of the reaction and on kinetic factors. A reaction is favorable if the Gibbs free energy of reaction, $\Delta_r G°$, at the relevant T is negative. A favorable interfacial reaction will proceed at some rate determined by interdiffusion rate and by reaction activation energy, E_a, and this reaction rate will therefore increase exponentially with T. The $\Delta_r G°$ of a solid-state reaction that forms only pure compounds and not solid solutions depends only on the relative Gibbs free energies of formation of the reactant and product compounds from the elements, $\Delta_f G°$. Partial pressures and mole fractions are not involved. $\Delta_f G°$ is available for many compounds [28, 29]; for elements, $\Delta_f G° = 0$ by definition. Often, only the $\Delta_f G°$ at 298 K is available, but the T dependence is small. When only the enthalpy contribution to G, $\Delta_f H°$, is available, it can be used for rough estimates, because the entropy contribution, $\Delta_f S°$, is small for condensed-phase reactions. As an example of predicting whether an interfacial compound will form, consider the deposition of Si onto ZnSe, where there is the possible reaction

$$Si + 2ZnSe \rightarrow 2Zn + SiSe_2$$

However, $\Delta_f H°$ of 2ZnSe is 2×(–163 kJ/mol) while $\Delta_f H°$ of $SiSe_2$ is –29 kJ/mol, giving a $\Delta_r H°$ of +297 kJ/mol; so it is highly unlikely that the $\Delta_r G°$ of this reaction could be negative at any T. For further analysis of interfacial equilibria involving three elements, $\Delta_f G°$ data can be used to construct ternary phase diagrams [30].

The best choice for a diffusion-barrier layer is a compound with a highly negative $\Delta_f G°$ so that it will not react, and with good microstructural integrity so that it will not allow diffusion. For electronic applications such as Al/Si, the barrier must also be electrically conductive. TiN is one material that meets all three requirements fairly well, although it does form Al_3Ti at >780 K [27]. Many metal silicides, borides, and nitrides are conductive and are also significantly more stable than the metals themselves.

Solid-state interfacial reactions can only proceed by diffusion of one or more of the reacting elements through the interfacial compound layer which is being formed. Eventually, diffusion rate drops off due either to the increasing interface thickness or to the formation of an impervious layer of compound. The latter development causes the reaction to "self-limit" at the thickness when the layer becomes impervious. The oxidation of Al is a familiar example of a self-limiting interfacial reaction. Clean Al reacts spontaneously and vigorously with air to form Al_2O_3; however, once a few nm of oxide has formed, the oxide acts as a diffusion barrier to prevent further Al and O_2 from reaching each other, so the reaction stops. (Conversely, metals that form porous oxides, such as Fe in the presence of water, continue to oxidize.) Self-

limiting interfacial reactions lead to the structure shown in Fig. 5.27*d*. Sometimes, a "sacrificial" diffusion barrier is employed whose effectiveness results from its reaction to an impervious compound with one or the other of the adjoining films. Thus, the Al/Ti/Si structure reacts to form $TiAl_3$, which inhibits Si diffusion into the Al.

One problem which can arise even with a self-limiting interfacial compound is that illustrated in Fig. 5.27*e*. Here, the underlying film material is diffusing through the interfacial compound, while the overlying one is not. If there is any lateral nonuniformity in the diffusion rate, as there would be with grain-boundary diffusion, material will become depleted from the fast-diffusing regions. This leads to the "Kirkendall voids" shown, which of course degrade the integrity of the structure.

When more than one compound appears on a binary phase diagram, the one that will form first at the interface between the two constituent elements can often be predicted using the concept of "effective" heat of formation [31]. First, it is assumed that upon initial interdiffusion of two elemental films A and B, the composition of the mixture will be the eutectic (minimum melting T, T_m) composition, A_xB_{1-x}, since diffusion rate is proportional to T/T_m [21]. Now, suppose that there exist two compounds, A_yB_{1-y} and A_zB_{1-z}, with heats of formation $\Delta_f H°_y$ and $\Delta_f H°_z$. The interdiffusing mixture cannot *completely* react into either compound, because the composition (x) is off-stoichiometric; so each $\Delta_f H°$ is reduced to an effective value, $\Delta_f H'$, determined by the fractional deficiency of one element. Thus, if y > x (A deficient),

$$\Delta_f H'_y = \left(\frac{x}{y}\right) \Delta_f H°_y$$

and if x > y (B deficient),

$$\Delta_f H'_y = \left(\frac{1-x}{1-y}\right) \Delta_f H°_y$$

The interfacial compound formed will be the one with the lower $\Delta_f H'$.

5.6 Stress

Interfacial bonding of film layers to each other and to the substrate causes physical interaction as well as the chemical interaction that was discussed above. That is, the films and substrate can be held under a state of compressive or tensile stress by each other by transmitting forces across the interfaces. A film's stress affects its perfor-

mance and also reveals information about the behavior of the deposition process. Stress varies widely with deposition conditions and with the physical properties of the film and substrate materials. After reviewing some basic physics of solids, we will examine first the effects of stress on film behavior and then the process factors that influence stress.

