

Thermomechanical processing treatments involve the simultaneous use of both heat and plastic deformation to achieve desired changes in both the external shape and the microstructure of a material. The *hot rolling* of steels in the range $T = 1200$ to 1300°C , for example, achieves several purposes: the reduction in cross section of a large steel ingot, the breaking down of the original coarse microstructure in the as-cast material, the reduction of compositional inhomogeneities, and the redistribution of impurities. As hot rolling is carried out at successively lower temperatures, the precipitation of carbides, nitrides, and carbonitrides occurs, leading to the pinning of grain boundaries. As a result, grain refinement (i.e., the achievement of lower average grain sizes) and dispersion strengthening can both occur during hot rolling, leading to significant increases in the yield strength of the steel.

The *welding* of steels to fabricate structural forms is often an unavoidable processing step which can cause unwanted changes in the microstructure and properties of the steel in the vicinity of the weld. Fusion welding involves the melting of the steel in regions near the weld, known as the *fusion zone*, as well as large increases of temperature in surrounding areas known as the *heat-affected zone*. Significant changes in the microstructure of the steel can occur in both zones, affecting both its corrosion resistance and strength. Many of the phase transformations and processes already described in this section occur in and near the weld. Honeycombe and Bhadeshia (1996, Chapter 13) present a brief summary of the important effects associated with the generation of weld microstructures in steels.

W21.11 Precipitation Hardening of Aluminum Alloys

Pure FCC Al metal has the following properties: a low density, $\rho \approx 2700 \text{ kg/m}^3$, and a low melting point, $T_m = 660^\circ\text{C}$; high electrical and thermal conductivities; high ductility in the annealed state; high corrosion resistance due to the thin coating of the protective oxide Al_2O_3 . Because of the relatively low strength of pure Al, its alloys with elements such as Cu, Si, and Mg have found a wider range of applications. The microstructures of these alloys are characterized by a solid-solution phase, $\alpha\text{-Al}$, and by intermetallic compounds such as CuAl_2 and Al_3Mg_2 .

Al alloys are typically strengthened by the mechanism of *precipitation* or *age hardening*. The precipitation-hardening process involves the use of heat treatments, which result in precipitation within the original matrix of a uniform dispersion of very small particles of a second phase. Although a heat-treatment process, precipitation hardening involves a distinctly different sequence of steps than occur in the heat treatment of steels, which results in the formation of martensite, for example. Two heat treatments are typically required, the first for creating a solid solution and the second for accelerating the process of precipitation or aging. The first heat treatment takes place at a temperature near T_e and for a time long enough to produce a solid solution. The alloy is then quenched to room temperature to obtain a supersaturated solid solution. The second heat treatment is then carried out at a lower T to allow the diffusion to occur which is necessary for formation of the precipitates of the second phase, which results in the strengthening of the alloy. Precipitation hardening is more commonly carried out in Al–Cu, Al–Si, Cu–Be, Cu–Sn, and Mg–Al alloys and in Ni_3Ti and Ni_3Al compounds than in ferrous alloys. Precipitation hardening in Ni_3Al is discussed in Section 12.8.

To illustrate a specific example of the precipitation-hardening process in Al alloys, consider the Al-rich side of the Al–Cu equilibrium phase diagram (Fig. W21.24). The

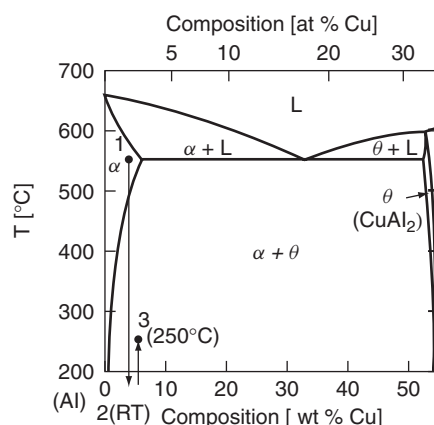


Figure W21.24. Al-rich side of the Al–Cu equilibrium phase diagram shown to illustrate the precipitation-hardening process. The two stable solid phases present are the solid-solution α -Al phase and the θ phase (i.e., the intermetallic compound CuAl_2). The sequence of treatments used for precipitation hardening of an Al–1.5Cu wt % alloy is also shown: 1, solid-solution heat treatment at $T \approx 550^\circ\text{C}$; 2, quench to room temperature; 3, precipitation heat treatment at $T \approx 250^\circ\text{C}$. (From ASM Handbook, 9th ed., Vol. 3, *Alloy Phase Diagrams*, ASM International, Materials Park, Ohio, 1992, p. 244.)

two stable solid phases present are α -Al, which is a solid solution of Cu in Al, and the θ phase corresponding to the intermetallic compound CuAl_2 . The solubility of Cu in α -Al reaches a maximum value of $x_e = 5.6$ wt % at $T_e = 548^\circ\text{C}$ and then decreases rapidly with decreasing T , reaching ≈ 0.02 wt % at room temperature. The initial heat treatment for obtaining a solid solution takes place near T_e for $\text{Al}_{1-x}\text{Cu}_x$ alloys with $x < x_e$. Following quenching to room temperature, the Al–Cu alloy then undergoes a precipitation heat treatment. If the alloy is left either at room temperature for a few days or is reheated to $T \approx 100$ to 150°C , the Cu atoms are not able to undergo sufficient diffusion to form precipitates of CuAl_2 . Instead, they rearrange themselves locally within the lattice on $\{100\}$ planes in two-dimensional platelets or disks known as *Guinier–Preston (GP) zones*. The first structures formed, known as *GP-1 zones*, are coherent with the Al lattice and are essentially randomly distributed in the alloy. They are typically 3 to 6 nm long with thicknesses of 0.5 to 1 nm. Their Cu contents are deficient with respect to $x = \frac{1}{3}$, the fraction found in CuAl_2 .

Additional aging of the alloy leads to the gradual growth of the GP-1 zones and then to the formation of a series of phases or precipitates. The larger *GP-2 zones*, also known as the θ'' phase, with lengths ≈ 10 nm, widths ≈ 1 to 4 nm, and Cu contents $x \approx \frac{1}{3}$ are formed next, followed by their conversion into an intermediate θ' phase, which is metastable and incoherent with the Al lattice. The stable θ equilibrium phase finally forms from the θ' phase when the aging temperature is raised to $T \approx 200$ to 250°C . The θ' and θ phases both have the CuAl_2 stoichiometry but have different crystal structures. The hardness and strength of precipitation-hardened Al–Cu alloys reach maximum values when the GP-2 zones (i.e., the θ'' phase) are formed and then decreases with further heat treatment as the θ' and then the θ phases appear.

The sequence of microstructures of the supersaturated α -Al solid solution and of the θ'' and θ phases are illustrated schematically in Fig. W21.25. Precipitation-hardened

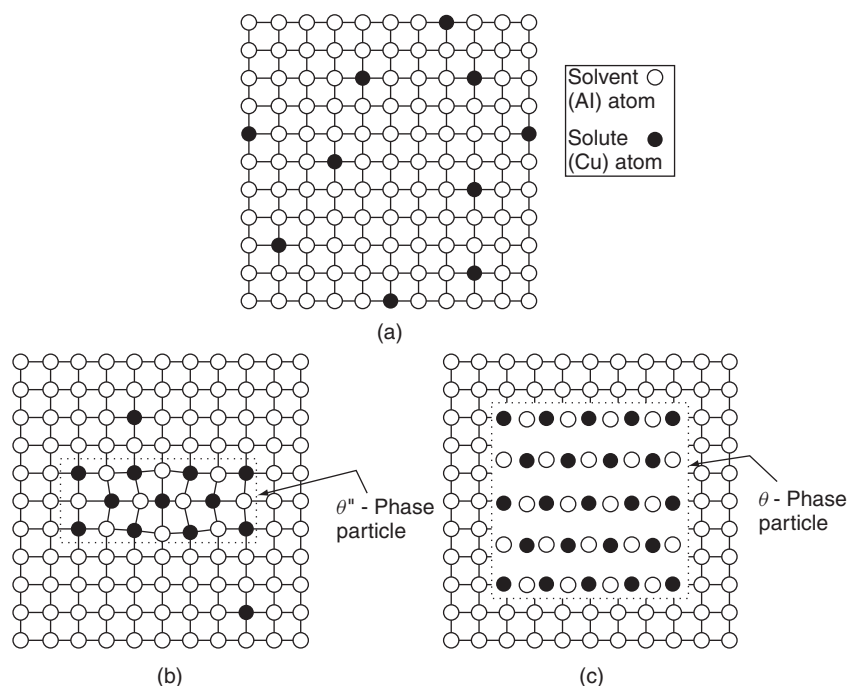


Figure W21.25. Microstructures of (a) the supersaturated α -Al solid solution and of (b) the θ'' and (c) the θ phases. The θ phase has the CuAl_2 stoichiometry. The actual particle or zone sizes are much larger than shown here. (From W. D. Callister, Jr., *Materials Science and Engineering*, 2nd ed., copyright 1991 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.)

Al alloys can in general have complicated microstructures corresponding to mixtures of the phases mentioned earlier. The strengthening of the alloy can be described by the Orowan expression, Eq. (W21.25), with Λ the average distance between precipitate particles. Strengthening is enhanced when significant lattice strain exists at the interface between the precipitates and the surrounding matrix. This lattice strain is particularly effective in impeding the motion of dislocations. When aging proceeds to the extent that the CuAl_2 precipitates become too large and too few in number, they are much less effective in impeding the motion of dislocations. When this happens, the strength of the alloy can actually decrease, a phenomenon known as *overaging*.

W21.12 Synthesis of Metals via Rapid Solidification

As the name indicates, *rapid-solidification processing* (RSP) of metals involves a rapid transition from the liquid to the solid state. RSP usually involves the cooling of liquid metals at sufficiently high rates, $\approx 10^3$ to 10^9 K/s, so that nonequilibrium compositions, phases, or microstructures that are not ordinarily obtainable at “normal” cooling rates of $\approx 10^{-2}$ to 10^2 K/s (≈ 10 to 10^5 K/h) can be synthesized. The amorphous or nanocrystalline microstructures often resulting from the RSP of metals have led to the use of the term *metallic glass*. It is ordinarily extremely difficult to produce elemental metals in an amorphous state due to the ease with which liquid metals crystallize due

to their low viscosities and high diffusivities and the ease with which solid metals recrystallize. By contrast, materials based on Si–O₄ tetrahedra, such as silicates, form glasses relatively easily on cooling due to the high viscosity of the liquid.

Metals that have been synthesized via RSP include hard and soft magnetic materials; high-strength Al, Mg, and Ti alloys; tool steels; shape-memory alloys; Ni-based superalloys and brazing materials. Some of the properties of metallic glasses are discussed in Chapter W12. The random close-packing model for the short-range order found in metallic glasses is discussed in Chapter 4.

Techniques that are used in RSP to obtain extremely high cooling rates include the following:

1. *Splat cooling.* A small, molten drop of metal is incident at high speed onto a metallic substrate (e.g., copper) held at room temperature or below. A related method involves the trapping of the molten drop between two cooled surfaces (e.g., a hammer and an anvil).
2. *Melt spinning.* A molten stream of metal is projected against a rapidly rotating surface.
3. *Twin-roller quenching.* A molten stream of metal is forced between a pair of rapidly rotating rollers.
4. *Plasma or flame spraying.* The metal in the form of a powder is introduced into a high-temperature plasma or flame and then sprayed onto a cooled substrate.
5. *Surface melting.* A source of thermal energy such as a laser, ion beam, or electron beam causes a thin surface layer of a metal to melt. The surface layer then undergoes rapid resolidification as soon as the source of heat is removed.

In the first three techniques listed above, and in similar techniques not mentioned here specifically, the rapid solidification is achieved by placing as thin a layer of molten metal as possible in contact with a cooled surface of high thermal conductivity to obtain as high a rate of heat extraction as possible from the molten metal. As a result, the materials are typically thin foils or thin, continuous ribbons. The small dimension of the rapidly solidified material is typically ≈ 25 to $50\ \mu\text{m}$.

Another technique for achieving the rapid solidification of a metal is through the use of strong undercooling of several hundreds of degrees celsius, as when small, molten metallic particles are cooled well below their normal melting point by avoiding nucleation of the solid phase. This RSP technique, known as *atomization*, involves breakup of a stream of molten metal into fine particles. In this case once a solid nucleus forms in a given particle, solidification occurs extremely rapidly due to the high velocity of the solid–liquid interface, which passes through the particle. The resulting solid powder usually needs additional processing (e.g., consolidation) before it can be used to form a solid object. Additional processing of RSP materials is often needed to develop microstructures with the desired mechanical properties. Strong undercooling can, of course, also occur during the rapid cooling processes listed above.

A necessary condition for obtaining nonequilibrium compositions via RSP is that the growth rate or solidification velocity v_{sl} be greater than the diffusive speed $v_d = D/d_a$ of the solute in the liquid metal. Here D is the thermal diffusivity, $\approx 10^{-9}\ \text{m}^2/\text{s}$, of the solute and d_a is the interatomic distance, $\approx 3 \times 10^{-10}\ \text{m}$. Other important materials parameters that influence the degree of solute incorporation in the solid phase include the solid–liquid interface energy density σ_{sl} and the latent heat ΔH_m and entropy

change ΔS_m for the liquid–solid transition. When $v_{sl} > v_d \approx 0.03$ m/s, it follows that solute can be trapped at above-equilibrium levels in the solidifying solvent. In the limit $v_{sl} \gg v_d$, the solute distribution coefficient K will approach 1. This has been observed in doped Si and in metallic alloys when $v_{sl} > 5$ m/s. For comparison, a typical value for the normal cooling of a steel ingot is $v_{sl} \approx 3 \times 10^{-5}$ m/s.

It is useful to discuss RSP in terms of the equilibrium phase diagram of the system in question even though the process of rapid solidification leads to nonequilibrium solid products. Consider the solid solution and eutectic binary phase diagrams shown schematically in Fig. W21.26. Indicated in each diagram is the curve of T_0 versus composition, where T_0 is the temperature at which the liquid and solid phases of the same composition have the same Gibbs free energy. For the eutectic system shown in the middle, where the two solid phases have the same crystal structure, there is a smooth T_0 curve. In the right-hand phase diagram where the two solid phases have very limited mutual solid solubilities, the T_0 curves do not intersect. In all three cases shown in Fig. W21.26, the solid formed will have the same composition as the liquid when cooling is rapid enough so that solidification occurs at $T < T_0$. Under these conditions the solidification rate can exceed the diffusion rate in the liquid so that the components cannot redistribute themselves in the liquid phase. The *glass-transition temperature* T_g is shown in the right-hand phase diagram. In a glass-forming system where T_g is so low that it cannot be readily reached via rapid solidification, a dispersion of particles of a second phase can then occur in the primary matrix.

Metastable phases can also be formed when cooling rates are sufficiently high. In addition to the important example of the Fe–C system, where Fe_3C is a metastable product, a wide variety of interesting icosohedral metastable phases of Al with fivefold rotational symmetry (e.g., $\text{Al}_6\text{Mn}_{1-x}$, $\text{Al}_6\text{Mn}_{1-x}\text{Fe}_x$, $\text{Al}_{12}\text{Fe}_{1-x}\text{Mo}_x$, and $\text{Al}_{62}\text{Cu}_{26}\text{Fe}_{12}$) have been prepared via RSP. An RSP phase diagram using information obtained by heating the surfaces of Al-rich Al–Mn alloys with a scanned electron beam is presented in Fig. W21.27. Here the solid phases obtained for a range of scan (i.e., solidification) velocities v_{sl} from 0.001 to 1 m/s and for Mn concentrations from 0 to 30 wt % are shown. Icosohedral (fivefold symmetry) and decagonal (tenfold symmetry) phases in the form of dendrites in an Al-rich matrix are obtained for v_{sl} greater than about 0.02 m/s and for more than ≈ 18 wt % Mn. The solid-solution phase α -Al extends up to ≈ 14 wt % Mn for v_{sl} greater than about 0.03 m/s, well beyond the equilibrium eutectic

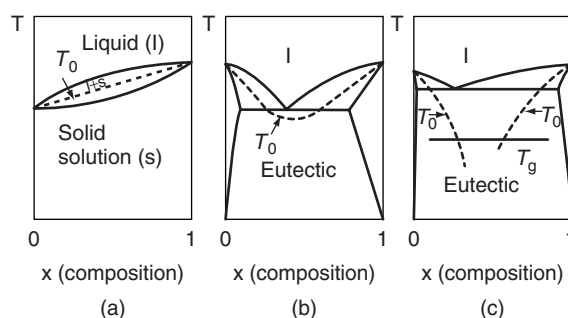


Figure W21.26. Solid-solution and eutectic binary phase diagrams are shown schematically, with the temperature T_0 at which the liquid and solid phases of the same composition have the same Gibbs free energy indicated.

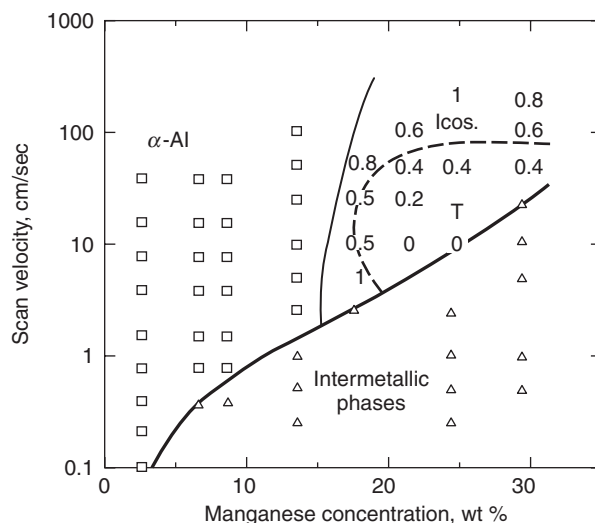


Figure W21.27. RSP phase diagram. The numbers indicate the relative fractions of the intermetallics that are icosahedral; from x-ray diffraction intensities. The region labeled T is a decagonal region. [From R. J. Schaefer et al., *Metall. Trans.*, **17A**, 2117 (1986).]

limit of 1.8 wt % at $T_e = 658^\circ\text{C}$. The possibility of obtaining metastable phases in Al–Mn alloys is enhanced due to the many different intermetallic compounds found in Al-rich alloys and also due to their relatively low growth velocities.

Despite the initial and continuing enthusiasm for the RSP technique, many of the hoped-for applications have not yet materialized, due in part, perhaps, to a lack of fundamental knowledge concerning the processes occurring during rapid solidification. It is, of course, an extremely difficult problem to control the microstructure, morphology, and stoichiometry of a rapidly solidified material under processing conditions that are so far from equilibrium. The consolidation of RSP-generated materials into useful forms without causing a degradation of their desirable as-synthesized properties has also proven to be difficult.

W21.13 Surface Treatments for Metals

Most pure metals are thermodynamically unstable with respect to oxidation and other environmental chemical reactions. As a result, a wide variety of physical and chemical processing procedures is used to modify the surface properties of metals in order to improve their corrosion resistance, wear resistance, and surface hardness. Some of these procedures have been mentioned in Chapter W12 and include electroplating, chemical reactions, vapor deposition, ion implantation, and thermal reactions. In addition, the electrolytic anodization of Al resulting in the formation of an oxide layer has been discussed in Section 19.11. Two additional surface-treatment procedures are discussed briefly here: surface carburizing and nitriding and the intense-pulsed-ion-beam (IPIB) surface treatment.

The *surface carburizing* and *nitriding* of metals are both processes that involve changing the chemical composition of the metal in a surface layer. They can be achieved using a variety of techniques for introducing C and N into the material.

Gas carburizing (in the austenite region near $T = 1000^{\circ}\text{C}$) and nitriding (in the ferrite region near $T = 500^{\circ}\text{C}$) of low-carbon steels typically involve heating the steel in C- or N-containing atmospheres (CH_4 or NH_3), which leads to the rapid diffusion of C or N atoms into the near-surface region known as the *case* (hence the use of the term *casehardening*). Other metallic substitutional alloying elements, such as Ni, Mn, and Cr, are not affected by this treatment, due to their much lower diffusivities in iron. Analogous processes known as *carbonitriding* (or *nitrocarburizing*) and *boronizing* can also be used for surface hardening.

The resulting spatial distribution of C in the steel depends on both the temperature and time of the carburizing process. The carbon concentration is given approximately by the solution of Fick's second law of diffusion [see Eq. (W6.2)]. Typical C concentrations obtained in the surface layer are ≈ 0.8 to 1 wt % (i.e., well below the solubility limit of C in austenite). As the steel is cooled from the carburizing temperature, the microstructure that develops varies with depth into the material due to the varying C concentration. Pearlite and cementite are formed at and just below the surface, then only pearlite when the C concentration has fallen to the eutectoid composition, followed by a mixture of pearlite and ferrite at greater depths. For most steels carburized for 5 to 10 h, the thickness of the carburized surface layer is from 0.5 to 2 mm.

Following the carburizing step, additional heat treatments known as casehardening are necessary to form precipitates of martensite, which result in the formation of a wear-resistant surface layer on the steel. This subsequent heat treatment usually takes place in the austenite phase near $T = 850^{\circ}\text{C}$ and is followed by rapid quenching to form martensite. A martensite tempering heat treatment is then carried out in the range $T = 150$ to 200°C to relieve stresses.

Surface nitriding procedures are ordinarily employed for steels containing the alloying elements Al, V, Cr, and Mo and result in surface layers which are harder than those which are obtained by carburizing. Nitriding is usually carried out in an NH_3 atmosphere and at lower temperatures, and therefore for longer times, than for the case of carburizing since the eutectoid temperature T_e in the Fe–N system is only $\approx 590^{\circ}\text{C}$. The possible microstructures appearing in the Fe–N system are more complicated than in the Fe–C system since more than one stable iron nitride (e.g., Fe_4N , Fe_3N , and Fe_2N) can exist in the nitrided surface layer, depending on the processing conditions. The relatively N-rich compound Fe_3N is typically found near the surface, while Fe_4N is found at a greater depth where the diffused N concentration is lower. In addition, precipitates of the nitrides of the alloying elements Al, V, Cr, and Mo are also found in the nitrided surface layer. As a result, the surface layer can be quite hard due to the dispersion-strengthening mechanism. In contrast to carburizing, no additional heat treatment is required to harden the nitrided surface layer.

In the case of surface hardening via carbonitriding or boronizing, carbonitrides and borides are formed instead of carbides or nitrides. The Fe_2B phase is preferred over the FeB phase because it is less brittle and also because the resulting casehardened surface is under compressive stress. Boronized layers on plain carbon steels are typically two or three times harder than carburized layers on the same steels.

The carburizing and nitriding of steels can also be carried out in $\text{CH}_4/\text{Ar}/\text{H}_2$ or CO_2 and NH_3 or N_2/H_2 plasmas, respectively, with the result that the necessary treatment times and temperatures can be greatly reduced. In addition, the plasma can clean the surface via sputtering, activate the chemical species so that they interact more readily with the surface to be hardened, and even heat the surface. Plasma nitriding is also used

to improve the surface hardness and wear resistance of Ti alloys containing Al and V. Four distinct layers can be found in the surface region following plasma nitriding at $T = 800^\circ\text{C}$ for 13 to 15 h: a 0.3 to 0.5- μm surface layer of FCC $\delta\text{-TiN}$, a 1.7 to 2- μm layer of tetragonal $\epsilon\text{-Ti}_2\text{N}$, a thin layer of Ti_2AlN , and then the diffusion zone containing nitrogen-stabilized $\alpha\text{-Ti}$. An alternative source of energy is employed in the laser nitriding of Fe and Ti in a N_2 atmosphere which leads to improved hardness and corrosion resistance.

The *intense-pulsed-ion-beam* (IPIB) surface treatment is a recently developed thermal process that causes rapid heating and melting of the surface layer of a metal, followed by extremely rapid cooling, $\approx 10^9$ K/s, of the layer. This procedure, which can be considered to be a type of rapid-solidification processing, results in nonequilibrium microstructures such as amorphous, metastable, or nanocrystalline layers in the surface region. Such surface layers on tool steels and high-temperature Ti alloys have greatly improved surface hardnesses and wear and corrosion resistances. The plasma-immersion ion-implantation (PIII) procedure used to implant dopant ions into semiconductors is also used to implant N into the surfaces of metals in order to improve wear resistance.

The intense pulsed ion beams are typically composed of H or heavier ions. A single ion pulse containing $\approx 10^{13}$ to 10^{14} ions/ cm^2 leads to the implantation of ionic species at the level of only $\approx 10^{-5}$ at % in the implanted surface region, which can be $\approx 10^2$ to 10^3cm^2 in area. The depth of the IPIB treatment can be ≈ 2 to 10 μm for H ions but a factor of 20 less than this for heavier ions. IPIB-induced shock waves due to the use of heavier ions such as N can lead to greatly improved mechanical and chemical properties to a depth of up to 100 μm .

As an example of the IPIB treatment, the surface cross section of a tool steel sample treated with a 40-ns-duration 10-J/ cm^2 pulsed beam of 0.5–1 MeV C and H ions is shown in Fig. W21.28. The treated depth is ≈ 5 μm . In this near-surface layer which originally consisted of ferrite and large cementite particles, the carbon has been dissolved into the Fe matrix during the melting. Following rapid resolidification of this region, 20-nm carbide grains have been observed.

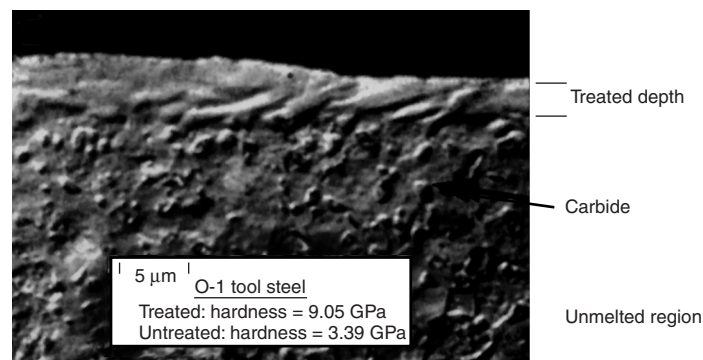


Figure W21.28. As an example of the intense-pulsed-ion beam (IPIB) treatment, the surface cross-section of a O1 tool steel sample treated with a 40-ns-duration 10-J/ cm^2 pulsed beam of 0.5- to 1-MeV C and H ions is shown. [From H. A. Davis et al. *Mater. Res. Soc. Bull.*, **21**(8), 58 (1996).]

W21.14 Chemical Vapor Deposition of Diamond

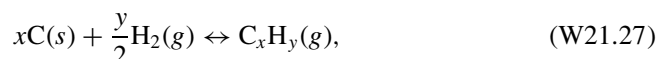
The synthesis of crystalline diamond films via CVD has become an important area of research over the last 15 to 20 years. The growth of diamond takes place either at atmospheric pressure (10^5 Pa), as in the case of the oxygen–acetylene or plasma torches, or at reduced pressures of about 10^3 to 10^4 Pa (7.6 to 76 torr) when microwave plasmas or hot filaments are used. The substrates employed are Si crystals, transition metals such as Mo and W, and ferrous-based materials such as tool steels. Substrate temperatures T_s are normally in the range 800 to 1100°C, although growth of diamond has been observed up to $\approx 1250^\circ\text{C}$ and down to $\approx 500^\circ\text{C}$. Graphite is deposited at higher T_s while amorphous carbon is deposited at lower T_s . Typical chemical compositions of the CVD environment as expressed by the ratios of the feedstock gas flow rates are $\text{H}_2/\text{CH}_4 \approx 100:1$ or $\text{H}_2/\text{CH}_4/\text{O}_2 \approx 100:4:0.4$ in the microwave plasma or the hot filament method and $\text{C}_2\text{H}_2/\text{O}_2 \approx 101:100$ (i.e., slightly carbon-rich) in the oxygen–acetylene torch.

An understanding of the growth of diamond under conditions where graphite is the thermodynamically stable form of carbon can be obtained by recognizing that the competing forms of solid carbon, graphite, and amorphous carbon have higher solubilities in the vapor phase relative to diamond in reactive environments containing large amounts of either atomic hydrogen or oxygen (or both). The thermodynamic *quasiequilibrium* (QE) model[†] has been applied to the carbon–hydrogen (C–H) and C–H–O systems to provide the basis for an analysis of the CVD of diamond. In this approach the dominant vapor species (H, C_xH_y , O) in equilibrium with either the diamond or graphite surfaces and also the deposition and etching rates of diamond or of graphite can be determined. When the kinetic effects associated with the enhanced etching of graphite by atomic hydrogen and oxygen are included in the model, regions in the CVD phase diagram of the C–H and C–H–O systems are predicted where diamond is the only stable form of solid carbon present.

The key assumption of the QE model is that *thermochemical equilibrium* exists between the solid carbon surface and the vapor species desorbed from it. Kinetic theory is employed to determine the rates at which vapor species arrive at and leave the carbon surface. The standard Gibbs free energies of formation $\Delta_f G^0(\text{C}_x\text{H}_y, T)$ of the vapor species are employed to obtain the needed equilibrium constants $K(\text{C}_x\text{H}_y, T)$ using the expression

$$K(\text{C}_x\text{H}_y, T) = \exp \left[-\frac{\Delta_f G^0(\text{C}_x\text{H}_y, T)}{RT} \right]. \quad (\text{W21.26})$$

These in turn provide the equilibrium vapor pressures of the $\text{C}_x\text{H}_y(g)$ species for the reactions



using

$$P_{\text{eq}}(\text{C}_x\text{H}_y, T) = K(\text{C}_x\text{H}_y, T)[P(\text{H}_2)]^{y/2}, \quad (\text{W21.28})$$

[†] J. C. Batty and R. E. Stickney, *J. Chem. Phys.*, **51**, 4475 (1969).

where $P(\text{H}_2)$ is the partial pressure of H_2 in the system. The pressures in this equation are expressed in atmospheres.

By requiring conservation of H atoms in the fluxes of atoms and molecules incident on and leaving either the diamond or the graphite surface, predictions for the evaporation rates $R_e(\text{C}_x\text{H}_y, T)$ can be obtained. Deposition rates are then obtained from

$$R_d(\text{C}) = I(\text{C}) - R_e(\text{C}), \quad (\text{W21.29})$$

where $I(\text{C})$ is the net flux of incident C atoms and $R_e(\text{C})$ is the net flux of C atoms leaving the surface. The evaporation rates $R_e(\text{C}_x\text{H}_y, T)$ and deposition rates R_d of diamond and graphite are presented as functions of temperature in Fig. W21.29 for a mixture of 1% CH_4 in H_2 at $P = 5 \times 10^3$ Pa. It can be seen that the evaporation rates of C_xH_y species are predicted to be higher above diamond (dashed curves) than above graphite (solid curves), as expected from the slightly higher free energy of formation of diamond relative to graphite. Under the conditions presented in Fig. W21.29, there exists an intermediate temperature range, from $T = 910$ to 2295 K, where diamond is stable relative to hydrogen. For $T < 910$ K diamond is etched via the formation of $\text{CH}_4(\text{g})$ while for $T > 2295$ K etching via the formation of $\text{C}_2\text{H}_2(\text{g})$ dominates.

The data presented in Fig. W21.29 can be used to construct the *CVD phase diagram* for the C–H system shown in Fig. W21.30. Here the regions of stability of solid carbon (i.e., diamond or graphite) are presented at 5×10^3 Pa as functions of temperature and reactant ratio $\text{C}/(\text{C} + \text{H})$. In this case there exists a region where diamond is predicted to be the only stable phase of solid carbon. This occurs because the phase boundary of graphite has been shifted to the right by taking into account the enhanced etching of graphite by atomic hydrogen. Experimental data points for the deposition of diamond

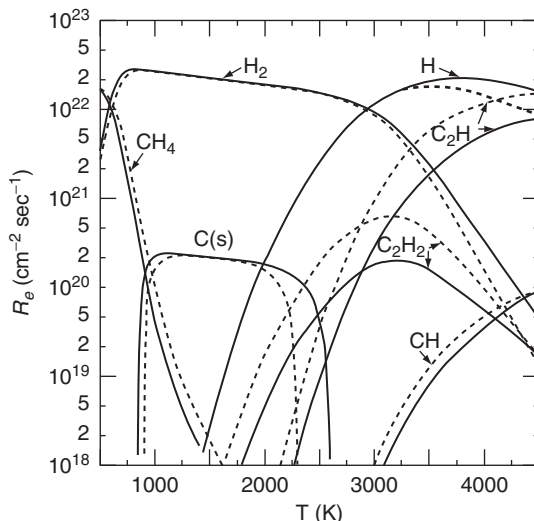


Figure W21.29. Predictions of the quasiequilibrium model for the evaporation rates $R_e(\text{C}_x\text{H}_y, T)$ of C_xH_y vapor species and the deposition rates $R_d(T)$ of either diamond or graphite are presented as functions of temperature for a mixture of 1% CH_4 in H_2 at $P = 5 \times 10^3$ Pa. [From M. Sommer and F. W. Smith, *High Temp. Sci.*, **27**, 173 (1989). Reprinted by permission of Humana Press, Inc.]

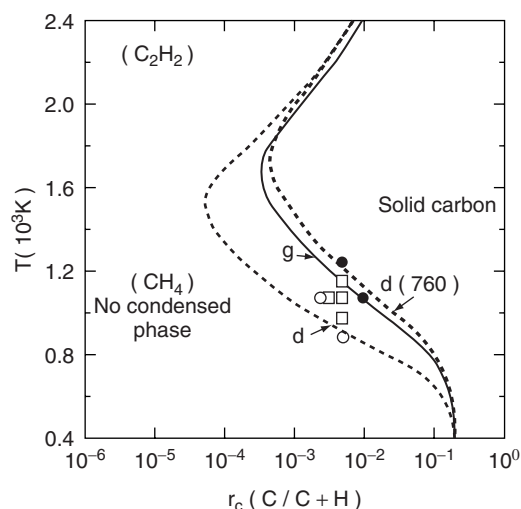


Figure W21.30. CVD phase diagram for the C–H system. The regions of stability of solid carbon (i.e., diamond or graphite) are presented at $P = 5 \times 10^3$ Pa as functions of temperature T and reactant ratio $C/(C + H)$. [Reprinted from M. Sommer, K. Mui, and F. W. Smith, *Solid State Commun.*, **69**, 775 (1989). Copyright 1989, with permission from Elsevier Science.]

are also presented and can be seen to be in very good agreement with the predicted region of stability of diamond. Similar predictions for the C–H–O system have been obtained for the deposition of diamond via the oxyacetylene torch.[†]

Problems remaining in the CVD of diamond films are related to obtaining films with fewer defects and with lower levels of nondiamond components, such as graphitic carbon, amorphous carbon, and impurities such as hydrogen and nitrogen. The successful preparation of *n*-type films is also an important goal for the eventual use of diamond as an active element in electronic devices. The *p*-type doping of diamond by substitutional B acceptors is well established.

In addition to the CVD of diamond films, the synthesis of diamond at high temperatures (≈ 2000 K) and pressures (≈ 60 atm) (i.e., under HPHT conditions) in the form of small single crystals or abrasive grains is a well-developed technology, with several tons of diamond being prepared yearly. Under these conditions, diamond is apparently thermodynamically stable with respect to graphite, although the phase boundary between diamond and graphite is still not very well known in the HPHT region. The HPHT method relies on the solubility of carbon in molten transition metals such as Ni at high T and P and its subsequent controlled precipitation as diamond crystals.

