\, d^3v$ i.e., the mean number of molecules with positions between r and $r + dr$, and velocities in the range v and $v + dv$. Since $v = p/m$, it is easily seen that

$$f(r,v)d^3r\, d^3v = C\exp(-\beta mv^2/2)\, d^3r\, d^3v,$$

where C is a constant of proportionality. This constant can be determined by the condition

$$\int_{(r)}\int_{(v)} f(r,v)d^3r\, d^3v = N :$$

i.e., the sum over molecules with all possible positions and velocities gives the total number of molecules, N. The integral over the molecular position coordinates just gives the volume V of the gas, since the Boltzmann factor is independent of position. The integration over the velocity coordinates can be reduced to the product of three identical integrals (one for v_x, one for v_y, and one for v_z), so we have

$$CV\left[\int_{-\infty}^{\infty}\exp(-\beta mv_z^2/2)dv_z\right]^3 = N.$$

Now,

$$\int_{-\infty}^{\infty}\exp(-\beta mv_z^2/2)dv_z = \sqrt{\frac{2}{\beta m}}\int_{-\infty}^{\infty}\exp(-y^2)dy = \sqrt{\frac{2\pi}{\beta m}},$$

so $C = (N/V)\,(\beta m/2\pi)^{3/2}$. Thus, the properly normalized distribution function for molecular velocities is written

$$f(v)d^3r\, d^3v = n\left(\frac{m}{2\pi kT}\right)^{3/2}\exp(-mv^2/2kT)d^3r\, d^3v.$$

Here, $n = N/V$ is the number density of the molecules. We have omitted the variable r in the argument of f, since f clearly does not depend on position. In other words, the distribution of molecular velocities is uniform in space. This is hardly surprising, since there is nothing to distinguish one region of space from another in our calculation.

The above distribution is called the Maxwell velocity distribution, because it was discovered by James Clark Maxwell in the middle of the nineteenth century. The average number of molecules per unit volume with velocities in the range v to $v + dv$ is obviously $f(v)\ d^3v$. Let us consider the distribution of a given component of velocity: the z-component, say. Suppose that $g(v_z)\ dv_z$ is the average number of molecules per unit volume with the z-component of velocity in the range v_z to $v_z + dv_z$, irrespective of the values of their other velocity components. It is fairly obvious that this distribution is obtained from the Maxwell distribution by summing (integrating actually) over all possible values of v_x and v_y, with v_z in the specified range. Thus,

$$g(vz)dv_z = \int_{(v_z)}\int_{(v_y)} f(v)d^3v.$$

This gives

$$g(v_z)dv_z = n\left(\frac{m}{2\pi kT}\right) \int_{(v_z)}\int_{(v_y)} \exp\left[-(m/2kT)\left(v_z^2 + v_v^2 + v_z^2\right)\right]dv_x dv_y dv_z$$

$$= n\left(\frac{m}{2\pi\, kT}\right)^{3/2} \exp\left(-mv_z^2/2kT\right)\left[\int_{-\infty}^{\infty} \exp\left(-mv_x^2/2k\,T\right)\right]^2$$

$$= n\left(\frac{m}{2\pi\, kT}\right)^{3/2} \exp\left(-mv_z^2/2\,kT\right)\left(\sqrt{\frac{2\pi\, kT}{m}}\right)^2,$$

or

$$g(v_z)\,dv_z = n\left(\frac{m}{2\pi\, kT}\right)^{1/2} \exp\left(-mv_z^2/2\,kT\right)dv_z.$$

Of course, this expression is properly normalized, so that

$$\int_{-\infty}^{\infty} g(v_z)dv_z = n.$$

It is clear that each component (since there is nothing special about the z-component) of the velocity is distributed with a Gaussian probability distribution, centred on a mean value

$$\overline{v_z} = 0,$$

with variance

$$\overline{v_z^2} = \frac{kT}{m}.$$

Equation implies that each molecule is just as likely to be moving in the plus z-direction as in the minus z-direction. Equation can be rearranged to give

$$\overline{\frac{1}{2}mv_z^2} = \frac{1}{2}kT,$$

in accordance with the equipartition theorem.