5.6.1 Physics

The basic physical behavior of solids is described by the stress-strain curve, Fig. 5.30. The force applied per unit of cross-sectional area is the stress, σ (N/m^2 or Pa), tensile being positive and compressive negative. Tensile stress along one direction, say x, causes the material to stretch along that direction by a fractional amount called the strain, $\varepsilon_x = \Delta x/x$. Up to the "yield point," the relationship is linear, and the material is said to be "elastic." The yield point is usually defined to be where the deviation from linearity reaches 0.2 percent. The slope in the elastic region is a measure of material stiffness and is called the elastic modulus or Young's modulus, Y; thus,

$$\sigma_x = Y\varepsilon_x \quad \text{or} \quad \varepsilon_x = \frac{\sigma_x}{Y} \tag{5.50}$$

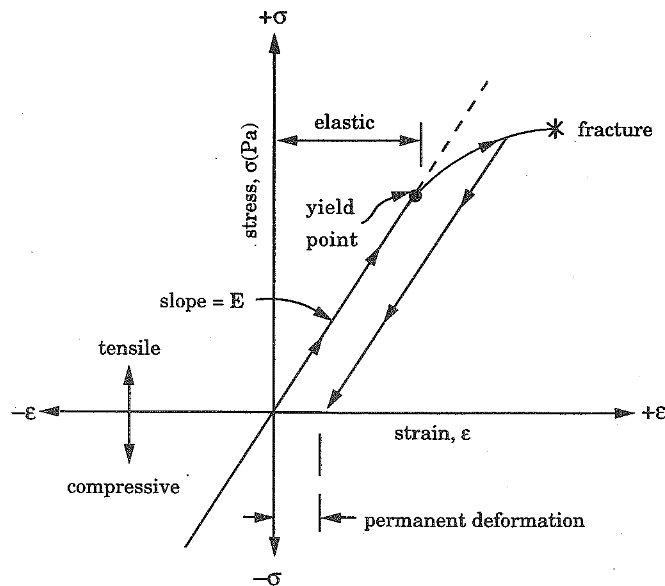

Figure 5.30 Characteristics of the stress-strain relationship.

This is Hooke's law in one dimension, and it is the same as the equation for a spring, F = kx. A typical value of Y is 10^2 GPa (10^{11} Pa) for hard materials. In the elastic region of the stress-strain curve, removing the stress returns the material to its original dimension. Applying stress higher than the yield point causes deformation which remains after removing the stress, as shown in Fig. 5.30. Finally, at the stress level called the "tensile strength," the material fractures. The curve shape is similar in the compressive quadrant. Tensile strengths range up to 1 GPa for hard materials, corresponding to a strain of about 1 percent.

The shape of the stress-strain curve for a given material depends on crystalline grain structure, T, and strain rate. When T is well below the melting point, T_m, there is usually a well defined elastic region, above which yield occurs by the "glide" of crystallographic dislocations (more in Sec. 6.6). At higher T, the material "creeps" by diffusion of atoms especially along grain boundaries, and the elastic region shrinks. For Pb, for example, it has been determined that creep becomes significant at $T/T_m > 0.4$, with T in K (Murakami, 1991). Then, the size of the elastic region depends on the *rate* of imposed stress change relative to the rate at which the stress can relax by creep.

"Uniaxial" tensile stress along x in a freestanding sample of material causes stretch along the x direction and also shrinkage along the unconstrained y and z directions, as shown in Fig. 5.31*a*; that is, ε_y and ε_z are negative. These three strains are related by Poisson's ratio, $\nu = -\varepsilon_y/\varepsilon_x = -\varepsilon_z/\varepsilon_x$. If there were no *volume* increase upon stressing, we would have $\varepsilon_y + \varepsilon_z + \varepsilon_x = 0$ and thus $\nu = 1/2$. This is the case for rubber, for example, and it is the upper limit of ν. However, for most materials, $\nu = 0.3 \pm 0.1$. Thin-film materials are always under *biaxial* stress, as shown in Fig. 5.31*b*, because they are being pulled on by the substrate in *two* dimensions. This changes the stress-strain relationship from that of Eq. (5.50). That is, σ_x causes a stretch along x which is σ_x/E, but σ_y causes a shrinkage along x which is $-\nu\sigma_y/E$, so the net strain is the sum of the two as illustrated by the following equation:

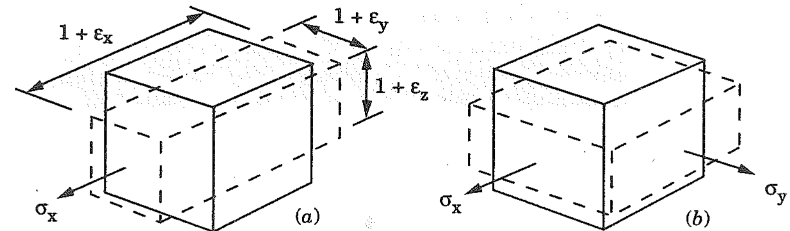

Figure 5.31 Strain resulting from (*a*) uniaxial and (*b*) biaxial stress applied to a cube of unit volume.

$$\varepsilon_x = \varepsilon_y = \varepsilon_{x,y} = \frac{(1-\nu)}{Y}\sigma_{x,y} = \frac{\sigma_{x,y}}{Y'} \quad (5.51)$$

where Y′ is sometimes known as the biaxial elastic modulus.

Thin films contain both intrinsic and extrinsic stress. Intrinsic stress is defined as that frozen in during deposition or during post-deposition treatment, and we will examine the causes of this shortly. Extrinsic stress is applied to the film by external forces, the most common one arising from differential thermal expansion with the substrate. Almost all materials expand upon heating, and the fractional linear expansion per unit ΔT is called the thermal expansion coefficient, α_T (units of K^{-1} or $°C^{-1}$). Usually, α_T decreases gradually with increasing T. For noncubic crystals or anisotropic microstructures, α_T can change with orientation as well. An extreme case is represented by materials with "layered" structures, which are discussed further at Figs. 5.7 and 6.22*b*. For example, pyrolytic BN has $\alpha_T = 37\times10^{-6}$/K perpendicular to the layers (*c* direction) and $\alpha_T = 1.6\times10^{-6}$/K parallel to them (*a* direction).