Cubic BN (c-BN) with the zincblende crystal structure is similar in many respects to diamond, having essentially the same lattice constant, a wide bandgap (≈ 6.4 eV) and also very high hardness and thermal conductivity. c-BN is actually superior to diamond for electronic applications due to the fact that it can be doped both *n*- and *p*-type with Si and Be, respectively. The ceramic c-BN also has excellent potential for use as a hard, wear-resistant coating for tools since its solubility in ferrous materials is much

[†] R. B. Wang, M. Sommer, and F. W. Smith, *J. Cryst. Growth*, **119**, 271 (1992).

lower than that of carbon. So far a successful technique for preparing single-phase c-BN in thin film or bulk form has not been developed.

W21.15 Synthesis of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$

Early methods of synthesizing the high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ (YBCO or 1:2:3) involved a solid-state self-flux reaction resulting in a metastable compound. Typically, a mixture of BaCO_3 , CuO , and Y_2O_3 with the molar ratios $\text{Y}/\text{Ba}/\text{Cu} = 1:4:10$ was mixed and ground in a zirconia crucible, pressed into pellets, and heated at 890°C for a day. The process was repeated a second time. Finally, the material was annealed at 1000°C while being subjected to flowing O_2 for three days. The cooling rates had to be slow to obtain crystals of size ≈ 1 mm. The parent compound is $\text{YBa}_2\text{Cu}_3\text{O}_7$, which is nonstoichiometric. This compound is enriched with oxygen as the O atoms intercalate into the crystal and order. The oxygen content of the crystals $(7 - x)$ was found to be a function of the oxygen partial pressure during annealing. Optimal values of T_c (≈ 90 K) were obtained for $x = 0.3$. To obtain crystal growth the temperature had to be sufficiently high to obtain a partial melt, yet sufficiently low so as not to decompose the crystals to more thermodynamically stable forms (such as Y_2BaCuO_5). A ternary phase diagram is given in Fig. W21.31.

The deposition of thin films of YBCO requires a different approach. Methods such as magnetron sputtering, pulsed excimer-laser ablation, and metal-organic chemical vapor deposition (MOCVD) have been developed. A proper choice of substrate has to be made so that epitaxial growth will occur. YBCO is an orthorhombic crystal with lattice constants a , b , and $c = 0.383$, 0.389 , and 1.169 nm. Suitable substrates for growing crystals with the c axis normal to the substrate surface are the (100) faces of SrTiO_3 ($a = 0.39$ nm) and LaAlO_3 ($a = 0.536$ nm $\approx a\sqrt{2}$). These substrates have high melting temperatures, $T_m = 2030^\circ\text{C}$ and 2110°C , respectively, and also have low microwave loss, which is important in designing superconducting microwave filters and cavities.

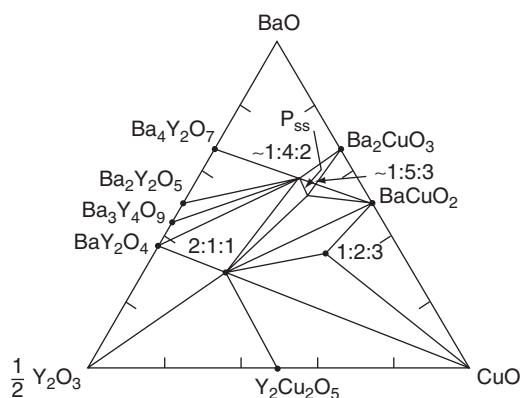


Figure W21.31. Ternary phase diagram for Y–Ba–Cu–O. The numbers $x:y:z$ refer to the Y:Ba:Cu stoichiometry. P_{ss} denotes a solid solution region. The temperature is $T \approx 975$ to 1000°C . (Adapted from L. F. Schneemeyer et al., Barium yttrium copper oxide crystals, in D. W. Murphy and L. V. Interrante, eds., *Inorganic Synthesis*, Vol. 30, Wiley, New York, 1995.)

In the MOCVD method the cations are bound to organic anions, and the resulting metal–organic precursor has a high vapor pressure at relatively low temperatures ($\approx 250^\circ\text{C}$). An inert carrier gas such as Ar is used. Precursors such as $\text{Ba}(\text{THD})_2$, $\text{Cu}(\text{THD})_2$, and $\text{Y}(\text{THD})_3$ are used, where THD is bis(2,2,6,6-tetramethyl-1,3,5-heptanedione). Oxygen is introduced. The vapor is transported to the heated substrate, where the organic components are pyrolyzed and the oxides of the cations are deposited. In this method the substrate is the hottest part of the system, so the deposition takes place only on the substrate, not on the walls of the system.

In pulsed laser deposition (PLD) a Kr–F excimer laser generates a $\lambda = 248$ nm pulse of duration 30 ns with a fluence of $\approx 5 \times 10^4$ J/m². The pulse is absorbed near the surface of a block of material that is to be deposited on a substrate some distance away. The pulse has sufficient energy to vaporize (and partially ionize) several hundred layers of atoms. A plume of ablated material is cast off primarily in the forward direction perpendicular to the target. The substrate temperature is in the range 500 to 700°C, which provides sufficient atomic mobility for crystal growth to occur. The deposited layer retains the chemical composition of the target. Unlike the MOCVD method, one is not dependent on all the precursors having a high vapor pressure.

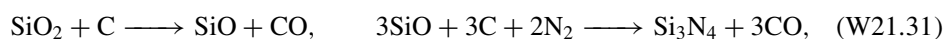
W21.16 Synthesis of Si_3N_4

There exist a variety of methods for synthesizing Si_3N_4 powders, but only three methods are used commercially. They are carbothermal reduction and nitridation, direct nitridation of silicon, and a liquid-phase process. Si_3N_4 exists in two phases, a low-temperature metastable α -phase and a high-temperature stable β -phase. Both phases have a hexagonal unit cell, but the stacking sequences of the planes along the c axis are different. The α -phase has the stacking sequence ABABABAB..., whereas the β -phase has the sequence ABCDABCD.... The α -phase can readily accommodate cations within its structure.

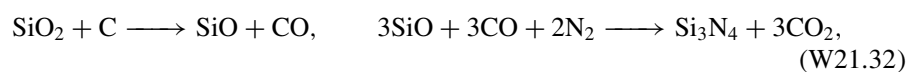
In the carbothermal reduction and nitridation process silica reacts with carbon in a nitrogen atmosphere according to the overall formula



The reaction occurs at temperatures in excess of 1420°C at atmospheric pressure and is endothermic with $\Delta H = 1270$ kJ/mol. To prevent the reverse reaction from occurring, the CO gas must be removed. Unless impurities are added the reaction strongly favors the production of the α -phase. The reaction proceeds in several stages. One possible route is to produce a supersaturated SiO gas and then have this gas react with carbon and nitrogen. Some possible pathways are



or

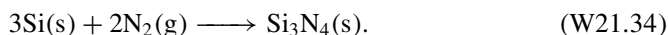


followed by



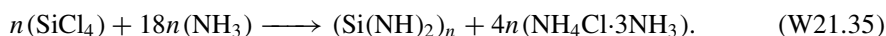
Other reactions are possible as well.

The direct nitridation of silicon involves “burning” solid silicon in a nitrogen atmosphere to promote the exothermic reaction

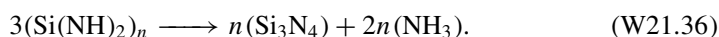


The enthalpy of formation is $\Delta_f H^\circ = -730$ kJ/mol at a temperature of 1320°C. This reaction produces a mixture of both the α -phase and the β -phase. It is possible to reaction-sinter the Si_3N_4 by slowly raising the temperature to $\approx 1400^\circ\text{C}$ and keeping it at that temperature for several days. Then the β -phase is produced.

The liquid-phase process involves pouring liquid silicon tetrachloride into liquid ammonia at $P = 230$ Pa and $T = -29^\circ\text{C}$. The SiCl_4 is dissolved in an organic solvent composed of cyclohexane and benzene. A polymer of silicon diimide is formed at the interface according to the reaction



The $(\text{Si}(\text{NH})_2)_n$ dissolves in the organic solvent. The solvent, ammonia, and NH_4Cl are then removed and the silicon diimide is heated to 1000°C to convert it to amorphous silicon nitride according to the reaction



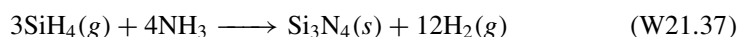
Further heating anneals the amorphous material to crystalline α - Si_3N_4 .

Densification of Si_3N_4 can be accomplished, for example, by liquid-phase sintering. The Si_3N_4 is mixed with silica and additives such as alumina and yttria which are used to lower the melting temperature of the silica. This is important because Si_3N_4 has a low dissociation temperature (1500°C) and it is desirable to keep the liquid temperature at around 1470°C . In the sintering process the silica and additives melt, some of the Si_3N_4 goes into solution and precipitates out, ultimately causing the grains to merge and to eliminate the intergranular void spaces. The microstructure that results is strongly influenced by the additives used.

It is also possible to densify Si_3N_4 powders by means of hot pressing, hot isostatic pressing, or gas pressure sintering. Oxides, such as MgO or Y_2O_3 , or BeSiN_2 are added as sintering aids.

It is possible to deposit Si_3N_4 films by means of CVD. The precursors are ammonia (NH_3) and dichlorosilane (SiCl_2H_2). The operating temperature is 700 to 800°C . Unfortunately, this is too high for application to electronic VLSI chips. PECVD is used to reduce the operating temperatures to below 450°C , in which case amorphous films also containing H are deposited.

Laser reactions may also be used to synthesize Si_3N_4 . A mixture of NH_3 and SiH_4 is irradiated with infrared radiation from a CO_2 laser. The SiH_4 is vibrationally excited and the net endothermic reaction



is able to proceed. Particles of size ≈ 20 to 100 nm are produced.

W21.17 Synthesis of SiC

At low temperatures ($T < 1800^{\circ}\text{C}$) one may sinter powders of Si and C to produce the β (zincblende) form of SiC via the reaction $\text{Si} + \text{C} \rightarrow \text{SiC}$. For temperatures higher than 2000°C , hot pressing may be used. Silicon carbide is most commonly synthesized using the Acheson process, which employs a resistance furnace. A mixture consisting of carbon, NaCl, SiO_2 sand, and some sawdust is placed around a core of graphite. An electrical current is passed through the graphite, heating it to a temperature of around 2600 to 2700°C (below the melting temperature of 2830°C). The reaction $\text{SiO}_2 + 3\text{C} \rightarrow \text{SiC} + 2\text{CO}$ is highly exothermic, with $\Delta H = 14,700 \text{ kJ/kg}$, and this helps create the high temperature. The α (wurtzite, high temperature) form of SiC grows around the graphite core. The NaCl helps to remove impurities from the material. The sawdust creates sufficient pore space so that the CO gas may escape.

The Lely process is another way of synthesizing SiC. Amorphous SiC granules are placed inside a hollow graphite tube and the combination is heated to $\approx 2500^{\circ}\text{C}$ in an inert gas such as Ar. Some of the SiC sublimates, forming a vapor. From this vapor SiC crystals nucleate on the granules and then continue to grow.

Silicon carbide may also be grown by chemical vapor deposition on a hot substrate. The temperatures are typically much cooler than used in the Acheson and Lely processes. The precursor gases that are used are silane, (SiH_4) and methane (CH_4) or propane (C_3H_8). Typical net reactions are $\text{SiH}_4 + \text{CH}_4 \rightarrow \text{SiC} + 4\text{H}_2$ or $3\text{SiH}_4 + \text{C}_3\text{H}_8 \rightarrow 3\text{SiC} + 10\text{H}_2$. Laser-induced reactions are also possible, such as $2\text{SiH}_4 + \text{C}_2\text{H}_4 \rightarrow 2\text{SiC} + 6\text{H}_2$. It is also possible to use single molecules called *carbosilanes*, containing Si and C in a 1:1 ratio, as the precursor. Included are molecules such as 1,3-disilacyclobutane. It is possible to produce β -SiC at temperatures $\approx 1000^{\circ}\text{C}$ and even lower. Other molecules in use include 1,3-disila-*n*-butane and methylsilane (CH_3SiH_3).

Silicon carbide powders may be formed into shapes using methods such as extrusion, injection molding, and hot isostatic pressing, among others. SiC may be sintered using the hot-pressing technique at temperatures in excess of 2000°C .

Of the various methods for preparing SiC, CVD produces the highest-quality crystals. For example, a thermal conductivity of $\kappa = 300 \text{ W/m}\cdot\text{K}$ at $T = 300 \text{ K}$ is attainable, compared with values in the range 15 to 120 for sintered SiC, 120 to 170 for reaction-bonded SiC, and 50 to 120 for hot-pressed SiC. CVD SiC also yields the material with the highest elastic modulus, $E = 466 \text{ GPa}$, and the lowest coefficient of thermal expansion, $2.0 \times 10^{-6} \text{ K}^{-1}$, at room temperature.

W21.18 Synthesis of the Zeolite ZSM-5

Sol-gel synthesis is also used to produce the zeolite ZSM-5, introduced in Section 13.6. This zeolite is an aluminosilicate in which the silicon-to-aluminum ratio is very high. A typical procedure is to first prepare NaAlO_2 by $\text{Al}_2\text{O}_3 + 2\text{NaOH} \rightarrow 2\text{NaAlO}_2 + \text{H}_2\text{O}$ and then put it in a solution of NaOH and H_2O . A second solution is prepared by dissolving a small amount of tetrapropylammonium bromide in H_2SO_4 and water. The solutions are combined with a sol consisting of silica, Na_2O , and water. The silica-to-alumina ratio can be kept high to make the resulting crystal almost entirely silica. The resulting solution is kept at 95°C for up to two weeks and the sol-gel reaction is monitored closely to see when crystallites of the zeolite form. When the crystallization is complete, the organic molecules can be slowly pyrolyzed in oxygen at elevated

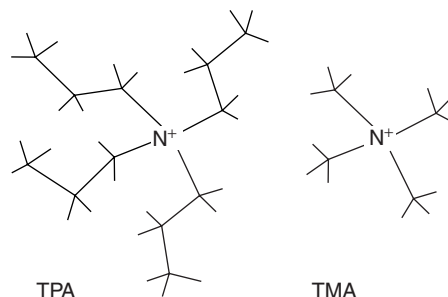


Figure W21.32. Tetrapropylammonium (TPA) ion and tetramethylammonium (TMA) ion. [Adapted from D. W. Lewis et al., *Nature*, **382**, 604(1996).]

temperatures. The resulting crystal consists mainly of tetrahedrally coordinated silica with aluminum ions incorporated into the framework. Sodium ions (equal in number to the Al ions for charge balance) reside outside the framework in the pore spaces. The zeolite serves as an ion exchanger, so other ions may be substituted for the sodium.

The tetrahedral molecule tetrapropylammonium (TPA) ion (Fig. W21.32) serves as a template molecule upon which the zeolite nucleates. The framework and pore size of the crystal are determined by the geometry of this ion. The steric hindrance presented by the ion guarantees a large pore size. Since the charge on the TPA ion is $+1e$, it serves to compensate for the valence deficit that occurs when an Al^{3+} ion replaces a Si^{4+} ion. Each of the TPA propyl groups extends into one of the four channels that emanate from each intersection, with the nitrogen atom residing at the junction. If other ions are used, the pore size will be different. This gives the chemist the opportunity to custom design zeolite structures based on the template molecule employed.

Recent observation of the formation and growth of a similar material, zeolite A ($Na_{12}[(AlO_2)_{12}(SiO_2)_{12}] \cdot 27H_2O$), identified the steps involved in the formation of crystals.[†] The monomers polymerized to form small amorphous clusters of aluminosilicate particles with diameters in the range 5 to 10 nm in solution. Tetramethylammonium (TMA) (see Fig. W21.32) is used as a template for zeolite A. When TMA is added to the solution, the solution becomes basic and the particles aggregate to form amorphous gel particles, with sizes in the range 40 to 80 nm. The aggregation is presumably due to the screening of the Coulomb repulsion between the particles by the ions in solution, allowing the long-range van der Waals forces to bring the particles together. After three days at room temperature, single crystals nucleate within the gel particles and grow to the size 10 to 30 nm. After a week there is complete conversion of the gel particles to the single crystals, of size 40 to 80 nm. Presumably the high supersaturation present in the amorphous gel particles is the driving force for the nucleation and growth of the crystals. If the temperature is then elevated to $80^\circ C$, there is transport through the solution and the crystals undergo Ostwald ripening. Larger crystals in the range 200 to 400 nm are formed within one day.

By using micelles as the templating agent it is possible to produce mesoporous films of transition metal oxides with variable pore sizes.[‡] The micelles are rodlike structures

[†] S. Mintova et al., *Science*, **283**, 958 (1999).

[‡] T. Sun and J. Y. Ying, *Nature*, **389**, 704 (1997).

that are self-assembled from hexylamine molecules, with the molecular axes directed perpendicular to the rod axis. The length of the hydrocarbon chain determines the radius of the rod.

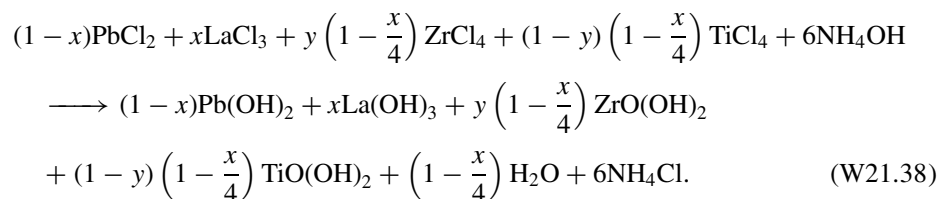
It has been reported[†] that zeolite-like materials with helical pores could be constructed by using inorganic cations or amines as templates. In place of the aluminosilicate structure of ZSM-5, the inorganic framework is based on zinc and beryllium arsenate and gallium germanate. Such structures could serve as chiral catalysts that would yield products with enantioselectivity (i.e., with a fixed handedness).

W21.19 Synthesis of the Perovskite PLZT

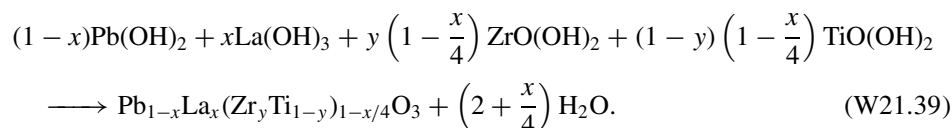
Lead zirconate titanate, $\text{Pb}(\text{Zr,Ti})\text{O}_3$ (PZT), is a ceramic material used in the electronics and optics industries. There are a number of ways of synthesizing it. Powders can be made by the solid-state reaction method or by wet chemical synthesis.

In the solid-state reaction method the reagent powders PbO , ZrO_2 , and TiO_2 are mixed, ground, and then heated at 850°C for about 3 hours, during which time crystallization takes place. To create the daughter compound $\text{Pb}_{1-x}\text{La}_x(\text{Zr}_y\text{Ti}_{1-y})_{1-x/4}\text{O}_3$ (PLZT) one uses a combination of La_2O_3 and ZrO_2 instead of pure zirconia.

Wet chemical methods include coprecipitation, hydrothermal synthesis, and sol-gel synthesis. An example of the coprecipitation method is to mix various salts together with ammonium hydroxide and water, for example,



The various hydroxides form a gel precipitate. The solution is washed to eliminate the ammonium chloride salt. It is then heated at a temperature of 550°C for an hour during which time the hydroxide groups are converted to water and the PLZT crystals are formed through the reaction

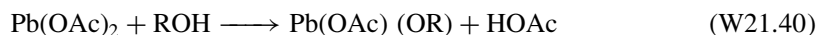


Hydrothermal synthesis allows the reaction to occur at lower temperatures (350°C), but at higher pressures.

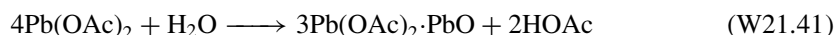
The sol-gel synthesis of PLZT utilizes precursors typically consisting of metal salts (lead acetate hydrate and lanthanum acetate hydrate) and alkoxides (zirconium *n*-propoxide and titanium isopropoxide). Salts are used because the alkoxides of lead and

[†] T. E. Gier et al., *Nature*, **395**, 154 (1998).

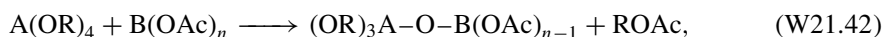
lanthanum are not soluble. The acetates are added to an alcohol such as methoxyethanol in water. Reactions such as



or



occur, with corresponding ones for La(OAc)_3 . Here $\text{R} = \text{C}_2\text{H}_4\text{OCH}_3$ and it is seen that the reaction replaces the OAc ion by an OR ion. Typical condensation reactions that can occur are



where $n = 2$ or 3 , $\text{A} = \text{Ti}$ or Zr , $\text{B} = \text{Pb}$ or La , and $\text{R} = \text{C}_3\text{H}_7$ or $\text{C}_2\text{H}_4\text{OCH}_3$.

Thin films of PLZT created by the sol-gel process may be spun onto silica or MgO substrates while still wet and then dried. The films may be processed further for various applications.

W21.20 Synthesis of Glasses: Pilkington Process

The synthesis of glass involves essentially three steps. In the first step a batch of raw materials is prepared. The principal ingredient is SiO_2 . Modifier oxides, such as Na_2O or K_2O , are added to lower the melting temperature. Other oxides, such as CaO , are added to provide chemical stability. If a glass-ceramic with controlled crystallinity is to be produced, Al_2O_3 is also added.

In the second stage the mixture is melted. For common glasses the temperature is elevated to 1300 to 1400°C , while for glass-ceramics the temperature range is 1400 to 1500°C . Volatile gases leave the liquid. The viscosity of the liquid decreases rapidly with increasing temperature, so the rate of escape of the gas bubbles is sensitive to temperature.

The final stage involves forming the glass into the desired shape. Techniques such as rolling, blowing, casting, pressing, and drawing are used. In creating ordinary glass the cooling rate is as fast as it can be without producing cracking. If it is too high, the temperature differential between the surface and interior portions of the glass produces stress fields that could lead to cracking. In creating glass-ceramics, slower cooling is required. The cooling rate is critical in determining the amount of crystallization that will occur. Residual stresses may be eliminated or reduced by annealing the glass.

In some cases the surface of the glass is tempered to enhance its mechanical properties. For example, one may heat the glass uniformly in a furnace, remove it, and then rapidly cool the outer surface. Due to the poor thermal conductivity of glass, the interior remains hot for some time. Viscoelastic relaxation allows the atoms in the interior to assume new configurations to relieve the stress. Upon further cooling the glass becomes so viscous that relaxation no longer can occur and the interior develops a tensile stress in response to the thermal contraction. Correspondingly, the surface region is put in a state of compressive stress. The existence of the internal stress field permits the glass to withstand larger flexural stresses that may be imposed on it.

In addition to thermal tempering, chemical tempering is also possible. For example, by exchanging the Na^+ ions for smaller Li^+ ions near the surface, the surface is placed

under compressive stress. It is also possible to remove Na by exposing the glass to gases such as SO_2 and H_2O . Effectively, the Na^+ ions are replaced by protons from the water, producing OH radicals.

The *Pilkington process*, or *float process*, represents more of a manufacturing process than a change in the microstructure of a material. It is included here because it shows the importance of surface tension in a practical manufacturing setting. The process provides an economical means for large-scale production of finished sheets of glass. Molten glass is poured onto a tub of molten metal and is allowed to float until it cools below the glass temperature, T_g . If T_m is the melting temperature of the metal, then if $T_m < T < T_g$, the solidified glass that forms will float on the molten metal and may readily be removed. Tin is usually used as the metal because it melts at a sufficiently low temperature ($T_m = 232^\circ\text{C}$).

Let ρ_m and ρ_g be the densities of the metal and glass. For flotation it is required that $\rho_m > \rho_g$. For tin and glass the specific gravities are 6.5 and 2.2, respectively. The interfacial surface tensions are denoted by γ_{mv} , γ_{gv} , and γ_{mg} , where the subscript v refers to the surrounding atmosphere (without oxygen). The thickness of the glass sheet will be denoted by t and its base area by A . The base area of the vat of metal is A' . The geometry is depicted in Fig. W21.33.

To find t one minimizes the total potential energy, consisting of gravitational and surface contributions,

$$U = \frac{1}{2}\rho_m g[A(y-h)^2 + (A' - A)y^2] + \rho_g gAt \left(y - h + \frac{t}{2}\right) + \gamma_{gv}A + \gamma_{mg}A + \gamma_{mv}(A' - A), \quad (\text{W21.43})$$

subject to the constraints of constant glass and metal volumes

$$V_m = A(y-h) + (A' - A)y, \quad (\text{W21.44})$$

$$V_g = At. \quad (\text{W21.45})$$

The surface energy associated with the vertical sides of the slab is small and is neglected. Introducing Lagrange multipliers λ and μ , one has

$$\delta(U - \mu V_m - \lambda V_g) = 0. \quad (\text{W21.46})$$

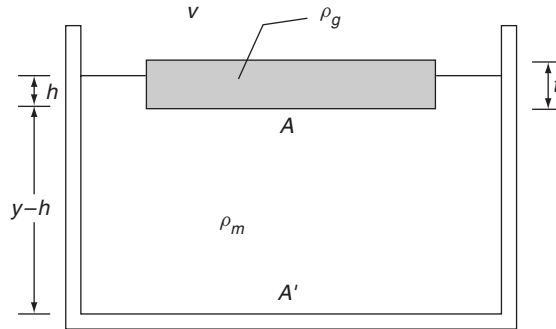


Figure W21.33. Slab of molten glass floating on a bath of molten metal in the Pilkington process.

The partial derivatives are taken independently with respect to the variables y , h , t , and A to obtain the four equations

$$\rho_m g(A'y - Ah) + \rho_g Agt - \mu A' = 0, \quad (\text{W21.47})$$

$$\rho_m g(y - h) + \rho_g gt = \mu, \quad (\text{W21.48})$$

$$\rho_g g(y - h + t) = \lambda, \quad (\text{W21.49})$$

$$\frac{1}{2}\rho_m g(h^2 - 2hy) + \rho_g gt \left(y - h + \frac{t}{2}\right) + \gamma_{gv} + \gamma_{gm} - \gamma_{mv} + \mu h - \lambda t = 0. \quad (\text{W21.50})$$

Eliminating the Lagrange multipliers results in

$$\rho_m h = \rho_g t, \quad (\text{W21.51})$$

which could have been deduced from Archimedes principle, and a formula for t ,

$$t = \sqrt{\frac{2\rho_m(\gamma_{gv} + \gamma_{gm} - \gamma_{mv})}{g\rho_g(\rho_m - \rho_g)}}, \quad (\text{W21.52})$$

independent of the volume of the glass. Note that it is necessary for $\gamma_{gv} + \gamma_{gm} > \gamma_{mv}$; otherwise, the glass would spread, with $A \rightarrow A'$. Since the interfacial surface tensions are dependent on T , one has some control over the thickness of the sheet by varying the temperature and the cooling rates.

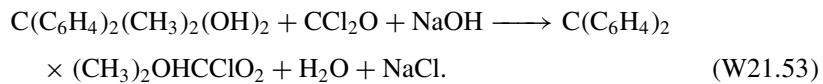
By applying a tensile stress to the sheet of glass while it is cooling, it is possible to stretch it and thereby make it thinner, according to the relation $t = V_g/A$. Let a constant external tension per unit thickness $\tau = S/t$ be introduced, where S is the tension. This is equivalent to adding the term $-\tau A$ to the potential energy, so replacing the term γ_{gv} by $\gamma_{gv} - \tau$ gives the result

$$t = \sqrt{\frac{2\rho_m(\gamma_{gv} + \gamma_{gm} - \gamma_{mv} - \tau)}{g\rho_g(\rho_m - \rho_g)}}. \quad (\text{W21.52}')$$

As the parameter τ increases, the thickness t decreases.

W21.21 Synthesis of Polycarbonate

Polycarbonate is synthesized by means of a polymerization reaction that occurs at the interface between two immiscible liquids. One liquid is an organic solvent (such as methylene chloride, CH_2Cl_2) and the other is a basic solution (such as NaOH in water) which acts as the initiator for the reaction. The starting material for the monomer from which the polymer is built is bisphenol-A, $\text{C}(\text{C}_6\text{H}_4)_2(\text{CH}_3)_2(\text{OH})_2$, and is soluble in the organic solvent. The organic solvent is dispersed into small globules in the alkaline solution. Phosgene gas, CCl_2O , is bubbled through the emulsion. The primary reaction is



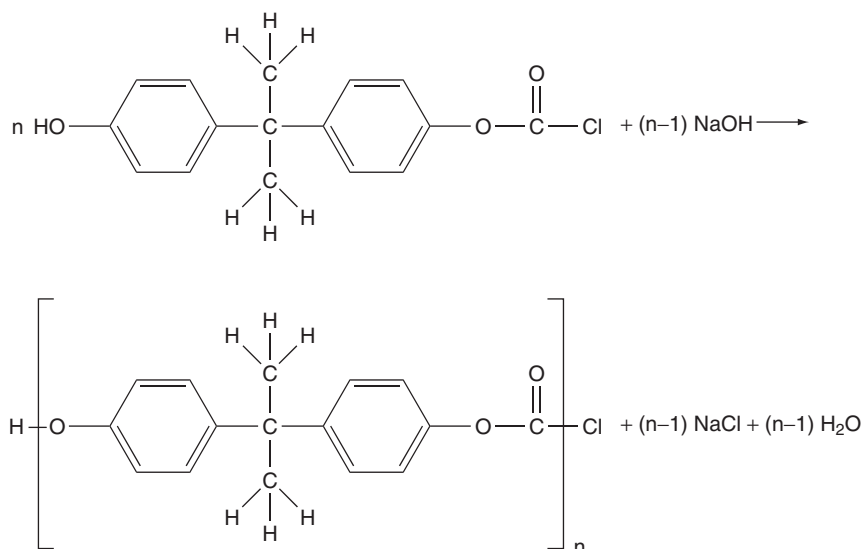
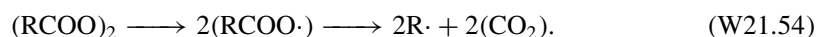


Figure W21.34. Synthesis of polycarbonate.

Successive reactions form the polymer, as indicated in Fig. W21.34. The polycarbonate that is formed is soluble in the organic solvent. Catalysts used to accelerate the reaction include triethylbenzyl ammonium chloride and various tertiary amines. The size (n) of the monomer is regulated by adding a monofunctional phenol such as *p*-*tert*-butyl phenol, which serves to terminate the chain. Since there is only one functional group on this terminator, it is utilized to attach to the polymer and to stop the polymer growth. The solvent globule size also serves to limit the size of the polymer. A typical synthesis sequence might include an initiation and reaction stage, followed by removal of the aqueous solution, removing the solvent (and recycling it), drying the product, and finally, extruding the polycarbonate residue.

W21.22 Synthesis of Polystyrene

Several methods are available for the synthesis of polystyrene (see Fig. 14.1). One may polymerize it in a solution, suspension, or in bulk. The monomer is styrene ($\text{C}_6\text{H}_5\text{CHCH}_2$), which will be denoted by M. In free-radical polymerization a free-radical initiator such as benzoyl peroxide $[(\text{RCOO})_2]$, with $\text{R} = \text{C}_6\text{H}_5$ is added. The initiator thermally dissociates into two free radicals according to the reaction



The free radicals are like ordinary molecules but have one unpaired electron. The free radical combines with the monomer to create a longer radical:



This in turn can link up to another monomer in a chain-growth process to create a still longer radical:



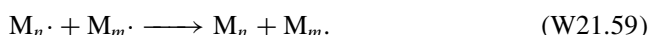
and the polymer continues to grow, one monomer at a time. After n steps one has



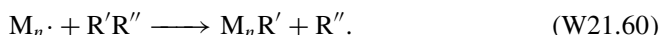
The growth eventually stops by a competing termination reaction. Two such reactions are chain coupling and disproportionation. In the chain-coupling reaction, two polymeric radicals combine, the electron spins become paired, and the growth is terminated:



In the disproportionation reaction a hydrogen atom is transferred from one radical to another. The net result is



This could involve changes of bond order within the resulting molecules. Alternatively, termination could be induced through a transfer reaction by introducing a molecule of the generic form $R'R''$ with an acceptor group R' :



Typical chain-transfer molecules include ethylbenzene, CCl_4 , and CBr_4 .

To prevent spontaneous polymerization, styrene is introduced into the processing environment with a retardant, such as benzoquinone. This retardant is stripped away before the styrene is introduced into the reactor. Also introduced are the initiator and the transfer molecule. After the polymerization is complete, the volatile chemicals are distilled out and the polymer is fed to an extruder, which produces the end product—pellets.

W21.23 Synthesis of Electro-active Polymers

Included among the electro-active polymers are *cis*- and *trans*-polyacetylene $[(\text{CH})_x]$, polythiophene (PT), polyaniline (PAN), polypyrrole (PPY), poly(*para*-pyridine) (PPPy), poly(*para*-phenylene) (PPP), and its decorated variant poly(2-methoxy-5-(2'-ethylhexyloxy)-1,4-phenylene vinylene) (MEH-PPP) (Fig. W21.35). They tend to have delocalized π electrons distributed over the polymer and filling the valence band. The conduction band consists of the antibonding π^* orbitals and is empty. The polymers are therefore one-dimensional semiconductors.

The electrical activity comes about when the polymers are doped. Unlike the case of solid-state semiconductors where the dopant atoms are introduced into the lattice, these polymers are doped by placing atoms into the space between polymer chains. If the dopant is a donor, it donates an electron to the π^* conduction band. The polymer is thereby reduced. Typical elemental donor atoms are the alkalis Li, Na, and K, and the alkaline earth Ca. If the dopant is an acceptor, it creates a hole in the π valence band.

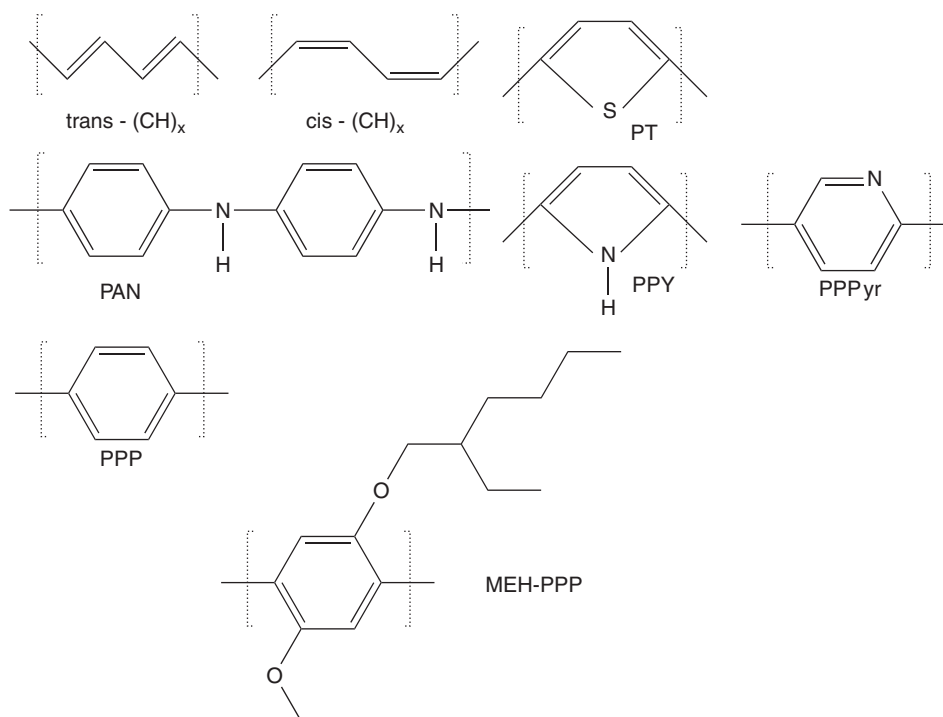


Figure W21.35. Some electroactive polymers.

The polymer is thereby oxidized. Typical elemental acceptors include the halogens Cl and I.

