13. Adamson, S., C. O'Carroll, and J.F. McGlip. 1989. "Monte Carlo Calculations of the Beam Flux Distribution from Molecular Beam Epitaxy Sources." *J. Vac. Sci. Technol. B* 7:487.
14. Shiralagi, K.T., A.M. Kriman, and G.N. Maracas. 1991. "Effusion Cell Orientation Dependence of Molecular Beam Epitaxy Flux Uniformity." *J. Vac. Sci. Technol. A*, 9:65.
15. Wasilewski, Z.R., G.C. Aers, A.J. SpringThorpe, and C.J. Miner. 1991. "Studies and Modeling of Growth Uniformity in Molecular Beam Epitaxy." *J. Vac. Sci. Technol. B* 9:120.
16. Manufactured by Leybold Inficon, Inc.
17. Benes, E. 1984. "Improved Quartz Crystal Microbalance Technique." *J. Appl. Phys.* 56:608.
18. Smith, D.L., and V.Y. Pickhardt. 1975. "Molecular Beam Epitaxy of II-VI Compounds." *J. Appl. Phys.* 46:2366.
19. Theeten, J.B., and D.E. Aspnes. 1981. "Ellipsometry in Thin Film Analysis." *Ann. Rev. Materials Science* 11:97.
20. Quinn, W. E., D.E. Aspnes, M.J.S.P. Brasil, M.A.A. Pudensi, S.A. Schwarz, M.C. Tamargo, S. Gregory, and R.E. Nahory. 1992. "Automated Control of III-V Semiconductor Composition and Structure by Spectroellipsometry." *J. Vac. Sci. Technol. B* 10:759.

4.12 Recommended Readings

Aspnes, D.E., and R.P.H. Chang. 1989. "Spectroscopic Ellipsometry in Plasma Processing." In *Plasma Diagnostics, vol. 2*, ed. O. Auciello and D.L. Flamm. Boston, Mass.: Academic Press.

Herman, M.A., and H. Sitter. 1989. *Molecular Beam Epitaxy*, Chap. 2. Berlin: Springer-Verlag.

1991. Proceedings of the Fifth International Symposium on Small Particles and Inorganic Clusters (Konstanz, Germany). *Zeitschrift fur Physik D*, vol. 19 and 20 (in English).

Sears, F.W. 1950. *An Introduction to Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics*. Cambridge, Massachusetts: Addison-Wesley.

Stoecker, W.F. 1989. *Design of Thermal Systems*, 3rd ed, Chap. 2. New York: McGraw-Hill.

Tompkins, H.G. 1993. *A Users's Guide to Ellipsometry*. Boston, Mass.: Academic Press.

Chapter

5 Deposition

We now come to the heart of the thin-film process sequence. Deposition may be considered as six sequential substeps, and we will examine these one by one in this chapter. The arriving atoms and molecules must first (1) adsorb on the surface, after which they often (2) diffuse some distance before becoming incorporated into the film. Incorporation involves (3) reaction of the adsorbed species with each other and the surface to form the bonds of the film material. The (4) initial aggregation of the film material is called nucleation. As the film grows thicker, it (5) develops a structure, or morphology, which includes both topography (roughness) and crystallography. A film's crystallography may range from amorphous to polycrystalline to single-crystal. The last is obtained by epitaxy—that is, by replicating the crystalline order of a single-crystal substrate. Epitaxy has special techniques and features which are dealt with separately in Chap. 6. Finally, (6) diffusional interactions occur within the bulk of the film and with the substrate. These interactions are similar to those of post-deposition annealing, since they occur beneath the surface on which deposition is continuing to occur. Sometimes, after deposition, further heat treatment of a film is carried out to modify its properties. For example, composition can be modified by annealing in a vapor, and crystal growth can be achieved by long annealing or by briefly melting. These post-deposition techniques will be mentioned only in passing in this book.

In the latter part of the chapter, we will examine three key practical aspects of deposition: the development of mechanical stress, adhesion of the film to the substrate, and substrate T control. The issue of composition control during deposition will be reserved for Chap. 6, since it is more easily studied under epitaxial growth conditions.

For the present, we will consider that only thermal energy is being supplied to the surface except where energy enhancement is specifically noted. Adding energy to the deposition process by nonthermal means is an important process technique which can activate deposition reactions and alter film structure. It was introduced in Sec. 4.5.4 and will be treated extensively in Chaps. 8 and 9.