Typically, a film is deposited at elevated substrate T and then cooled to room T. If a film having a lower α_T than the substrate were *not* bonded to the substrate, the film after cool-down would be wider than the substrate by an amount δ as shown in Fig. 5.32*a*. The same situation would occur if the film contained intrinsic compressive stress, since this would make it want to expand, as indeed it would if it were not bonded to the substrate. However, when the film and substrate are bonded, they are constrained to the same lateral dimension. Thus, stresses develop so as to satisfy the force balance shown in Fig. 5.32*b*, where right-facing arrows indicate tensile σ and left-facing ones compressive σ. In the following, we will assume that $h_f << h_s << L$ as shown. For unit width in y, the force balance in the x direction, neglecting x subscripts and using Eq. (5.51), is

$$F_f = -F_s \quad \text{or} \quad \sigma_f h_f = -\sigma_s h_s \quad \text{or} \quad \left(\frac{Y}{1-\nu}\right)_f \varepsilon_f h_f = -\left(\frac{Y}{1-\nu}\right)_s \varepsilon_s h_s \quad (5.52)$$

Since $h_f << h_s$, it follows that $\varepsilon_f >> \varepsilon_s$ (unless $Y_s << Y_f$). That is, essentially all of the strain appears in the film, and the film's lateral dimensions are determined entirely by those of the substrate. In a multilayer stack of films, the lateral dimensions of *all* of the films are determined by those of the substrate. It is sometimes stated in the literature that a "buffer" film was deposited underneath the desired film to reduce film stress resulting from thermal mismatch to the substrate. But this clearly cannot work unless the buffer (1) has $h_f \geq L$; (2) is capable of gross deformation without de-adhering, as is the case

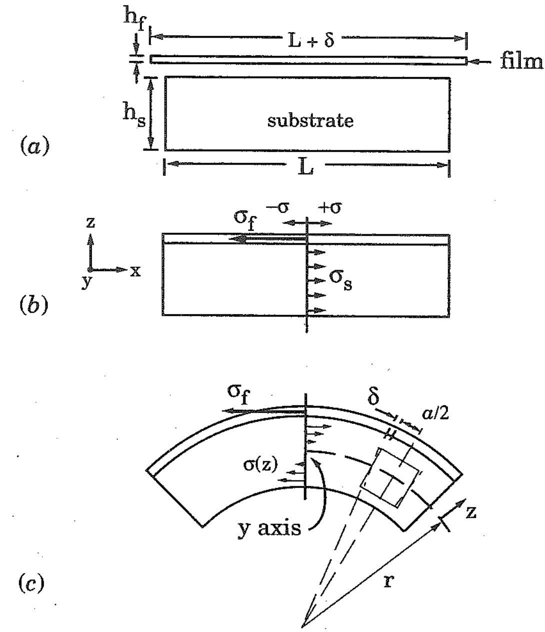

Figure 5.32 Stresses between film and substrate: (*a*) film unbonded to remove stress, (*b*) expansion component of substrate stress, and (*c*) bending component.

to some degree with indium; or (3) changes the intrinsic stress of the overlying film by affecting its grain structure.

The substrate strain is usually negligible compared to the film strain, ε_f, because $h_s >> h_f$. Then, ε_f is given directly by the differential thermal expansion:

$$\int_{T_o}^{T_s} (\alpha_{Tf} - \alpha_{Ts})\, dT \approx (\bar{\alpha}_{Tf} - \bar{\alpha}_{Ts})(T_s - T_o) = \varepsilon_f = \left(\frac{1-\nu}{Y}\right)_f \sigma_f \quad (5.53)$$

where T_s is the deposition T, T_o is the T after cool-down, and the $\bar{\alpha}_T$ represents the average α_T over the T range. For films deposited at high T, the ε_f resulting from this thermal mismatch can approach the fracture point upon cool-down. For example, if the α_T difference is 10^{-5}/K and $T_s = 1300$ K, then $\varepsilon_f = 10^{-2}$. Note that Eq. (5.53) predicts only the *extrinsic* component of film stress. To estimate the *intrinsic* component, the total stress must be measured and the extrinsic component, calculated from Eq. (5.53), must be subtracted.

Total film stress is easily measured by the curvature it produces in the substrate. We neglected this curvature in Fig. 5.32*b* while examining only the *uniform* component of substrate stress, σ_s, that resulted from the force balance. Now, in Fig. 5.32*c*, we neglect σ_s while examining only the component of substrate stress that is due to bending. Bending results from the exertion of film force only on the front face of the substrate. This bending produces a torque or bending moment, M_f, about the y axis, which we have placed in the center of the substrate for convenience. For unit width in y,

$$M_f\,(\mathrm{N\cdot m}) = F_f z = \sigma_f h_f (h_s/2) \tag{5.54}$$

This moment must be balanced by the moment generated by substrate bending. The latter moment is illustrated by the stress arrows in Fig. 5.32*c*, where the bending stress, $\sigma(z)$, is seen to vary from tensile on the front face ($z = +h_s/2$) to compressive on the back ($z = -h_s/2$). Stress $\sigma(z)$ can be related to the substrate radius of curvature, r, by examining the trapezoidal distortion, which is illustrated to the right of the figure. In a rectangle of substrate material having variable half-height z and arbitrary half-width $a/2$, the expansion at the top edge due to the bending is δ, and the strain is $\varepsilon(z) = \delta/(a/2)$. This is the same construction as Fig. 5.24, and again we can write, by similar triangles, that $(a/2)/r = \delta/z$. Thus,

$$\varepsilon(z) = \frac{z}{r} = \left(\frac{1-\nu}{Y}\right)_s \sigma(z) \tag{5.55}$$