The doping may be introduced simply by exposing the polymer to a vapor or solution containing the dopant and allowing the dopant to diffuse into the bulk. Alternatively, electrochemical injection could be used. A “battery” is constructed with one electrode being the polymer and the other electrode containing the dopant. For example, to dope *t*-(CH)_x *p*-type one takes one electrode to be *t*-(CH)_x and the other electrode to be Li. An electrolyte such as LiClO₄ dissolved in propylene carbonate could be used. One proceeds to charge this battery. The Li electrode is biased sufficiently negative to reduce Li⁺ ions. The battery pumps *n* electrons from the (CH)_x, leaving it as positively charged (CH)_x^{*n*+}, thereby oxidizing it. For each polymer molecule *n* ClO₄[−] ions drift through the electrolyte over to the (CH)_x^{*n*+} and diffuse into it to preserve charge neutrality. The reactions for *p*-doping may be summarized as $n\text{Li}^+ + ne^- \rightarrow n\text{Li}$ and $(\text{CH})_x \rightarrow (\text{CH})_x^{n+} + ne^-$.

Light-emitting diodes can be fabricated by placing a film of polymer between two electrodes. The electrodes are chosen so that the anode Fermi level lies slightly above the π valence band of the polymer and the cathode Fermi level lies below the π^* conduction band. Schottky barriers are formed. Application of a forward bias then allows electrons to tunnel via Fowler–Nordheim tunneling from the cathode into the conduction band. Similarly, electrons tunnel from the valence band into the anode, leaving behind holes. The electrons and holes drift in opposite directions and radiatively combine in the interior of the film. A typical polymer that is used is MEH-PPV. Indium

tin oxide (ITO) is used as an anode because of its transparency to light. A typical cathodic material is Ca. A pattern of ITO anodes can be deposited on a substrate using photolithographic techniques. The polymer is then spin cast onto the device. Finally, the cathodic metal can be deposited from a vapor.

In light-emitting diodes the color of the light is determined by the bandgap. By using other polymers or adding side groups onto the polymer, one may tune this bandgap to other values. For example, MEH-PPV emits red light, whereas its mother, PPV, emits green light. The polymers PVK [poly(*N*-vinylcarbazole)] and PPP emit blue light.

W21.24 Spin Coating

The rheology of polymers puts severe constraints on possible processing methods. Above the glass-transition temperature, T_g , the polymer is a viscoelastic fluid that cannot simply be poured into molds. Of course, the temperature could be elevated to reduce the viscosity, but this also entails the risk of breaking the polymer chains or thermally inducing other unwanted chemical reactions. The methods used to form the materials into useful shapes involve various forms of extrusion, injection molding, blow molding, compression molding, coating, fiber spinning, thermoforming, and calendering. These manufacturing processes are not considered here. Instead, the focus of attention will be on the technique of spin coating, which has proven to be valuable in fabricating microelectronic circuits.

For a simple Newtonian fluid the stress and the strain rate are proportional to each other. To obtain a high strain rate, one must apply a large stress. If one wished to push the fluid through a die, one could do so by maintaining a pressure differential ΔP across its ends. For a die in the shape of a circular pipe of radius R and length Δz , Poisseuille's law gives the flow rate Q (volume/time) as

$$Q = -\frac{\pi R^4 \Delta P}{8\eta \Delta z}, \quad (\text{W21.61})$$

where η is the viscosity. Clearly, a high-pressure differential favors a rapid throughput of material.

For polymers above T_g , the stress and the strain rates are not linearly related. An empirical formula is

$$\sigma = \lambda \left(-\frac{\partial u}{\partial r} \right)^n, \quad (\text{W21.62})$$

where u is the axial velocity component, r the radial distance from the axis, and λ a constant depending on the polymer. The exponent n can be as small as 0.1. One may regard this as a formula defining a strain-rate-dependent viscosity $\eta = \lambda(-\partial u/\partial r)^{n-1}$. As the strain rate increases, the polymer becomes less viscous. By balancing forces on a coaxial cylindrical section of length Δz and radius r , one finds that $\sigma = -r\Delta P/(2\Delta z)$. Combining this with the previous formula gives

$$u(r) = -\frac{1}{2\lambda} \frac{\Delta P}{\Delta z} \frac{n}{n+1} (R^{1+1/n} - r^{1+1/n}). \quad (\text{W21.63})$$

The flow rate is found to be

$$Q = \frac{n\pi}{3n+1} \left(-\frac{1}{2\lambda} \frac{\Delta P}{\Delta z} \right)^{1/n} R^{3+1/n}. \quad (\text{W21.64})$$

For $n = 1$ this reduces to the Poisseuille law, with $\lambda = n$. For $n \neq 1$ it is a suitable generalization.

In the spin-coating process the substrate is mounted on a rotating turntable that is spun at angular velocity ω (Fig. W21.36). A polymer is poured on the axial region and it flows radially outward as it revolves around the axis. The centripetal force on any volume of polymer is due to the viscous force.

The radial component of the velocity is determined by solving the equation

$$\frac{\partial}{\partial z} \left[\eta \left(\frac{\partial v_r}{\partial z} \right) \right] = -r\rho\omega^2, \quad (\text{W21.65})$$

where v_r is zero at the surface of the substrate ($z = 0$) and reaches a maximum value at the surface of the coating. Integration leads to

$$v_r(r, z) = \frac{n}{n+1} \frac{r\omega^2\rho}{\lambda} [h^{1+1/n} - (h-z)^{1+1/n}], \quad (\text{W21.66})$$

where h is the thickness of the coating. The volume flow rate, Q , is given by

$$Q(r) = \int_0^h v_r 2\pi r dz = \frac{n}{2n+1} \frac{2\pi r^2 \omega^2 \rho}{\lambda} h^{2+1/n}, \quad (\text{W21.67})$$

so

$$h(r) = \left(\frac{2n+1}{n} \frac{\lambda Q}{2\pi \rho r^2 \omega^2} \right)^{n/(2n+1)}. \quad (\text{W21.68})$$

When the spinning stops, the surface tension will flatten the surface so as to minimize the surface energy. The time scale for this relaxation may be expressed using dimensional analysis in the form $\tau = \eta D / \sigma$, where D is a characteristic dimensionless scale length formed from R and the volume of the film.

Suppose that there is a uniform coating of thickness h_0 at time $t = 0$. Allow the spinning to occur so that an excess of polymer will flow over the edge. Apply continuity

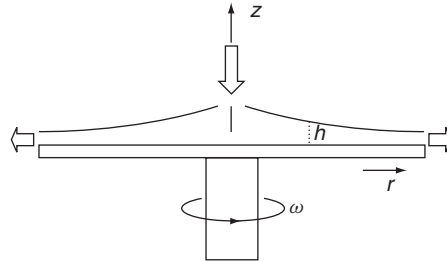


Figure W21.36. Spin-coating process.

concepts to develop an equation for h as a function of time. Consider the rate of change of the volume above the annulus lying between r and $r + dr$:

$$\frac{\partial(dV)}{\partial t} = 2\pi r dr \frac{\partial h}{\partial t}. \quad (\text{W21.69})$$

This change would be caused by a net flux into the cylinder. The flux coming from the inside is

$$\Phi(r) = \int_0^h v_r(r, z) 2\pi r dz = 2\pi r^2 \frac{n}{2n+1} \frac{\omega^2 \rho}{\lambda} h^{(2n+1)/n}. \quad (\text{W21.70})$$

The continuity equation states that

$$\Phi(r) - \Phi(r + dr) = \frac{\partial(dV)}{\partial t}, \quad (\text{W21.71})$$

so

$$\frac{\partial h}{\partial t} = -\frac{n}{2n+1} \frac{\omega^2 \rho}{\lambda} \frac{1}{r} \frac{\partial}{\partial r} (r^2 h^{(2n+1)/n}). \quad (\text{W21.72})$$

A solution of this equation $h(t)$, independent of r , is

$$h(t) = \left(h_0^{-(n+1)/n} + \frac{2n+2}{2n+1} \frac{\rho \omega^2}{\lambda} t \right)^{-n/(n+1)}, \quad (\text{W21.73})$$

where $h(0) = h_0$. In the case of a Newtonian fluid, $n = 1$ and $\lambda = \eta$. The volume flow per unit time over the outer rim of the disk, of radius R , is

$$\Phi(R) = \int_0^h v_r(R, z) 2\pi R dz = 2\pi R^2 \frac{n}{2n+1} \frac{\omega^2 \rho}{\lambda} h^{(2n+1)/n}. \quad (\text{W21.74})$$

W21.25 Microwave and Plasma Processing of Polymers

Microwave Processing. Often, it is advantageous to heat a material to remove defects, alter the crystallinity, control the morphology, promote drying, accelerate a polymerization reaction, and so on. The problem with polymers, however, is that they tend to be poor thermal conductors. The temperature that needs to be reached is usually close to the ceiling temperature T_c , the temperature at which there is irreversible chemical damage to the polymer. Heating the surface of the material leads to thermal gradients, so the heating rate must be made very slow to avoid having parts of the material with $T > T_c$. Microwaves offer the advantage of being able to heat a sample uniformly, at least over a skin depth, making it possible to achieve much more rapid heating rates.

Microwave radiation couples to molecular groups on the polymer that possess electric-dipole moments, such as OH, CN, or NH₂. Once local vibrations, librations, or rotations of these groups are excited, they dissipate the energy to the rest of the polymer chain by radiating one-dimensional phonons. The radiation also couples to ions and produces ac currents. The moving ions collide with the polymer chain, also creating phonons. In the liquid phase the dipole moments are able to reorient themselves readily, and this gives rise to strong coupling to the microwave field. As the temperature is

lowered and the polymer becomes more rubbery, the groups lose some of their flexibility and the absorption is reduced. If the liquid is cooled below the glass-transition temperature T_g , more of the degrees of freedom are removed and the dipolar coupling becomes still less. In the crystalline phase the steric hindrance becomes even more severe and the coupling drops even further. Often, additives are added to enhance the dipolar absorption during processing. These include metallic particles, carbon black, or carbon fibers. If conducting particles are added and these percolate throughout the material, eddy currents may be established by the microwave fields. These currents produce ohmic heating.

The skin depth δ is given in terms of the ac conductivity σ by the formula $\delta = 1/\sqrt{2\omega\mu_0\sigma}$. The value of σ is determined by the concentration of carbon black, so the value of δ may be chosen so it is larger than the sample thickness. Within the sample the intensity of the radiation falls off according to Beer's law, $I(z) = I_0 \exp(-z/\delta)$. In the limit of a material of small thickness, obeying $d \ll \delta$, the power absorbed per unit area is given by $P/A = \sigma|E_0|^2 d/2$, where E_0 is the electric field in the medium.

An important application of microwave radiation is in curing the polyimide thin films used in the design of electronic chips. Microwaves may also be used to ensure a uniformity of epoxy resins that are thermoset. In composite materials uniform curing is important to improve the fracture properties. Thus microwave radiation provides a valuable processing tool for polymers.

Plasma Processing. Whereas microwaves are most useful in processing the bulk of polymers, plasma processing is used to modify the surfaces of polymers. The plasma may be used to clean the polymer surface. It may be used selectively to remove unwanted polymers of low molecular weight (oligomers). It is often used to remove photoresist from electronic chips as part of the lithography process. It may serve to etch new topographic features into the surface or to sputter away old ones. It may be employed to alter the surface chemically, such as by adding new chemical groups to the polymer. It may also be used to deposit protective coatings or coatings on the surface that modify its physical properties.

The reason for the vulnerability of small molecules to plasma etching has to do with their inability to dissipate energy via phonons. The longer-chain molecules may conduct thermal energy along their chains to the interior of the polymer. The shorter chains are not able to do so. Instead, the vibrational temperature of the oligomers is increased by exposure to the plasma until they dissociate. Frequently, gaseous products such as CH_4 and H_2 or free radicals such as CH_2 are produced. For the sputtering process Ar^+ plasmas are used. Since Ar is inert, it does not contaminate the underlying surface by chemically bonding to it.

Surfaces may be reactively etched by using plasmas with ions such as O_2^+ , SF_6^+ , or CF_4^+ . The microstructure of the surface may be suitably changed. One may roughen a surface to increase its surface area so that a coating will adhere to it better.

At times the surface is chemically altered. It is first exposed to process gases that adsorb on the surface and create free radicals. Gases such as O_2 , N_2 , NH_3 , and CF_4 are used. Subsequent exposure to the plasma provides the activation energy that permits the polymer chain to acquire new functional groups.

An example of a coating that may be deposited using plasmas is PMMA. The polymerization reaction, in which gaseous methylmethacrylate molecules are polymerized

into long chains of polymethylmetacrylate (PMMA), could be initiated by exposure to a plasma.

REFERENCES

Synthesis and Processing Procedures

- Venables, J. A., in C. B. Duke, ed., *Surface Science: The First Thirty Years*, North-Holland, Amsterdam, 1994, p. 798.
- Venables, J. A., G. D. T. Spiller, and M. Hanbucken, *Rep. Prog. Phys.*, **47**, 399 (1984).
- Voorhoeve, R. J. H., Molecular beam deposition of solids on surfaces: ultrathin films, in N. B. Hannay, ed., *Treatise in Solid State Chemistry*, Vol. 6A, Plenum Press, New York, 1976.

Synthesis and Processing of Semiconductors

- Lieberman, M. A., and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing*, Wiley, New York, 1994.
- Maly, W., *Atlas of IC Technologies: An Introduction to VLSI Processes*, Benjamin-Cummings, Menlo Park, Calif., 1987.
- Meyerson, B. S., Low-temperature Si and Si:Ge ultrahigh-vacuum/chemical vapor deposition: process fundamentals, *IBM J. Res. Dev.*, **14**, 806 (1990).
- Pamplin, B. R., ed., *Molecular Beam Epitaxy*, Pergamon Press, Oxford, 1980.
- Shimura, F., ed., *Oxygen in Silicon*, Vol. 42 of R. K. Willardson, A. C. Beer, and E. R. Weber, eds., *Semiconductors and Semimetals*, Academic Press, San Diego, Calif., 1994.
- Stringfellow, G. B., *Organometallic Vapor-Phase Epitaxy: Theory and Practice*, Academic Press, San Diego, Calif., 1989.
- Weissler, G. L., and R. W. Carlson, eds., *Vacuum Physics and Technology*, Vol. 14 of *Methods of Experimental Physics*, Academic Press, San Diego, Calif., 1979.
- Wolf, S., and R. N. Tauber, *Silicon Processing for the VLSI Era*, Vol. 1, *Process Technology*; S. Wolf, *ibid.*, Vol. 2, *Process Integration*, Lattice Press, Sunset Beach, Calif., 1986 (Vol. 1), 1990 (Vol. 2).

Synthesis and Processing of Metals

- Honeycombe, R. W. K., and H. K. D. H. Bhadeshia, *Steel: Microstructure and Properties*, 2nd ed., Edward Arnold, London, 1996.
- Jacobson, L. A., and J. McKittrick, Rapid solidification processing, *Mater. Sci. Eng.*, **R11**, 355 (1994).

Silicon Nitride

- Weimer, A. W., ed., *Carbide, Nitride and Boride Materials Synthesis and Processing*, Chapman & Hall, London, 1997.

PLZT

- Beltram, T., M. Kosec, and S. Stavber, *Mater. Res. Bull.*, **28**, 313 (1993).

Plasma Processing of Polymers

- Coates, D. M., and S. L. Kaplan, Modification of polymeric surfaces with plasmas, *Mater. Res. Soc. Bull.*, Aug. 1996, p. 43.

Microwave Processing of Polymers

Lewis, D. A., and J. M. Shaw, Recent development in the microwave processing of polymers, *Mater. Res. Soc. Bull.*, Nov. 1993, p. 37.

PROBLEMS

- W21.1** Calculate the equilibrium constant K for the reaction $\text{SiO}_2(s) \leftrightarrow \text{Si}(s) + \text{O}_2(g)$ at $T = 1300$ K and also the $\text{O}_2(g)$ equilibrium vapor pressure $P_{\text{eq}}[\text{O}_2(g), -1300 \text{ K}]$. Use $\Delta_f G^\circ[\text{SiO}_2(s), 1300 \text{ K}] = -678.535 \text{ kJ/mol}$ for high cristobalite, the stable form of $\text{SiO}_2(s)$ at $T = 1300$ K. This problem illustrates that $\text{SiO}_2(s)$ is very stable against decomposition into its elemental components $\text{Si}(s)$ and $\text{O}_2(g)$.
- W21.2** The heat or enthalpy of formation $\Delta_f H^\circ$ of $\text{Si}(g)$ vapor is essentially independent of temperature from $T = 0$ K up to the melting temperature $T_m = 1414^\circ\text{C}$ and is approximately equal to $(448 \pm 2) \text{ kJ/mol}$. Convert this result from kJ/mol to eV/atom and compare the result with the Si–Si bond energy given in Chapter 2, [i.e., $E(\text{Si–Si}) = 2.34 \text{ eV}$]. From what you know about the crystal structure and bonding in $\text{Si}(s)$, what can you conclude about the vaporization of $\text{Si}(s)$ from this comparison?
- W21.3** Consider the equilibrium phase diagram for the Si–A system shown schematically in Fig. W21.8, where the liquidus and solidus curves are both nearly straight lines for low concentrations of element A in Si, with negative slopes s_L and s_S . Show that the distribution coefficient K_A of A in Si is given by the ratio of the slopes of these two lines [i.e., by Eq. (W21.14)]. (*Hint*: Review the discussion of equilibrium binary phase diagrams in Section 6.5.)
- W21.4** For the thermal oxidation of Si using dry O_2 :
- Derive the expression for the thickness $x(t)$ of the growing a-SiO₂ layer given in Eq. (21.53).
 - Derive the expressions for $x(t)$ given in Eq. (21.54) for thin layers and in Eq. (21.55) for thicker layers.
- W21.5** Calculate the mass densities of $\alpha\text{-Fe}$ at $T = 20^\circ\text{C}$, $\gamma\text{-Fe}$ at $T = 912^\circ\text{C}$, $\delta\text{-Fe}$ at $T = 1394^\circ\text{C}$, and Fe_3C at $T = 20^\circ\text{C}$ (four formula units per orthorhombic unit cell), and of martensite at $T = 20^\circ\text{C}$ for $x = 0.02$ and 0.04 using the structural data given in Table W21.5.
- W21.6** Propose and explain a mechanism by which the alternating layers of ferrite and cementite found in the pearlite shown in Fig. 21.11 can be formed when austenite is slowly cooled through the eutectoid temperature T_e .
- W21.7** For the situation illustrated in Fig. W21.28, where an ion beam deposits 10 J/cm^2 of energy into the surface of a tool steel, calculate the thickness d of the surface layer that can be melted. To simplify the calculation, assume that the steel is pure Fe, its specific heat is constant from $T = 300$ K up to $T_m = 1538$ K, and the energy is deposited uniformly within the thickness d that is melted.

Characterization of Materials

22.1 Introduction

The characterization of materials requires obtaining detailed information about the spatial arrangement of the atoms and identifying precisely which atoms occupy which particular sites in the crystal structure. It also includes the specification of imperfections, impurities, inhomogeneities, and so on. Often, it involves measuring some particular electronic or optical properties. In this chapter we describe a set of tools that the materials scientist has at his or her disposal to characterize materials physically. Some of these tools may be thought of as “cameras” that produce pictures of the material in real space. These include the scanning-electron microscope (SEM), the transmission-electron microscope (TEM), the scanning-tunneling microscope (STM), and the atomic-force microscope (AFM). Also included is the high-resolution transmission-electron microscope (HRTEM) and the low-energy electron microscope (LEEM). Other complementary tools take pictures in momentum space or wave-vector space rather than in real space. Included among these are apparatuses that study x-ray diffraction (XRD), low-energy electron diffraction (LEED), neutron diffraction, and reflection high-energy electron diffraction (RHEED). As with any still camera, information is most readily obtained about the static structure of the material, although the blurriness also conveys some dynamical information.

Additional tools provide information about the dynamical response of a material. Again, this information could be obtained as a function of the real time variable or as a function of the complementary frequency variable. In real time one may use time-dependent luminescent studies to capture the dynamical evolution of a system. Included among the numerous techniques available in frequency space are the optical spectroscopies: infrared, visible, and ultraviolet spectroscopy and light scattering, ellipsometry, infrared absorption, Raman scattering, photoluminescence, and nonlinear optical spectroscopy. Spectroscopy may also be performed with electrons in such techniques as electron energy-loss spectroscopy (EELS) and Auger emission spectroscopy (AES). Inelastic neutron scattering is often used. There are also spectroscopic techniques in which both the electrons and photons play a significant role, such as extended x-ray absorption fine-structure spectroscopy (EXAFS), x-ray photoemission spectroscopy (XPS), and ultraviolet photoemission spectroscopy (UPS). Included among the spectroscopic tools are a variety of resonance techniques: nuclear magnetic resonance (NMR), electron-spin resonance (ESR), nuclear-quadrupole resonance (NQR), and the Mössbauer effect.

Further information about the solid may be obtained using ionic probes, such as in secondary-ion mass spectrometry (SIMS) and Rutherford backscattering (RBS).

Positron beams are used in positron annihilation spectroscopy (PAS) and positive muon beams are used in the technique of muon-precession spectroscopy (μ PS).

The chapter also touches briefly on transport measurements of electrical resistivity, the Hall effect, thermal conductivity, thermopower, and the Peltier coefficient. It describes some magnetic characterization tools, such as the Foner magnetometer, the Faraday balance, and the ac bridge. The SQUID magnetometer is discussed in the textbook in Section 16.7.[†]

Not all methods of characterization are of equal importance. Such techniques as XRD and NMR are more universally employed than others such as LEEM, EXAFS, and HRTEM. Therefore, more space is devoted in the chapter to the former than to the latter techniques. Nevertheless, all the methods in the chapter (as well as others) are used to characterize materials and so should be understood.

DIFFRACTION TECHNIQUES

In this section various diffraction techniques are studied. The most important is x-ray diffraction, which provides information about the long-range order in the bulk of the material. Low-energy electron diffraction provides similar information for the surface of the material. Reflection high-energy electron diffraction and neutron diffraction are also very useful in determining the structure. In particular, neutron scattering is sensitive to the magnetic ordering of a solid.

W22.2 X-ray Diffraction

When a beam of x-rays interacts with an arbitrary material its atoms may scatter the rays into all possible directions. In a crystalline solid, however, the atoms are arranged in a periodic array and this imposes strong constraints on the resulting diffraction pattern. It will be assumed for now that the temperature is sufficiently low that the atoms may be regarded as being frozen in position. Diffraction was introduced in Chapter 3, where the emphasis was on the kinematical aspects of the diffraction. In Section 3.4 the Bragg and von Laue points of view were stated and compared.

In the Bragg description, x-ray diffraction (XRD) is brought about by the constructive interference of waves scattered from successive lattice planes in the crystal. Each plane actually scatters from 10^{-4} to 10^{-3} of the incident wave. Referring to Fig. 3.6, let an incident beam of wave vector \mathbf{k} impinge on a set of lattice planes, the rays making an angle θ with respect to the planes. Attention is restricted to the case of specular elastic scattering, so the outgoing scattered beam, of wave vector \mathbf{k}' , also makes an angle θ with these planes and

$$k' = k. \quad (\text{W22.1})$$

The angle of deviation between the outgoing and incident rays is $\phi = 2\theta$. The separation between neighboring planes is denoted by d . The Bragg condition is given by

[†] The material on this home page is supplemental to *The Physics and Chemistry of Materials* by Joel I. Gersten and Frederick W. Smith. Cross-references to material herein are prefixed by a “W”; cross-references to material in the textbook appear without the “W.”

Eq. (3.52). Constructive interference between successive paths occurs when the path difference equals an integer number of wavelengths.

Von Laue regarded x-ray diffraction as coming about due to the scattering of photons from the periodic lattice. Since the crystal possesses a discrete translational symmetry, there is only wave-vector conservation modulus a reciprocal lattice vector \mathbf{G} . The incident and outgoing wave vectors have the same magnitude and are related by Eq. (3.54). It follows that

$$G^2 + 2\mathbf{G} \cdot \mathbf{k} = 0 \quad (\text{W22.2})$$

As a result, only very specific directions of the incident wave vector \mathbf{k} will result in diffracted beams.

There are at least four ways that one may perform x-ray diffraction experiments:

1. *Using a broadband (nonmonochromatic) x-ray source and looking at the back-reflection.* By utilizing a broadband source such as is produced by bremsstrahlung, there will be a spread of frequencies and hence a spread of wave-vector magnitudes. Even if the angle of incidence is held fixed, there will be some values of k for which Laue backscattering will occur.
2. *Using a diverging (noncollimated) beam of x-rays.* Similarly, by using a beam with a spread of angles, it is possible for the Bragg formula to be satisfied even if k is held fixed.
3. *Using a monochromatic and collimated source but rotating the crystal until the diffraction condition is met.* It is also possible to keep the beam unidirectional and monochromatic but to rotate the sample through some angular trajectory. Whenever the angle is such that the Bragg condition is met, diffraction will occur.
4. *Diffracting the monochromatic x-rays from a powder.* In a powder there are mesoscopic-sized crystals oriented in arbitrary directions. If the x-ray beam impinges on such a powder there will be some orientations for which θ will satisfy the Bragg condition. Having fixed the direction that \mathbf{k} makes with the normal to the crystal, any rotation of the crystal around \mathbf{k} will still satisfy the Bragg condition. Such rotations will cause the Bragg spots to sweep out circles. Since there are a huge number of orientations present in a powder, a circular diffraction pattern is produced.

According to the Heisenberg uncertainty relation, a finite size Δs for a crystal fragment implies an uncertainty in the \mathbf{G} vectors that give rise to diffraction maxima; that is,

$$\Delta G \Delta s \approx 1. \quad (\text{W22.3})$$

This means that the diffraction lines are not perfectly sharp but rather have an angular width on the order of

$$\Delta \theta \approx \frac{\tan \theta}{G \Delta s}. \quad (\text{W22.4})$$

This helps in satisfying the Bragg condition in a powder. It also permits a quantitative estimate of the degree of long-range order to be made by examining the width of the diffraction spots or lines.

Having determined the allowed directions for x-ray scattering from simple conservation laws, one proceeds to obtain expressions for the intensities of the various diffracted beams. X-ray energies are much larger than typical energies of electrons in the conduction band (e.g., the Fermi energy) or the energies of electrons in the upper valence bands (characterized by the energy gaps and bandwidths). However, the x-ray energy may be less than the binding energies of some of the deep-core electrons, particularly in the heavier elements. One may classify the electrons into two categories, which will be termed active and deep-core. *Active electrons* are the electrons in the conduction and upper valence bands; *deep-core electrons* lie in the deep bands. To a first approximation the active electrons may be treated as if they were free. The deep-core electrons are tightly bound to the nuclei and, aside from special resonance situations, are essentially inert.

The dynamics of a free electron interacting with an electromagnetic field follows from Newton's second law:

$$-e\mathbf{E}_0 \sin \omega t = m\mathbf{a}(t). \quad (\text{W22.5})$$

The total instantaneous power radiated by the accelerating charge is given by Larmor's radiation formula:

$$P(t) = \frac{2}{3} \frac{e^2 a^2(t)}{4\pi\epsilon_0 c^3}. \quad (\text{W22.6})$$

The time-averaged radiated power is thus

$$\langle P \rangle = \frac{e^4 E_0^2}{12\pi\epsilon_0 m^2 c^3}. \quad (\text{W22.7})$$

The incident intensity (power per unit area) of the x-ray field is given by the product of the speed of light and the energy density in the field

$$I = \frac{c}{2} \left[\epsilon_0 E^2(t) + \frac{B^2(t)}{\mu_0} \right]. \quad (\text{W22.8})$$

The time-averaged intensity is obtained by noting that the electric and magnetic energy densities are the same, so

$$\left\langle \epsilon_0 E^2(t) + \frac{B^2(t)}{\mu_0} \right\rangle = \epsilon_0 E_0^2, \quad (\text{W22.9})$$

$$\langle I \rangle = \frac{\epsilon_0 c}{2} E_0^2. \quad (\text{W22.10})$$

The cross section for x-ray scattering is the ratio of the scattered power to the incident intensity:

$$\sigma = \frac{\langle P \rangle}{\langle I \rangle} = \frac{8\pi}{3} r_0^2. \quad (\text{W22.11})$$

This is the Thomson cross section for x-ray scattering. The quantity $r_0 = e^2/4\pi\epsilon_0 mc^2 = 2.818 \times 10^{-15}$ m is called the *classical radius of the electron*.

The scattered radiation is not emitted isotropically (i.e., equally in all directions). Consider first a linearly polarized incident electromagnetic wave. An electron oscillating back and forth constitutes a microscopic antenna. The angular distribution of this antenna is given by the dipole distribution

$$P \approx (\hat{k}' \times \hat{E}_0)^2. \quad (\text{W22.12})$$

The polarization vector \mathbf{E}_0 is perpendicular to the wave vector of the incident beam \mathbf{k} . If \mathbf{E}_0 lies in the scattering plane (the plane containing \mathbf{k} and \mathbf{k}' ; see Fig. 3.6) the function above is proportional to $\cos^2 \phi$. If it is perpendicular to the scattering plane, the function above is 1. For unpolarized radiation, which consists of an equal admixture of the two polarization states, the factor is $(1 + \cos^2 \phi)/2$. The differential scattering cross section for scattering radiation into a given solid angle $d\Omega$ centered around angle θ is thus

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2}(1 + \cos^2 2\theta). \quad (\text{W22.13})$$

This has been normalized so that when integrated over all solid angles, the previously obtained formula for the total cross section is regained.

Having derived the cross section for x-ray scattering from a single electron using classical mechanics, this result may now be generalized to the quantum-mechanical case. Two points need to be considered. First, the electron is to be described by a wavefunction whose magnitude squared gives the local probability density for finding the electron at a point in space. If space were decomposed into small volume elements, each element has a probability for containing the electron and hence will contribute to the total x-ray scattering signal. Second, each element radiates coherently to produce the scattered x-ray beam. Determination of the phase of the scattering is simple. Suppose that the element is located at position \mathbf{r} . The incoming field arrives at this position with phase $(i\mathbf{k} \cdot \mathbf{r})$. For the outgoing beam the radiation is created at position \mathbf{r} and emerges with wave vector \mathbf{k}' . Therefore, the outgoing field has a phase factor $\exp(-i\mathbf{k}' \cdot \mathbf{r})$. The scattering amplitude has a phase factor $\exp[i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}]$. The atomic form factor is the Fourier transform of the electron probability distribution:

$$f(\mathbf{q}) = \int n(\mathbf{r}) \exp[i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}] d\mathbf{r} = \int n(\mathbf{r}) \exp[-i\mathbf{q} \cdot \mathbf{r}] d\mathbf{r}. \quad (\text{W22.14})$$

where $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ is the wave-vector transfer (proportional to the momentum transfer), $n(\mathbf{r})$ is the probability density for the electrons, and the integral extends over the volume of the crystal. The classical differential scattering cross section derived previously is multiplied by the absolute square of this factor and becomes

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2}[1 + \cos^2(2\theta)]|f(\mathbf{q})|^2. \quad (\text{W22.15})$$

One may extend this result immediately to the case of x-ray scattering by an atom by interpreting $n(\mathbf{r})$ as the electron number density of the atom. Note that the nucleus, although electrically charged, does not contribute to the x-ray signal because of its heavy mass. As mentioned earlier, the deep-core electrons of the heavier elements also

are not effective in scattering x-rays, so Eq. (W22.15) should only be regarded as being approximate.

The case of x-ray scattering from a crystal may now be investigated. The scattering amplitude $F(\mathbf{q})$ for the crystal is given by Eq. (3.31). It may be expressed as the product of an atomic-form factor and a geometric-structure factor, as in Eq. (3.45). For a monatomic crystal the electron number density is taken to be a superposition of atomic densities and to be of the form

$$n(\mathbf{r}) = \sum_{\mathbf{R}} n_{\text{atom}}(\mathbf{r} - \mathbf{R}). \quad (\text{W22.16})$$

In cases where there are several atoms per unit cell, the electron density is

$$n(\mathbf{r}) = \sum_{\mathbf{R}} \sum_j n_j(\mathbf{r} - \mathbf{R} - \mathbf{s}_j). \quad (\text{W22.17})$$

In place of Eq. (3.34), one obtains

$$F(\mathbf{q}) = S(\mathbf{q}) \sum_j f_j(\mathbf{q}) \exp(-i\mathbf{q} \cdot \mathbf{s}_j). \quad (\text{W22.18})$$

The final formula for the differential cross section becomes

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2} N^2 (1 + \cos^2 2\theta) \sum_{\mathbf{G}} \delta_{\mathbf{q}, \mathbf{G}} \left| \sum_j f_j(\mathbf{G}) \exp(-i\mathbf{G} \cdot \mathbf{s}_j) \right|^2 \exp(-2W), \quad (\text{W22.19})$$

where N is the number of unit cells in the crystal. The factor $\exp(-2W)$, called the *Debye-Waller factor*, takes into account thermal fluctuations. It is introduced in Section W5.2. The existence of the N^2 factor points to the fact that x-ray Bragg scattering is a coherent effect.

In particular experimental implementations of x-ray diffraction, additional angular-dependent terms may enter. For example, in the rotating-crystal method there is a factor $1/\sin 2\theta$ that arises from the time the crystal spends satisfying the Bragg condition. If the crystal were to rotate with an angular velocity ω_c , the time integral of the von Laue momentum constraint would be

$$\int dt \delta(G^2 - 2Gk \sin \theta) = \frac{1}{2Gk\omega_c \cos \theta} = \frac{1}{2\omega_c k^2 \sin 2\theta}. \quad (\text{W22.20})$$

This enters as an additional factor multiplying the differential cross-section formula. For the powder-diffraction method there is a different angular factor.

In Fig. W22.1 a Laue back-reflection diffraction pattern for x-rays backscattered from Si(111) is presented. In Fig. W22.2 an x-ray diffraction pattern from a powdered sample of β -SiC is presented. In this figure the intensities of the diffracted x-ray cones are plotted as a function of the scattering angle, 2θ . This type of graph conveys more

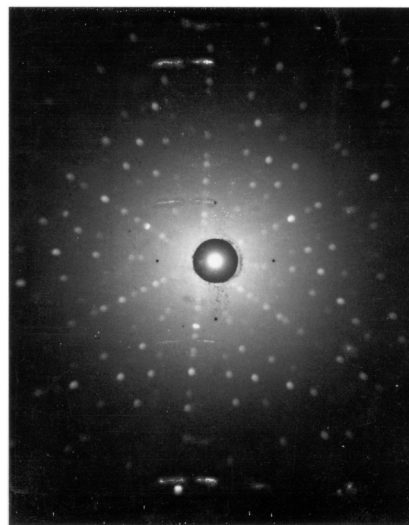


Figure W22.1. Laue back-reflection x-ray diffraction pattern for Si(111). The threefold rotational symmetry of the Si(111) planes is apparent.

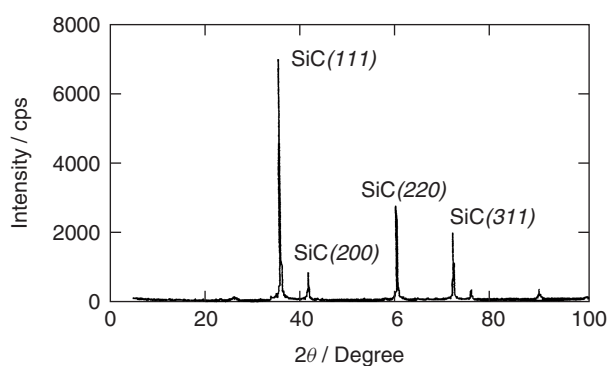


Figure W22.2. X-ray diffraction pattern of sintered SiC fiber-bonded ceramic powders. [Reprinted with permission from T. Ishikawa et al., *Science*, **282**, 1295 (1998). Copyright 1998, American Association for the Advancement of Science.]

information than the powder x-ray diffraction pattern that is presented in Fig. 6.16 in that the relative contributions from the different diffraction peaks are presented. In addition, the widths of the diffraction peaks are related to the quality of the crystallites. The larger and more perfect the crystallites are, the sharper the diffraction pattern will be.