5.1 Adsorption

Consider a molecule approaching a surface from the vapor phase, as shown in Fig. 5.1. Upon arriving within a few atomic distances of the surface, it will begin to feel an attraction due to interaction with the surface molecules. This happens even with symmetrical molecules and with inert gases, neither of which have dipole moments. It happens because even these molecules and atoms act as *oscillating* dipoles, and this behavior creates the induced-dipole interaction known as the van der Waals force or London dispersion force. Polar molecules, having permanent dipoles, are attracted more strongly. The approaching molecule is being attracted into a potential well like the one that was illustrated in Fig. 4.2 for condensation. Condensation is just a special case of adsorption in which the substrate composition is the same as that of the adsorbant. This is sometimes the case in thin-film deposition and sometimes not. In either case, the molecule accelerates down the curve of the potential well until it passes the bottom and is repelled by the steeply rising portion, which is caused by mutual repulsion of the nuclei (more on this in Sec. 8.5.2.2). If enough of the

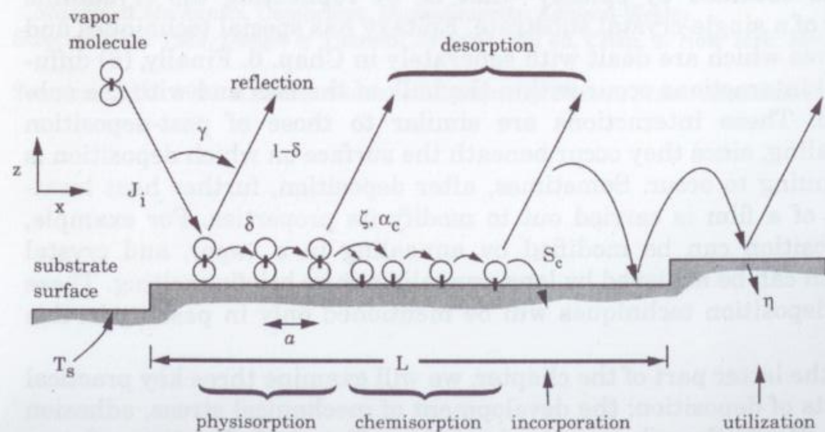


Figure 5.1 Adsorption processes and quantities. α_c is used only for condensation (adsorption of a material onto itself). A vertical connecting bar denotes a chemical bond.

molecule's perpendicular component of momentum is dissipated into the surface during this interaction, the molecule will not be able to escape the potential well after being repelled, though it will still be able to migrate along the surface. This molecule is trapped in a weakly-adsorbed state known as physical adsorption or physisorption. The fraction of approaching molecules so adsorbed is called the trapping probability, δ, and the fraction escaping (reflecting) is (1 − δ) as shown in Fig. 5.1. The quantity δ is different from the thermal accommodation coefficient, γ, which was defined by Eq. (2.33). In general, a molecule is at least partially accommodated thermally to the surface temperature, T_s, even when it is reflected without having been trapped.

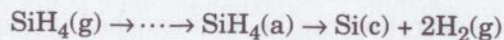
The physisorbed molecule is mobile on the surface except at cryogenic T, so it is shown hopping (diffusing) between surface atomic sites on Fig. 5.1. It may desorb after a while by gaining enough energy in the tail of the thermal energy distribution, or it may undergo a further interaction consisting of the formation of chemical bonds with the surface atoms; that is, chemisorption. If both adsorption states exist, the physisorbed state is called the precursor state. Chemisorption involves the sharing of electrons in new molecular orbitals and is much stronger than physisorption, which involves only dipole interactions. These two types of adsorption can be distinguished in almost all vapor-surface combinations, so they constitute a valuable model with which to analyze any surface process. This model has long been applied to heterogeneous catalysis, thin-film deposition, and condensation of molecular vapors. Recent theory indicates that even the condensation of a monatomic vapor such as Al can involve both adsorption states, the precursor state in that case being an Al-Al dimer whose bonding to the bulk Al is inhibited by the existence of the dimer bond [1]. In such a case, and in the case of condensing *molecular* vapors such as As_4, the vapor would not be considered actually *condensed* until it had become fully incorporated into the solid phase by chemisorption. Thus, the condensation coefficient, α_c, defined by Eq. (4.18) is that fraction of the arriving vapor that becomes not only trapped but also chemisorbed, as indicated in Fig. 5.1. However, the term α_c is not used in the case of chemisorption on a foreign substrate. Then, we speak of the chemisorption reaction probability, ζ, which will be derived later. The precursor model may also be applied to cases where both of the adsorption states involve chemical bonding, but where the bonding in one state is weaker than in the other.