Integrating over the substrate thickness using this expression for $\sigma(z)$ gives the substrate bending moment:

$$M_s = \int_{-h_s/2}^{+h_s/2} \sigma(z)\cdot z\,dz = \left(\frac{Y}{1-\nu}\right)_s \cdot \frac{h_s^3}{12r} = \left(\frac{Y}{1-\nu}\right)_s I_z \cdot \frac{1}{r} \tag{5.56}$$

where I_z is the "moment of inertia" of the substrate for unit width in y. Setting this equal to M_f from Eq. (5.54) gives the desired result relating total film stress to substrate curvature:

$$\boxed{\sigma_f = -\left(\frac{Y}{1-\nu}\right)_s \frac{h_s^2}{6rh_f} = -\left(\frac{Y}{1-\nu}\right)_s \frac{Kh_s^2}{6h_f}} \tag{5.57}$$

where K (m^{-1}) is the curvature, with positive being convex on the film face and negative being concave. The same curvature occurs in the y

dimension of the substrate due to bending moments along the x axis, so that the substrate is deformed into a dome. The dome may be approximated as spherical, and the curvature is independent of the size or shape of the substrate, $L_x \times L_y$, as long as L_x, $L_y \ll r$. Note that for a given σ_f, $K \propto 1/h_s^2$, so that thin substrates considerably increase sensitivity to σ_f measurement.

Very small curvatures down to 10^{-4} m^{-1} can be measured by the change in angle of reflection of a laser beam as it is scanned across the substrate. Here the beam is acting as an optical lever. Alternatively, optical interference fringes against a reference flat can be measured, in which case any variation in K over the substrate is also readily detected. The laser-beam technique can also monitor the buildup of intrinsic stress *during* deposition with the use of a ribbon-shaped substrate suspended from one end only as shown in Fig. 5.33. By the way, single-crystal Si wafers are ideal substrates to use for studying the effects of process conditions on intrinsic stress, because they are: (1) thin and well-controlled in h_s to maximize sensitivity; (2) very flat; (3) low-cost and often reusable by etching off the film; and (4) very well characterized in Y, ν, and α_T. $Y/(1-\nu)$ is 181 GPa for Si(100) and 229 for Si(111), and α_T is 2.60×10^{-6}/K at 300 K (3.84 at 600 K and 4.38 at 1200 K). Most materials have higher α_T, although quartz and Invar (an Fe-Ni alloy) have α_T near zero.

We assumed above that σ_f is constant through the thickness of the film. In fact, σ_f can vary with z, but the curvature measurement gives only the average value across h_f. For example, we will see in Sec. 5.6.3 that intrinsic stress arises both from the growth process occurring near the surface and from any annealing which may be occurring throughout the bulk of the film during deposition at elevated T_s. Annealing will have more time to proceed during film deposition for that portion of the film deposited first. The annealing effect complicates the

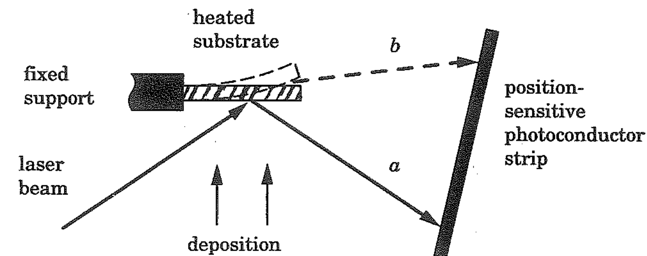

Figure 5.33 Technique for measuring intrinsic-stress build-up during deposition: (*a*) light beam before deposition and (*b*) after deposition of a compressive film.

determination of σ_f versus z using the Fig. 5.33 technique, because σ_f at a given z will continue to change after it has been measured, as deposition proceeds. Also, σ_f can vary with z whenever σ_f becomes so high that the film yields or creeps during deposition. This motion will always proceed in the direction of stress relaxation. Since it can happen more easily further from the constraining influence of a rigid substrate, it will cause σ_f to decrease with increasing z. Thus, σ_f can vary either way with z depending on conditions.

There is also *surface* stress, which can become a significant component of σ_f for very thin films [32]. Surface stress is the sum of the surface energy, γ, of Eq. (5.35) and an elastic-strain term, $d\gamma/d\varepsilon$. The latter arises from the fact that for the surface monolayer of crystalline solids, the equilibrium bond length parallel to the surface (Fig. 4.2) may be larger or smaller than that in the bulk due to charge rearrangement at the surface [33]. This means that the surface monolayer will want to expand or contract laterally, though it cannot do so because it is bonded to the bulk. The resulting strain is often *partially* relaxed by spontaneous atomic reconstruction of the surface (Sec. 6.5.3). Note that the energy stored in this strain is in *addition* to the surface energy arising from the creation of new surface by bond-breaking.