W22.3 Low-Energy Electron Diffraction

C. J. Davisson and L. H. Germer, *Phys. Rev.*, **30**, 705 (1927), directed a monoenergetic beam of electrons at the surface of a solid and found that the reflected electrons

consisted of a set of diffracted beams. This was consistent with the de Broglie hypothesis that, associated with electrons of momentum p , there is a wave with wavelength given by $\lambda = h/p = 2\pi/k$. The momentum of a free electron is related to the energy by $p = (2mE)^{1/2}$. Thus the wave vector of the electron is

$$k = \frac{2\pi}{\lambda} = \frac{1}{\hbar} \sqrt{2mE}. \quad (\text{W22.21})$$

The solid-state crystal provides a microscopic diffraction grating for these electrons.

The wavelength of a 100-eV electron is 0.124 nm, a distance comparable to the spacing between atoms in a solid. The wavelength may be conveniently adjusted by varying the accelerating voltage of the electrons. This method of studying the crystal is called low-energy electron diffraction (LEED). Since the mean free path of electrons in crystals is short (typically, around 1.0 nm for 100 eV), the penetration distance is short. LEED is therefore a tool that provides information about the surface and the first few atomic layers of a solid.

The projectile electron interacts with the ion cores and electrons of the solid. Assume that the surface is flat on a distance scale large compared with the interatomic spacing. The interaction with the ion cores is primarily coulombic, whereas the interaction with the electrons includes an exchange contribution. The net result is that the potential energy is given by some function $V(\mathbf{R}, z)$, where \mathbf{R} is a vector along the surface and z is the coordinate normal to the surface. In most cases of interest $V(\mathbf{R}, z)$ is a periodic function of \mathbf{R} and may be expanded in a Fourier series

$$V(\mathbf{R}, z) = \sum_{\mathbf{G}} V_{\mathbf{G}}(z) \exp(i\mathbf{G} \cdot \mathbf{R}). \quad (\text{W22.22})$$

Here the \mathbf{G} vectors constitute a set of two-dimensional vectors called the *surface reciprocal net*. They play the same role in two-dimensional periodic systems as the reciprocal lattice plays in three dimensions. Note that the Fourier coefficients are themselves functions of z . The periodicity in the z direction is broken for two reasons. First, the crystal is terminated by the surface. Second, there is lattice-plane relaxation as discussed in Chapter 19. In many instances surface reconstruction occurs, in which the surface layer has a translational symmetry parallel to the surface which is not the same as the atoms in the bulk. The unit net of the reconstructed surface is in registry with the underlying bulk lattice and can include several bulk unit-cell projections.

In describing the kinematics of LEED there are two conservation laws operating. The first is conservation of wave vector parallel to the surface, modulus a reciprocal lattice vector

$$\mathbf{K}' = \mathbf{K} + \mathbf{G}. \quad (\text{W22.23})$$

The second law is conservation of energy,

$$\frac{\hbar^2 k'^2}{2m} = \frac{\hbar^2 k^2}{2m}. \quad (\text{W22.24})$$

Here the wave vector \mathbf{k} is expressed as the sum of a vector lying in the surface plane, \mathbf{K} , and a vector perpendicular to the surface:

$$\mathbf{k} = \mathbf{K} + k_z \hat{z}. \quad (\text{W22.25})$$

Similarly, for the outgoing electron,

$$\mathbf{k}' = \mathbf{K}' + k'_z \hat{z}. \quad (\text{W22.26})$$

The scattering geometry is presented in Fig. W22.3. Note that the vector \mathbf{K} has been drawn twice for presentation purposes. Let the angle the incident electron makes with the surface be θ and the corresponding angle for the outgoing electron be θ' . The conservation laws relate these angles:

$$\sin^2 \theta' = \sin^2 \theta - \frac{G^2 + 2\mathbf{K} \cdot \mathbf{G}}{k^2} = \sin^2 \theta - \frac{G^2 + 2Gk \cos \theta \cos \psi}{k^2}, \quad (\text{W22.27})$$

where ψ is the angle between \mathbf{K} and \mathbf{G} . Thus for a given incident angle there will be a set of outgoing angles corresponding to the different values of \mathbf{G} . Naturally, the value of the right-hand side of Eq. (W22.27) must lie between 0 and 1 or the diffracted beam will be suppressed. The surface components of the electron wave vectors make an angle ϕ with respect to each other given by

$$\cos \phi = \frac{k^2(\cos^2 \theta + \cos^2 \theta') - G^2}{2k^2 \cos \theta \cos \theta'}. \quad (\text{W22.28})$$

A simple geometric interpretation of the result above is obtained by referring to Fig. W22.3. Since \mathbf{k} and \mathbf{k}' have the same magnitude, they may be regarded as both touching a sphere (the Ewald sphere) of radius k centered around the origin. Their respective shadows in the plane of the surface must differ by a surface reciprocal-lattice vector in order to produce a diffraction peak. One may imagine a set of parallel rods extending upward from the surface piercing the Ewald sphere with the intersections determining the diffraction directions.

As mentioned in Chapter 19, in two dimensions there are five possible Bravais nets tiled with: squares, hexagons, rectangles, centered rectangles, and parallelograms. These are illustrated in Fig. 19.2. The primitive unit mesh vectors, \mathbf{u}_1 and \mathbf{u}_2 , (as defined in Table 19.1), with their corresponding reciprocal net vectors, \mathbf{G}_1 and \mathbf{G}_2 , are

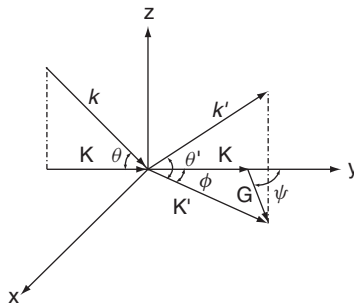


Figure W22.3. LEED scattering geometry: an incident electron with wave vector \mathbf{k} is scattered to an outgoing state with wave vector \mathbf{k}' .

TABLE W22.1 Five Bravais Nets and Their Properties

Bravais Net	Primitive Vectors	Reciprocal Lattice Vectors
Square	$a\hat{i}$ $a\hat{j}$	$(2\pi/a)\hat{i}$ $(2\pi/a)\hat{j}$
Rectangular	$a\hat{i}$ $b\hat{j}$	$(2\pi/a)\hat{i}$ $(2\pi/b)\hat{j}$
Centered rectangular	$a\hat{i}$ $(a\hat{i} + b\hat{j})/2$	$(2\pi/a)\hat{i} - (2\pi/b)\hat{j}$ $(4\pi/b)\hat{j}$
Hexagonal	$a\hat{i}$ $a(\hat{i} + \hat{j}\sqrt{3})/2$	$(2\pi/a)[\hat{i} - \hat{j}/\sqrt{3}]$ $(4\pi/a)\hat{j}/\sqrt{3}$
Oblique	$a\hat{i}$ $b(\hat{i}\cos\phi + \hat{j}\sin\phi)$	$(2\pi/a)[\hat{i} - \hat{j}\cot(\phi)]$ $(2\pi/b)\hat{j}\csc(\phi)$

given in Table W22.1. They are related by

$$\mathbf{G}_1 = -2\pi \frac{\hat{k} \times \mathbf{u}_2}{\hat{k} \cdot \mathbf{u}_1 \times \mathbf{u}_2}, \quad \mathbf{G}_2 = 2\pi \frac{\hat{k} \times \mathbf{u}_1}{\hat{k} \cdot \mathbf{u}_1 \times \mathbf{u}_2}. \quad (\text{W22.29})$$

It is usually necessary to fit the observed LEED intensities to a model of the surface and near-surface region to obtain a detailed picture of the surface atomic structure. An example of a typical LEED pattern is given in Fig. W22.4. It shows the reconstruction of an Ir(100) surface with a 5×1 superstructure. The reciprocal lattice vectors are of the form $\mathbf{G} = h\mathbf{G}_1 + k\mathbf{G}_2$, with h and k being integers. The spots may be enumerated by these integers in the figure.

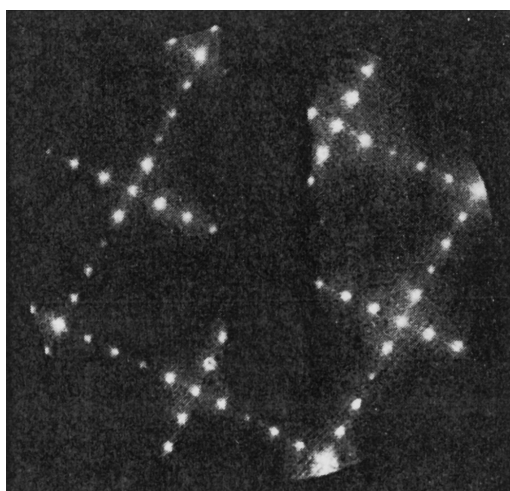


Figure W22.4. LEED pattern for a reconstructed Ir (100) surface. [Reprinted from K. Heinz, *Surf. Sci.*, **299/300**, 433 (1994), Copyright 1994 with permission from Elsevier Science.]

W22.4 Reflection High-Energy Electron Diffraction

In reflection high-energy electron diffraction (RHEED) a high-energy beam of electrons is directed at grazing angles of incidence onto the solid. The electron energy is in the range 10 to 100 keV and the angles are in the range 0.1° to 5° . The scattering mechanism becomes more Coulomb-like, with the dominant scattering in the near-forward direction. It is particularly sensitive to the surface structure of the solid.

Referring to the kinematic formulas developed for LEED and the corresponding figure (Fig. W22.3) illustrating the scattering geometry, the following simplifications are made. First, it is assumed that θ is small, so that $\cos \theta \approx 1$. For electrons in the energy range 10 to 100 keV the wave vector k is in the range $5.1 \times 10^{11} \text{ m}^{-1}$ to $1.6 \times 10^{12} \text{ m}^{-1}$. This is typically two orders of magnitude larger than the reciprocal net vectors, G . One may therefore make a series expansion in powers of G/k and retain the lowest-order terms. Thus the numerical values for θ , θ' , and ϕ and G/k are all small. The results are

$$\theta'^2 = \theta^2 - \frac{2G}{k} \cos \psi - \frac{G^2}{k^2}, \quad (\text{W22.30})$$

$$\theta^2 = \frac{G^2}{k^2}. \quad (\text{W22.31})$$

If a spherical screen is located a distance R from the sample along the y axis, spots will appear at the points

$$x = R \cos \theta' \sin \phi \approx R\phi \approx \pm \frac{RG}{K}, \quad (\text{W22.32a})$$

$$y = R \cos \theta' \cos \phi \approx R, \quad (\text{W22.32b})$$

$$z = R \sin \theta' \approx R\theta' = R\sqrt{\theta^2 - \frac{2G}{k} \cos \psi - \frac{G^2}{k^2}}. \quad (\text{W22.32c})$$

Note that $G \cos \psi = G_y$. Thus the spots lie on a circle whose radius is

$$r = \sqrt{x^2 + z^2} = R\sqrt{\theta^2 - \frac{2G_y}{k}}. \quad (\text{W22.33})$$

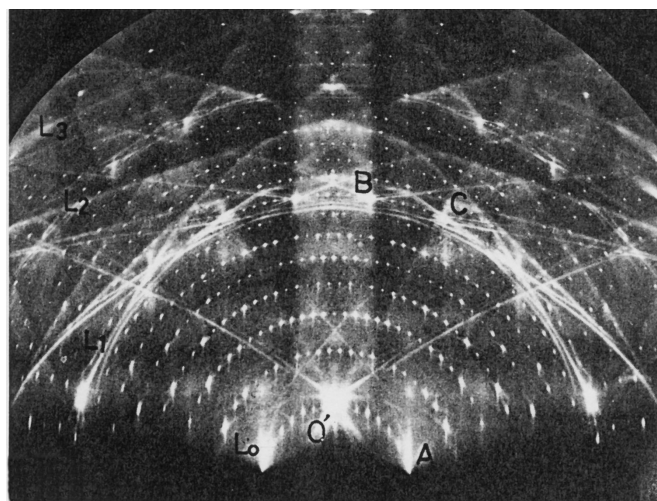
If the surface of the solid consists of a square mesh of side a then the components of \mathbf{G} are

$$(G_x, G_y) = \left(\frac{2\pi n_x}{a}, \frac{2\pi n_y}{a} \right), \quad (\text{W22.34})$$

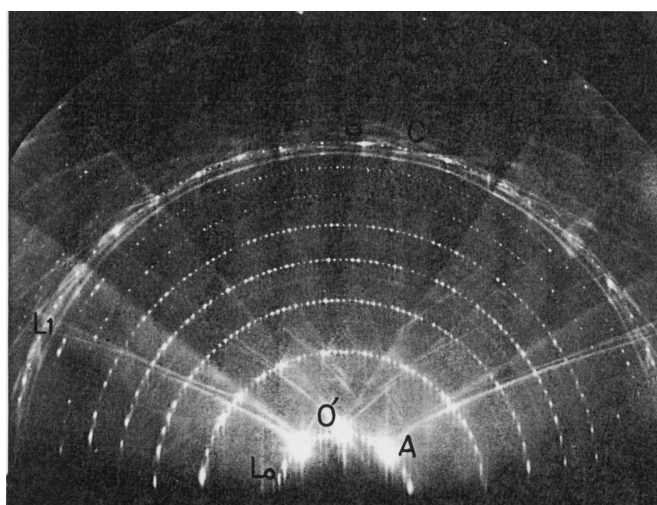
where n_x and n_y are integers. The radius of the circle is

$$r(\theta, n_y) = R\sqrt{\theta^2 - \frac{4\pi n_y}{ka}}. \quad (\text{W22.35})$$

Corresponding to a given value of n_y is a circle of a given radius. The location of points along the circle is determined by n_x .



(a)



(b)

Figure W22.5. RHEED patterns taken with 20-keV electrons from Si (111) with the 7×7 structure. The upper pattern (a) is for $[\bar{1}2\bar{1}]$ and the lower pattern (b) is for $[0\bar{1}1]$ incidence. [From S. Ino, *Jpn. J. Appl. Phys.*, **16**, 891(1977). Copyright 1977 by the Japanese Journal of Applied Physics.]

Examples of RHEED patterns are given in Fig. W22.5. RHEED is used to monitor atomic layer-by-layer thin-film growth in MBE. Oscillations of spot amplitudes yield growth rates and provide information on the growth mode.

W22.5 Neutron Scattering

Neutron scattering involves directing a beam of neutrons from a nuclear reactor at a solid and monitoring the scattered neutrons with a detector. There are both elastic

and inelastic contributions. The elastic contribution produces a diffraction pattern and provides useful structural information such as interatomic spacings and symmetries of the crystal. Inelastic scattering provides information concerning the phonon spectrum and magnons in magnetic materials. Neutrons scatter off the nuclei via the strong interaction and off electrons by magnetic interactions. The cross sections are small, on the order of several square fermis $[(10^{-15} \text{ m})^2]$, so the mean free path is long. Therefore, information is obtained concerning the bulk sample. It makes a difference whether the sample is magnetized or not, and neutron scattering is therefore a valuable tool for probing the magnetic ordering of matter. Studies may be made as a function of temperature, external magnetic field, pressure, stress, or other thermodynamic variables. It provides a powerful tool for obtaining phase diagrams of materials.

The neutrons that emerge from the reactor are collimated into a narrow beam by placing absorbing shields in all but the desired beam direction. For thermal neutrons ($T = 300 \text{ K}$) the mean wave-vector magnitude is $k = 4.3 \times 10^{10} \text{ m}^{-1}$. A monoenergetic source is constructed from the thermal source by Bragg scattering the neutrons off a crystal and further collimating the beam. By varying the angle of incidence the neutron energy is tunable through some range of values. The typical wave vector is several times greater than typical reciprocal lattice vectors G of samples to be studied. Scattering from magnetized crystals can produce a polarized neutron beam.

The discussion begins by studying the case of neutron diffraction (i.e., elastic scattering from the crystal). Suppose that the incident neutron wave vector is \mathbf{k} and the scattered wave vector is \mathbf{k}' . The wave-vector transfer is

$$\mathbf{q} = \mathbf{k}' - \mathbf{k}. \quad (\text{W22.36})$$

If the lattice is periodic in three dimensions, momentum is conserved modulus a reciprocal lattice vector, implying that

$$\mathbf{q} = \mathbf{G}. \quad (\text{W22.37})$$

The energy conservation condition is

$$\frac{\hbar^2 k'^2}{2M_n} = \frac{\hbar^2 k^2}{2M_n}, \quad (\text{W22.38})$$

where M_n is the neutron mass. It follows that

$$q = 2k \sin \frac{\theta}{2}, \quad (\text{W22.39})$$

where θ is the scattering angle between vectors \mathbf{k} and \mathbf{k}' . The sample is rotated, and whenever $\mathbf{q} = \mathbf{G}$ is satisfied, a diffraction peak occurs. For single crystals the actual value of the \mathbf{G} vector is determined from the knowledge of \mathbf{k} and \mathbf{k}' .

As with x-ray diffraction, at finite temperatures the crystal may be regarded as an admixture of ordered and disordered phases. The disordered part does not produce a diffraction pattern but rather, produces a thermal diffuse background scattering. When the crystal temperature is raised so that it finally melts, there is only diffuse scattering. Additional contributions to the diffuse background are due to imperfections, impurities, the magnetic moments of unaligned nuclei, and isotopic inhomogeneities.

Inelastic neutron scattering is also possible. In this case one or more phonons or magnons are either absorbed or emitted by the neutron in the course of scattering. It is possible to measure the dispersion curve of these elementary excitations [i.e., to determine the dependence of their frequency $\omega(\mathbf{Q})$ on their wave vector \mathbf{Q}]. For absorption of a single phonon (or magnon) the conservation laws are

$$\mathbf{k}' = \mathbf{k} + \mathbf{Q} + \mathbf{G}, \quad \frac{\hbar^2 k'^2}{2M_n} = \frac{\hbar^2 k^2}{2M_n} + \hbar\omega(\mathbf{Q}). \quad (\text{W22.40})$$

The corresponding formulas for the emission of a single excitation are

$$\mathbf{k}' = \mathbf{k} - \mathbf{Q} + \mathbf{G}, \quad \frac{\hbar^2 k'^2}{2M_n} = \frac{\hbar^2 k^2}{2M_n} - \hbar\omega(\mathbf{Q}). \quad (\text{W22.41})$$

Phonons and magnons are excitations that obey Bose–Einstein statistics. The number of excitations in mode \mathbf{Q} is given by the Bose–Einstein distribution function:

$$n(\mathbf{Q}, T) = \frac{1}{\exp[\hbar\omega(\mathbf{Q})/k_B T] - 1}. \quad (\text{W22.42})$$

The relative contribution of occupied to unoccupied states should be in the ratio of the Boltzmann factor, $\exp[-\hbar\omega(\mathbf{Q})/k_B T]$. One therefore expects to see a spectrum (called

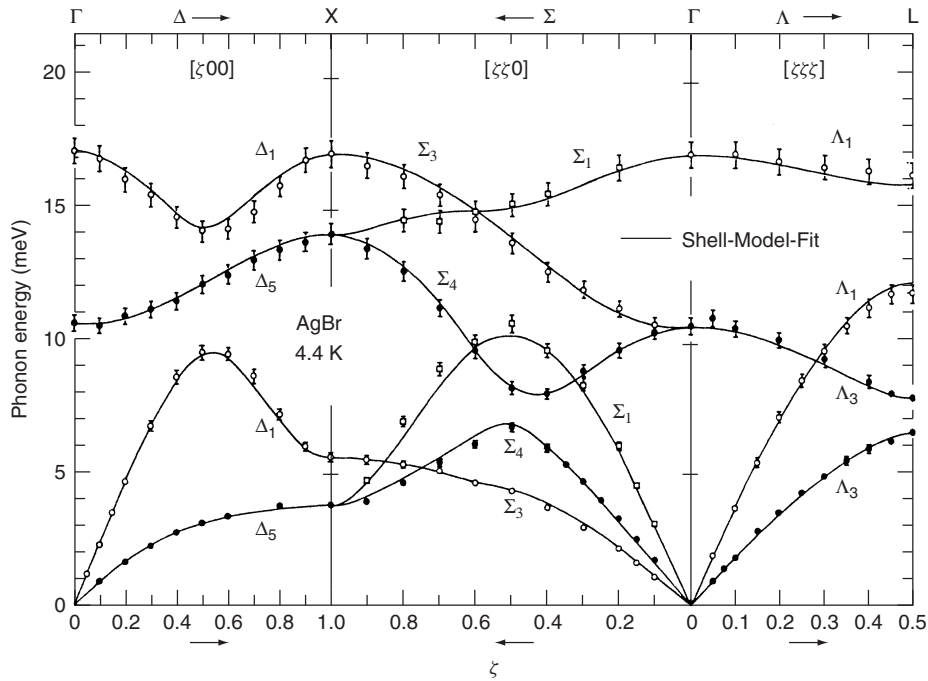


Figure W22.6. Phonon dispersion curves for AgBr at $T = 296$ K measured by neutron scattering. [From Y. Fujii et al, Phys. Rev. B, **15**, 358 (1977). Copyright 1977 by the American Physical Society.]

the *dynamical structure factor*) of the form

$$S(\mathbf{Q}, \omega) = e^{-2W} \sum_{\text{type}, \mathbf{Q}} |M(\mathbf{Q})|^2 [[n(\mathbf{Q}, T) + 1]\delta(\omega + \omega(\mathbf{Q})) + n(\mathbf{Q}, T)\delta(\omega - \omega(\mathbf{Q}))], \quad (\text{W22.43})$$

where M is a matrix element describing the coupling of the neutron to the particular excitation and is defined in terms of the energy transfer:

$$\hbar\omega = \frac{\hbar^2 k'^2}{2M_n} - \frac{\hbar^2 k^2}{2M_n}. \quad (\text{W22.44})$$

The sum over “types” includes the various branches of the phonon spectrum (TA, LA, TO, LO) as well as magnons, if they are present.

If the energy of the neutrons is sufficiently high, it is also possible to have emission or absorption of more than one elementary excitation. Thus two-phonon, three-phonon, . . . processes are possible.

In summary, neutron scattering provides the means to measure the dispersion curves of the low-energy elementary excitations of the crystal. An example of a set of such curves is given in Fig. W22.6, where data for AgBr are presented along principal symmetry directions of the reciprocal lattice.

OPTICAL SPECTROSCOPY

In the following sections various optical methods to acquire information about solids are explored. First, optical spectroscopy is discussed for electromagnetic radiation in the infrared, visible, and ultraviolet regions of the spectrum. The solid is excited from its ground state to some excited state and information concerning such transitions may be obtained from the reflectance or transmittance. This is followed by a discussion of ellipsometry, which provides a technique for systematically mapping out the electromagnetic properties of materials by examining the reflected light. The methods of Fourier transform infrared spectroscopy and Raman spectroscopy provide complementary techniques for further characterizing the bulk of materials in the infrared. Luminescence provides additional information about the excited states of the solid as they relax toward lower-energy states by radiative emission. Finally, nonlinear optical spectroscopy provides important information on how strong electromagnetic fields interact with solids.

W22.6 Optical Spectroscopy in the Infrared, Visible, and Ultraviolet

The propagation of light through a material is governed by a complex index of refraction $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega)$. The real part of this index determines the speed of propagation through the medium

$$v = \frac{c}{n(\omega)}. \quad (\text{W22.45})$$

The imaginary part determines the absorption coefficient that appears in the law of attenuation, Beer's law:

$$\frac{I(x)}{I_0} = \exp[-\alpha(\omega)x]. \quad (\text{W22.46})$$

Here $I(x)/I_0$ is the fraction of the initial light intensity entering the material that remains after traversing a distance x and

$$\alpha(\omega) = 2\frac{\omega}{c}\kappa(\omega). \quad (\text{W22.47})$$

One of the central goals of solid-state physics over the years has been to try to explain the optical properties of materials in terms of the response of the electrons and the lattice.

The index of refraction is measured directly using a reflectance apparatus. The reflection coefficient, R , is the ratio of the reflected intensity to the incident intensity and is given, for the case of normal incidence on a thick sample [see Eq. (8.16)], by

$$R(\omega) = \frac{[n(\omega) - 1]^2 + [\kappa(\omega)]^2}{[n(\omega) + 1]^2 + [\kappa(\omega)]^2} \quad (\text{W22.48})$$

(For thin samples the expression is more complicated due to interference effects caused by reflection from the back surface of the film.) Selected data for $n(\omega)$ and $\kappa(\omega)$ for the semiconductor GaAs are shown in Fig. W22.7. There is a sharp electronic band-edge apparent at $\approx 1 \mu\text{m}$ and a sharp optical-phonon resonance at $\approx 30 \mu\text{m}$.

In attempting to understand the optical properties, the various contributions to the index of refraction need to be identified. These include the effect of optical phonons, conduction electrons, interband transitions, and excitons, as well as trapped electrons. These effects are studied in Chapter 8.

Typical ionic vibrational frequencies lie in the infrared region of the spectrum, and this is true for the phonons as well. From Chapter 8 the optical-phonon contribution to the complex index of refraction is given by the formula

$$[n(\omega) + i\kappa(\omega)]^2 = \epsilon_r(\infty) \left(1 + \frac{\omega_L^2 - \omega_T^2}{\omega_T^2 - \omega^2 - i\omega\gamma} \right). \quad (\text{W22.49})$$

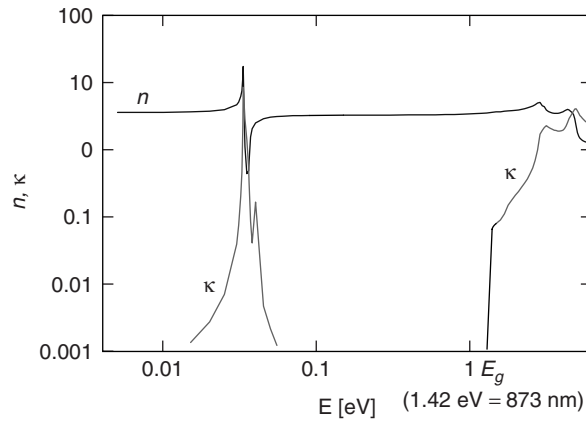


Figure W22.7. Optical constants n and κ for GaAs as a function of energy E . (Data taken from D. E. Aspnes and A. A. Studna, *Optical Constants of Solids*, E. D. Palik, ed., Academic Press, San Diego, Calif., 1985).

For GaAs the fit is made to the infrared spectrum and the fitting parameters are $\bar{\nu}_L = 292 \text{ cm}^{-1}$, $\bar{\nu}_T = 269 \text{ cm}^{-1}$, $\epsilon_r(\infty) = 11.0$, and $\gamma = 2.4 \text{ cm}^{-1}$, where $\omega = 2\pi\bar{\nu}c$.

In the case of a metal, one must take into account the plasma oscillations in describing the optical region of the spectrum, as in Section 8.4. The index of refraction would then be given by

$$[n(\omega) + i\kappa(\omega)]^2 = \epsilon_r(\infty) \left[1 - \frac{\omega_p^2}{\omega(\omega + i/\tau)} \right], \quad (\text{W22.50})$$

where the plasma frequency is given by [see Eq. (8.11)]

$$\omega_p = \sqrt{\frac{ne^2}{m\epsilon(\infty)}}. \quad (\text{W22.51})$$

This is often used to fit the dielectric function in the region of the plasma resonance. The lifetime τ is determined by the collisions of the electrons within a given band with phonons (intraband transitions).

Other channels for absorption open when the photon energy is larger than the gap between an occupied and an unoccupied band. A semiconductor is a case in point. When the photon energy exceeds the bandgap, electrons may be excited from the top of the valence band to the bottom of the conduction band. If these extrema coincide in \mathbf{k} space, the transition is called direct. If not, phonon emission or absorption can make up for the momentum mismatch and the transition is called indirect. One therefore finds a sharp increase in $\kappa(\omega)$ when the photon energy exceeds the bandgap. The functional form depends on the density of states in the two bands. To the extent that the smooth variation of the radiation matrix elements are neglected, $\kappa(\omega)$ is proportional to

$$\kappa(\omega) \approx \int_{-\infty}^{\mu} dE \rho_v(E) \rho_c(E + \hbar\omega). \quad (\text{W22.52})$$

Thus it is given by the joint density of states between the valence band (v) and the conduction band (c).

If excitons are present, then, as seen in Section 8.10, sharp absorption lines are observed below the bandgap energy. These correspond to the energy of the formation of the exciton (bandgap energy less the binding energy) and of its excited states.

One may also liberate electrons trapped in impurity levels by photon absorption and monitor their presence by studying the behavior of $\kappa(\omega)$ below the energy gap.

The chemical composition of a solid may readily be determined by sputtering excited atoms from the surface. The resulting atomic emission lines are identified using a monochromator and photodetector. The technique is called *atomic-emission spectrometry*. Typically, a beam of high-energy argon ions is used for the excitation. Before an accurate quantitative analysis may be made, however, one must know the sputtering and excitation probabilities for the individual atoms when subjected to the argon ions at a given energy. These may be obtained from sputtering experiments using samples with predetermined compositions.

It is also possible to look at the atomic absorption spectrum of the resulting plasma to obtain a quantitative analysis. The technique is then called *atomic-absorption spectrometry* (AAS).

Other methods, beside Ar^+ ion bombardment, are available for producing sputtering, including laser irradiation, electric arcs, and electron-stimulated desorption (ESD).

W22.7 Ellipsometry

Ellipsometry provides a means for characterizing the optical properties of solids, interfaces, thin films, and multilayers. The method employs polarized monochromatic light which is reflected from the surface and analyzed. A typical experimental apparatus with an arrangement referred to as the *PCSA geometry* is depicted in Fig. W22.8. A collimated light beam, with wave vector \mathbf{k} , passes through a polarizer P, and a compensator C, is obliquely reflected from a sample S, passes through an analyzer A, and the beam is finally detected with wave vector \mathbf{k}'' . The angle of incidence (and reflection) is ϕ .

Two basis states of the polarization geometry are shown in Fig. W22.9. The polarization plane is defined as the plane containing the vectors \mathbf{k} and \mathbf{k}'' and it also contains the normal to the solid surface, \hat{n} . In the *p*-polarization case the incident and reflected electric field vectors, \mathbf{E} and \mathbf{E}' , lie in the polarization plane, as shown in Fig. W22.9a. In the *s*-polarization case, shown on the right, the electric field vectors are perpendicular

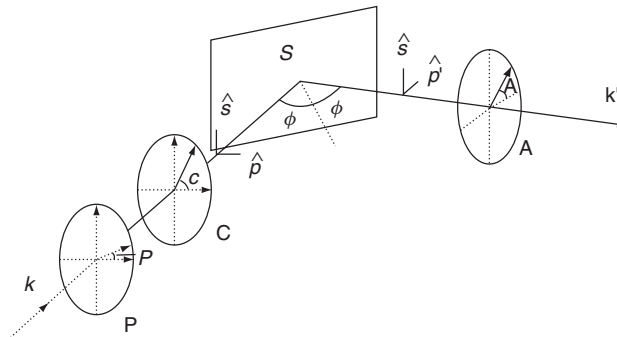


Figure W22.8. Geometry of an ellipsometry experiment.

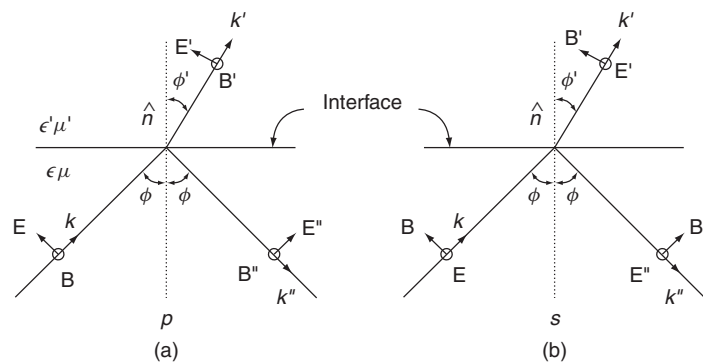


Figure W22.9. Electric and magnetic fields for the (a) *p*-polarization and (b) *s*-polarization cases. (Adapted from J. D. Jackson, *Electrodynamics*, 2nd ed., Wiley, New York, 1975, Figs. 7.6 and 7.7.)

to the polarization plane. The unit basis vectors shown in Fig. W22.8, \hat{s} , \hat{p} , and \hat{p}' , are appropriate for expressing the incident and reflected fields. The electric permittivity of the solid, ϵ' , will be a complex function of frequency, in general.

After passing through the polarizer, whose polarization axis makes an angle P with the polarization plane, the electric field is $\mathbf{E}_1 = E_0(\hat{p} \cos P + \hat{s} \sin P)$. The compensator has a polarization axis defined by the unit vector \hat{C} , making an angle C with the polarization plane. It has different transmittance for electric fields along \hat{C} and perpendicular to \hat{C} . It may be described by the transmittance dyadic $\vec{T} = T_{\parallel} \hat{C} \hat{C} + T_{\perp} (\vec{I}_2 - \hat{C} \hat{C})$, where the two-dimensional unit dyadic may be expressed as $\vec{I}_2 = \hat{p} \hat{p} + \hat{s} \hat{s}$. The field emerging from the compensator is

$$\begin{aligned} \mathbf{E} &= [T_{\parallel} \hat{C} \hat{C} + T_{\perp} (\vec{I}_2 - \hat{C} \hat{C})] \cdot \mathbf{E}_1 \\ &= E_0 \{ [(T_{\parallel} - T_{\perp}) \cos(P - C) \sin C + T_{\perp} \sin P] \hat{s} \\ &\quad + [(T_{\parallel} - T_{\perp}) \cos(P - C) \cos C + T_{\perp} \cos P] \hat{p} \}. \end{aligned} \quad (\text{W22.53})$$

The ratio of E_s to E_p may be written as

$$\frac{E_s}{E_p} = \frac{\tan C + \rho_c \tan(P - C)}{1 - \rho_c \tan C \tan(P - C)} \quad (\text{W22.54})$$

where $\rho_c \equiv T_{\perp}/T_{\parallel}$. In the case where the compensator is a quarter-wave plate one has $\rho_c = i = \sqrt{-1}$.

By matching boundary conditions for plane-wave reflection and transmission at the interface, Fresnel derived expressions for the reflection coefficients for the amplitudes of p - and s -polarized light. For the case of reflection from a semi-infinite solid, Fresnel found that (Jackson, 1975, Sec. 7.3)

$$r_s = \frac{\tilde{n} \cos \phi - \sqrt{\tilde{n}'^2 - \tilde{n}^2 \sin^2 \phi}}{\tilde{n} \cos \phi + \sqrt{\tilde{n}'^2 - \tilde{n}^2 \sin^2 \phi}}, \quad (\text{W22.55a})$$

$$r_p = \frac{\tilde{n}'^2 \cos \phi - \tilde{n} \sqrt{\tilde{n}'^2 - \tilde{n}^2 \sin^2 \phi}}{\tilde{n}'^2 \cos \phi + \tilde{n} \sqrt{\tilde{n}'^2 - \tilde{n}^2 \sin^2 \phi}}, \quad (\text{W22.55b})$$

where it is assumed that the magnetic permeabilities are equal (i.e., $\mu' = \mu$). In the general case, either \tilde{n} or \tilde{n}' , the indices of refraction for the incident and reflecting media, respectively, may be complex functions of frequency. In the case where light is incident from vacuum, $\tilde{n} = 1$.