Since some of the physisorbed species eventually escape back into the vapor phase, a third term, called the sticking coefficient, S_c, is used to denote that fraction of the arriving vapor that remains adsorbed *for the duration of the experiment*. Since this duration is arbitrary, S_c has less of a fundamental meaning than δ and α_c, which are

determined solely by chemistry and energy. Nevertheless, S_c is very useful in thin-film deposition, since it is equal to the fraction of arriving vapor which becomes incorporated into the film. That is, this fraction becomes adsorbed and then buried before it can desorb. Note that in the limiting case of vapor-solid equilibrium, $S_c \rightarrow 0$ whereas $\alpha_c > 0$. One more fraction, which is useful in CVD practice, is the utilization fraction, η, of a chemical vapor. Chemical-vapor molecules diffuse around in the deposition chamber and can hit the film surface many times before finally being swept downstream in the flow, as we will see in Chap. 7 and as shown in Fig. 5.1. Thus, η can approach unity even when S_c is very low.

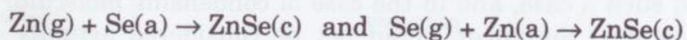
There are many examples of precursor adsorption in thin-film technology. The precursor phase exists any time there is both a weak and a strong bonding state of the adsorbing vapor with the substrate or film surface. Here are some examples:

1. In most CVD reactions, the feed vapors adsorb as molecules and then undergo the reactions which break their molecular bonds and form new bonds to the film surface; thus,



$$SiH_4(g) \rightarrow \cdots \rightarrow SiH_4(a) \rightarrow Si(c) + 2H_2(g)$$

   where (g), (a), and (c) denote the gas, adsorbed, and condensed phases, and where $(\cdots)$ denotes a series of intermediate reaction steps.
2. In deposition of compounds from separate vapor sources of each element, adsorbing vapor bonds much more strongly to those surface sites occupied by the other element; thus,



$$Zn(g) + Se(a) \rightarrow ZnSe(c) \quad \text{and} \quad Se(g) + Zn(a) \rightarrow ZnSe(c)$$

3. Silicon that is chemically passivated by an atomic layer of H reacts with adsorbates mainly at those few sites that are missing an H atom. On the H-passivated sites, adsorbates remain only physisorbed.
4. Atomically flat surfaces often bond more strongly with adsorbates at atomic steps such as those shown at a spacing of L in Fig. 5.1.
5. Upon adsorption, atoms of low-reactivity metals often bond much less easily to nonmetallic substrates than to those sites containing another metal atom.

We will revisit these examples after examining the energetics of the precursor adsorption model in more detail.

Consider a hypothetical diatomic gas-phase molecule $Y_2(g)$ adsorbing and then dissociatively chemisorbing as two Y atoms. Figure 5.2

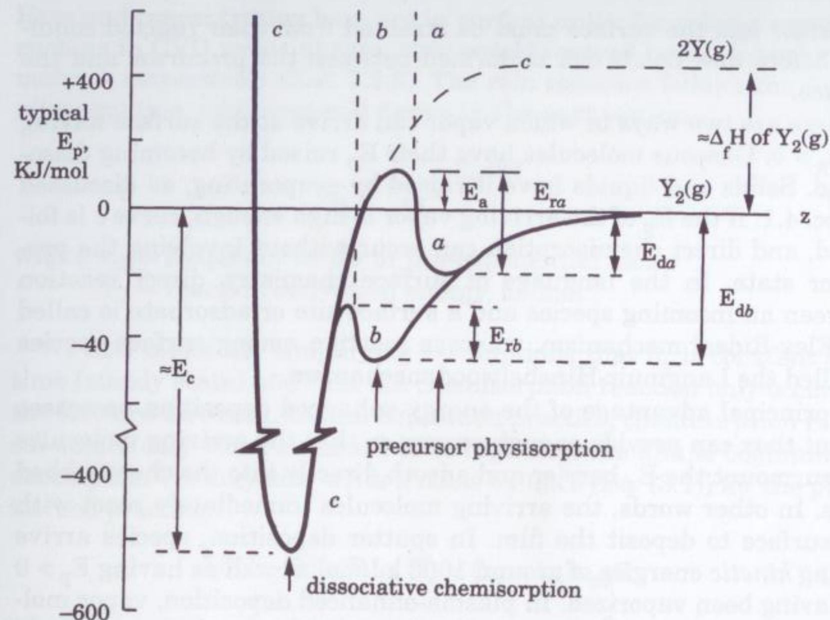


Figure 5.2 Energetics of the precursor adsorption model. Energy scale is typical only.