We also assumed above that $h_f << L$. However, when films are patterned into fine lines for integrated circuitry or are deposited over such lines, the line width becomes the relevant L, and one often has $h_f \rightarrow L$ as shown in Fig. 5.34. For example, a typical metal interconnect line on a memory chip is 1 µm wide and 1 µm thick. This is still another cause of σ_f variation with z. Recall that stress is maintained in a film by force transmitted across the interface from the constraining substrate. When $h_f << L$, the resulting stress in the film is parallel to the substrate as shown in Fig. 5.32*b*. However, the edge of a film can have no stress parallel to the substrate, because the edge is not connected to anything. Thus, a tensile film relaxes inward as shown qualitatively in Fig. 5.34*a*, and also upward at the edges in accordance with Poisson's ratio. At lateral positions $>h_f$ away from either edge, the edge effect has attenuated [34], and the film in this middle region behaves as in the $h_f << L$ case, with the lateral tensile stress distributed through its thickness and with a resulting compressive strain downward. The force $\sigma_f h_f$ in this middle region must be transmitted to the substrate at the edge of the film, as shown by the arrows. But at the edge, the stress is distributed only over the lower portion of the film thickness, so there is a stress *concentration* at this corner both in the tensile stress of the film and in the shear stress of the interface. When continuous films are deposited over steps in underlying layers, similar stress concentrations develop at the step corners. These concentrations can produce adherence failure.

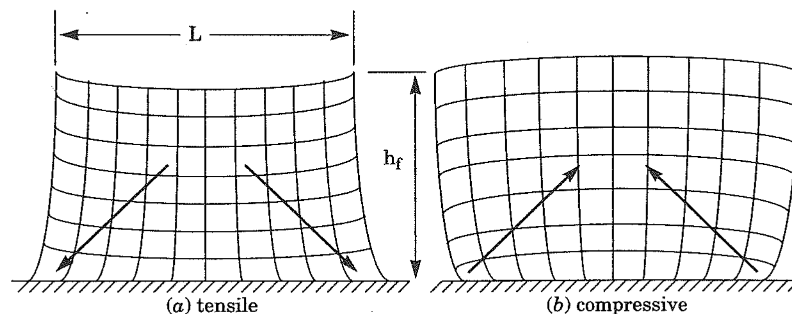

Figure 5.34 Qualitative cross-sectional strain distributions near the edges of patterned film lines that are under stress.

5.6.2 Problems

Whether or not film stress causes problems in the application at hand depends on the circumstances and on the level of the stress. Actually, a small level of compressive stress can strengthen a film, because it reduces the chances of the film being put under sufficient tensile stress to cause fracturing in severe mechanical applications such as tool-bit coatings. Corrosion resistance is also improved by avoiding tensile stress. Small levels of strain of either sign can improve the properties of epitaxial structures in electronic applications, as we will see in Chap. 6. However, high stresses usually lead to problems. The upper limit to stress is the point of catastrophic failure, which is illustrated in cross section in Fig. 5.35. Tensile stress failure is characterized by cracking, which appears as a mosaic pattern when viewed from the top. The cracked film may then peel away from the substrate at the crack edges, where the stress is concentrated. Compressive stress failure is characterized by de-adherence and buckling, which from the top appears sometimes as domes or bubbles and sometimes as an undulating meander pattern looking like a mole tunnel.

At lower stress levels, other problems can still arise. The curvature induced in the substrate is unacceptable in applications where flatness is important, such as mirrors. When a film is stressed into plastic

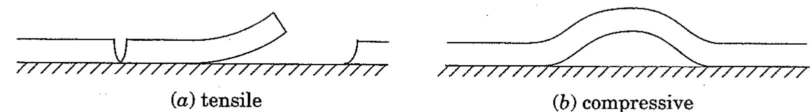

Figure 5.35 Catastrophic failure from film stress.

deformation, its structure degrades. In polycrystalline films, yield at grain boundaries weakens bonding there and thus aggravates the various grain-boundary problems that were discussed in Sec. 5.4.3. In epitaxial films, yield occurs by the generation and glide of dislocations within the crystal lattice (see Sec. 6.6); these defects are especially detrimental to electronic properties.

One particularly dramatic and difficult problem occurs upon compressively stressing polycrystalline films of soft metals beyond the yield point at $T/T_m > 0.4$ or so. The stress relaxes itself by transporting film material to the surface and growing "hillocks" of it there. The transport apparently occurs by diffusion of the film material along grain boundaries ("Coble creep"). Diffusion is known to be faster there than in the bulk of the grains, and in the case of Pb, the known value of the grain-boundary diffusion coefficient was found to agree with the rate of hillock growth [34]. Another clue is that hillocks tend to occur over grain boundaries. An unusually steep hillock in Al is illustrated in Fig. 5.36. Hillocks cause light scattering in optical applications and short circuits through overlying insulator films in electronic applications. The reverse problem of void formation also occurs in the same soft metals when they are stressed in *tension* beyond the yield point. Then, material is selectively transported out of certain regions of the film to relax the stress elsewhere.