The reflected wave is given by $\mathbf{E}'' = (r_p \hat{p}' \hat{p} + r_s \hat{s} \hat{s}) \cdot \mathbf{E}$, that is,

$$\mathbf{E}'' = E_0(x_s \hat{s} + x_p' \hat{p}'), \quad (\text{W22.56})$$

where

$$x_s = r_s [(T_{\parallel} - T_{\perp}) \cos(P - C) \sin C + T_{\perp} \sin P] \equiv |x_s| e^{i\Delta_s}, \quad (\text{W22.57a})$$

$$x_p' = r_s [(T_{\parallel} - T_{\perp}) \cos(P - C) \cos C + T_{\perp} \cos P] \equiv |x_p'| e^{i\Delta_{p'}}. \quad (\text{W22.57b})$$

The ratio $(E_{p'}/E_s)_{\text{out}}$ is often expressed as

$$\left(\frac{E_{p'}}{E_s}\right)_{\text{out}} = \frac{x_{p'}}{x_s} \equiv \tan \psi e^{i\Delta}, \quad (\text{W22.58})$$

where $\Delta = \Delta_{p'} - \Delta_s$.

After emerging from the analyzer, which makes an angle A with the polarizing plane, the electric field is $\mathbf{E}_{\text{out}} = \hat{A} \hat{A} \cdot \mathbf{E}''$. The intensity of the outgoing light is proportional to $|\mathbf{E}_{\text{out}}|^2$,

$$I(A) = I_0(1 + \alpha \cos 2A + \beta \sin 2A), \quad (\text{W22.59})$$

where $I_0(\propto |E_0|^2)$ is the mean intensity amplitude,

$$\alpha = \frac{|x_{p'}|^2 - |x_s|^2}{|x_{p'}|^2 + |x_s|^2} = \frac{\tan^2 \psi - 1}{\tan^2 \psi + 1}, \quad (\text{W22.60a})$$

$$\beta = \frac{2|x_{p'}||x_s| \cos \Delta}{|x_{p'}|^2 + |x_s|^2} = \frac{2 \tan \psi \cos \Delta}{\tan^2 \psi + 1}. \quad (\text{W22.60b})$$

Solving for ψ and Δ gives

$$\psi = \tan^{-1} \left(\sqrt{\frac{1+\alpha}{1-\alpha}} |\tan P| \right), \quad 0 \leq \psi \leq \pi/2, \quad (\text{W22.61a})$$

$$\Delta = \pm \cos^{-1} \left(\sqrt{\frac{\beta^2}{1-\alpha^2}} \right), \quad -\pi < \Delta \leq \pi. \quad (\text{W22.61b})$$

By fitting $I(A)$ expressed by Eq. (W22.59) to the measured data, values for α and β may be determined. One may then adjust the real and imaginary parts of $\tilde{n}' = n' + i\kappa'$ to obtain an optimum fit to the measured values of α and β . By carrying out this procedure as a function of photon frequency, the real and imaginary parts of the index of refraction, $n'(\omega)$ and $\kappa'(\omega)$, may be obtained.

The method is readily extended to the case where there is a film on a substrate. The dielectric properties of the film are often modeled using an effective-medium approximation. The sensitivity of ellipsometry is very high and the effects of monolayer or even submonolayer films may be detected.

The growth of an hydrogenated amorphous carbon (a-C:H) film on a Mo substrate is illustrated in Fig. W22.10, which shows the evolution of $\langle \epsilon_1 \rangle$ and $\langle \epsilon_2 \rangle$, the real and imaginary parts of the pseudodielectric function $\langle \epsilon_r \rangle$. The pseudodielectric function corresponds to a hypothetical uniform and isotropic material with a clean and abrupt surface boundary that gives the same ellipsometric data as the actual sample being studied.[†] The start point corresponds to the bare Mo substrate and the end point to $a \approx 123$ nm thick a-C:H film. The evolution of $\langle \epsilon_1 \rangle$ and $\langle \epsilon_2 \rangle$ has been modeled in

[†] D. E. Aspnes, *Thin Solid Films*, **89**, 249 (1982).

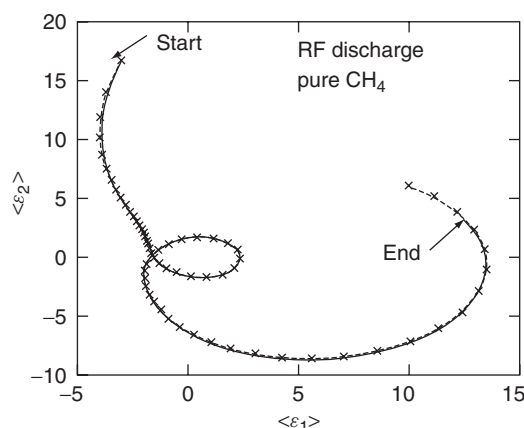


Figure W22.10. Pseudodielectric function for a-C:H film growth on Mo. “START” refers to a bare Mo substrate and “END” to a film of 123 nm covering the substrate. [From R. W. Collins, *Appl. Phys. Lett.*, **52**, 2025 (1988). Copyright 1988, American Institute of Physics.]

this case by the layer-by-layer growth of a uniform a-C:H film with $\epsilon_1 = 2.84$ and $\epsilon_2 = 0.425$ at $\lambda = 388$ nm.

W22.8 Fourier Transform Infrared Spectroscopy

It is possible to measure the entire infrared-transmission spectrum of a solid at one time using Fourier transform infrared spectroscopy (FTIR). The method is not limited to the infrared or to absorption spectroscopy, although it is often used there. The technique makes use of an interferometer with a moving mirror. The technique is to create a replica of the infrared spectrum at much lower frequencies and to make the measurements there.

The interferometer is illustrated in Fig. W22.11. Broadband infrared radiation $E_0(\omega)$ (usually blackbody radiation) is produced by a filament or globar O and is directed toward a “half-silvered” mirror m . The beam is split into two parts, each one traveling along a different arm of the interferometer. The first beam strikes mirror M and the second beam strikes the moving mirror M' . The beams are recombined at

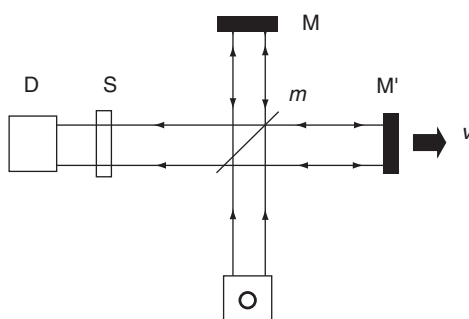


Figure W22.11. FTIR apparatus showing the interferometer.

the “half-silvered” mirror, are directed through sample S, and are finally detected at detector D (usually, a bolometer). A recording is made of the intensity as a function of time, which is then Fourier analyzed.

Let d be the distance from m to mirror M and $d' + vt$ be the distance from m to mirror M'. The amplitude of the recombined wave is the superposition of the amplitudes of the two beams

$$E(t) = \frac{E_0(\omega)}{2} \exp(2ikd) + \frac{E_0(\omega)}{2} \exp[2ik(d' + vt)]. \quad (\text{W22.62})$$

The intensity incident on the sample is proportional to the absolute square of E . The detected intensity is

$$I(t) = 2 \int_0^\infty d\omega I_0(\omega) T(\omega) \left[1 + \cos \left(2 \frac{\omega}{c} (d' - d + vt) \right) \right]. \quad (\text{W22.63})$$

where $I_0(\omega) = |E_0(\omega)|^2$, $T(\omega)$ being the transmission coefficient for the sample. Now take the Fourier transform of this to obtain

$$\begin{aligned} I(\Omega) &= \int_{-\infty}^\infty \frac{dt}{2\pi} I(t) \exp(i\Omega t) \\ &= \int_0^\infty d\omega I_0(\omega) T(\omega) \left[2\delta(\Omega) + \delta \left(\Omega - 2\frac{v}{c}\omega \right) \exp(i\eta) \right. \\ &\quad \left. + \delta \left(\Omega + 2\frac{v}{c}\omega \right) \exp(-i\eta) \right], \end{aligned} \quad (\text{W22.64})$$

where $\eta = 2\omega(d - d')/c$. Focusing attention on the resonant (second) term and computing its amplitude gives

$$|I(\Omega)| \simeq \frac{c}{2v} I_0 \left(\frac{c}{2v} \Omega \right) \left| T \left(\frac{c}{2v} \Omega \right) \right|. \quad (\text{W22.65})$$

In the ideal case, since the blackbody spectrum is known, the functional dependence of $I_0(\omega)$ is known. Therefore, a measurement of $I(\Omega)$ permits the determination of $|T(c\Omega/2v)|$. Since v will typically be on the order of 1 mm/s, the ratio $c/2v$ will be 1.5×10^{11} . Thus a measurement in the frequency range of $\Omega \approx 1$ kHz is used to determine the spectrum in the range of 10^{14} Hz! A replica of the infrared spectrum has been produced at low frequencies.

In reality, the situation is more complicated, since the source is not a blackbody. Usually, a baseline spectrum is taken without a sample. In this way the output can be normalized to the response of the system, including the source spectrum and detector sensitivity.

FTIR permits one to obtain data simultaneously over a large frequency range and over a large collection angle. Multiple scans are used to improve the signal-to-noise ratio. The technique is readily extended to other forms of spectroscopy, such as Raman spectroscopy.

The FTIR spectrum for diamond is presented in Fig. W22.12. The spectrum clearly shows various critical points and combinations of critical points in the phonon spectrum.

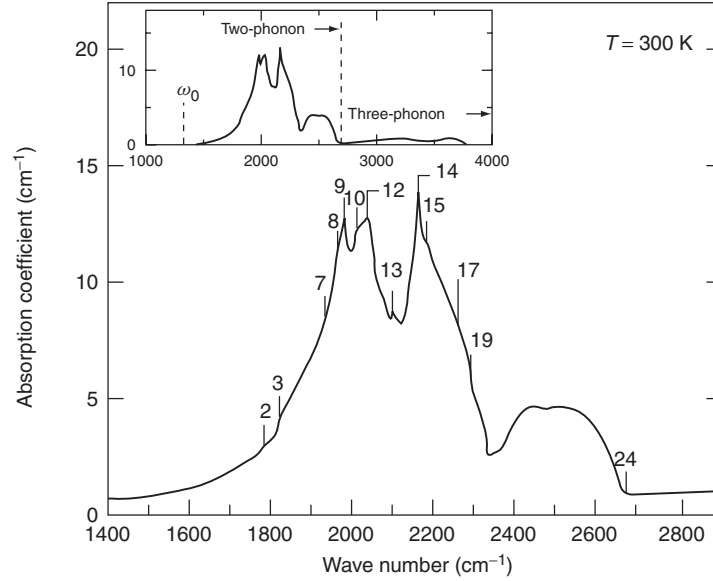


Figure W22.12. FTIR spectrum for diamond at $T = 300$ K. [From R. Vogelgesang et al., *Phys. Rev. B*, **58**, 5408 (1998). Copyright 1998 by the American Physical Society.]

This spectrum may also be contrasted with the Raman spectrum given in the following section. The spectrum should be compared with the phonon density of states presented in Fig. 5.9.

W22.9 Raman Spectroscopy

The Raman effect was originally discovered in molecular physics. Monochromatic light of frequency ω was directed at a gas sample, and the scattered light was passed through a monochromator and onto a photodetector. The scattered light consisted mainly of radiation at frequency ω (Rayleigh scattering), but also possessed sidebands at lower (Stokes shifted) and higher (anti-Stokes shifted) frequencies. The displacement of the sidebands is characteristic of the type of molecule under study and is related to the vibrational frequencies associated with nuclear motion. The angular momentum selection rules $\Delta J = 0, \pm 2$ are obeyed, where J is the total angular momentum, consistent with what is expected for scattering of a spin 1 particle, the photon. This differs from the absorption case where the selection rules are $\Delta J = 0, \pm 1$.

A simple classical theory provides a heuristic explanation of the effect, although a quantum-mechanical treatment is required to understand the effect quantitatively. Let the molecule be described by a polarizability tensor $\tilde{\alpha}$, defined in Chapter 8. Incident light provides an electric field with amplitude E_0 , which induces an oscillating electric dipole

$$\mu = \tilde{\alpha}(\omega) \cdot \mathbf{E}_0 \exp(-i\omega t). \quad (\text{W22.66})$$

This dipole will radiate in accordance with the Larmor radiation formula. The energy emitted per frequency interval $d\omega$ is

$$U(\omega) = \frac{\omega^4}{12\pi\epsilon_0 c^3} |\tilde{\alpha}(\omega) \cdot \vec{E}_0|^2. \quad (\text{W22.67})$$

This is elastically scattered light and is called *Rayleigh scattering*. Now suppose that the molecule is allowed to vibrate in a particular normal mode with a vibrational frequency Ω (which is much less than ω). The polarizability tensor will also fluctuate at this frequency. Let Q be the normal-mode coordinate displacement associated with Ω . Then, to a first approximation,

$$\alpha(\omega, t) = \alpha_0(\omega) + \frac{\partial \alpha(\omega)}{\partial Q} Q \cos \Omega t. \quad (\text{W22.68})$$

The oscillating dipole now produces sidebands at frequencies $\omega + \Omega$ and $\omega - \Omega$, in addition to the oscillation at ω . The emission at these frequencies constitutes the Raman anti-Stokes and Stokes radiation, respectively. Rayleigh scattering occurs at frequency ω .

In the quantum-mechanical description the molecule is originally in the ground-electronic state in some vibrational state, and the light causes it to make a virtual transition to an excited-electronic state. This is followed by the molecule radiating a photon and falling into any vibrational state associated with the ground-electronic state. If the state happens to be the original one, it produces Rayleigh scattering. If it is to a higher-energy state, it is Stokes Raman scattering, whereas if it is to a lower-energy state, it is anti-Stokes Raman scattering. In Raman scattering the outgoing photon is either lowered in energy or raised in energy by the vibrational quantum $\hbar\Omega$. In order for anti-Stokes scattering to occur, there must be population in the excited vibrational state to begin with, which arises from thermal excitation. Stokes scattering can always occur. The ratio of the anti-Stokes to the Stokes scattering is given by the Boltzmann factor:

$$\frac{I_{\text{anti-Stokes}}}{I_{\text{Stokes}}} = \exp\left(-\frac{\hbar\Omega}{k_B T}\right). \quad (\text{W22.69})$$

Raman scattering is useful in condensed matter physics and chemistry in several instances. In solids the vibrational motions of the molecules are coupled and the excitations spread out in energy. In crystals they assume the character of phonons and are delocalized over the entire crystal. In highly disordered materials they may remain as localized oscillations extending over many nearby neighbors. The phonons may be categorized as being optical or acoustic. Raman scattering from the acoustic phonons is called *Brillouin scattering*.

For example, consider the scattering by conduction electrons in a lightly n -doped semiconductor. An electron may be virtually excited to some higher energy band and then reemit a different photon in returning to the original band. However, the wave vector of the photon is small compared with the size of the Brillouin zone. Therefore, there cannot be much of a change in the wave vector of the electron. It could emit an optical phonon with $\mathbf{k} = 0$, selection rules permitting. It could also produce Brillouin scattering. If anharmonic effects are taken into account, however, terms involving the simultaneous excitation of two phonons are also present. In terms of the simple classical model introduced earlier,

$$\alpha(t) = \alpha_0(\omega) + \sum_{i=1}^2 \frac{\partial \alpha(\omega)}{\partial Q_i} Q_i \cos \Omega_i t + \frac{1}{2} \sum_{ij} \frac{\partial^2 \alpha(\omega)}{\partial Q_i \partial Q_j} Q_i \cos \Omega_i t Q_j \cos \Omega_j t. \quad (\text{W22.70})$$

Sidebands now include terms with frequencies $\omega - \Omega_1 - \Omega_2$, among others. Extending this concept to solids implies that two-phonon production is possible. The net wave vector carried off by a pair of optical phonons may be small (i.e. $\mathbf{k}_1 + \mathbf{k}_2 = 0$). Thus light is able to create such a state with little momentum transfer.

Surface-enhanced Raman scattering (SERS) has emerged as a powerful tool for studying adsorbed species on the surfaces of solids. The Raman cross section for adsorbed species is found to be enhanced by as much as six orders of magnitude over the gaseous cross sections. Much of this enhancement is due to the increase in the strength of the local electromagnetic field at the surface over its value in free space. The amplification occurs because of local surface roughness, which creates miniature “lightning rods,” and also because of particular electronic resonances of the solid, such as surface plasmons. At frequencies approaching these resonances the surface acts as a high- Q resonator and has high-frequency (ac) electric fields due to the incident and outgoing radiation. There is also considerable evidence that the formation of the chemical bond between the adsorbed molecule and the substrate enhances the value of the Raman tensor, $\partial\alpha/\partial Q$.

An example of a Raman spectrum is given in Fig. W22.13. The intensity of the Raman scattering for diamond is plotted as a function of the frequency shift (in wave numbers). The Raman spectrum may be contrasted with the infrared absorption spectrum given in Fig. W22.12. The Raman spectrum is due to both a single-phonon process at ω_0 (as shown in the inset to Fig. W22.13) and to much weaker two-phonon processes. The one-phonon Raman peak at $\omega_0 = 1332.4 \text{ cm}^{-1}$ corresponds to the zone-center optic mode at $\approx 2.5 \times 10^{14} \text{ rad/s}$ of Fig. W22.12. Note that Raman scattering

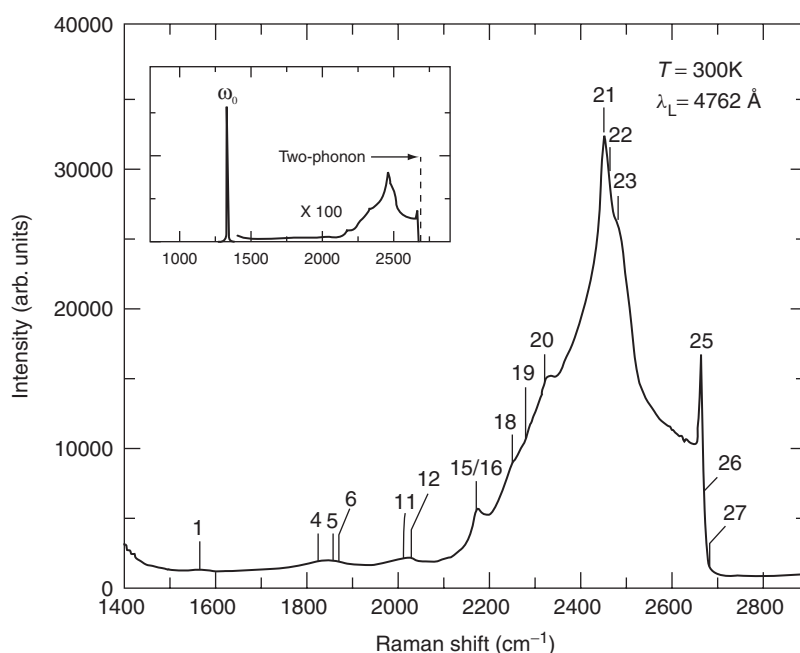


Figure W22.13. Raman spectrum for diamond at $T = 300 \text{ K}$. The incident light is polarized in the (111) plane. The backscattered light is in the [111] direction. [From R. Vogelgesang et al., *Phys. Rev. B*, **58**, 5408 (1998). Copyright 1998 by the American Physical Society.]

provides a much higher precision measurement of mode frequencies than does neutron scattering. See also Fig. 11.21, which gives a Raman spectrum for a Si–Ge alloy.

W22.10 Luminescence

Light is absorbed by materials and a fraction of the light is reemitted, usually with photons of lower frequencies. The process is called *luminescence*. The light may come out promptly, on a time scale of the order of a nanosecond, in which case the process is called *fluorescence*. It may come out on a much longer time scale, in which case it is called *phosphorescence*. Just how much light comes out depends on the nature of the competing channels for nonradiative decay. The quantum efficiency for luminescence may be defined as the ratio of the number of output photons per unit time to the number of input photons per unit time:

$$\eta = \frac{\dot{N}_{\text{output}}}{\dot{N}_{\text{input}}} \times 100\%. \quad (\text{W22.71})$$

In metals, where the excitation of electrons–hole pairs requires no activation energy, the nonradiative decay mechanism is probable and the quantum efficiency is very small. In semiconductors, where there is a substantial energy gap, the quantum efficiency may be quite large.

In Fig. W22.14 a typical luminescence process for a semiconductor is illustrated. An incident photon is absorbed by the solid, promoting an electron from a filled valence-band state (v) to a vacant conduction-band state (c). The photon must, in most instances, have an energy that exceeds the energy gap, E_g . The notable exception is the case where excitons exist just below the bottom of the conduction band. The processes above, in which an electron jumps from one band to the other band, is called an *interband process*. A hole is left behind in the valence band. The electron is generally produced in an excited state of the conduction band. By a sequence of phonon-emission processes the electron relaxes to the bottom of the band. Similarly, the hole migrates to the top of the valence band by such intraband processes. The time scale for these transitions is typically picoseconds or less. Luminescence takes place when the electron makes a radiative-decay transition from the bottom of the conduction band to the top of the valence band. The radiative lifetime is longer than a nanosecond.

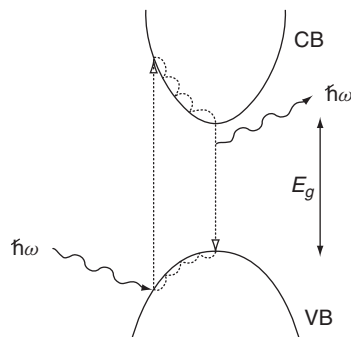


Figure W22.14. Luminescence in a direct-gap semiconductor.

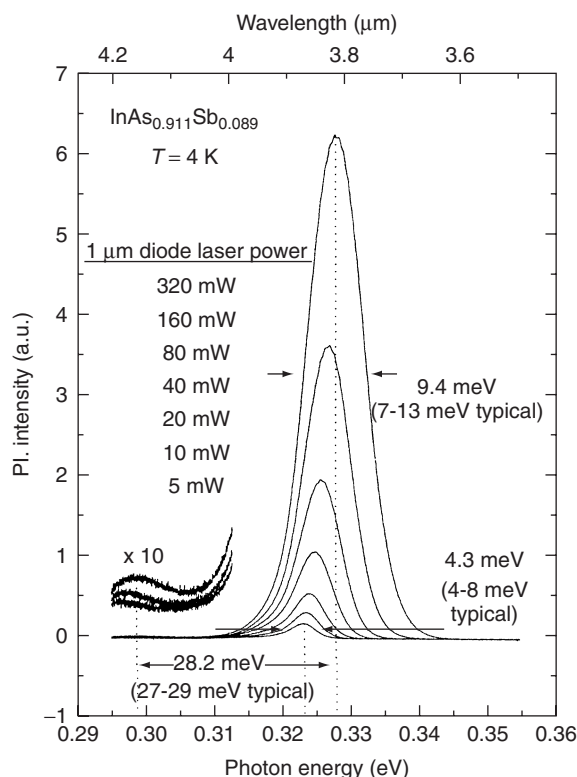


Figure W22.15. Photoluminescence spectra for MBE-grown $\text{InAs}_{0.911}\text{Sb}_{0.089}$ on a GaSb substrate at $T = 4$ K. [From M. A. Marciniak et al., *J. Appl. Phys.*, **84**, 480 (1998). Copyright 1998 by the American Institute of Physics.]

Hot luminescence occurs when the radiative recombination occurs not from the bottom of the band but from some excited state in the conduction band. If the relaxation occurs primarily with optical-phonon emission, a series of bumps will be seen in the emission spectrum, corresponding to the photon energy less some number of optical-phonon energies.

It is possible to study luminescence either in the frequency domain or in the time domain. In the latter case the procedure is called a *time-resolved luminescence study*. Luminescence may also be used to study defects. Cathodoluminescence is produced by an electron beam striking the surface of a solid.

An example of a photoluminescence spectrum is given in Fig. W22.15 for a film of $\text{InAs}_{0.911}\text{Sb}_{0.089}$ lattice-matched to a GaSb substrate. In addition to the main luminescence peak, there is a sideband lowered by the energy of a single LO phonon (≈ 28.2 meV). The narrow line width (≈ 5 meV) indicates that the material is of high quality.

W22.11 Nonlinear Optical Spectroscopy

With the advent of the laser it has become very easy to generate high-intensity electromagnetic fields. Materials no longer necessarily respond in a linear manner to

these fields, and it is important to understand their nonlinear properties. A number of phenomena are associated with nonlinear optics, such as second- and third-harmonic generation, three- and four-wave mixing, parametric excitation, self-focusing, self-phase modulation, and self-induced transparency, etc. Closely related to the pure nonlinear optical properties are the electro-optical and acousto-optical properties of materials. One is often interested in knowing how the optical properties of a material can be altered by applying electric fields or sound waves.

Attention will be focused on the polarization vector induced when an electric field exists in a medium. For a linear isotropic medium,

$$\mathbf{P}(\omega) = \epsilon_0 \chi(\omega) \mathbf{E}(\omega) \quad (\text{W22.72})$$

where $\chi(\omega)$ is the electric susceptibility. For an anisotropic linear material the corresponding formula is [see Eq. (8.44)]

$$\mathbf{P}(\omega) = \epsilon_0 \hat{\chi}(\omega) \cdot \mathbf{E}(\omega) \quad (\text{W22.73})$$

where $\hat{\chi}(\omega)$ is the electric susceptibility tensor. The anisotropy of this tensor is responsible for birefringence (i.e., the variation of the speed of light in a material with the polarization direction).

In nonlinear optics there are also nonlinear susceptibilities that may be defined. For example, there is the second-order susceptibility defined by [see Eq. (8.46)]

$$P_\alpha(\omega) = \epsilon_0 \sum_{\beta\gamma} \int d_{\alpha\beta\gamma}^{(2)}(\omega_1, \omega_2; \omega) E_\beta(\omega_1) E_\gamma(\omega_2) d\omega_1 d\omega_2. \quad (\text{W22.74})$$

This process describes the interaction of two photons of frequencies ω_1 and ω_2 in a material to create a photon of frequency ω . Energy conservation requires that

$$\omega_1 + \omega_2 = \omega. \quad (\text{W22.75})$$

For this process to proceed it is also necessary to guarantee wave vector conservation, that is,

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}. \quad (\text{W22.76})$$

This is called *phase matching*. The concept appears in Section W8.1, where the index ellipsoid is introduced. Methods for achieving phase matching in inhomogeneous media were discussed in Sections W20.6 and W20.8. Using the dispersion formula it implies that

$$\omega_1 \frac{n(\omega_1)}{c} + (\omega - \omega_1) \frac{n(\omega - \omega_1)}{c} = \omega \frac{n(\omega)}{c}. \quad (\text{W22.77})$$

This will, in general, not be valid for arbitrary frequencies. By rotating the crystal and making use of the different indices of refraction for the ordinary and extraordinary waves, however, it is possible to achieve phase matching.

A particular application of the second-order nonlinearity is in the process of second-harmonic generation. In that case [see Eq. (8.46)]

$$P_\alpha(2\omega) = \epsilon_0 \sum_{\beta\gamma} d_{\alpha\beta\gamma}^{(2)}(\omega, \omega; 2\omega) E_\beta(\omega) E_\gamma(\omega). \quad (\text{W22.78})$$

Depending on the symmetry of the crystal, there will only be a small number of independent components of $d_{\alpha\beta\gamma}^{(2)}$. The various components of the second-order polarization may be measured by focusing lasers of various polarizations onto a crystal volume and measuring the amount of second-harmonic light that is generated. Values for the $d_{\alpha\beta\gamma}^{(2)}$ components for various materials are given in Table 8.4.

The second-order polarizability exists only in crystals without inversion symmetry. The polarization vector \mathbf{P} , being a vector, should reverse its direction under a reflection operation, as should \mathbf{E} . But this is inconsistent with Eq. (W22.78), since the left-hand side changes sign but the right-hand side does not. In crystals with inversion symmetry $d_{\alpha\beta\gamma}^{(2)}$ is zero.

The third-order nonlinearity is described in terms of a fourth-order polarizability tensor defined analogously as

$$P_\alpha(\omega) = \epsilon_0 \sum_{\beta\gamma\delta} \int \int d_{\alpha\beta\gamma\delta}^{(3)}(\omega_1, \omega_2, \omega - \omega_1 - \omega_2; \omega) \times E_\beta(\omega_1) E_\gamma(\omega_2) E_\delta(\omega - \omega_1 - \omega_2) d\omega_1 d\omega_2, \quad (\text{W22.79})$$

where the phase-matching condition is

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = \mathbf{k} \quad (\text{W22.80})$$

and energy conservation requires that

$$\omega_1 + \omega_2 + \omega_3 = \omega. \quad (\text{W22.81})$$

The tensor $d_{\alpha\beta\gamma\delta}^{(3)}(\omega, \omega, \omega; 3\omega)$ may be determined by performing a third-harmonic generation experiment. Values for it appear in Table 8.5.

The application of an electric field to a crystal may alter the linear index of refraction of the crystal. This is of considerable technological importance since it implies that laser beams may be deflected electronically and at electronic frequencies. The degree to which the index of refraction changes when an electric field is applied to the crystal is determined by the electro-optic tensor (see Section 18.8).

The effective index of refraction for light propagating in a given direction \mathbf{k} with a given polarization vector $\hat{\epsilon}$ is defined in terms of the index ellipsoid. One constructs an imaginary ellipsoid in space (see Eq. (W8.12)):

$$\left(\frac{x}{n_x}\right)^2 + \left(\frac{y}{n_y}\right)^2 + \left(\frac{z}{n_z}\right)^2 = 1 \quad (\text{W22.82})$$

where x , y , and z define the coordinates in which the index of refraction tensor (related to the polarization tensor) is diagonal, and n_x , n_y , and n_z are the corresponding indices

of refraction. Draw a plane through the center of the ellipsoid perpendicular to \mathbf{k} . The plane intercepts the ellipsoid in an ellipse. The length of the vector from the center of the ellipsoid to the ellipse in the direction of $\hat{\varepsilon}$ is the value of n for that light ray.

Now introduce an electric field \mathbf{E} . The index ellipsoid will become stretched or compressed and will be rotated relative to the coordinates above. The new equation becomes

$$\left[\frac{1}{n^2}\right]_1 x^2 + \left[\frac{1}{n^2}\right]_2 y^2 + \left[\frac{1}{n^2}\right]_3 z^2 + 2\left[\frac{1}{n^2}\right]_4 yz + 2\left[\frac{1}{n^2}\right]_5 xz + 2\left[\frac{1}{n^2}\right]_6 xy = 1. \quad (\text{W22.83})$$

The dependence of these coefficients on \mathbf{E} is, for weak fields, a linear one. Thus

$$\left[\frac{1}{n^2}\right]_\alpha = \begin{cases} \frac{1}{n_\alpha^2} + \sum_{\beta=1}^3 r_{\alpha\beta} E_\beta, & \alpha = 1, 2, 3, \\ \sum_{\beta=1}^6 r_{\alpha\beta} E_\beta, & \alpha = 4, 5, 6. \end{cases} \quad (\text{W22.84a})$$

$$(\text{W22.84b})$$

The electro-optic tensor coefficients $r_{\alpha\beta}$ may be measured by passing a laser beam through a crystal with various orientations, applying an electric field, and measuring the beam deflection produced.

Using similar ideas, it is possible to study the photoelastic tensor, which is a tensor describing the variation of the index of refraction when a strain is introduced.

ELECTRON MICROSCOPY

Conventional optical microscopy is limited in its ability to resolve structure smaller in size than the wavelength of visible light, λ . The Rayleigh criterion is

$$\sin \theta \approx 1.22 \frac{\lambda}{d}, \quad (\text{W22.85})$$

which relates the acceptance angle of the microscope, θ , and the distance, d , between two points that can be resolved. Since visible light has wavelengths in the range 400 to 700 nm, light cannot be used to see individual atoms, whose size is typically 0.1 nm. If electromagnetic radiation is to be used to study materials, one may improve matters in two ways. The first is to use shorter-wavelength radiation. X-rays would be ideal, since their wavelength can be chosen to be comparable to the size of an atom. Another approach is to use very fine optical fibers tapered to a “point” whose size is ≈ 10 nm and then bring this fiber close to the surface of the material to be probed. The coupling is done through the near field of the electromagnetic field. Using this technique, 10-nm resolution can be achieved simply and inexpensively. The method is called *near-field scanning optical microscopy* (NSOM).

Another approach is to use electrons instead of light. The relativistic expressions for the wavelength of an electron are

$$\lambda = \frac{h}{p} = \frac{hc}{\sqrt{E^2 - (mc^2)^2}} = \frac{hc}{\sqrt{K(K + 2mc^2)}} = \frac{hc}{\sqrt{eV(eV + 2mc^2)}}, \quad (\text{W22.86})$$

where p is the momentum, E the total energy, K the kinetic energy, and V the potential difference through which the electron is accelerated to achieve this kinetic energy. By using 20-kV potentials, wavelengths of 0.009 nm are obtained, smaller than an atom. Thus resolution is no longer a limitation, but other factors, such as aberrations, prevent this fine resolution from being realized.

Electrons may be focused using electrostatic or magnetostatic lenses. The focal lengths of these lenses may be varied at will by changing the potentials and currents, respectively. It is therefore possible to construct electron microscopes in much the same way as optical microscopes are constructed. The main difference is that in electron microscopy the distance from the lenses to the sample is held fixed while the focal lengths are changed. In optical microscopy, of course, it is the other way around. The image in electron microscopy is usually obtained by rastering the beam across the sample and having the electrons collected by a detector. After amplification, the processed image is displayed on a fluorescent screen. High-vacuum conditions are needed for the electron beam to avoid collisions with gas molecules.

When high-energy electrons strike a material, they excite it and thereby lose energy. Bulk and surface plasmons can be excited. Interband transitions occur and electron-hole pair excitations are produced. There are also core-electron knock-out processes, which are followed by x-ray emission or Auger deexcitation. The Auger process is a multielectron process in which one electron fills an inner-shell vacancy, and one or more other electrons are ejected from the atom. Intraband transitions occur in metals. The net result is that copious amounts of secondary electrons are produced. In addition, there are backscattered primary electrons. Light may be emitted from the material when the electron-hole pairs recombine. If the sample is thin enough, a beam of electrons will be transmitted through the sample.

There are several methods for observing the sample. These include scanning-electron microscopy (SEM), transmission-electron microscopy (TEM), high-resolution transmission-electron microscopy (HRTEM), and low-energy electron microscopy (LEEM). These cases are discussed individually.

A number of typical electron micrographs using these techniques have appeared in Chapter 4. Figure 4.1*d* showed nanocrystalline diamond with a resolution of ≈ 100 nm. Figure 4.1*e* was a micrograph with atomic-scale resolution of the interface between crystalline Si and amorphous SiO₂. Figure 4.6 displayed nanocrystalline Au clusters embedded in an amorphous matrix. Figure 4.7 presented various morphologies of colloidal α -Fe₂O₃ particles. Figure 4.3 gave an HRTEM micrograph of a PbTiO₃-SrTiO₃ superlattice. Figure 4.9 showed the microstructure of a quasicrystal. Figures 4.20 and 4.21 presented images of a stacking fault and a twinned structure, respectively. These micrographs attest to the versatility of electron microscopy as a tool for studying the microstructure of materials.