shows a diagram of the potential energy versus molecular distance, z, from the surface. This is similar to Fig. 4.2 for condensation except that we have changed from the molecular (ε_p) to the molar (E_p) quantities of potential energy which are more conventional in chemistry. The energy scales shown represent typical bond strengths. Three curves are shown: two alternate ones for the precursor state (*a* and *b*), and one for the chemisorbed state (*c*). By convention, the zero of E_p is set at the E_p of the element Y in its thermodynamic standard state, which we specify for this element to be the diatomic molecule in the gas phase. In fact, all gaseous elements except the inert gases have diatomic standard states. Note that lifting *atomic* Y out of its potential well along curve *c* results in a much higher E_p in the gas phase, which corresponds roughly to the heat of formation, $\Delta_f H$, of 2Y(g) from $Y_2(g)$. [$\Delta_f H$ usually can be found in thermodynamic tables (Refs. 1–3, Chap. 4).] The result of this high E_p for Y(g) is that curves *a* and *c* intersect at positive E_p, meaning that there is an activation energy, E_a, to be overcome for $Y_2(g)$ to become dissociatively chemisorbed. For the deeper precursor well, *b*, chemisorption is not "activated," though there still is a barrier, as shown. The level of E_{ra} or E_{rb}, and hence of E_a, is determined by the degree to which the bonds within both the

precursor and the surface must be strained from their relaxed condition before new bonds can be formed between the precursor and the surface.

There are two ways in which vapor can arrive at the surface having an $E_p > 0$. Gaseous molecules have their E_p raised by becoming dissociated. Solids and liquids have it raised by evaporating, as discussed in Sec. 4.1. If the E_p of the arriving vapor is high enough, curve *c* is followed, and direct chemisorption can occur without involving the precursor state. In the language of surface chemistry, direct reaction between an incoming species and a surface site or adsorbate is called the Eley-Rideal mechanism, whereas reaction among surface species is called the Langmuir-Hinshelwood mechanism.

A principal advantage of the energy-enhanced deposition processes is that they can provide enough energy so that the arriving molecules can surmount the E_a barrier and adsorb directly into the chemisorbed state. In other words, the arriving molecules immediately react with the surface to deposit the film. In sputter deposition, species arrive having *kinetic* energies of around 1000 kJ/mol as well as having $E_p > 0$ by having been vaporized. In plasma-enhanced deposition, vapor molecules become dissociated in the plasma and thus arrive along curve *c*, above the E_a barrier. Thus, an energy-enhanced process can supply E_a to the arriving species either as kinetic energy of accelerated molecules or as potential energy of dissociated ones.

Conversely, in thermally controlled deposition processes such as evaporation and CVD, the vapor often adsorbs first into the precursor state; that is, it falls to the bottom of the well on curve *a* or *b*. Thence, it may either chemisorb by overcoming the barrier $E_{r(a,b)}$ shown in Fig. 5.2, or it may desorb by overcoming the heat of physisorption, which is roughly $E_{d(a,b)}$. The competition between these two reactions results in a net rate of chemisorption whose behavior we would like to describe, since it is the basic film-forming reaction. We start with the conventional expression for the rate of a first-order chemical reaction, first-order meaning that rate is proportional to the concentration of one reactant; thus,

$$R_k = k_k n_s = k_k n_{so} \Theta \tag{5.1}$$

where R_k = rate of the kth surface reaction per unit surface area, $mc/cm^2{\cdot}s$
k_k = rate constant, s^{-1}
n_s = surface concentration of reactant, mc/cm^2
n_{so} = monolayer surface concentration, mc/cm^2
Θ = fractional surface coverage by reactant

Rate and concentration here are in surface units; for volume reactions such as in CVD vapor phases, they would instead be in $mc/cm^3{\cdot}s$ and mc/cm^3, respectively (Sec. 7.3.2). The rate constant follows the Arrhenius equation, which we will derive in the next section:

$$k_k = \nu_{ok} e^{-E_k/RT} \tag{5.2}$$

where ν_{ok} = frequency factor or pre-exponential factor
E_k = reaction activation energy, kJ/mol

We now make the simplifying assumptions that n_s is constant over time (steady state) and that the chemisorption reaction only occurs in the forward direction. In film deposition practice, chemisorption reversal occurs only when surface T is so high that the film is beginning to decompose. We may now write a mass balance [Eq. (3.1)] for the physisorbed precursor:

$$J_i \delta (1 - \Theta) = R_r + R_d = (k_r + k_d) n_{so} \Theta \tag{5.3}$$

where J_i = molecular impingement flux, $mc/cm^2{\cdot}s$ [Eq. (2.18)]
R_r = reaction (chemisorption) rate
R_d = desorption rate
n_s = surface concentration of the precursor