The initial stress of a film after cool-down from the deposition T can often be kept below the elastic limit by minimizing deposition T and

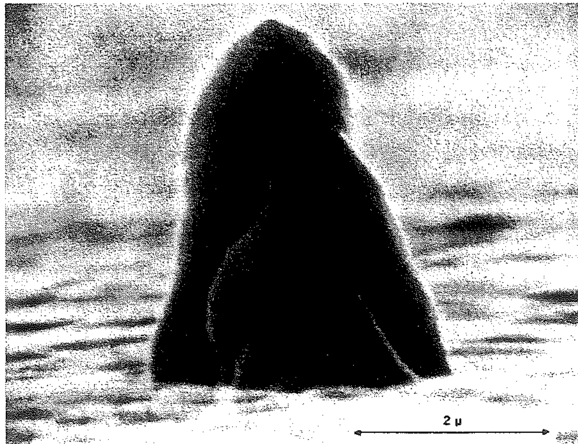

Figure 5.36 Scanning electron micrograph of a 3 μm high hillock in a 0.7 μm thick Al film. (Source: Reprinted from Ref. 35 by permission.)

adjusting other deposition conditions, thus avoiding hillocks and other problems of the yield regime. However, when subsequent processing or use of the film involves thermal cycling, as in integrated-circuit manufacture, thermal mismatch to the substrate can increase stress beyond the yield point. Figure 5.37 is adapted from a thermal cycling study [36] between room T and 450° C of Al-1% Si and Al-2% Cu films sputter-deposited onto oxidized Si(100) at room T. It illustrates the typical hysteresis behavior that is observed in thermal cycling of films beyond their yield point. The exact shape of the curve will depend on the material, its thermal history, and the T ramp rate, but some or all of the regimes shown will be observed. This particular film contains intrinsic tensile stress after room-T deposition. Upon subsequent heating, the high α_T of the film relative to the substrate causes film stress, σ_f, to become compressive, and the slope in this elastic regime is determined by Eq. (5.53). The first decrease in slope with further T increase occurs in this film not because of yield but because of the onset of recrystallization and grain growth, which reduces the disorder frozen into the film during deposition. This amounts to a transition from a Z1 or ZT structure to a Z2 or Z3 structure (see Sec. 5.4). Since the crystallites have higher density than disordered material, these crystallization processes cause the film to want to contract. This produces a tensile stress component that partially cancels the thermal mismatch compression and thus reduces the slope. Further T increase causes the slope to reverse upon relaxation of the compressive stress due to yield. Then, at the start of the down ramp in T, elastic behavior is first observed, because the stress is now below the yield point even though the T is high. (Elastic behavior can also be observed beyond the yield point when the rate of strain change is much higher than the rate of relaxation by yield.) With further cooling, the film crosses into tension, and when the tensile stress becomes high enough, the film again yields, and the slope decreases. Finally, the slope increases back to-

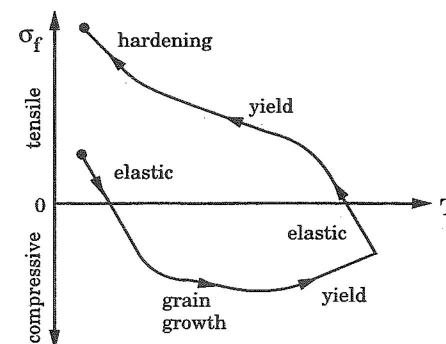

Figure 5.37 Typical hysteresis behavior during thermal cycling of metal films when α_T is higher for the film than for the substrate.

ward the elastic value near room T due to hardening of the film against yielding at the lower T. Information about how various materials yield as a function of strain and T is available in deformation-mechanism maps [37].

5.6.3 Intrinsic stress

Intrinsic stress is incorporated into the film during deposition or post-deposition treatment; that is, it is intrinsic to these processes. Here, we will focus on stress arising from the deposition process. Intrinsic stress is commonly observed and has at least three origins: chemistry, microstructure, and particle bombardment.

Chemical reactions occurring in the deposition process can produce stress whenever they continue to occur to some extent beneath the growth surface, where the film structure is beginning to become frozen. Reactions which add material to this structure produce compressive stress, and those which remove it produce tensile stress, as one would expect. For example, chemically-reactive metals such as Ti which are deposited in poor vacuums or with O_2 background gas deliberately added can develop compressive stress [38] due to oxidation proceeding beneath the surface. Conversely, plasma-deposited silicon nitride (SiN_xH_y) made using SiH_4 and NH_3 gas develops high tensile stress because the triaminosilane precursor radical, $Si(NH_2)_3$, continues to evolve NH_3 gas from beneath the growth surface as it chemically condenses toward Si_3N_4 [39] (more in Sec. 9.6.4.2). These chemical processes can also modify stress during post-deposition treatment.

The microstructure of the film and its evolution with time beneath the growth surface can produce tensile stress. In terms of the zone structure discussed in Sec. 5.4, films that are well into Z1 have little stress, because stress cannot be supported across the microvoids which separate the columns of material. However, as the film moves toward the dense ZT or Z2 structures, the microvoids collapse enough to allow atomic attraction across them. Then, tensile-strained bonds develop, and the resulting tensile stress cannot relax if the material is within its elastic limit. Additional tensile stress can develop when recrystallization of disordered material or grain growth is proceeding beneath the surface of ZT or amorphous films, due to densification as mentioned in the discussion of Fig. 5.37. At higher T/T_m well into Z2 or Z3, yield occurs more easily and partially or completely relaxes these microstructural tensile stresses. Both densification and yield can further occur during post-deposition annealing.

Bombardment of the film surface by ions or energetic neutrals can produce compressive stress both by implanting these particles into the film and by momentum transfer to surface atoms. Momentum transfer

forces the surface atoms into closer proximity to each other than their relaxed bond lengths, and they become frozen in this compressed state when T is low. This is similar to the shot-peening and ball-peen-hammering processes which are used to compressively stress the surfaces of bulk metals, and it has thus been termed “ion peening” [38]. The energetic bombardment available in energy-enhanced film-deposition processes is a very effective way to counteract the tensile stress which arises from chemical or microstructural effects. In some cases, it can be controlled to just neutralize the stress.