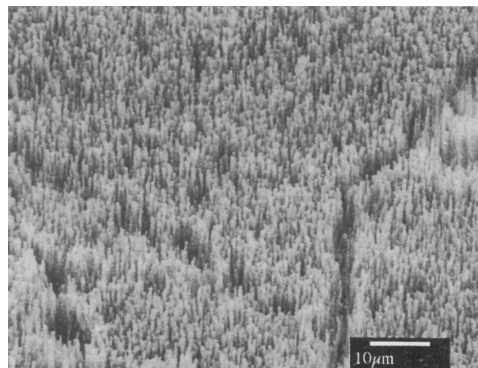
W22.12 Scanning-Electron Microscopy

The scanning-electron microscope (SEM) collects the backscattered and secondary electrons that are emitted from the surface of the material. Typically, a focused 5-nm-diameter beam with a current of 10^{-11} A is directed at the surface and penetrates the material. At first, when the electron is moving fast, high-energy processes such as Auger excitation are possible. Secondary electrons are produced, but backscattering is improbable at first because of the small Rutherford cross sections at high energies.

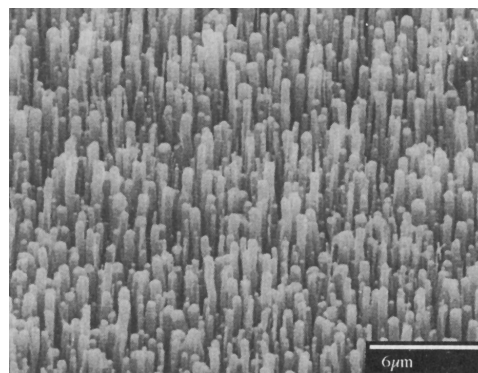
Note that single electron–electron collisions will not produce backscattering, whereas electron–ion collisions will. At high energies, where core-electron knockout is possible, it will often be followed by characteristic x-ray emission. Superimposed on this will be the brehmsstrahlung spectrum due to the rapid deceleration of the electrons. The energy of the beam spreads out and is deposited over a region $\approx 1\text{ }\mu\text{m}$ into the surface.

The easiest electrons to detect are the secondary electrons because they are numerous and all of low energy. Backscattered electrons are particularly useful because the cross section for backscattering depends on the atomic number of the target, in accordance with our understanding of Rutherford scattering. Thus materials with different Z will produce different amounts of backscattering. This provides a means for contrasting one type of atom with another.

In the scanning-electron microscope, one may deflect the electron beam so that it focuses on different parts of the sample. As it rasters over the surface, one detects the electrons, optical luminescence, or x-rays emitted as a function of the position where the beam is when these are produced. This often involves the use of a scintillation counter and photomultiplier tube. The data may be processed and a visual image of



(a)



(b)

Figure W22.16. SEM micrographs of carbon nanotubes on polycrystalline Ni substrates. [From Z. P. Huang et al., *Appl. Phys. Lett.*, **73**, 3845 (1998). Copyright 1998 by the American Institute of Physics.]

the surface presented. The depth of field and magnification are controlled by varying the focal lengths. Resolutions ≈ 10 nm are achievable.

Instead of gathering the electrons or radiation from the incident surface of the crystal, it is also possible to use a thin film and gather them from the opposite side of the crystal. This technique is called *scanning-transmission electron microscopy* (STEM).

Figure W22.16 illustrates a scanning electron micrograph of highly oriented, multi-walled carbon nanotubes on a polycrystalline Ni substrate. The nanotubes were grown by chemical vapor deposition. It is found that growth takes place within the grain surface but not along the grain boundaries.

W22.13 Transmission-Electron Microscopy

In the transmission-electron microscope (TEM) a thin slice of the material to be studied (0.1 to $0.5\ \mu\text{m}$ thick) is used and an energetic electron beam (20 to $100\ \text{keV}$) is passed directly through the sample. One may view either the image of the sample or the diffraction pattern that is produced on a fluorescent screen. Magnetic lenses are usually employed because it is possible to achieve short focal lengths with them. A schematic drawing of the TEM is presented in Fig. W22.17. The filament f heats the needle like cathode C , which emits electrons thermionically. The beam passes through a hole (called a *wehnelt cap*), which causes the beam to converge on and cross the optic axis (or reach some minimum spot size of radius $r'_c \approx 20\ \mu\text{m}$ for thermionic emitters). The corresponding point serves as a point source of electrons. The beam is then accelerated by an anode a and passes through a series of condenser lenses C_i and apertures A_i . It passes through a small portion of the sample and then a series of magnifying lenses M_i before it is projected on the fluorescent screen S and recorded on a photographic plate P . The sample is usually covered with a conductive coating in order to enable it to discharge electrically.

Typical cathode materials include W and LaB_6 . Their work functions are 4.5 and $2.7\ \text{eV}$, and their operating temperatures are 2800 and $\approx 1700\ \text{K}$, respectively. In some TEMs field emitters are used instead of thermionic emitters. They make use of Fowler–Nordheim tunneling from very fine cathode tips. Materials used are W and W covered with ZrO_2 . They may be operated at considerably lower temperatures, so the thermal spread of electron energies is considerably smaller than the $\approx 3\ \text{eV}$ for thermionic emitters. This allows one to obtain a much better minimum spot size (≈ 5 to $50\ \text{nm}$). Field-emission sources are brighter than thermionic sources. The respective brightnesses are typically $\approx 10^9$ and $\approx 10^{12}\ \text{A/m}^2\cdot\text{steradian}$. Typical vacuums for thermionic-emitter systems range from 10^{-2} to $10^{-4}\ \text{Pa}$. For field-emitter systems ultrahigh vacuums are established, typically in the range 10^{-7} to $10^{-8}\ \text{Pa}$.

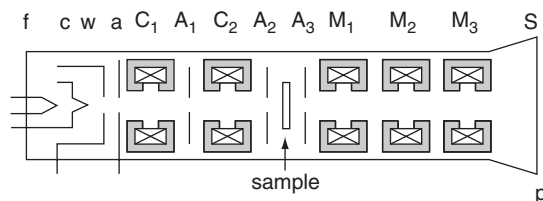


Figure W22.17. Schematic view of a transmission-electron microscope.

The focal length of a magnetic lens is given by the formula

$$\frac{1}{f} = \frac{e}{8mV} \int_{-\infty}^{\infty} B^2(z) dz, \quad (\text{W22.87})$$

where $B(z)$ is the magnetic field along the axis and eV is the energy of the electron as it passes through the lens. The magnetic lens also rotates the image by an angle ϕ given by

$$\phi = \sqrt{\frac{e}{8mV}} \int_{-\infty}^{\infty} B(z) dz. \quad (\text{W22.88})$$

The magnetic lenses have aberrations associated with them. These include the important spherical aberrations, chromatic aberrations, and astigmatism, as well as the less important coma and barrel or pincushion distortions. Spherical aberrations are due the fact that the rays scattered from an object point at small angles β through the lens are focused at a different image point than rays scattered at larger angles. The net result is that an object point along the axis is imaged to a disk of radius r_s rather than to a point. It may be shown that the size of this disk is given by the formula $r_s = C_s \beta^3$, where C_s is ≈ 1 mm. Unlike in the optical-lens case, there is no simple way to correct for spherical aberrations.

Chromatic aberrations are due to a variation of focal length with beam energy, and may be understood in terms of Eqs. (W22.87) and (W22.88). Inelastic losses of $\Delta V \approx 5$ to 50 eV are common for electrons passing through the sample, due to the excitation of electron-hole pairs, plasmons, phonons, and so on. The net result is that an object point is imaged to a disk-of-confusion whose radius is r_c . It may be shown that $r_c = C_c \beta \Delta V / V$, where C_c is a distance characterizing the axial variation of the magnetic field in the lens.

The imprecision with which magnetic lenses can be constructed leads to asymmetries being present. Astigmatism is caused by having a different focal length for electrons deflected in the x direction than in the y direction, the difference being Δf_A . A point source is focused to a disk of minimum size $r_A = \beta \Delta f_A$. It may usually be corrected by employing a balancing astigmatic lens.

The Rayleigh diffraction criterion gives an estimate for the minimum separation of two resolvable points, $\Delta r_d = 0.61\lambda/\beta$, where λ is the wavelength of the electrons. One may obtain an estimate for the instrument resolution by assuming that the spherical aberration and diffraction dominate. Then

$$\Delta r(\beta) = \sqrt{\left(\frac{0.61\lambda}{\beta}\right)^2 + (C_s \beta^3)^2}. \quad (\text{W22.89})$$

There is a competition between the diffraction of the beam and the spherical aberrations of the lenses. The minimum value of Δr occurs when $\beta = (0.61\lambda/C_s\sqrt{3})^{1/4}$ and its value is $\Delta r_{\min} = 0.91C_s^{1/4}\lambda^{3/4}$. The smaller λ can be made, the better the resolution will be. Typical optimal instrumental resolutions are of the order of several tenths of a nanometer. For 100-keV beams $\Delta r = 0.33$ nm is possible, but the samples must be less than 5 nm in thickness. For 400-keV beams, a resolution of 0.17 nm has been obtained. By using beams of incoherent rather than coherent electrons and by underfocusing the

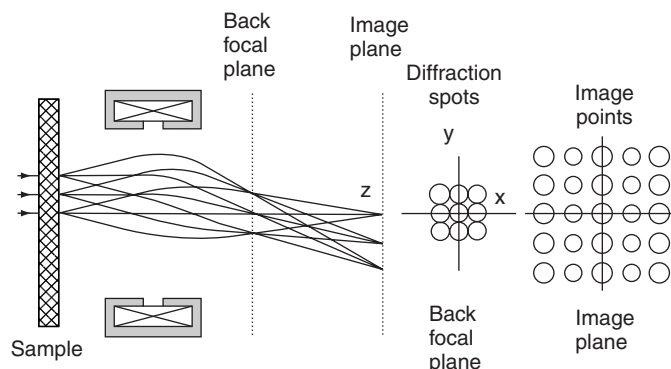


Figure W22.18. A forward-directed beam and diffracted beams are produced when the incident electron beam passes through the sample. The right figures show the diffraction spots in the back focal plane and the sample image points in the image plane.

electron beam to compensate partially for spherical aberrations, it has recently been possible to obtain 0.078-nm resolution with a 300-keV electron beam.[†]

After passing through the sample, the diffraction spots create images on the back-focal plane of the object lens (e.g., C_2 of Fig. W22.17). One may use the magnifying lenses to view this diffraction pattern directly on the screen and to produce a transmission-electron diffraction (TED) micrograph. Alternatively, a real image of the sample is produced on the image plane of the object lens. This real image may be magnified and projected on the fluorescent screen.

Emerging from the sample is both a forward-directed transmitted beam of electrons and diffracted beams of electrons, as illustrated in Fig. W22.18. The aperture A_3 of Fig. W22.17 may be used to choose one or more of these beams selectively and block the others. The technique is called *selective-area diffraction* (SAD). In the method of bright-field (BF) imaging one allows the forward-directed beam to be imaged on the screen. In the method of dark-field (DF) imaging, a diffracted beam is selected instead. The diffraction pattern in the back-focal plane and the image of the sample are depicted in the right-hand side of Fig. W22.18. The central diffraction spot corresponds to the forward-directed beam. The other spots correspond to various diffracted beams for a case where there is fourfold symmetry. As will be seen in Section W22.14, the more beams that are accepted by the aperture, the higher the resolution will be.

Kikuchi Lines. Diffraction of electrons differs from that of x-rays in two important respects. First, the samples are thin, so there are only a finite number of atomic layers, N_z , perpendicular to the beam. Second, inelastic processes are much more important for electrons than for x-rays.

Due to the finite value of N_z (≈ 1000), instead of there being a pure spot diffraction pattern, the spots are elongated into streaks of length $\Delta G_z \approx 2/N_z d$, where d is the lattice spacing. Thus it is still possible to see a diffraction spot even when the von Laue condition is not exactly satisfied. One defines the mismatch reciprocal vector \mathbf{s}

[†] P. D. Nellist and S. J. Pennycook, *Phys. Rev. Lett.*, **81**, 4156 (1998).

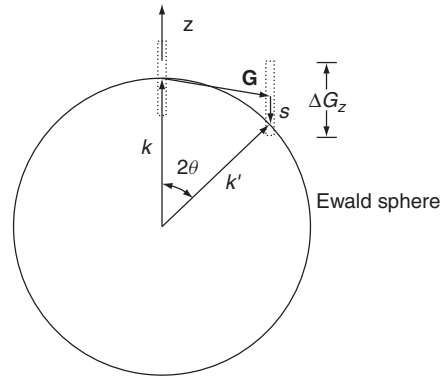


Figure W22.19. Ewald sphere and streaking of diffraction spots.

through the relation $\mathbf{k}' - \mathbf{k} = \mathbf{G} + \mathbf{s}$. This is illustrated in Fig. W22.19. The vectors \mathbf{k} and \mathbf{k}' , making an angle 2θ with each other, have a common origin and have the same length. The head of vector \mathbf{k} touches the origin of reciprocal-lattice space. The Ewald sphere passes through this origin. A diffraction spot will still be produced if the head of vector \mathbf{k}' also touches the Ewald sphere, but is located a distance $s \leq \Delta G_z$ in the z direction away from the nearest reciprocal lattice point, \mathbf{G} . The extent of ΔG_z is denoted by the dashed rectangle in the figure.

Suppose the electron is incident on the lattice in a direction that is far removed from satisfying the von Laue diffraction condition. Electrons can suffer inelastic collisions, thereby losing energy and scattering into various directions centered around the incident direction. Energy losses of up to 50 eV are common. The probability for scattering through a given angle relative to the forward direction falls rapidly with increasing angle. An effect of this is illustrated in Fig. W22.20, where an electron suffers an inelastic collision at point p . For some scattered beams the energy and direction will be just right to satisfy the Bragg diffraction condition for a set of lattice planes. Two such planes L and L' are illustrated in Fig. W22.20. The beams B and B' are able

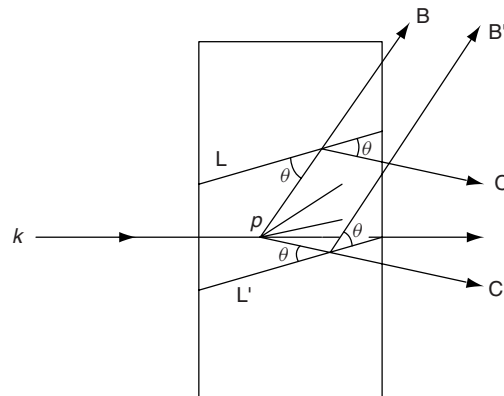


Figure W22.20. Formation of Kikuchi lines by Bragg diffraction and constructive interference of beams of inelastically scattered electrons.

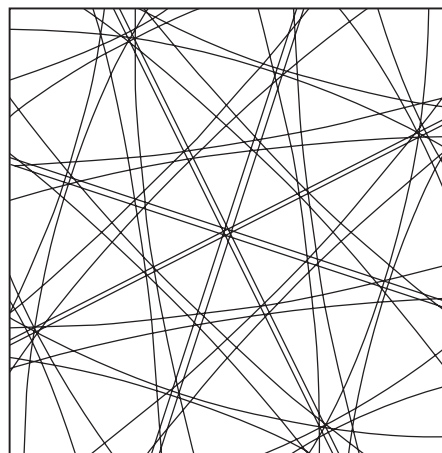


Figure W22.21. Kikuchi line pattern. (Generated using the computer program KOQUA2.2 written by S. Weber and C. Schetelich.)

to interfere with each other, as are the beams C and C'. If the Bragg condition is satisfied for the direction of beams B and B', a bright spot will be produced in that direction. The intensity comes at the expense of radiation that was in the near-forward direction, so there will be a corresponding dark spot established along the direction of beams C and C'. The angular spread between the lines C and B is 2θ , independent of the angles through which the electrons are scattered. When one takes into account all possible angles of incidence for the electrons, the beams B and C sweep out the surfaces of cones. These cones intersect the projection of the Ewald sphere in a pair of hyperbolas. For high-energy electrons the radius of the Ewald sphere is large and it looks approximately planar. The intersections of the cones with the plane then produce a pair of approximately parallel lines. This gives rise to the *Kikuchi line pattern*. To each set of lattice planes there is a pair of Kikuchi lines, one dark and one bright. Furthermore, these lines are parallel to the intersection of the corresponding lattice planes with the Ewald sphere.

The Kikuchi line pattern shifts if the crystal is tilted. This permits one to orient a crystal precisely. An example of a Kikuchi line pattern, together with a set of diffraction spots, is illustrated in Fig. W22.21. Furthermore, at the intersection of Kikuchi lines, one finds spot patterns. Thus the Kikuchi patterns are used as maps to locate the orientations of the crystal, which give rise to diffraction-spot patterns.

W22.14 High-Resolution Transmission-Electron Microscopy

In high-resolution transmission-electron microscopy (HRTEM) a beam of high-energy electrons is passed through a thin sample and focused on an image plane. Suppose that the sample is a crystal. Recall that the diffraction condition is

$$\mathbf{G} \cdot \mathbf{k} + \frac{G^2}{2} = 0. \quad (\text{W22.90})$$

If k is much larger than G , this equation can be satisfied only for those lattice planes whose \mathbf{G} vectors are almost perpendicular to \mathbf{k} , for that is the only way to keep the

magnitude of the first term comparable to that of the second term. The angles between \mathbf{k} and \mathbf{k}' are given by

$$\sin \theta = \frac{G}{2k} \quad (\text{W22.91})$$

and will be small.

As mentioned earlier, magnetic lenses have associated with them spherical aberrations. A beam far from the optical axis will not focus at the same point as a beam near the optical axis. In conventional microscopy, therefore, it is usually not possible to focus the Bragg-reflected beams along with the directly transmitted beam. This puts a limitation on the resolution. However, in HRTEM the Bragg peaks are coming off at small angles and therefore remain paraxial (i.e., close to the optic axes). The spherical aberrations are therefore not of major concern and it is possible to focus several Bragg orders together (Fig. W22.22).

In the figure, the sample is labeled S, the lens, L, the focal plane, F, and the image plane, I. An electron beam is incident on the sample from the left and is focused on the image plane. Note that the various Bragg-reflected beams combine with the direct beam in the image plane. This causes a sharply defined focal spot. The reason for this is that slightly away from the focal spot the various beams start to interfere with each other destructively and the intensity decays rapidly with distance away from the spot. The more diffracted beams that can be collected, the sharper the image. Once the image is formed on the image plane, further magnification is possible by the use of additional lenses, as in the case of TEM.

The angular sharpness of the image varies inversely as the number of diffracted beams in a given direction that may be focused. This is illustrated with a two-dimensional example. Consider Fig. W22.23, in which a set of N beams passes through the lens at positions y_i and is focused at the lower point on the screen. At this point all the beams arrive in phase with each other:

$$\psi = \phi_i + kl_i, \quad i = 1, \dots, N, \quad (\text{W22.92})$$

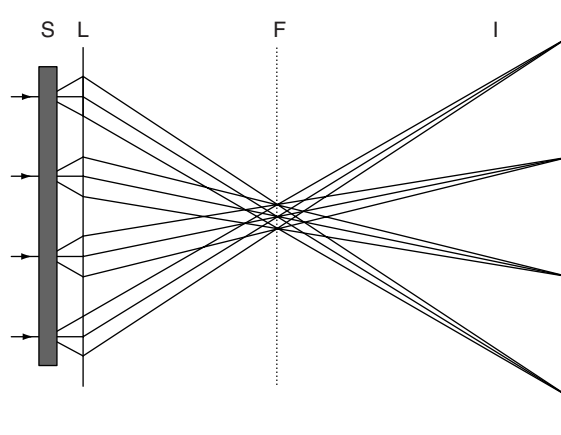


Figure W22.22. A large number of diffracted beams are focused by the lens L and are combined to form a high-resolution image of the sample S on the image plane I.

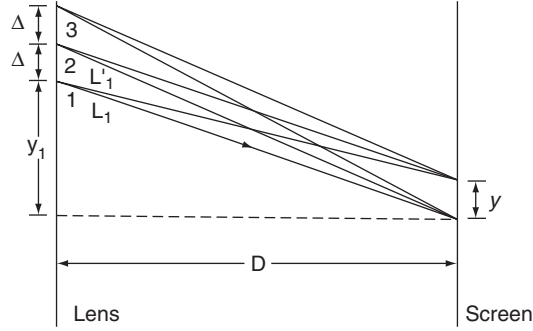


Figure W22.23. Estimation of the angular sharpness of the image.

where the $(\phi_i + kl_i)$ differ by integer multiples of 2π . The total amplitude at the point is therefore

$$A = \sum_{i=1}^N \exp(i\psi) = N \exp(i\psi). \quad (\text{W22.93})$$

Next consider a point on the screen a distance y above the original point. The phase that each beam arrives with is now different:

$$\psi_i = \phi_i + k\sqrt{(y_i - y)^2 + D^2} \simeq \psi - k\frac{yy_i - y^2/2}{D}, \quad (\text{W22.94})$$

where it is assumed that D is much larger than y_i . The amplitude at the upper point is therefore

$$A' = \sum_{n=1}^N \exp \left[i\psi - \frac{ik}{D} \left(yy_n - \frac{y^2}{2} \right) \right]. \quad (\text{W22.95})$$

The points y_n on the lens are separated from each other by an arbitrary distance Δ . This amplitude will fall to zero when the phases are spread uniformly over a circle, that is, when

$$\frac{kyN\Delta}{D} = 2\pi, \quad (\text{W22.96})$$

which shows that y is inversely proportional to N :

$$y = \frac{\lambda D}{N\Delta}. \quad (\text{W22.97})$$

In the three-dimensional case the diffracted beams will be spread out over an area. The total number of beams will be denoted by N^2 . Thus the size of the image falls off as the inverse of the square root of the number of focused diffracted beams.

It is now possible to use HRTEM to obtain spatial resolution approaching 0.1 nm, if beams of 1 MeV are used, although most conventional HRTEM applications use lower-energy beams and settle for more modest resolution goals. It is an ideal tool for studying line defects, planar defects, and interfaces.

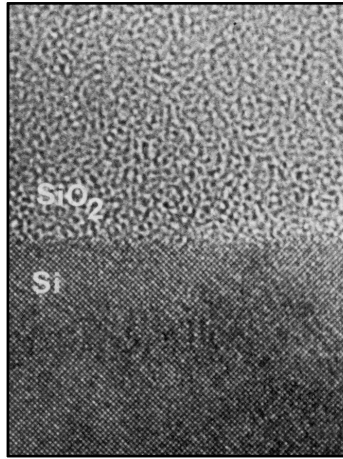


Figure W22.24. HRTEM image of the Si/a-SiO₂ interface. (From J. M. Gibson, High resolution transmission electron microscopy, *Mater. Res. Soc. Bull.*, Mar. 1991, p. 27.)

An HRTEM image of a Si/a-SiO₂ interface is shown in Fig. W22.24. The individual atoms of the Si crystal appear in the lower half of the image. The upper half shows the image of the amorphous silica.

W22.15 Low-Energy Electron Microscopy

Like LEED, low-energy electron microscopy (LEEM), involves forming a diffraction pattern using elastically scattered electrons from a crystalline surface. Unlike LEED, however, most of the pattern is discarded. Usually, one diffracted beam is extracted and imaged on a screen using conventional electron-microscope lenses. Typical energies used are in the range 100 eV to 3 keV. The resolution is as fine as 2 nm. LEEM is useful for seeing structure on a mesoscopic size scale. This includes surface steps, dislocations, imperfections, islands of adsorbates, superlattice structure, grains, and surface inhomogeneity.

In LEEM the electron beam is directed at the sample with near-normal incidence. Electrons are emitted from an electron gun from an oblique direction to the surface, and a bending magnet is used to change the direction to normal incidence. The same bending magnet is used to redirect the reflected electrons in another oblique direction toward the image plane. In Fig. W22.25 the basic imaging scheme is displayed for the case where LEEM is used to image a surface step, S, on the left. For simplicity the injection, bending magnet, and extraction are not shown, and it is simply assumed that the electron beam is incident from the right. Four reflected beams are illustrated, each with a pair of diffracted beams. The beams pass through an electron lens, L, and then through a screen, A, with an aperture in it. The aperture is placed in the focal plane of the lens. Only the specularly reflected rays are allowed through, the other diffracted beams are blocked. The rays that pass through the aperture illuminate the screen I. In passing through the aperture, the beams undergo Fresnel diffraction and are broadened into cones. Thus beam 1 is spread into cone 1, beam 2 into cone 2, and so on. The regions illuminated by these cones are denoted 1', 2', and so on.

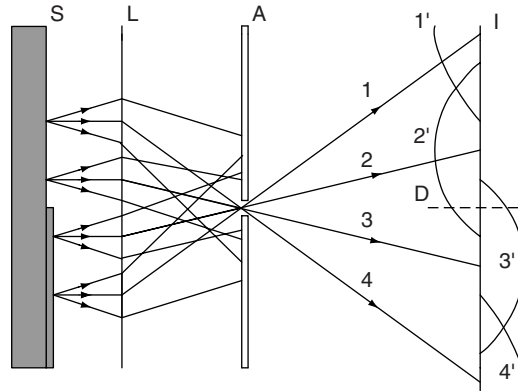


Figure W22.25. Optics of the formation of a LEEM image.

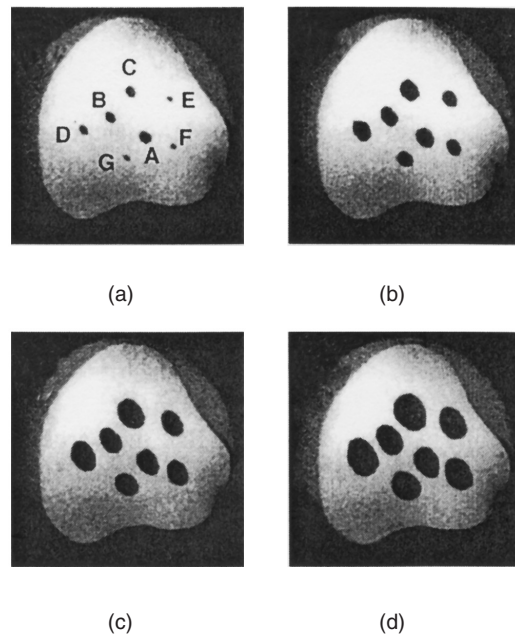


Figure W22.26. LEEM micrograph of the etching of a terrace on Si(100) by oxygen at $T = 1235$ K. The images are those of vacancy islands at (a) 1 s, (b) 10 s, (c) 20 s, and (d) 30 s after nucleation. [From J. B. Hannon et al., *Phys. Rev. Lett.*, **81**, 4676 (1998). Copyright 1998 by the American Physical Society.]

The figure illustrates the situation where rays 1 and 2 are reflected from the part of the step closer to the lens, while rays 3 and 4 come from the part farther from the lens. Positive amplitude is indicated to the left of the image plane and negative amplitude to the right of the image plane. Suppose that the energy of the electrons is adjusted so that the step size is one-fourth of a wavelength. This would cause rays 1 and 2 to be half a wavelength out of step with rays 3 and 4 when they hit the image plane, I. At

point D on the image plane, the net amplitude is zero. It will therefore show up as a dark line. This line is the phase-contrast image of the step on the surface.

Although there are other means of creating the phase contrast, such as defocusing by a small amount, the foregoing scheme illustrates the basic method of how a surface may be imaged using low-energy electrons in a microscopy arrangement. In practice a small area of the sample is illuminated with the incident beam. Information from the diffraction pattern is then processed. The beam is rastered over the sample and data are stored for presentation. The spatial resolution is a function of the electron energy used, varying from 60 nm at 250 eV to 2 nm at 30 keV.

An illustration of a LEEM micrograph is given in Fig. W22.26. The dark-field micrographs show various stages of the nucleation of vacancy islands formed during the etching of a 10 μm terrace on the Si(100) surface in an oxygen atmosphere.

ELECTRON SPECTROSCOPY AND ION SCATTERING

In the following sections we describe methods for obtaining the energy distribution of charged particles. These distributions provide important information about the elementary excitations of the solid. In photoemission experiments a beam of electromagnetic radiation is used to produce energetic electrons that are emitted from the surface and are analyzed and detected. Both ultraviolet radiation and x-rays are used. Low-energy electron beams are scattered from solids to provide information concerning the surface and adsorbates on the surface. Extended x-ray absorption fine structure may be used to obtain accurate information about short-range order in solids. Auger emission spectroscopy is an important tool for obtaining quantitative information concerning the chemical composition on or near surfaces. Secondary-ion mass spectrometry and Rutherford backscattering provide additional information regarding the chemical composition and defect structure.

W22.16 Photoemission

Photoemission involves the absorption of a photon by a material and the immediate emission of an electron into vacuum. It has been studied in some detail in Section 19.9. The energy spectrum and photoelectron yield are measured, often as a function of photon energy. Photoemission may be carried out with ultraviolet radiation, in which case it is called *ultraviolet photoemission spectroscopy* (UPS), or with x-rays, in which case it is called *x-ray photoelectron spectroscopy* (XPS) or *electron-spectroscopy for chemical analysis* (ESCA). Since the mean free path of electrons is limited in materials, photoemission provides information concerning the surface region of the solid, especially in the case of UPS. Photoemission may be used to study either crystalline or amorphous solids. It is not useful for liquids because of the need to have a good vacuum present, so that electrons may reach the detector without making collisions with gas molecules.

Ultraviolet Photoemission Spectroscopy (UPS). In UPS electrons are promoted from occupied states below the Fermi level to states above the vacuum level. The photon's energy must exceed the work function $e\Phi$ of the material being studied. The maximum energy the electron may have is given by a famous formula of Einstein:

$$E = \hbar\omega - e\Phi, \quad (\text{W22.98})$$

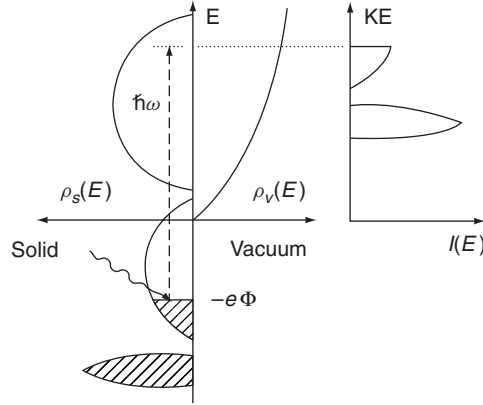


Figure W22.27. Photoemission from a metal with an occupied valence band and a partially occupied conduction band. The density of electron states in the solid and vacuum, and the energy-distribution curve $I(E)$ are shown.

where $\hbar\omega$ is the energy of the incident photon. Since the energy of the ultraviolet photon is relatively small, electrons are extracted from the conduction band and the upper valence bands. A schematic of the photoemission process is given in Fig. W22.27. Three quantities are sketched in this figure. The left-hand side shows the density of states in the solid, $\rho_s(E)$. The vacuum level is taken to be the zero of energy. The Fermi level lies at energy $-e\Phi$. Those states below the Fermi level are occupied and are shaded on the diagram. The density of states in the vacuum $\rho_v(E)$ is also sketched in the figure. It corresponds to that of a free electron. On the right-hand side of the figure the energy distribution curve of the emitted electrons, $I(E)$, is sketched. Ideally, this curve is (aside from possibly smoothly varying distortions due to the energy dependence of the dipole matrix elements) a replica of the density of states of the solid below the Fermi energy. More realistically, there are significant contributions due to secondary electrons.

A formula for the energy-distribution curves may be derived from Fermi's golden rule. The rate of absorption of photons is

$$\Gamma(\omega) = \frac{2\pi}{\hbar} \sum_{i,f} \sum_s |M|^2 \delta(E_f - E_i - \hbar\omega) f(E_i, T) [1 - f(E_f, T)], \quad (\text{W22.99})$$

where M is the dipole matrix element of the interaction of the photon with the electron, and i and f refer to the initial and final states of the electron, respectively. There is a sum over the two spin states, s , of the electron. The radiation interaction preserves spin projection. There are also Fermi–Dirac distribution function factors introduced in Chapter 7 and Appendix WB,

$$f(E, T) = \frac{1}{\exp[\beta(E - \mu)] + 1}. \quad (\text{W22.100})$$

Here μ is the chemical potential (approximately equal to the Fermi energy, E_F , at low temperatures). The first Fermi factor guarantees that there is an electron in state i , the

second factor guarantees that state f is empty, so a transition can occur. Introduce the electron density of states $\rho(E)$ as in Eq. (7.67). The absorption rate may be expressed as

$$\Gamma = \int \Gamma(E') dE', \quad (\text{W22.101})$$

where $\Gamma(E') dE'$ is the rate of absorption of photons leading to excited electrons within the energy band E' to $E' + dE'$. This rate is given by

$$\Gamma(E') = \frac{\pi}{\hbar} \overline{|M|^2} \rho_v(E') \rho_s(E' - \hbar\omega) f(E' - \hbar\omega, T) [1 - f(E', T)], \quad (\text{W22.102})$$

where an average squared matrix element is used as an approximation. The rate of producing photoemitted electrons is

$$I(E') = \Gamma(E') P(E'), \quad (\text{W22.103})$$

where $P(E')$ is the probability that if a photoelectron is produced, it will emerge from the surface.

The graph of $I(E')$ versus E' is called the *energy-distribution curve* (EDC). The previous formulas show that $I(E')$ is proportional to the product of the density of states for the initial and final states. If the photon energy is sufficiently high, the final density of states may be approximated by a free-electron density of states $\rho_v(E') \propto (E')^{1/2}$. The energy-distribution curve may then be used to determine the density of states $\rho_s(E' - \hbar\omega)$ below the Fermi surface.

The total photoelectric current divided by the incident current of photons is called the *photoelectric yield*. It is seen to be proportional to the joint density of states,

$$I(\omega) \sim \frac{\pi}{\hbar} \overline{|M|^2} P \int \rho_s(E') \rho_v(E' - \hbar\omega) f(E' - \hbar\omega, T) [1 - f(E', T)] dE', \quad (\text{W22.104})$$

where an average escape probability factor P has been extracted from the integral.

As the electron leaves the solid, it can undergo inelastic-scattering processes with other electrons. Some of these other electrons emerge as secondary electrons. One therefore finds a large number of low-energy secondary electrons emerging from the solid as well as the photoemitted electron. The energy-distribution curve therefore rises at low energies.

In some experiments the angular distribution of the emitted electrons is analyzed as well as the energy distribution. The study is called *angular-resolved photoemission spectroscopy* (ARPES). This is particularly useful for obtaining information about the surface layer of the solid or atoms or molecules adsorbed on the surface. Different orbitals in these atoms or molecules point in different directions, and this influences the emission pattern. For example, those orbitals pointing perpendicular to the surface are more likely to photoemit electrons in a direction perpendicular to the surface. This can reveal interesting information regarding the nature of the chemical bonds or the particular bonding sites of adsorbed species.

An example of a UPS spectrum is given in Fig. W22.28 for sputter-deposited $\text{Ge}_{100-x}\text{Ag}_x$ with $0 \leq x \leq 39.6$ at room temperature. The spectra were taken with 21.2-eV photons from a He I ultraviolet light source. The Ge 4*p* valence band extends from

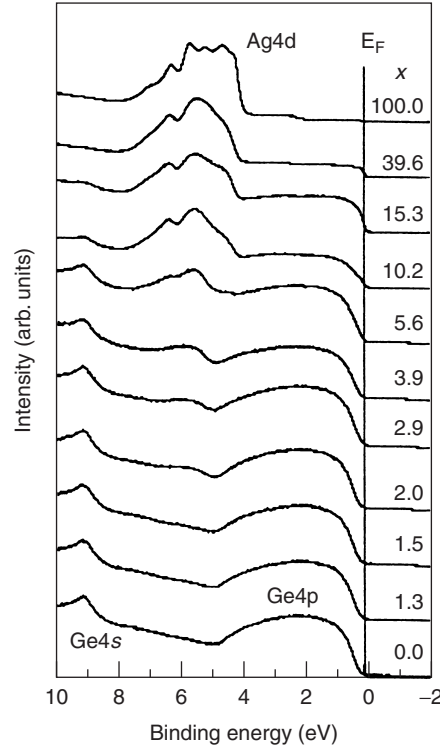


Figure W22.28. Ultraviolet photoemission spectrum of sputter-deposited $\text{Ge}_{100-x}\text{Ag}_x$ for $0 \leq x \leq 39.6$ [From A. Suzuki and K. Tanaka, *Jpn. J. Appl Phys.*, **37**, 4872 (1998). Copyright 1998 by the Japanese Journal of Applied Physics.]

a binding energy of 4.5 to 0 eV, and the 4s band is at a binding energy of 9 eV. The peak that develops at 5.5 eV is due to the Ag-4d band. For $0 \leq x \leq 5.6$ the spectra show that Ag is dissolved in a Ge matrix, since a single Ag-4d peak appears. For $x \geq 5.6$, phase separation occurs as silver clusters begin to form and the UPS spectrum evolves toward that of bulk Ag, shown at the top of the figure.