Here, we have also made the reasonable assumption that adsorption does not occur on the area already occupied by adsorbate (Θ). Rearranging this expression, we have

$$\Theta = \frac{J_i \delta / n_{so}}{J_i \delta / n_{so} + k_r + k_d} \tag{5.4}$$

and substituting into the chemisorption rate expression [Eq. (5.1)],

$$R_r = k_r n_{so} \Theta = \frac{J_i \delta k_r}{J_i \delta / n_{so} + k_r + k_d} \tag{5.5}$$

With this, we may now define the sticking coefficient in film deposition more precisely as

$$S_c = R_r / J_i \tag{5.6}$$

Note that S_c depends on both J_i and the rate constants; this dependency has consequences for CVD film conformality over topography, as will be discussed further in Sec. 7.3.3.

In the case of small J_i and thus small Θ, Eq. (5.5) simplifies (Weinberg, 1991) to

$$R_r = \frac{J_i \delta}{1 + k_d/k_r} = J_i \left[\frac{\delta}{1 + \frac{v_{od}}{v_{or}} e^{-(E_r - E_d)/RT_s}} \right] = J_i \zeta \tag{5.7}$$

This equation defines ζ, the chemisorption reaction probability—the fraction of that vapor impinging on *bare* $(1 - \Theta)$ sites that becomes chemisorbed rather than being reflected or desorbed. For $\Theta << 1$, $S_c \approx \zeta$; but for larger Θ, $S_c < \zeta$. It is important not to confuse these two terms. For the special case of film deposition from a single vapor having the same composition as the film, $\zeta \equiv \alpha_c$, the condensation coefficient. For the more complicated case of compound-film deposition from multicomponent vapors, the assumption of first-order kinetics on which Eqs. (5.5) and (5.7) depend is not always valid, because more than one reactant is involved. This case will be discussed more in Sec. 7.3.3.

The quantity R_r governs the rate of film deposition when k_r is the same from site to site along the surface and when T_s is not so high that decomposition or re-evaporation of the film is occurring. By Eq. (5.7), R_r is going to increase as the exponential energy term, $(E_r - E_d)$, decreases. If this term is positive, there is an activation energy, E_a, for chemisorption, as shown for curve *a* in Fig. 5.2, where $E_a = E_{ra} - E_{da} > 0$. If E_a is high enough, the film will fail to deposit unless T_s is raised to make the exponential term smaller. On the other hand, when chemisorption is not activated ($E_r < E_d$) as on curve *b*, R_r *decreases* with increasing T_s. Thus, R_r can go either way with T_s depending on the energetics at the surface. The activated case is very common in CVD. For example, Si deposits from silane gas (SiH_4) at elevated T but not at room T. Of course, if T_s becomes *too* high, the evaporation flux of the Si itself [J_v from Eq. (4.17)] will exceed R_r, and Si again stops depositing. The net deposition flux of Si is thus given by

$$J_r = R_r - J_v \tag{5.8}$$

For this case, there is a T_s *window* between reaction activation and re-evaporation, within which deposition can be achieved. The opposite case of R_r increasing with *decreasing* T_s is more difficult to identify, because other factors can cause J_r to increase with decreasing T_s even after J_v vanishes, such as increasing *nucleation*.

Nucleation is a complication that must often be added to the above model, which assumed identical kinetics for all surface sites. We will examine the nucleation process in more detail in Sec. 5.3, but it needs some introduction now. When nucleation is important, $J_r > 0$ in Eq. (5.8) only on certain active substrate-surface sites called nucleation sites or on nuclei of film material which have spontaneously accumulated. There are many ways in which nucleation sites may arise. In deposition examples (3) and (5) mentioned earlier in this section, R_r was higher at unpassivated Si surface atoms and at metal sites, respectively, because of a lower E_r there. In example (2) on the deposition of compounds from separate vapors, chemisorption only occurred at sites containing the other element, and in the atomically-flat-surface example (4), it only occurred at atomic steps. In examples such as (2) through (4), exploiting the nucleation phenomenon can result in films which have less roughness and defects, as we will see on many occasions later on. Note that the precursor bonds more readily to the surface than to itself in all three of these latter examples.