Figure 5.38 shows the general behavior of film stress, σ_f, with process pressure, p, in sputter deposition (Thornton, 1986). Some effects of p on microstructure were discussed in Sec. 5.4.1. The transition from Z1 to ZT with decreasing p in sputtering is due both to a decrease in the spread of incident angle of depositing particles and to an increase in particle kinetic energy. This transition causes tensile stress to rise as the microvoids collapse. The height of the maximum tensile stress correlates with decreasing deposition/melting T, T_s/T_m, for many metals (Thornton, 1989), presumably because the stress is less able to anneal itself out at lower T_s/T_m. At still lower p, stress drops again due to compaction by energetic bombardment, and for some materials it becomes compressive. Compressive stress can also be obtained by using a negative bias on the substrate to increase ion-bombardment energy. Ion bombardment has the same effect in other energy-enhanced processes, and is discussed in a general context in Sec. 8.5.3.

Despite the above discussion, the origin of intrinsic stress in a given deposition situation is frequently not known, and much work remains to be done in this area to understand and control this stress.

5.7 Adhesion

Loss of film adhesion requires both high stress and weak bonding at the interface to the adjoining layer or substrate. Then, the interfacial

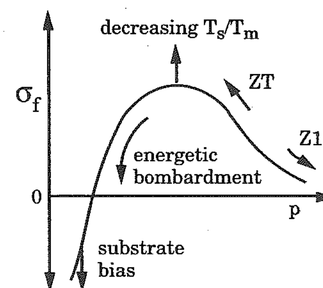

Figure 5.38 Behavior of film stress with sputter-deposition conditions.

bond will fail by either peeling or buckling as shown in Fig. 5.35, depending on the sign of the film stress. The stresses contributing to de-adherence include not only the intrinsic and thermal-mismatch stresses, but also the stress applied to the film in its application. For example, coatings on cutting tools are subjected to high local stress at the point of contact with the material being cut. Coatings on high-power laser optics are subjected to high thermal-mismatch stress during the laser pulse ("thermal shock"). Wire bonding to metal films on integrated-circuit chips generates high stress whether the thermocompression or the ultrasonic bonding method is used. In such applications, strong interfacial bonding is crucial. Conversely, if external stresses are not present in the intended application and if the initial film stress is low, weak interfacial bonding may not result in de-adherence. Also, since the shear force at the interface is proportional to film thickness in accordance with Eq. (5.52), thicker films fail more easily under a given stress than do thinner ones. Gas evolution from films can also cause de-adherence if the gas becomes trapped at the interface so that it builds up pressure there. Sputter-deposited films often contain a percent or two of Ar, and plasma-CVD films can contain tens of percent of H.

Sometimes, one *wants* to remove the film from the substrate to produce a free-standing membrane for use as, say, an x-ray window or a target in a high-energy-physics experiment, in which case interfacial bonding needs to be minimized. But more commonly, good adherence is desired. We have already examined the factors which affect interfacial stress. Here, we examine first the factors that inhibit bonding and then the process remedies used to promote bonding.

Since chemical bonding forces extend only a few tenths of nm, only one monolayer of poorly-bonded contaminant can be sufficient to prevent bonding of the depositing film material to the substrate. We saw in Sec. 3.4.2 that ordinary surfaces tend to be contaminated with water, oil, and salts. Most of this contamination can be removed by solvent degreasing followed by rinsing in deionized water, but the last monolayer or two usually remains. Even if chemical etching [40] is employed as a final step, upon exposure to the air most surfaces will readsorb a monolayer or two of water, organic vapors, and CO_2. These species will physisorb on any surface even at partial pressures well below their saturation vapor pressures whenever their bonding is stronger to that surface than to their own condensed phase, as we discussed in Sec. 5.1. In special cases such as H on Si and on GaAs (Sec. 6.3), surfaces can be chemically passivated against contaminant adsorption, but then one has a monolayer of adsorbed passivant to deal with.

The adsorbed contaminants end up sandwiched between the film material and the substrate and can block chemical bonding between

the two materials. Thus, unless both the substrate and the film bond *chemically* to the adsorbate, the interfacial bonding remains weak. In some cases, chemical bonding does occur. For example, metals that form strongly bonded surface oxides can be deposited one upon the other with good adherence because the interfacial oxide forms a strong bridge of chemical bonding between the metals. Indeed, film adhesion correlates well with the free energy of formation of the film metal's oxide [41]. The metals Ti, Zr, Cr, and Al have particularly strong oxides, while Zn, Cu, and the noble metals have weak ones. Adsorbed water, organics, and CO_2 also are likely to form strong oxide and carbide bonds at the surfaces of reactive metals.

Even if the adsorbates are removed in the deposition chamber so that deposition is carried out on an atomically-clean surface, interfacial bonding will be weak if the film and substrate materials have such different bonding character (covalent, ionic, or metallic) that they do not easily bond to each other. This was discussed in Sec. 5.3.1 following Eq. (5.36) on the wetting criterion. There, the poor adherence of Au to SiO_2 was improved by inserting 10 nm or so of the "glue" layer Ti, so named because it forms both strong metallic bonds to the Au and strong covalent bonds to the oxide. Conversely, even an active metal like Ti can be peeled easily from the ionically-bonded material CaF_2. A film of CaF_2 is therefore useful as a "parting layer" in making free-standing metal membranes.