X-ray Photoemission Spectroscopy (XPS or ESCA). Often, x-rays rather than UV light are used in a photoemission experiment. The high energy of the x-ray permits the observation of photoemitted core electrons of the solid. The bandwidths of the core electrons are very narrow and the levels may be approximated as having a single energy, E_{core} . The energy of the emitted electron is

$$E' = \hbar\omega - E_{\text{core}}. \quad (\text{W22.105})$$

For a given x-ray photon energy $\hbar\omega$ there will be a sharp peak in the EDC.

The precise value of the core energy is sensitive to the distribution of valence electrons surrounding the core. To photoionize the core electron, the electron must exit the atom by passing through the valence shells. There is a difference of potential between the core and the outside world determined by the charge distribution of the valence electrons. To get a qualitative feeling for this, consider a simple example.

Suppose that a distribution of valence electrons is described by a charge distribution $\rho(\mathbf{r})$, which will be taken to be spherically symmetric, for the sake of simplicity. Poisson's equation gives the potential

$$\nabla^2 V(r) = \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial}{\partial r} V(r) \right] = -\frac{\rho(r)}{\epsilon_0}. \quad (\text{W22.106})$$

Here $V(r)$ is the contribution to the potential from the valence electrons. The contribution to the potential due to the nucleus is fixed, and will be ignored. Taking the position of the core to be approximately at $r = 0$, this gives a difference of potential

$$V(\infty) - V(0) = -\frac{1}{\epsilon_0} \int_0^\infty \frac{1}{r^2} \int_0^r \rho(r') r'^2 dr' dr. \quad (\text{W22.107})$$

For example, suppose that the valence-electron charge distribution is given by

$$\rho(r) = -Q \frac{\mu^3}{8\pi} \exp(-\mu r), \quad (\text{W22.108})$$

so that the total valence charge is $-Q$. The parameter μ in this model represents the inverse of the length over which the valence charge distribution decays outside the atom in question. Then the difference of potential will be

$$V(\infty) - V(0) = Q \frac{\mu}{8\pi\epsilon_0}. \quad (\text{W22.109})$$

The energy of an electron residing in the core may be written as the sum of a constant plus the difference in potential energy between the electron at the core position and the electron at infinity:

$$E_{\text{core}} = \text{constant} - e[V(0) - V(\infty)]. \quad (\text{W22.110})$$

For the model above, therefore,

$$E_{\text{core}} = \text{constant} + \frac{eQ\mu}{8\pi\epsilon_0}. \quad (\text{W22.111})$$

For more compact charge distributions μ will be larger and the core level will be shifted upward (i.e., less tightly bound). Correspondingly, for more spread-out valence charge distributions, the core level will be lowered. In forming chemical bonds, the electron distribution around atoms is distorted. This gives rise to core-level shifts characteristic of the particular bonds that are formed. By measuring the difference between the energy of the incident photon and the emitted electron, the energy of the core level may be found.

Examples of x-ray core-level spectra are given in Fig. W22.29. Data for $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ are taken at $T = 300$ K, where it is semiconducting, and $T = 80$ K, where it is superconducting. The spectrum focuses on the Cu $2p_{3/2}$ state. The data provide evidence for a change of valence state with temperature.

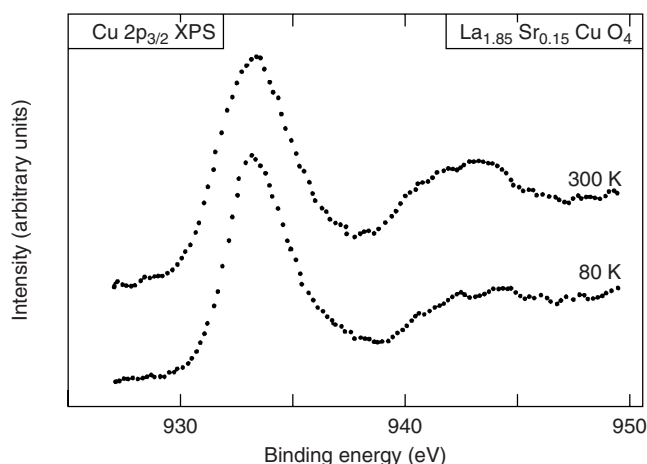


Figure W22.29. X-ray core-level spectroscopy of $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ at $T = 300\text{ K}$ and $T = 80\text{ K}$. [From D. D. Sarma, *Phys. Rev. B*, **37**, 7948 (1988). Copyright 1988 by the American Physical Society.]

W22.17 Low-Energy Electron Loss Spectroscopy

As in LEED, the technique of low-energy electron loss spectroscopy (LEELS) involves directing a beam of electrons at a surface. In LEELS, however, the energy loss of the electron is studied rather than the elastic scattering. Electrons of energy E impinge on a solid, making an angle θ with respect to the surface and come off at a variety of angles. A detector is positioned so it accepts electrons that emerge at an angle θ' and an azimuthal angle ϕ (Fig. W22.30). The current of the scattered beam, I , is then analyzed as a function of the energy of the electron, E' . LEELS data generally can consist of a table of $I(E', \theta', \phi)$ as a function of E and θ , but more often are presented as an angular-integrated function $I(E')$, showing loss peaks. As with LEED, LEELS provides information primarily about what is occurring on or near the surface.

When the electron scatters from the surface, it may emit (or absorb) an elementary excitation from the solid. This excitation is usually a phonon, but other types of

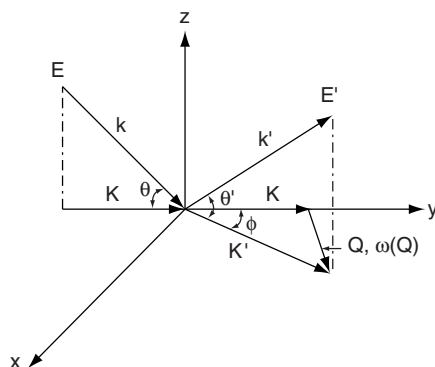


Figure W22.30. Scattering geometry for a LEELS experiment.

excitations, such as two-dimensional plasmons associated with charged layers on the surface, are also possible. The excitation carries with it both energy and momentum. In general, the LEELS spectrum consists of energy-loss peaks from three origins: bulk excitations of the substrate, surface excitations of the substrate, and excitations of adsorbed species on the surface. Because of the limited penetration of electrons into the solid, LEELS is particularly useful for studying the latter two surface excitations.

Surface excitations of the substrate are characterized by having a wave vector parallel to the surface, \mathbf{Q} , and a frequency $\omega(\mathbf{Q})$. For the case of a periodic lattice there is conservation of wave vector in the plane of the surface, modulus a reciprocal-net vector (i.e., surface reciprocal-lattice vector):

$$\mathbf{K}' = \mathbf{K} + \mathbf{Q} + \mathbf{G}, \quad (\text{W22.112})$$

where \mathbf{K} and \mathbf{K}' are the surface components of \mathbf{k} and \mathbf{k}' . In the case of surface adsorbates, unless the adsorbates form an ordered net, there will be no wave-vector conservation.

In the following, attention will be restricted to the case where there is energy loss. Energy gain, however, is possible if the temperature of the surface is high enough for a thermal excitation to be present and absorbed by the electron. The basic equation of LEELS is the energy conservation condition:

$$E' = E - \hbar\omega. \quad (\text{W22.113})$$

For example, in the case of the excitation during inelastic scattering from an adsorbed molecule, the energy of the electron will be reduced by the difference in energy between two vibrational levels of the adsorbed molecule. It is also possible to study the vibrational spectrum of the adsorbate bonded to the surface. As an analytical tool one may make a quantitative analysis of the adsorbates, since the vibrational frequencies of each molecule are a unique fingerprint for that molecule. The strength of the LEELS signal is proportional to the number of adsorbed molecules.

Suppose that a substrate surface excitation is excited. It is possible to obtain the dispersion curve of the excitation [i.e., to find $\omega(\mathbf{Q})$]. The procedure follows from the energy conservation law:

$$E' = E - \hbar\omega(\mathbf{Q}). \quad (\text{W22.114})$$

Attention will be restricted to the case of near-specular scattering (i.e., let $\mathbf{G} = 0$). Using the following expressions for the wave-vector components (see Fig. W22.30),

$$K = \frac{\sqrt{2ME}}{\hbar} \cos \theta, \quad K' = \frac{\sqrt{2ME'}}{\hbar} \cos \theta', \quad (\text{W22.115})$$

and the law of cosines

$$Q^2 = K^2 + K'^2 - 2KK' \cos \phi, \quad (\text{W22.116})$$

the following expression for the wave-vector transfer is found:

$$Q = \frac{1}{\hbar} \sqrt{2m(E' \cos^2 \theta' + E \cos^2 \theta - 2\sqrt{EE'} \cos \theta \cos \theta' \cos \phi)}. \quad (\text{W22.117})$$

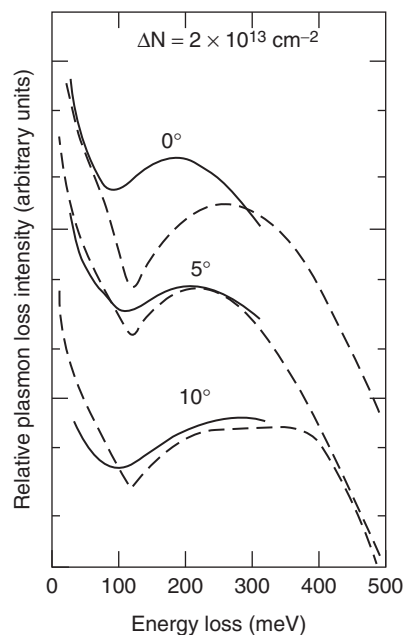


Figure W22.31. LEELS spectra for ZnO for several scattering angles. [Reprinted From Y. Goldstein et al., *Surf. Sci.*, **98**, 599 (1980), Copyright 1980, with permission from Elsevier Science.]

Since E' is measured and E is known, the value of $\omega(\mathbf{Q})$ may be determined. Equation (W22.117) gives Q in terms of E , E' , θ , θ' and ϕ . Thus the dispersion relation for the excitation can be measured.

An example of a LEELS spectrum is presented in Fig. 19.17 for n -type GaAs. The spectrum shows phonon loss and gain peaks as well as a surface-plasmon loss peak. In Fig. W22.31 data for angular-resolved LEELS are presented for electrons scattering from a ZnO surface with an accumulation layer. The data are interpreted in terms of the excitation of two-dimensional plasmons in the accumulation layer. From this data, using Eq. (W22.117), it is possible to obtain information about $\omega(\mathbf{Q})$ for the two-dimensional plasmon. The breadth of the peaks is due to the large dispersion of the two-dimensional plasmon.

W22.18 Extended X-ray Absorption Fine Structure

An accurate determination of interatomic distances in a crystal may be obtained by carefully studying the x-ray absorption spectrum. The absorption spectrum exhibits oscillatory structure that comes about due to an interference effect involving the electrons. The method is called *extended x-ray absorption fine-structure* (EXAFS) *spectroscopy*.

When x-rays pass through a sample of thickness d the intensity of the emerging beam, I , is related to the intensity of the incident beam, I_0 , through Beer's law:

$$I(d) = I_0 \exp(-\alpha d), \quad (\text{W22.118})$$

where the very small surface reflection of the x-rays is neglected. The attenuation constant, α , has contributions arising from both the absorption of x-rays and the Bragg scattering of x-rays out of the incident beam (extinction). In this section attention centers on the absorption contribution.

Absorption comes about when an electron is photoionized from an atom. The electron is promoted from some low-lying state to a state in the conduction band. In the case of deep-core levels the bandwidths are very narrow and there is a threshold absorption energy from a given band equal to the difference in energy between the Fermi energy, E_F , and the core-level energy, E_{core} . For simplicity's sake, restrict the discussion to the case of a parabolic conduction band. When the excited electron travels through the crystal it has a wave vector

$$k = \frac{1}{\hbar} \sqrt{2m[\hbar\omega - (E_c - E_{\text{core}})]}, \quad (\text{W22.119})$$

where E_c is the energy of the bottom of the conduction band. Thus the wave vector is a function of the x-ray frequency.

The rate at which photon absorption takes place depends on how probable it is to find the *excited* electron at the position of the nucleus. Technically, this comes about because the rate depends on a matrix element of the radiation operator between wavefunctions governing the initial and final states of the electron. In particular, it is sensitive to the magnitude of the final-state wavefunction at the position of the atom. If this magnitude were somehow to increase, the absorption would increase, whereas if it were to decrease, the absorption would decrease.

Upon absorption of the photon a spherically outgoing electron wave is created with the wave vector above. This wave may scatter off neighboring atoms in the crystal a distance a_j away. The waves reflected interfere with the wave emitted as in Fig. W22.32. What is of primary interest is the situation at the location of the ionized atom. If there is *constructive interference*, the amplitude of the final-state electron wavefunction will be maximum. If there is *destructive interference*, the amplitude will be minimum. The condition for constructive interference is

$$2ka_j + \delta_j = 2\pi N. \quad (\text{W22.120})$$

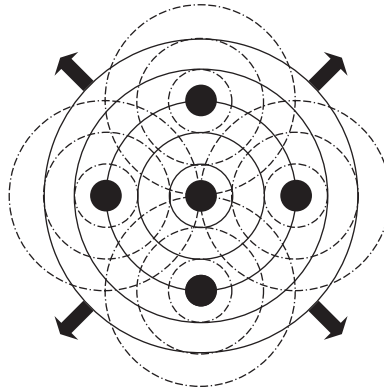


Figure W22.32. Spherically outgoing excited electron waves scatter off neighboring atoms and these reflected waves interfere with the emitted wave.

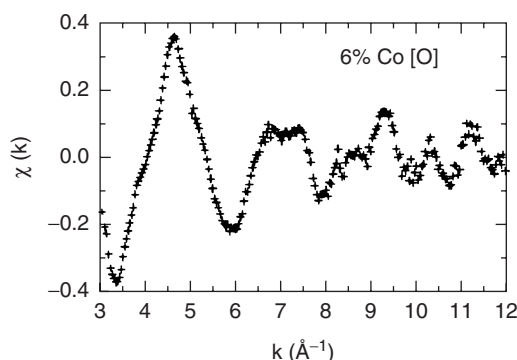


Figure W22.33. EXAFS oscillations for $\text{YBa}_2(\text{Cu}_{1-y}\text{Co}_y)_3\text{O}_{6+x}$. [From H. Renevier et al., *Phys. Rev. B*, **47**, 11398 (1993). Copyright 1993 by the American Physical Society.]

Here $2a_j$ is the round-trip distance to atom j and δ_j is a phase shift characteristic of the scattering of the electrons from the atoms. One expects the phase shift to be a slowly varying function of electron energy. Thus interference oscillations in the x-ray absorption spectrum are expected. The separation between neighboring interference maxima (Fig. W22.33) provides a measurement of the various distances to shells of nearby atoms. Thus

$$\Delta k = \frac{\pi}{a_j} = \frac{1}{\hbar} \Delta \sqrt{2m[\hbar\omega - (E_c - E_{\text{core}})]}. \quad (\text{W22.121})$$

In practice, the absorption spectrum is Fourier analyzed as a function of k and the peak positions in r space appear directly. Separate peaks may be identified with NNs, next-NNs, and so on.

An example of EXAFS oscillations appears in Fig. W22.33 for excitation of a Co core level. The data are for the compound $\text{YBa}_2(\text{Cu}_{1-y}\text{Co}_y)_3\text{O}_{6+x}$. The quantity $\chi(k)$ is the modulated part of the absorption constant. It is defined by $\chi(k) = [\alpha(k) - \alpha_0(k)]/\alpha_0(k)$, where $\alpha(k)$ is the absorption coefficient, including its oscillations, and $\alpha_0(k)$ is obtained by averaging $\alpha(k)$ (a smoothly varying function of k) over the oscillations. By using the oscillations to determine the NN distance, it is possible to determine that the Co ion has a valence state of +3. It is also possible to determine the coordination number (5) of the Co ions to the oxygen ions.

In addition to EXAFS there is a technique called SEXAFS, which is surface EXAFS. Grazing-incidence x-rays are used so that the radiation does not penetrate the solid deeply and the surface region of the solid is probed. A technique closely related to EXAFS is XANES (*x-ray absorption near-edge structure*).

W22.19 Auger Emission Spectroscopy

A useful tool for characterizing the chemical composition of a solid in the vicinity of the surface is Auger emission spectroscopy (AES). A monoenergetic beam of high-energy (1 to 10 keV) primary electrons impinges on the surface of the solid and causes collisional ionization events to occur. Some of these events result in deep core-level electrons being knocked out. In light elements (Be to Si), typically a K-shell electron

is ionized, leaving a K-shell hole behind. In intermediate atomic-number elements (Al–Nb) the core hole might be in the L shell, and in still heavier elements (Zr–Au) in the M shell. (The various shells are actually themselves split into subshells by both fine-structure splitting and crystal-field splitting. Thus one may refer to the L-I, L-II, L-III subshells, etc.)

Upon formation of the hole, the ion is left in an excited state. An electron from some higher-energy shell (which may be broadened into a band) can fill the vacancy, but first it must get rid of its excess energy. Suppose, for example, that a K-shell hole is created and is to be filled by an electron falling from the L shell. There are two methods by which the L-shell electron can shed its excess energy. One is by emitting an x-ray, whose energy is given by

$$\hbar\omega = E_L - E_K. \quad (\text{W22.122})$$

The second method is by having the electron make a Coulomb collision with another electron [e.g., also from a subshell of the L shell (denote it by L')] and transfer the energy to that electron. The energy of the L' electron will then be elevated to

$$E = E_{L'} + E_L - E_K. \quad (\text{W22.123})$$

If this energy exceeds the vacuum level, some fraction of the Auger-excited L' electrons will be emitted from the solid (Fig. W22.34). Since for the inner shells the energies E_K , E_L , and $E_{L'}$, are all well defined and vary from atom to atom, the energy E will also be well defined and will be characteristic of the particular atom involved.

The intensities of the Auger peaks provide quantitative information about the chemical abundance of those elements present. The location of the peaks in the energy distribution and their line shapes also provide information about their chemical bonding. In Table W22.2 some characteristic Auger-transition energies are listed.

For light atoms the Auger process is the dominant mode of filling the core hole. For heavy atoms x-ray emission becomes appreciable. Other possible Auger transitions involve additional shells and/or subshells of the atom. Thus one has K–L–M, K–M–M, L–M–M, N–O–O, L–M–N processes, and so on. For the upper valence bands, however, where the band width is large, there will be a broad band of electron energies emitted and the technique loses its value as an analytical tool.

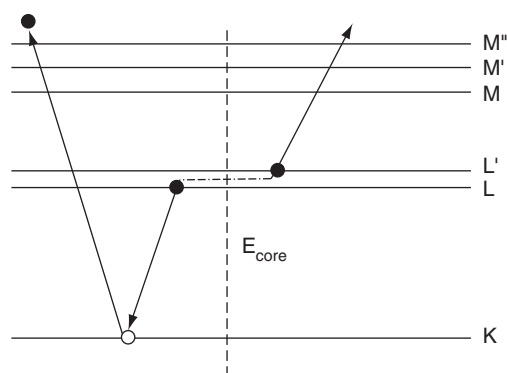


Figure W22.34. Auger process. An electron from the L-shell fills the K-shell vacancy and causes an L' electron to be emitted.

TABLE W22.2 Typical Auger Transitions and Their Energies

Atom	Transition	Auger Electron Energy (eV)
Ag	M–N–N	351
Si	K–L–L	1619
Al	K–L–L	1396
Mg	K–L–L	1186
Cu	L–M–N	920
Si	L–M–M	92.5
Al	L–M–M	68
Mg	L–M–M	45

The reason that AES is regarded as a surface technique has to do with the mean free path of electrons in solids. The electrons lose energy by a variety of processes, including plasmon emission, electron–hole pair excitations, and phonon emission. This limits the range in which it is possible to get Auger electrons out of the solid to the vicinity of the first few surface layers.

Auger spectra are usually presented as derivative spectra. This makes the spectra less sensitive to drifts in the electrical current. The derivatives are obtained by superimposing a weak ac component to the incident current and taking the difference in the Auger current electronically. An example of an Auger spectrum for galvanized steel exposed to atmospheric corrosion for four days is presented in Fig. W22.35. In the energy range of interest there are features due to Zn and also atmospheric components such as O and C present.

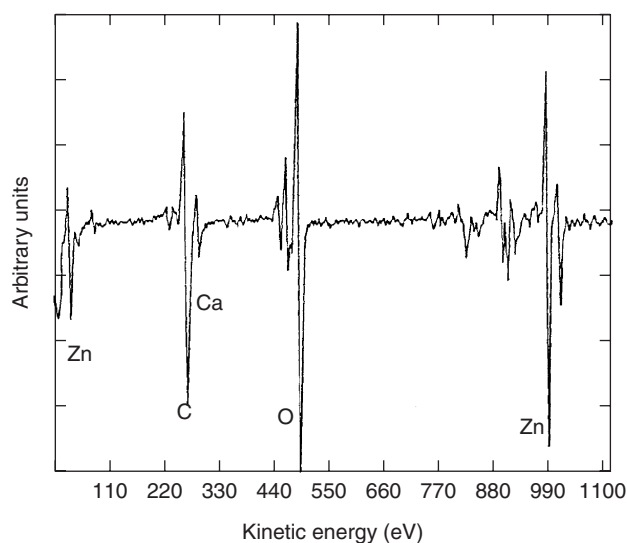


Figure W22.35. Auger electron emission spectrum for galvanized steel undergoing atmospheric corrosion. (From C. Beltran et al., in F. A. Ponce and M. Cardona, eds., *Surface Science*, Springer–Verlag, Berlin, 1991.)

Sometimes, instead of using atomic notation such as L–M–M one denotes the process by L–V–V, indicating that the valence bands (V) are considerably broader than the atomic levels.

W22.20 Secondary-Ion Mass Spectrometry

Sputtering is the process whereby a beam of energetic ions is directed at the surface of a solid and atomic and molecular fragments of the solid are ejected. The fragments may be electrically charged or neutral. In secondary-ion mass spectrometry (SIMS) a quantitative analysis of the emerging ion constituents is undertaken using a mass spectrometer. Often, the emerging neutrals are ionized by external means before the analysis is made. SIMS provides a powerful technique to study the profile of composition versus depth in a sample.

SIMS is capable, in principle, of detecting all elements present in the range of parts per million or even parts per billion. It has a dynamic range of nine orders of magnitude, meaning that it may detect dominant atoms as well as impurity atoms present in low concentrations. It can distinguish different isotopes. Typical depth resolution is on the order of 10 nm, whereas the focused beam size can be made as small as 100 nm. Sputtered holes as deep as 30 μm may be bored in the sample. It is therefore possible to create three-dimensional images of a heterogeneous structure by methodically sputtering away the outer layers. Sputtering is also used in conjunction with AES for depth profiling.

Typically, the energy of the incident ion is in the range 1 to 20 keV. The most often used ions are O_2^+ and Cs^+ . The oxygen ion is used when the sample is electropositive, whereas the cesium ion is used when the sample is electronegative.

In the sputtering process the incident ion makes Coulomb collisions with the ions of the material. Since the energy of the incident ion is fairly high, to a first approximation, one may regard the collisions as if they take place between free particles. This permits the use of conservation laws to analyze the process. Consider the collision of two ions of masses M_1 and M_2 , respectively. Suppose that particle 1 has momentum \mathbf{p}_1 ; particle 2 is at rest. After the collision the momenta are \mathbf{p}'_1 and \mathbf{p}'_2 . Momentum conservation requires that

$$\mathbf{p}_1 = \mathbf{p}'_1 + \mathbf{p}'_2. \quad (\text{W22.124})$$

Energy conservation gives

$$\frac{p_1^2}{2M_1} = \frac{p_1'^2}{2M_1} + \frac{p_2'^2}{2M_2}. \quad (\text{W22.125})$$

Let the angle that \mathbf{p}'_2 makes with \mathbf{p}_1 be ϕ . Then it follows that

$$E_2' = \frac{4M_1M_2}{[M_1 + M_2]^2} E_1 \cos^2 \phi. \quad (\text{W22.126})$$

Let the angle between vectors \mathbf{p}_1 and \mathbf{p}'_1 be denoted by θ . The final energy of particle 1 is then

$$E_1' = E_1 \left[\frac{\cos \theta + \sqrt{M_2^2/M_1^2 - \sin^2 \theta}}{1 + M_2/M_1} \right]^2. \quad (\text{W22.127})$$

In general, the collisions are not elastic and there is some degree of excitation and ionization taking place. For the hard collisions (i.e., collisions involving substantial momentum transfer) that are responsible for sputtering, however, the energy transfer involved in the moderation of the incident ions is large compared with the ionization energy. The effects of the weaker collisions responsible for ionization may be studied separately.

A 10-keV O_2^+ ion has a speed of 2.5×10^5 m/s, which greatly exceeds the speed of sound in solids. The lattice is unable to carry the energy away as phonons. A cascade of collisions occurs in the region where the incident ion strikes the surface. The energy of the ion is distributed among the atoms in that region. If the energy per atom exceeds the cohesive energy of the solid, these atoms are likely to evaporate from the surface. Some of them will emerge as ions, although most will come out as neutrals. Some of the emerging ions will be reneutralized on the way out. The probability that a given species will leave as an ion is very chemical dependent as well as a function of the nature of the sputtering ion. It is known, for example, that a cesiated surface has a low work function, whereas an oxygenated surface has a high work function. This could easily affect the reneutralization probabilities for the emerging ions, since electrons will have to tunnel out from the solid across a vacuum barrier to reach the emitted ions as they leave the solid.

Once the ions emerge from the sample, the mass spectrometry may be carried out in one of three ways. One may use an accelerating cathode to speed up the ions and then inject them into a uniform magnetic field. Alternatively, one may use a quadrupole mass spectrometer. Finally, one may make a time-of-flight measurement. The first method will be examined.

The speed of the positive ion as it passes through the cathode depends on the cathode voltage V , relative to the sample:

$$v = \sqrt{\frac{2qV}{M}}, \quad (\text{W22.128})$$

where the initial velocity of the ion as it leaves the solid is negligible. The diameter of the resulting circular orbit in the magnetic field is

$$D = \frac{2Mv}{qB}, \quad (\text{W22.129})$$

where q and M are the charge and mass of the ion and B is the strength of the magnetic induction. Thus the mass-to-charge ratio is

$$\frac{M}{q} = \frac{B^2 D^2}{8v}. \quad (\text{W22.130})$$

A typical SIMS spectrum of Si exposed to oxygen is presented in Fig. W22.36, where the number of counts in a detector is plotted as a function of the mass-to-charge ratio M/Z and where $q = Ze$. Note that the species ejected reflect the bonding in the solid and, in particular, that an SiO_2 fragment is not ejected.

W22.21 Rutherford Backscattering

A powerful technique for compositional depth profiling of a solid is Rutherford backscattering (RBS). Usually, an α -particle source is used with its energy on the

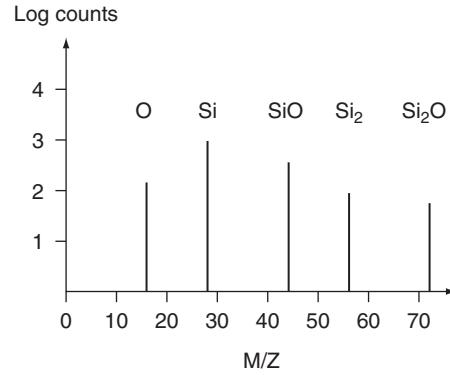


Figure W22.36. SIMS spectrum for SiO_2 .

order of 1 MeV. The α -particle is directed normal to the surface and, when scattered through an angle $\theta > \pi/2$, exits through the same surface that it entered. The energy of the α -particle is measured and the energy loss is determined. This energy loss depends on how far the particle penetrated the solid and the type of atom responsible for its deflection.

As a fast charged particle passes an atom it loses energy, primarily by electronically exciting or ionizing the atom. In a solid, phonon processes or other elementary excitations also come into play. These processes lead to a steady decrease in the energy of the particle and may be described by an energy loss per unit length. Bethe gave an approximate theoretical formula for the energy loss per unit distance due to electronic excitation and ionization:

$$\frac{dE}{ds} = -\frac{2\pi n Z_2}{E} \left(\frac{Z_1 e^2}{4\pi\epsilon_0} \right)^2 \ln \frac{2mv^2}{IE}, \quad (\text{W22.131})$$

where Z_1 is the charge state of the projectile (2 for α -particles), Z_2 the atomic number of the target nucleus, E is the energy of the projectile and v the corresponding speed, n the concentration of target atoms, m the mass of an electron, and IE the ionization energy of the target atom. The energy loss is a slowly varying function of the energy and may be assumed to be constant if the energy-loss range under consideration is sufficiently small. The precise form of the energy-loss function varies from material to material and may be determined experimentally by passing beams through thin samples and measuring the resulting energy loss. It presumably also contains corrections due to phonon losses.

In addition to the mechanism above, there exists the possibility of energy loss resulting from hard Coulomb collisions between the α -particle and the target nuclei (i.e., Rutherford scattering). The cross section for these collisions is on the order of a barn (10^{-28} m^2). The differential scattering cross section in the laboratory frame is

$$\frac{d\sigma}{d\Omega} = \left(\frac{Z_1 Z_2 e^2}{8\pi\epsilon_0 E_1} \right)^2 \frac{[\cos\theta + \sqrt{1 - x^2 \sin^2\theta}]^2}{\sin^4\theta \sqrt{1 - x^2 \sin^2\theta}}, \quad (\text{W22.132})$$

where E_1 is the energy of the α -particle just prior to scattering and $x = M_1/M_2$.

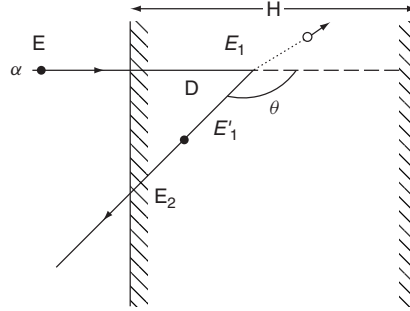


Figure W22.37. Rutherford backscattering geometry.

Suppose that the α -particle enters the solid with energy E at normal incidence and travels a distance D before undergoing Rutherford backscattering. It will arrive at the target nucleus with energy E_1 :

$$E_1 = E - \int_0^D ds \frac{dE}{ds} \simeq E - D \left(\frac{dE}{ds} \right)_1, \quad (\text{W22.133})$$

where the subscript indicates that the energy-loss function is to be evaluated at an average energy for the inward journey. The detector is set to measure the backscattered current at a scattering angle θ , as in Fig. W22.37. The energy of the projectile just after the backscattering event is

$$E'_1 = F(\theta)E_1, \quad (\text{W22.134})$$

where it was found in Eq. (W22.127) that

$$F(\theta) = \left[\frac{\cos \theta + \sqrt{1/x^2 - \sin^2 \theta}}{1 + 1/x} \right]^2 \leq 1. \quad (\text{W22.135})$$

The projectile then travels an additional distance $-D \sec \theta$ before emerging from the solid. The final energy is

$$E_2 = E'_1 - \int_0^{-D \sec \theta} \frac{dE}{ds} ds \simeq E'_1 + D \sec \theta \left(\frac{dE}{ds} \right)_2. \quad (\text{W22.136})$$

Here dE/ds is evaluated for the backscattered journey. Thus

$$E_2 = EF(\theta) - D \left[\left(\frac{dE}{ds} \right)_1 F(\theta) - \left(\frac{dE}{ds} \right)_2 \sec \theta \right] \equiv EF(\theta) - aD. \quad (\text{W22.137})$$

The current entering the detector per unit solid angle per unit energy is

$$\frac{d\dot{N}}{d\Omega dE_2} = I \int_0^H dD n \frac{d\sigma}{d\Omega} \delta(E_2 + aD - EF(\theta)), \quad (\text{W22.138})$$

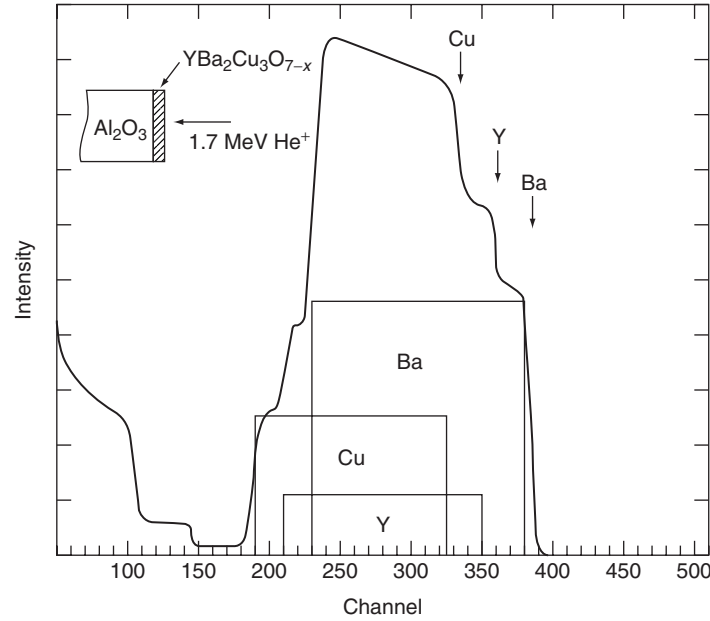


Figure W22.38. Rutherford backscattering spectrum for 1.7-MeV He^+ ions incident on a $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ film on an Al_2O_3 substrate. (From H. J. Gossmann and L. C. Feldman, *Mater. Res. Soc. Bull.*, Aug. 1987, p. 26.)

where I is the incident particle current, a was defined in Eq. (W22.137), H is the sample thickness, and n is the concentration of target atoms. The delta function ensures the correct energy relation. Carrying out the integral gives

$$\frac{d\dot{N}}{d\Omega dE_2} = I \frac{n}{a} \frac{d\sigma}{d\Omega} \Theta(EF(\theta) - E_2) \Theta(E_2 - (EF(\theta) - aH)). \quad (\text{W22.139})$$

The Θ function is 1 for positive argument and 0 for negative argument. It implies the existence of a high- and a low-energy cutoff in the energy spectrum. The high-energy cutoff corresponds to scattering from atoms on the front surface of the sample. The low-energy cutoff corresponds to scattering from atoms at the depth H (i.e., at the back surface of the sample). Since $F(\theta)$ is unique to each target atom, the locations of these cutoffs permits the identification of the presence of a particular type of atom. The size of the step is proportional to the concentration, n . A typical RBS spectrum is given in Fig. W22.38 for a thin film of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. The spectrum consists of a superposition of rectangles, one for each element, and each with its characteristic width aH , and energy E_2 extending from $EF(\theta) - aH$ to $EF(\theta)$.