Note that Eq. can be rewritten

$$\frac{f(v)d^3v}{n} = \left[\frac{g(v_x)dv_x}{n}\right]\left[\frac{g(v_y)dv_y}{n}\right]\left[\frac{g(v_z)dv_x}{n}\right],$$

where $g(v_x)$ and $g(v_y)$ are defined in an analogous way to $g(v_z)$. Thus, the probability that the velocity lies in the range v to $v + dv$ is just equal to the product of the probabilities that the velocity components lie in their respective ranges. In other words, the individual velocity components act like statistically independent variables.

Suppose that we now want to calculate $F(v)\,dv$: i.e., the average number of molecules per unit volume with a speed $v = |v|$ in the range v to $v + dv$. It is obvious that we can obtain this quantity by adding up all molecules with speeds in this range, irrespective of the direction of their velocities. Thus,

$$F(v)dv = \int f(v)d^3v,$$

where the integral extends over all velocities satisfying

$$v < |v| < v + dv.$$

This inequality is satisfied by a spherical shell of radius v and thickness dv in velocity space. Since $f(v)$ only depends on $|v|$, so $f(v) \equiv f(v)$, the above integral is just $f(v)$ multiplied by the volume of the spherical shell in velocity space. So,

$$F(v)dv = 4\pi f(v)v^2 dv,$$

which gives

$$F(v)dv = 4\pi f(v)\left(\frac{m}{2\pi kT}\right)v^2 \exp(-mv^2/2kT)\,dv.$$

This is the famous Maxwell distribution of molecular speeds. Of course, it is properly normalized, so that

$$\int_0^\infty F(v)dv = n.$$

Note that the Maxwell distribution exhibits a maximum at some non-zero value of v. The reason for this is quite simple. As v increases, the Boltzmann factor decreases, but the volume of phase-space available to the molecule (which is proportional to v^2) increases: the net result is a distribution with a non-zero maximum.

The mean molecular speed is given by

$$\overline{v} = \frac{1}{n}\int_0^\infty F(v)v\,dv.$$

Thus, we obtain

$$\overline{v} = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2}\int_0^\infty v^3 \exp(-mv^2/2kT)\,dv,$$

or

$$\overline{v} = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2}\left(\frac{2kT}{m}\right)^2\int_0^\infty y^3 \exp(-y^2)\,dy$$

Now

$$\int_0^\infty y^3 \exp(-y^2)\,dy = \frac{1}{2},$$

so

$$\overline{v} = \sqrt{\frac{8}{\pi}\frac{kT}{m}}.$$

A similar calculation gives

$$v_{\text{rms}} = \sqrt{\overline{v^2}} = \sqrt{\frac{3kT}{m}}.$$

However, this result can also be obtained from the equipartition theorem. Since

$$\overline{\frac{1}{2}mv^2} = \overline{\frac{1}{2}m\left(v_z^2 + v_y^2 + v_z^2\right)} = 3\left(\frac{1}{2}kT\right),$$

then Eq. follows immediately. It is easily demonstrated that the most probable molecular speed (i.e., the maximum of the Maxwell distribution function) is

$$\overline{v} = \sqrt{\frac{2kT}{m}}.$$

The speed of sound in an ideal gas is given by

$$c_s = \sqrt{\frac{\gamma kT}{m}},$$

where γ is the ratio of specific heats. This can also be written since $p = nkT$ and $\rho = nm$. It is clear that the various average speeds which we have just calculated are all of order the sound speed (i.e., a few hundred meters per second at room temperature). In ordinary air ($\gamma = 1.4$) the sound speed is about 84% of the most probable molecular speed, and about 74% of the mean molecular speed. Since sound waves ultimately propagate via molecular motion, it makes sense that they travel at slightly less than the most probable and mean molecular speeds.

The Maxwell velocity distribution as a function of molecular speed in units of the most probable speed. Also shown are the mean speed and the root mean square speed.

It is difficult to directly verify the Maxwell velocity distribution. However, this distribution can be verified indirectly by measuring the velocity distribution of atoms exiting from a small hole in an oven.

The velocity distribution of the escaping atoms is closely related to, but slightly different from, the velocity distribution inside the oven, since high velocity atoms escape more readily than low velocity atoms. In fact, the predicted velocity

distribution of the escaping atoms varies like $v^3 \exp(-mv^2/2kT)$, in contrast to the $v^3 \exp(-mv^2/2kT)$ variation of the velocity distribution inside the oven. The measured and theoretically predicted velocity distributions of potassium atoms escaping from an oven at 157°C. There is clearly very good agreement between the two.