There are several ways to remove physisorbed molecules once the substrate is in the deposition chamber where it will not become recontaminated. Heating either desorbs them or activates their chemisorption, per Eq. (5.5). In the chemisorbed state, they may no longer inhibit interfacial bonding. The progress of desorption can be followed by the accompanying pressure burst or with a mass spectrometer (Sec. 3.5). Alternatively, exposure to H_2 plasma reduces many oxides and removes them as H_2O, whereas O_2 plasma oxidizes organics and removes them as CO_2 and H_2O. Bombardment with ions of >100 eV or so from a plasma or from an ion gun can remove any surface species by sputter erosion. However, a small fraction of this material is sputtered forward instead and is thus embedded beneath the surface—the so-called "knock-on" effect. Plasma operation is discussed in Chap. 9.

Energy input to the substrate surface from ions, electron beams, or UV light can desorb contaminants and can also break bonds within the surface, thereby activating the surface toward bonding to the film material. Irradiation *after* film deposition can also improve adherence as long as the film is thin enough so that the radiation penetrates to the interface [42, 43]. In energy-enhanced deposition processes such as sputtering, activation energy for interfacial bonding is carried by the depositing material itself, and this results in consistently better

adhesion than is achieved using thermal-evaporative deposition of the same material.

More severe measures can be taken to further improve adherence if interface abruptness is not important. Mechanical or chemical roughening of the substrate improves adherence by increasing the bondable surface area and also by mechanically interlocking the materials on a microscopic scale. Deliberately grading the interface composition improves adherence in several possible ways: dispersion of interfacial contaminants, increase in number of bonds between the two materials, and inhibition of fracture propagation along the interface. Generally, it is not clear which mechanism dominates. Gradation can be achieved by thermal interdiffusion (Fig. 5.27*b*) or by "ion mixing" (Sec. 8.5.3). However, interdiffusion can *weaken* the interface if voids develop (Fig. 5.27*e*). Ion mixing occurs when ion bombardment in the keV range is present at the start of deposition to cause significant knock-on mixing of the film material into the substrate. When the ions are of film material itself, they mix also by their own shallow implantation. These effects typically extend over a range of a few nm. Between successive film layers, gradation can alternatively be achieved by gradually switching from the deposition of one material to the other.

One can see that the various adherence remedies each have their own problems and constraints. One should first establish that there is an adherence problem for the application at hand before taking remedial measures. The choice of remedy will usually be determined by constraints imposed by the materials involved, the interfacial structure desired, and the equipment available. Sometimes, changing the deposition conditions to reduce film stress is a better solution.

5.8 Temperature Control

It is clear by now that substrate temperature, T_s, is a very important variable in the deposition process, having profound effects on the structure and composition of films and interfaces. Since T is such a common quantity, its measurement is often treated casually. Unfortunately, T_s measurement is very difficult in the vacuum or partial-vacuum environment of most thin-film deposition processes. This is because thermal coupling to the substrate in vacuum is poor; yet one usually does not want to attach a T sensor directly to the substrate, because the substrate would become contaminated and/or needs to be moved during the process. There are noncontact techniques for measuring T, but they have shortcomings which we will examine later.

Sometimes substrates are suspended in the deposition chamber—especially when they are irregularly shaped. In this case, T control is

particularly difficult unless the entire chamber is heated to T_s so that the substrate is immersed in an isothermal enclosure. Flat substrates such as glass plates or Si wafers can be placed on a heated platform. The platform T, T_h, can be measured accurately and can be controlled closely by the feedback techniques discussed in Sec. 4.5.3. It is often assumed in the literature that $T_s = T_h$ in such a situation. The fallacy of this assumption is made clear by examining the substrate-platform interface on a microscopic scale, which is done schematically in Fig. 5.39. Because most surfaces are not atomically flat, intimate atomic contact occurs only at a few points, and these points add up to a negligibly small fraction of the macroscopic interface area. Even if the substrate is clamped around the periphery to a platform which is slightly domed so that the clamping force is distributed over the whole area, the contact area will still be negligible unless one of the two surfaces is so soft that it deforms and thus conforms to the other surface. Conformable thermal-contact materials can be used at the interface, but then contamination is a concern. Vacuum grease is one such material, but its vapor pressure rises steeply with T. Ga-In eutectic alloy is liquid at room T and still has low vapor pressure at high T (see Appendix B), but it requires chemical etching to remove, and Ga also alloys with most metals. Pure In is an alternative which melts at 156° C, and there are many refractory metals with which it does not alloy [26]. However, in the majority of cases when conformable or liquid contacting layers cannot be used, contact area between substrate and platform will be negligible.

The good heat conduction provided by the solid phase can only occur where there is *atomic* contact to transfer the heat by phonon vibrations and, in the case of metals, by electrons. Therefore, for the geometry of Fig. 5.39 where atomic contact area is negligible, heat transfer under vacuum can occur only by radiation, which results in a substantial difference between T_h and T_s. When pressure is higher, gas-phase

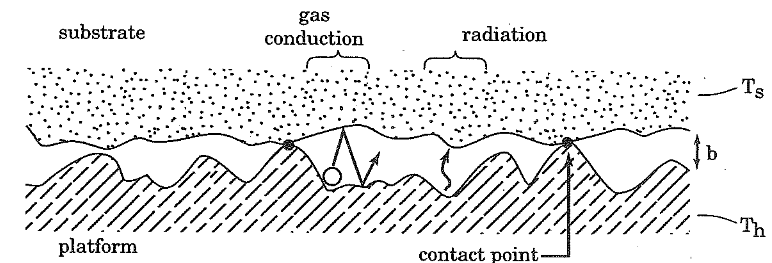

Figure 5.39 Microscopic schematic of the typical interface between a substrate and its heated platform.