SURFACE MICROSCOPY

The next three sections are concerned with scanning surface microscopy. The atomic-force microscope, the scanning-tunneling microscope, and the lateral-force microscope

are studied. A mobile probe is passed over the surface in a rastering fashion and a time-dependent voltage signal is sent by the microscope to a computer, where an image of the surface is constructed. In the atomic-force microscope the signals are proportional to the interatomic force between the tip of the probe and the surface. In the scanning-tunneling microscope it is proportional to the electron current that tunnels between the probe and the conducting surface. The lateral force microscope rubs the tip over the surface and measures both the normal force and the frictional force between the solids.

There are numerous extensions of scanning-probe microscopy. The near-field scanning optical microscope (NSOM) uses a tipped optical fiber to transmit light to a surface and to collect the scattered light, providing information concerning the reflectivity variations of the surface. The scanning-capacitance microscope employs the probe and substrate as the plates of a capacitor and measures the variation of capacitance due to variations in the surface height or due to dielectric deposits on the surface. The scanning-thermal microscope rasters a thermocouple over the surface to measure differences in local temperature. The scanning magnetic-force microscope probes the local magnetic structure on the surface by means of a magnetic tip. Numerous other physical effects are also used as the basis for microscopy.

W22.22 Atomic-Force Microscopy

Two objects brought in proximity will exert forces on each other. This is true of atoms and molecules and is also true of mesoscopic objects. At the most fundamental level, this force is of electromagnetic origin (neglecting the extremely weak gravitational force), although it usually appears in the guise of weak chemical bonding forces. These include van der Waals forces, the interaction of electric multipole moments with each other, and possibly magnetic forces as well. The atomic-force microscope (AFM) uses this force in a controlled way to determine surface structure.

Figure W22.39 is a sketch of the essential elements of the atomic-force microscope. A sample is mounted on a stage that is capable of being moved in three independent directions, x , y , and z . A conducting cantilever beam L with a stylus S at the end is brought close to the surface and the sample is moved in a rastering motion beneath it. Above the cantilever is a plate which, together with the cantilever, forms a capacitor. As the sample is moved back and forth, the force on the stylus varies with time. When the stylus is attracted to the sample, the gap size of the capacitor is increased and the capacitance decreases. If this capacitor is part of an LC circuit, the resonance frequency

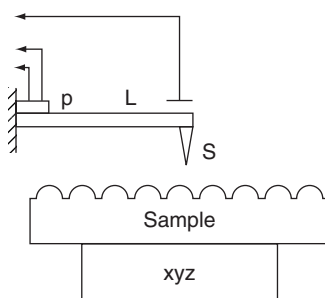


Figure W22.39. Atomic-force microscope.

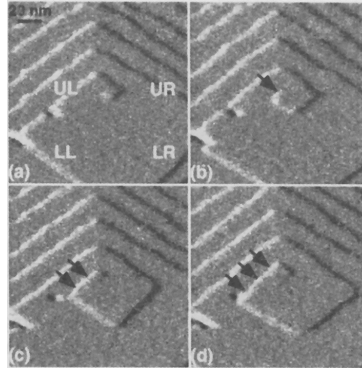


Figure W22.40. Atomic-force microscope micrographs for a growth spiral. [From G. T. Palocz et al., *Appl. Phys. Lett.*, **73**, 1658 (1998). Copyright 1998, American Institute of Physics.]

may be monitored as a function of time. In another mode of operation, a piezoelectric crystal, p , attached to the cantilever, can be sent a feedback signal to keep the height of S above the surface constant. The voltage across the piezoelectric crystal needed to maintain this constancy then becomes the signal. Other ways of detecting the stylus motion are possible, such as interferometry.

It is important that the microscope be immune to vibrations of the surrounding environment. In addition to vibration isolation, such immunity may be obtained by using a cantilever that has a high natural vibration frequency (in the tens of kilohertz) and by rigidly attaching it to the sample stage. Then, to a first approximation, the entire microscope will vibrate as a rigid body and the separation between the stylus and the sample surface will remain approximately constant.

Since interatomic forces tend to be short ranged, the tip of the stylus provides the dominant force in its interaction with the sample. The stylus is particularly sensitive to forces produced by the sample's dangling bonds, steps, and surface imperfections.

A state-of-the-art atomic-force microscope has recently been constructed with a cantilever consisting of a single crystal of silicon of dimensions $95\text{ }\mu\text{m}$ long by $0.6\text{ }\mu\text{m}$ thick. The resonant frequency is 77 kHz and it is sensitive to forces smaller than 10^{-11} N . A typical scanning velocity is 200 nm/s .

An example of the application of the AFM to the study of a growth spiral is presented in Fig. W22.40. Sequential images are shown for the outward growth of steps from a screw dislocation. It is found that when steps reach a critical length, new steps at right angles to them begin to grow. This is a result of the competition between step-length energy and layer-area energy. The surface is that of calcite.

W22.23 Scanning-Tunneling Microscope

The scanning-tunneling microscope (STM) uses electrons that tunnel from a conducting solid to a conducting probe electrode (stylus) to map the topography of the surface of a solid. The construction is almost identical to that of the atomic-force microscope except that a potential difference, V , is imposed between the stylus and the surface. A tunneling current is established, and this current depends sensitively on the distance between the stylus and the sample. The stylus is made as sharp as possible. Tunneling through the vacuum favors the most direct path, so the characteristic region of the surface

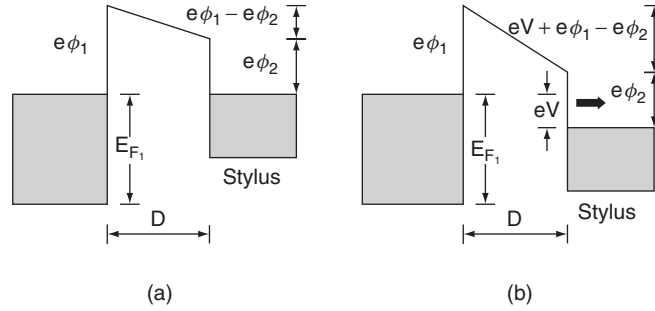


Figure W22.41. Tunneling process: (a) unbiased; (b) biased.

contributing tunneling electrons is somewhat smaller than the radius of curvature of the stylus tip.

As the surface is rastered past the stylus, the distance D between the stylus and the surface will fluctuate and this will cause the tunneling current to vary. As in the case of the AFM, it is common practice to supply a feedback voltage to the piezoelectric crystal to keep the surface at a constant distance below the stylus. This prevents the tip of the stylus (the “head”) from crashing into the surface, thereby destroying the stylus. The variation of this feedback voltage with time (and hence stylus location) provides the signal needed to reconstruct the image of the surface.

It is fairly simple to derive an approximate expression for the tunneling current in a one-dimensional approximation. Consider Fig. W22.41, which shows two cases where the stylus is in proximity to the surface, one without external bias and one with a bias voltage V . For the sake of definiteness, assume that the sample is on the left and the stylus is on the right in each case. Let ϕ_1 be the work function potential of the sample and ϕ_2 be the work function potential of the stylus. When the metals are brought into contact, or near contact, the Fermi levels will rapidly equilibrate by having some charge flow from the metal with the smaller work function potential. This establishes the contact potential difference. (This effect is the basis for what is called the Kelvin probe technique for measuring work function changes associated with adsorption.) Next, suppose that an external bias voltage V is imposed on the system. The Fermi levels are no longer the same and a tunneling current of electrons can flow from one metal to the other. In the case of the diagram it flows from the sample to the stylus.

The particle current per unit area is given by an integral over the Fermi sea of the left-hand conductor:

$$J_z = -2 \int \frac{d\mathbf{k}}{(2\pi)^3} v_z P \Theta(v_z) f(E, T) [1 - f(E - eV, T)], \quad (\text{W22.140})$$

where $f(E, T)$ is the Fermi–Dirac distribution function and P is the probability for tunneling through the barrier. The quantity P is given by

$$P = \frac{v'_z}{v_z} \exp \left[-\frac{2}{\hbar} \int_0^D \sqrt{2m[U(z) - E]} dz \right], \quad (\text{W22.141})$$

where v'_z is the velocity on the right and v_z is the velocity on the left. The Fermi–Dirac factors guarantee that the tunneling electron will come from an occupied sample state and tunnel into a vacant stylus state. The form of the barrier potential energy is

$$U(z) = E_{F_1} + e \left[\phi_1 + (\phi_2 - \phi_1 - V) \frac{z}{D} \right], \quad (\text{W22.142})$$

where $z = 0$ at the sample surface and image potential corrections are neglected. At low temperatures the Fermi factors may be replaced by unit step functions (i.e., Θ functions). If the Θ functions are expanded to first order in V , the expression becomes

$$J_z = \frac{2}{(2\pi)^3} \int d\mathbf{k} v_z P \Theta(v_z) eV \frac{1}{2} \delta(E - E_{F_1}), \quad (\text{W22.143})$$

which may be expressed in terms of the density of states at the Fermi level:

$$J_z = \frac{eV v_F}{2} \rho(E_{F_1}) \langle \cos \theta \rangle P = \frac{e^2 V v_F}{4} \rho(E_{F_1}) P. \quad (\text{W22.144})$$

Here v_z has been replaced by $v_F \cos \theta$ and the average value of $\cos \theta$ in the forward direction is equal to $\frac{1}{2}$. The tunneling integral is readily computed, and finally, a formula for the particle current density is obtained:

$$J_z(V) = \frac{eV}{4} \rho(E_F) v'_F \exp \left[-\frac{4D\sqrt{2m}}{3\hbar} \frac{(e\phi_1)^{3/2} - (e\phi_2)^{3/2}}{e\phi_1 - e\phi_2} \right]. \quad (\text{W22.145})$$

The quantity v'_F is the Fermi velocity for the tunneling probe. The exponential falloff with tunneling distance is expected as well as the dependence on some average barrier height.

The actual value of the electric current is given by $I = -eAJ_z$, where A is a characteristic area. For the case of a stylus tip with radius of curvature R , one may expect $A \approx \pi R^2$. Equation (W22.145) is not completely correct. In reality, one should use the local density of states rather than the bulk density of states. The local density of states varies from position to position in directions parallel to the surface and reflects the variations in local charge density of the surface bonds. Therefore, as one rasters the surface under the tip, the tunneling current will vary from position to position.

An example of an STM picture of the surface of Si(100) is presented in Fig. W22.42. It shows, with atomic resolution, a Si(100) 2×1 surface with a Na overlayer.

W22.24 Lateral-Force Microscope and Surface Force Apparatus

A variant of the atomic-force microscope, called the lateral-force microscope (LFM), can measure the shear stress on a microscopic stylus that is slid across a surface (Fig. W22.43). It is sensitive to forces as small as 1 pN. The stylus, which constitutes one of the solids (commonly diamond or Si_3N_4), is supported by a flexible cantilever that can be deflected as the stylus rubs against the other surface. By measuring the bending of the cantilever, one may determine the normal force exerted on it by the stylus. By measuring the torsion of the cantilever, information concerning the frictional force is obtained. These measurements are made by reflecting a beam of light from the back of the cantilever and recording the position of the reflected spot on a screen. The

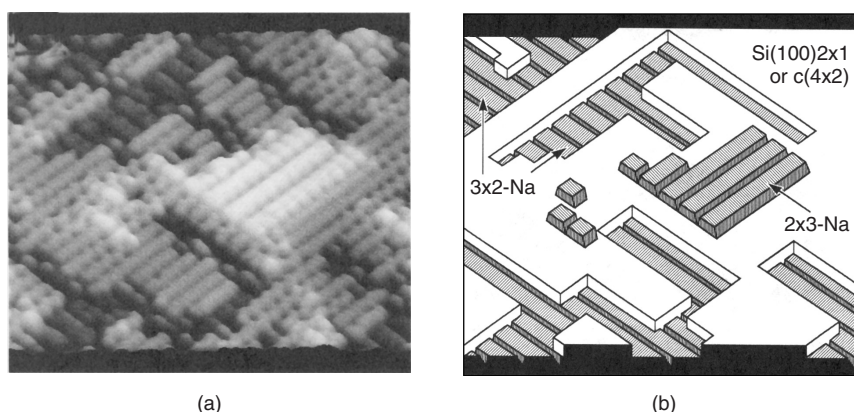


Figure W22.42. Micrograph of the Si(100) surface with an overlayer of Na atoms. [From A. A. Saranin et al., *Phys. Rev. B*, **58**, 4972 (1998). Copyright 1998 by the American Physical Society.]

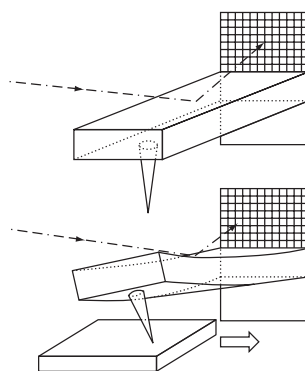


Figure W22.43. Lateral force microscope (LFM). A light beam is reflected off a cantilever and strikes an array of photodetectors. The flexure and torsion of the cantilever are related to the normal and friction forces.

solid lying beneath the stylus is attached to a piezoelectric crystal stage. By applying time-varying potential differences across this crystal, the sample may be rastered back and forth beneath the stylus. Thus a friction map may be generated.

By coating the stylus with a self-assembled monolayer (SAM) of organic molecules, it is possible to sensitize the stylus so that it will respond differently to different adsorbates on the other solid. This is because the chemical specificity of the intermolecular interactions determines the friction force. It was recently found that friction can be both anisotropic and asymmetric when the monolayer consists of tilted molecules.[†] The asymmetry refers to moving the stylus in the direction of the molecular tilt compared with against it.

The surface force apparatus (SFA) is a device with two atomically flat parallel mica plates. The width of the separation may be reduced to nm dimensions. Lubricants are

[†] M. Liley et al., *Science*, **280**, 273 (1998).

placed in the gap and the plates are slid past each other. Evidence for the formation of well-defined liquid layers is found when the gap distance is smaller than 10 nm. The viscosity rises very rapidly as the gap distance is reduced. One finds evidence for the formation of two-dimensional glassy solids within these layers if the layers are very thin (e.g., four molecules thick). As the lubricant is sheared, these layers are deformed elastically and then may release the strain energy by slipping or melting when a critical shear stress is reached. The resulting stick-slip motion is reminiscent of the mechanism believed to produce seismic earthquakes. The sudden slip motions are also believed to peel material off the surfaces, thereby producing delamination wear. Wear is the general term given to the change of geometry of the surfaces and the removal of material from them as a result of friction. In polymer applications it is found that the threshold for substantial wear is correlated with the product Pv , in accordance with Eq. (W19.39). The value is referred to as the Pv limit. For example, for polycarbonate and Teflon, the Pv limits are 0.01 and 0.06 MPa·m/s, respectively, for $v = 0.5$ m/s at room temperature.

Another device that is used to study the friction of lubricants is the quartz-crystal microbalance (QCM). The damping of vibrations (i.e., the Q of the quartz crystal plate) is influenced by the viscosity of the lubricant with which it is in contact.

TRANSPORT MEASUREMENTS

In Chapter 7 electrical-transport properties such as the electrical resistivities and the Hall coefficients of materials have been introduced. Some elementary thermal and thermoelectric properties are also discussed. In the following two sections some methods for measuring these properties are reviewed.

W22.25 Electrical Resistivity and Hall Effect

The simplest method for measuring resistivity involves the use of a cylindrical sample of material of length L and cross-sectional area A . The resistance R is measured and the resistivity is given by $\rho = RA/L$. The accuracy of the measurement is limited by the geometric measurements and the ability to control fringing fields. A simple geometrical arrangement for measuring the Hall coefficient is given in Fig. 7.1 and discussed in Section 7.3.

For a large sample of material with a planar surface, the four-contact method may be employed to measure ρ . Suppose that the material occupies the half-space $z < 0$. Place four contacts at four points on the surface at the locations defined by the vectors \mathbf{r}_A , \mathbf{r}_B , \mathbf{r}_C , and \mathbf{r}_D . The contacts are placed close together so the distance between them is much less than the distance to the edges of the surface. If a current I is injected into contact A, it will set up an electrostatic potential field $\phi_A(\mathbf{r}) = \rho I / (2\pi|\mathbf{r} - \mathbf{r}_A|)$ within the material. Similarly, if one draws a current I out of contact B, the potential field is given by $\phi_B(\mathbf{r}) = -\rho I / (2\pi|\mathbf{r} - \mathbf{r}_B|)$. When the current is injected at A and removed at B, these potentials are superimposed to give $\phi(\mathbf{r}) = \phi_A(\mathbf{r}) - \phi_B(\mathbf{r})$. The difference in potential is then measured between points C and D. The resistance is

$$R_{CD,AB} = \frac{V_{CD,AB}}{I} = \frac{\rho}{2\pi} \left(\frac{1}{r_{CA}} - \frac{1}{r_{CB}} - \frac{1}{r_{BA}} + \frac{1}{r_{DB}} \right), \quad (\text{W22.146})$$

where $r_{CA} = |\mathbf{r}_C - \mathbf{r}_A|$, and so on.

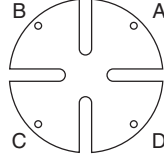


Figure W22.44. Clover-shaped sample for the van der Pauw method of measuring the resistivity or the Hall coefficient.

The van der Pauw method[†] extends this method to two dimensions and permits one to measure the resistivity and Hall coefficient for a thin sample of material. It will be assumed that there is isotropy in the plane of the slab. Four fine electrodes, labeled A, B, C, and D, are attached to the boundaries of a slab of thickness d . The shape of the sample is unimportant, as long as it has no holes in it (i.e., it must be simply connected). (This may be proven by the method of conformal transformations using complex-variable theory. It will not be derived here.) A typical geometry that is used is illustrated in Fig. W22.44. In the resistivity case two measurements are made. First a current I is driven from C to D and the voltage $V_{AB,CD}$ is measured across electrodes A and B. The resistance $R_{AB,CD}$ is computed by the formula $R_{AB,CD} = V_{AB,CD}/I$. The measurement is repeated with a current driven from D to A and the voltage measured across B and C. The resistivity is given implicitly by the formula

$$\exp(-\pi R_{AB,CD}d/\rho) + \exp(-\pi R_{BC,DA}d/\rho) = 1. \quad (\text{W22.147})$$

The method may be generalized to anisotropic samples.[‡]

The Hall coefficient R_H is determined by measuring the change in the resistance $R_{BD,AC}$ when a magnetic induction B is imposed perpendicular to the slab. The formula is

$$R_H = d \frac{\Delta R_{BD,AC}}{B}. \quad (\text{W22.148})$$

From a measurement of the Hall voltage the sign of the carrier may be determined.

W22.26 Thermopower, Peltier Coefficient, and Thermal Conductivity

A system in thermal equilibrium obeys the first law of thermodynamics, given by Eq. (WA.1), $T dS = dU + P dV - \mu dN$. When the system is driven slightly out of equilibrium, current densities are produced. These include the particle-current density, \mathbf{J} , and the energy-current density, \mathbf{J}_U . Consider the case where the charged carriers are electrons, so the particle current density is proportional to the electrical-current density (i.e., $\mathbf{J}_E = -e\mathbf{J}$). The driving forces for \mathbf{J}_E include the electric field, $\mathbf{E} = -\nabla\phi$, as well as the gradient in the chemical potential and the gradient in the temperature. The same forces drive \mathbf{J}_U . In place of the energy-current density, the first law of thermodynamics is used to define the heat-current density, \mathbf{J}_Q , in terms of the chemical potential:

$$\mathbf{J}_Q = T\mathbf{J}_S \equiv \mathbf{J}_U - \mu\mathbf{J}, \quad (\text{W22.149})$$

[†] L. J. van der Pauw, *Philips Res. Rep.*, **13**, 1 (1959).

[‡] L. J. van der Pauw, *Philips Res. Rep.*, **16**, 195 (1961).

where \mathbf{J}_S is interpreted as an entropy-current density. For weak driving forces the current densities are expressed as linear combinations of the driving forces:

$$\mathbf{J}_E = e \frac{L_{11}}{T} \nabla(\mu - e\phi) + eL_{12} \nabla \frac{1}{T}, \quad (\text{W22.150a})$$

$$\mathbf{J}_Q = \frac{L_{21}}{T} \nabla(\mu - e\phi) + L_{22} \nabla \frac{1}{T}, \quad (\text{W22.150b})$$

where L_{ij} are coefficients. Onsager proved (in general) that $L_{12} = L_{21}$ so there are three independent coefficients. An example of the Onsager relations has been encountered when the transport properties of metals were studied in Section 7.5.

The significance of the L_{ij} coefficients may be obtained by examining special cases:

1. If T and μ are constant in space, then

$$\mathbf{J}_E = e^2 \frac{L_{11}}{T} \mathbf{E}, \quad (\text{W22.151})$$

so $\sigma = e^2 L_{11}/T$. The coefficient L_{11} is therefore proportional to the electrical conductivity.

2. If the heat current is measured for the case where there is no electric current (i.e., $\mathbf{J}_E = 0$), it is found that

$$\mathbf{J}_Q = - \frac{L_{11}L_{22} - L_{12}^2}{L_{11}T^2} \nabla T = -\kappa \nabla T, \quad (\text{W22.152})$$

where κ is the thermal conductivity.

3. In the absence of an electric current, an electric field is established in the sample, that is,

$$-e\mathbf{E} = \nabla\mu - \frac{L_{12}}{TL_{11}} \nabla T. \quad (\text{W22.153})$$

The electromotive force is given by

$$\varepsilon = \oint \mathbf{E} \cdot d\mathbf{l} = -\frac{1}{e} \oint d\mathbf{l} \cdot \nabla\mu - \oint Q d\mathbf{l} \cdot \nabla T, \quad (\text{W22.154})$$

where $Q = -L_{12}/eTL_{11}$ is called the absolute thermoelectric power of the material. (The symbol Q is used here rather than S so as not to confuse it with the entropy.) The first term on the right-hand side may be written as $\oint d\mu$ and is zero. The second term may be written as $-\oint Q dT$.

Consider an experimental arrangement such as is shown in Fig. W22.45, consisting of two conductors, labeled A and B, with absolute thermoelectric powers Q_A and Q_B , respectively. Let a voltmeter be inserted in one of the conductors to measure the electromotive force $\Delta\varepsilon$. Label the temperatures at the left and right junctions T_J and $T_J + \Delta T$, respectively, and the temperature at the voltmeter $T_J + \Delta T'$. It is assumed that $\Delta T \ll T_J$ and $\Delta T' \ll T_J$. Then

$$\begin{aligned} \Delta\varepsilon &= -Q_A[(T_J + \Delta T) - (T_J + \Delta T')] - Q_B[T_J - (T_J + \Delta T)] \\ &\quad - Q_A[(T_J + \Delta T') - T_J] \\ &= (Q_B - Q_A) \Delta T. \end{aligned} \quad (\text{W22.155})$$

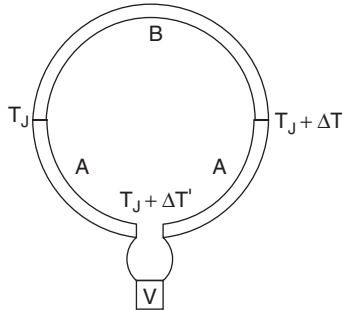


Figure W22.45. Arrangement for measuring the absolute thermopower.

Thus the difference in the thermopowers is the voltage per unit temperature difference:

$$Q_B - Q_A = \frac{\Delta \varepsilon}{\Delta T}. \quad (\text{W22.156})$$

A four-probe technique is used to measure the thermopower. Thermocouple thermometers are placed at the left and right junctions to measure T_J and $T_J + \Delta T$, and the difference of the temperatures is taken to obtain ΔT . The voltage leads are placed across the gap, as shown in Fig. W22.45. In measuring the thermopower one places both the sample and thermometer in vacuum, to eliminate convective heat channels. They are also shielded with highly reflecting surfaces to minimize radiative losses. (The same techniques are used in the design of a thermos bottle.) The voltage could be measured using a potentiometer connected to a sensitive galvanometer. Very small thermocouples, connected to very fine leads, are employed as thermometers.

Thermocouples are thermometers that produce an electromotive potential related to the temperature at the junction. A typical thermocouple is illustrated in Fig. W22.46. Two conductors, A and B, form a junction that acts as the temperature probe. The other wires are each connected to identical conductors, labeled C. The AC and BC junctions are each held at the same standard temperature, T_0 . A mixture of ice and water at atmospheric temperature is often used to set $T_0 = 0^\circ\text{C}$. The other ends of the C wires are connected to a galvanometer and a potentiometer at room temperature. Typical thermocouples involve the use of copper versus constantan, chromel versus alumel, chromel versus constantan, iron versus constantan, and platinum versus platinum–rhodium.

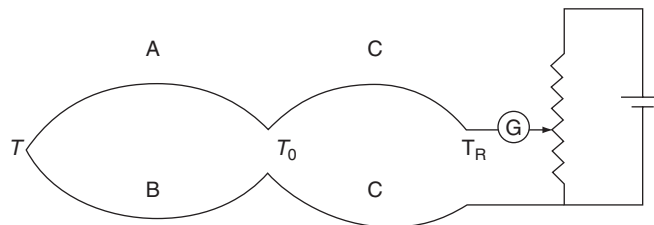


Figure W22.46. Thermocouple arrangement.

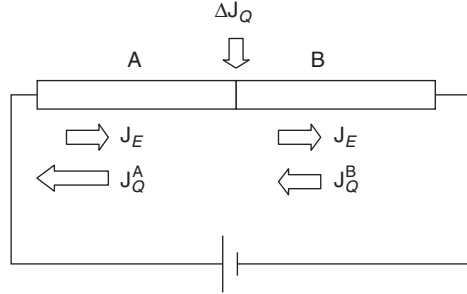


Figure W22.47. Peltier effect.

The Peltier effect involves creating a junction of two dissimilar conductors and passing an electrical current through it. The electrical current is the same in each conductor. Assume that the temperature is held constant. The situation is illustrated in Fig. W22.47. In the absence of a temperature gradient term, the heat current is proportional to the electric current:

$$\mathbf{J}_Q = \frac{L_{12}}{eL_{11}} \mathbf{J}_E = -QT \mathbf{J}_E. \quad (\text{W22.157})$$

Since Q is discontinuous from one conductor to the other, this implies that a heat transfer must take place at the junction. The heat extracted at the junction from the environment is given by

$$\Delta J_Q = J_Q^B - J_Q^A = -T(Q_B - Q_A)J_E \equiv \pi_{BA}J_E, \quad (\text{W22.158})$$

where π_{BA} is called the *Peltier coefficient*. Thus the Peltier coefficient is defined as the heat extracted per unit current. It may be determined from a measurement of the thermopower through the relation

$$\pi_{BA} = -T \frac{\Delta \varepsilon}{\Delta T}. \quad (\text{W22.159})$$

There are a number of ways to measure the thermal conductivity. They often may be classified as transient measurements or steady-state measurements. An example of a transient measurement is the following. Take a rod of length L initially at temperature T_0 . At $t = 0$ place the left end of the rod in contact with a thermal bath at temperature T_1 . Measure the temperature of a point on the rod at position x for times $t > 0$. The thermal diffusion equation is

$$\nabla \cdot \mathbf{J}_Q + \frac{\partial u}{\partial t} = -\nabla \cdot (\kappa \nabla T) + \frac{\partial(\rho c T)}{\partial t} = -\kappa \frac{\partial^2 T}{\partial x^2} + \rho c \frac{\partial T}{\partial t} = 0, \quad (\text{W22.160})$$

where it is assumed that κ is independent of T . The solution to Eq. (W22.160) is

$$T(x, t) = T_1 - (T_1 - T_0) \operatorname{erf} \left(\frac{x}{2\sqrt{at}} \right), \quad (\text{W22.161})$$

where ρ is the density, c the specific heat, and the thermal diffusivity is $a = \kappa/\rho c$. The error function $\text{erf}(x)$ is defined in Chapter W6. The rise of $T(x,t)$ with time at a fixed x is compared to this formula, and a value for a is determined. The value of c is obtained from a calorimetry experiment.

In the steady-state measurements simple geometrical arrangements are chosen and heat is supplied to the material at a known rate. The temperature differential is measured. For example, if a rod of length L is connected to a heater supplying a known heat flux J_Q , and the temperature difference ΔT is measured between two points along the rod a distance Δx apart, then $\kappa = J_Q \Delta x / \Delta T$.

A preferable geometrical arrangement involves the use of concentric cylinders. A cylindrical heater of length L and radius R_1 is surrounded by a hollow sample of material of the same length, with inner radius R_1 and outer radius R_2 . Heat is delivered by the electrical heater at a known rate, H . Thermocouples are used to measure the temperature difference ΔT between the inner and outer surfaces of the sample. The thermal conductivity is then given by

$$\kappa = \frac{H}{2\pi L \Delta T} \ln \frac{R_2}{R_1}. \quad (\text{W22.162})$$

MAGNETIC MEASUREMENTS

The magnetic properties of materials are discussed in Chapter 9, and a number of magnetic materials are studied in Chapter 17. In this section some of the measurement techniques for characterizing magnetic materials are described. They include use of the Foner magnetometer, the Faraday balance, and the ac bridge. The SQUID magnetometer is discussed in Chapter 16.

W22.27 Foner Magnetometer

The Foner magnetometer is used to measure the magnetization of a small sample of magnetic material. When measuring the saturation magnetization the shape of the sample is not important. For nonsaturation conditions a spherical sample is used so that the orientation of the sample is not relevant. The sample is placed on a reed and is made to vibrate in the presence of a coil of wire. For this reason the apparatus is also known as the *vibrating-sample magnetometer* (VSM). Alternatively, the coil may be vibrated in the presence of the magnetic sample. In either case an ac electromotive force is established in the coil which is readily measured. From this measurement the magnetization may be determined.

A formula for the EMF may be obtained by considering a coil with a current I in the neighborhood of the sample and neglecting resistance effects. Let L be the inductance of the coil in the absence of the sample. The energy of the system is

$$U = \frac{1}{2}LI^2 - \mu_0 \mathbf{m} \cdot \mathbf{H}, \quad (\text{W22.163})$$

where H is the magnetic field intensity and \mathbf{m} is the magnetic moment of the sample. It will be assumed that $\mathbf{H} = H\hat{k}$ and that $\mathbf{m} = \mathbf{M}V$, where \mathbf{M} is the magnetization and

V is the volume of the sample. The energy of the system will be constant, so

$$\frac{dU}{dt} = 0 = LI \frac{dI}{dt} - \mu_0 MV \frac{dH}{dt}. \quad (\text{W22.164})$$

Use $LI = N\Phi$, where N the number of turns in the coil and Φ is the magnetic flux through the coil. Assume that $H = H(z)$ and write $dH/dt = v_z dH/dz$, where v_z is the z component of the velocity of the sample. From Faraday's law the EMF is given by $\varepsilon = -N d\Phi/dt = -L dI/dt$. Thus

$$\varepsilon = -\mu_0 MV v_z \frac{1}{I} \frac{dH}{dz}. \quad (\text{W22.165})$$

For a harmonic oscillation of the sample, $z = A \cos \omega t$, where A is the amplitude (typically ≈ 1 mm) and ω is the frequency (typically corresponding to ≈ 100 Hz). Therefore,

$$\varepsilon(t) = \frac{\mu_0 \omega M V A}{I} \frac{dH}{dz} \sin \omega t. \quad (\text{W22.166})$$

From a measurement of the amplitude of the EMF and the mechanical motion, together with knowledge of the sensitivity of the instrument, $|(dH/dz)/I|$, and the volume of the sample, one may determine the magnetization of the sample. The sensitivity function depends on the geometry. For example, consider the ideal case of two coils of wire of radius R separated by a coaxial distance $2D$. A sketch of the Foner magnetometer is given in Fig. W22.48. Some external source (not shown), such as a loudspeaker, is used to establish vibrations in the reed to which the sample is attached. The coils are wound so that the currents generated in the coils will flow in opposite directions. Near the center of symmetry one finds the sensitivity

$$\frac{1}{I} \frac{dH_z}{dz} = -\frac{3NDR^2}{2(R^2 + D^2)^{5/2}}. \quad (\text{W22.167})$$

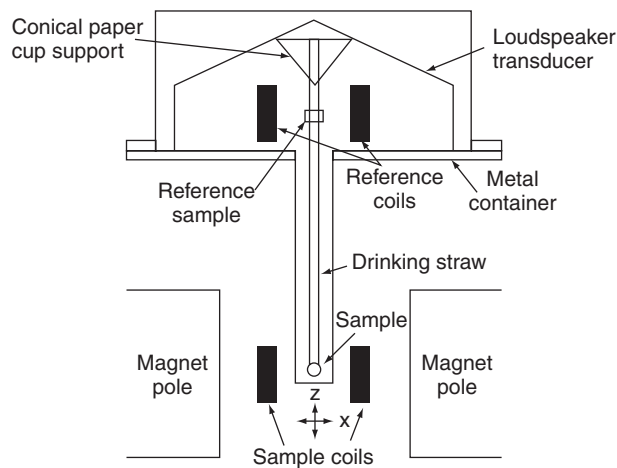


Figure W22.48. Foner magnetometer. (Adapted from S. Foner, J. Appl. Phys., **79**, 4740 (1996). Copyright 1996 by the American Institute of Physics.)

The maximum sensitivity occurs when $R = 2D$ and has the value $|(dH/dz)/I| = 96N/(5R^2\sqrt{5})$. The sensitivity grows with the number of turns (which could typically be $\approx 25,000$) and falls off inversely as the square of the radius.

The Foner magnetometer readily measures magnetic moments on the order of 10^{-10} A·m² at liquid-nitrogen temperatures, to reduce the thermal noise. The instrument is generally calibrated in terms of a known ferromagnetic material, such as Ni. Magnetizations are measured relative to the calibration standard.

W22.28 Faraday Balance

The Faraday balance permits one to measure the magnetization of a sample in a magnetic field. The technique is illustrated in Fig. W22.49. A solenoidal superconducting magnet establishes a magnetic field intensity H_0 in the axial direction which magnetizes the sample, the magnetization being $M(H_0)$. Note that this uniform magnetic field does not produce a net force on the sample. Weights are placed on the right-hand side of the balance equal to the weight of the sample to maintain equilibrium. Then an inhomogeneous magnetic field H is established by the smaller pair of coils. The coils are arranged as shown in Fig. W22.49. The magnetic force in the axial direction is given by

$$F_z = \frac{\partial(\mathbf{m} \cdot \mathbf{B})}{\partial z} = M(H_0)V\mu_0 \frac{\partial H}{\partial z} = W, \quad (\text{W22.168})$$

where V is the volume of the sample. The additional weight W is placed on the right-hand side to counterbalance the magnetic force. In practice, an analytical microbalance is adapted to serve as the balance. The field gradient is vertical. The radius, R , equals the separation between the coils, D , as in the Helmholtz coil arrangement, but the currents are in opposite directions so that a uniform gradient dH/dz is established.

W22.29 AC Bridge

The complex frequency-dependent magnetic permeability of a material, $\mu_r(\omega) = \mu_1(\omega) + i\mu_2(\omega)$, may be measured by means of the ac bridge method. One prepares a

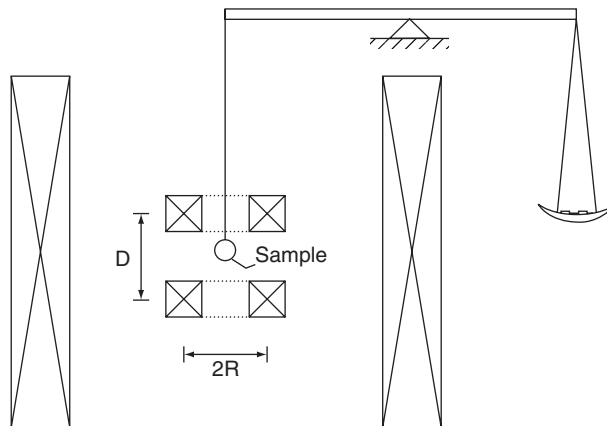


Figure W22.49. Faraday balance.