In other cases, where the precursor bonds more readily to itself than to the surface, nucleation limitations result in undesirably nonuniform deposition. One such case is example (5) for deposition of metals on a nonmetallic substrate, particularly weakly reacting metals such as Zn and Cd on glass or on an ionically bonded substrate such as NaCl or CaF_2. The activation energies (E_a) for chemical bonding of Zn and Cd to these substrates are very high because of the high bond strength between the elements making up the substrate. Consequently, the metals remain in the physisorbed state, from which they easily desorb, unless they encounter other adsorbed metal atoms with which to bond and form a nucleus. When the nucleus becomes large enough to behave like bulk metal, it will not evaporate as long as T_s is low enough so that J_v for the bulk metal is negligible in Eq. (5.8). Thus, at T_s well below where bulk Zn and Cd evaporation becomes significant, such as room T, one still obtains a most undesirable "splotchy" deposit on such substrates, consisting of islands of metal separated by areas devoid of deposition where nuclei have not yet had a chance to form. These bare areas represent a situation where $R_r << J_i\delta$ in Eq. (5.7).

Note that we have now seen two opposite examples for Zn vapor: Zn adsorbing onto a Se-covered surface, in which the bonding to the surface is much stronger than that of Zn to itself; and Zn on glass, in which the bonding is much weaker than that of Zn to itself. It is important to keep in mind that the re-evaporation rate of any species adsorbing onto a foreign surface will usually be vastly different from the evaporation rate of that species in its pure bulk form, because the rate is dominated by the degree of interaction with the foreign surface.

Faster adsorption onto a foreign surface is exploited in atomic-layer epitaxy (Sec. 6.5.5), and slower adsorption is exploited in selective deposition (Sec. 7.3.3).

Another dramatic example of nucleation-induced nonuniformity is diamond-film CVD, which is carried out using carbon-containing precursors such as methane (CH_4). Figure 5.3 shows diamond nuclei growing on the edges of etch pits in a Si substrate. The deposition rate on the diamond nuclei is high, and elsewhere it is zero. It is not clear why diamond nucleation occurs only at these particular sites, and this is a subject of intense current research. In many other cases, energy-enhanced techniques can activate bonding between adsorbing species and substrates; this enhances both nucleation and film adhesion to the substrate. Thus, sputter-deposited films often have better adhesion than those deposited from thermal evaporation sources.

The existence of certain sites which are more active for adsorption than the rest of the surface is common in thin-film processes. In such cases, access of the precursor to these favored sites can dominate the deposition kinetics. Access sometimes occurs mainly from the vapor

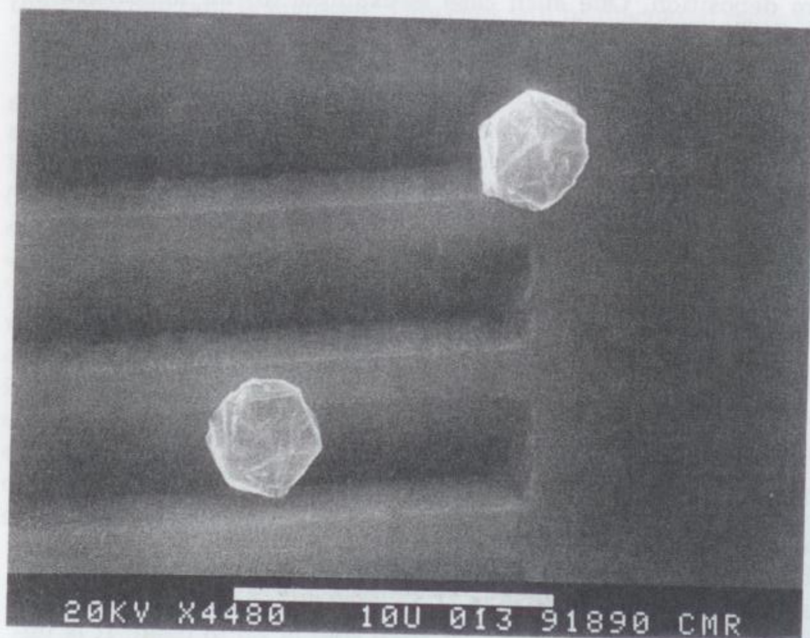


Figure 5.3 SEM photograph of two diamond nuclei growing on a patterned single-crystal Si substrate. The CVD of diamond from 1% CH_4 in H_2 at 4000 Pa was activated by a ≈2000° C Ta filament positioned 8 mm above the 900–1000° C substrate. (Previously unpublished photo courtesy of Paul A. Dennig from the laboratory of David A. Stevenson, Stanford University.)

phase (Eley-Rideal mechanism), but often surface diffusion (Langmuir-Hinshelwood mechanism) is the dominant route. Surface diffusion and nucleation will be addressed in the next two sections. These two complicating factors along with the possibility of non-first-order CVD reactions make the kinetic analysis of thin-film deposition very difficult, and consequently, such analysis has not often been performed. Nevertheless, simple models such as the one presented above for adsorption can provide a useful framework within which to think about deposition processes, provided that one remains aware of the limitations of the models being employed.

5.2 Surface Diffusion

Surface diffusion is one of the most important determinants of film structure because it allows the adsorbing species to find each other, find the most active sites, or find epitaxial sites. Various methods have been applied to measuring surface diffusion rates of adsorbed molecules, but most of this work has been done on chemical systems relevant to heterogeneous catalysis rather than to thin-film deposition. The role of surface diffusion in thin films has mainly been inferred from observations of film structure. However, the recent advent of the scanning tunneling microscope (STM) gives us the extraordinary power to directly observe individual atoms on surfaces in relation to the entire array of available atomic surface sites (Lagally, 1993). STM observation of the diffusion of these atoms should ultimately provide a wealth of data relevant to thin-film deposition.

We will develop here an expression for the rate of surface diffusion using absolute-reaction-rate theory. Although this approach cannot provide a quantitative estimate of the diffusion rate, it will provide valuable insight into what factors determine this rate. Figure 5.2 showed that adsorbed atoms or molecules reside in potential wells on the surface, but it did not consider the variation in well depth with position, x, along the surface. Figure 5.4*a* shows that this depth is periodic, or corrugated, with a potential-energy barrier of height E_s between surface sites. The top of the barrier is considered to be the "transition state" between surface sites, in the language of reaction rate theory. Figure 5.4*b* illustrates a typical adsorbate situation leading to this corrugation. It is a hexagonally close-packed surface lattice on which the adsorption sites are the centers of the triangles of surface atoms, and the transition state is the "saddle point" between them. Other bonding situations can lead to the adsorption sites being other points, such as the centers of the surface atoms. The surface diffusion process requires partly breaking the bond(s) between the adsorbate and the surface site so that the adsorbate may move to the

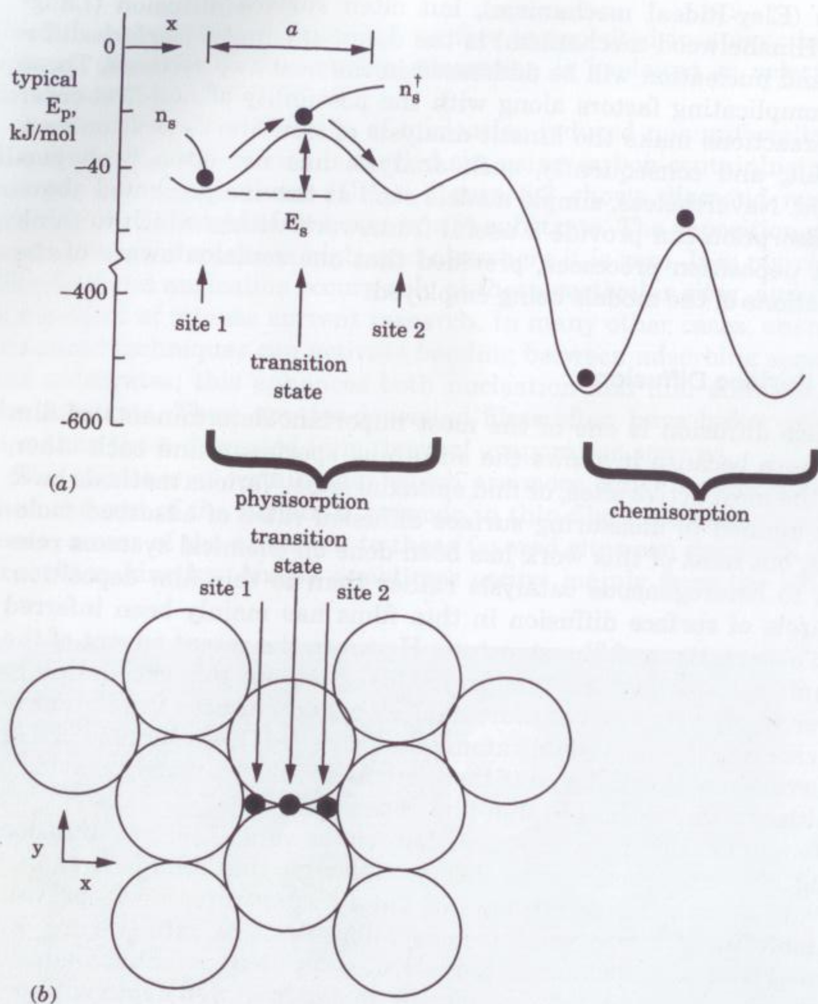


Figure 5.4 Surface diffusion: (*a*) potential energy vs. position x along the surface, and (*b*) typical adsorption sites on a surface lattice.

neighboring surface site and form new bonds there. This process may be viewed as an elementary form of chemical reaction, because any reaction involves the partial breaking of reactant bonds and partial formation of product bonds during motion of the atoms through a transition state. Thus, the principles to be discussed below apply to any chemical reactions, including those occurring in CVD.

There will be some flux, J_s (mc/cm·s), of adsorbate across the E_s barrier between sites 1 and 2 in the x direction of Fig. 5.4*b*. The flux here

is in surface units, which are per linear cm of crosswise distance, y, instead of the previously encountered volume flux units, which are per cm² of cross-sectional area. If the distance between sites is a, then the *rate* of barrier crossing by transition-state molecules, per unit area of surface, is

$$R_s = J_s/a \ (\text{mc/cm}^2\text{·s}) \tag{5.9}$$

Considering the adsorbate to be a two-dimensional gas at thermal equilibrium, the Maxwell-Boltzmann distribution applies to these translating molecules. Thus, we may use Eq. (2.7) for the flux of molecules impinging on the barrier and Eq. (2.3) for the mean speed. Here, for simplicity, we ignore the small changes in the numerical proportionality factors that arise in going from a three-dimensional to a two-dimensional situation. (It turns out that these factors cancel each other, anyway.) Inserting these equations into Eq. (5.9), we have

$$R_s = \frac{1}{4} n_s^\dagger \bar{c}/a = \frac{n_s^\dagger}{a}\sqrt{\frac{RT}{2\pi M}} = \frac{n_s^\dagger}{a}\sqrt{\frac{k_B T}{2\pi M}} \tag{5.10}$$

where $n_s^\dagger$(mc/cm²) denotes the surface concentration of adsorbate residing in the transition state. Now we must find the relation between $n_s^\dagger$ and n_s, the latter being the concentration of molecules in adsorption sites. At thermal equilibrium, statistical mechanics says that the concentration of molecules in a given state is proportional to the total number of ways of distributing the available thermal energy around a large system of molecules in that state. For each type of kinetic energy contributing to the thermal energy, the number of ways, Z, is equal to the sum over all of the quantized energy levels, ε_j, of the following products: the Boltzmann factor for each energy level times the number of ways of distributing energy at that level (the degeneracy of the level, g_j). Thus,

$$Z = \sum_j g_j e^{-\varepsilon_j/k_B T} \tag{5.11}$$

This type of summing was done for translational kinetic energy in the discussion of the Maxwell-Boltzmann distribution function (Sec. 2.2). The quantity Z is called the "partition function," and the product of the Z quantities for all the types of energy involved in a given state will be proportional to the concentration of molecules in that state; thus, for *any* two states,

$$\frac{n_s^\dagger}{n_s} = \frac{Z_r^\dagger Z_v^\dagger Z_t^\dagger}{Z_r Z_v Z_t} e^{-E_s/RT} \quad (5.12)$$

where r, v, and t denote the rotational, vibrational, and translational kinetic energies, each of which has various directional components. The final Boltzmann factor in Eq. (5.12) accounts for the *potential* energy difference between the adsorption-site state (n_s) and the transition state ($n_s^\dagger$), as shown in Fig. 5.4*a*. We neglect electronic excitation (Z_e), which occurs only at extremely high T.

We now need to evaluate the Z ratio in order to know $n_s^\dagger$. Because the molecule's rotational modes are eliminated or at least "frustrated" in their free motion by the adsorption, we may write $Z_r^\dagger = Z_r = 1$. For vibrational energy, the partition function derived from quantum mechanics for a harmonic oscillator is

$$Z_{vk} = \frac{1}{1 - e^{-\mathbf{h}\nu_k/k_B T}} \quad (5.13)$$

where $\mathbf{h}$ = Planck's constant = 6.63×10^{-34} J·s

ν_k = frequency of the kth vibrational mode

[Often, Z_{vk} is written with the oscillator's "residual" or "zero-point" energy factor, $\exp(-\mathbf{h}\nu_k/2k_B T)$, included in the numerator, but here this is accounted for in the potential energy factor of Eq. (5.12).] Optical absorption-band wavelengths for the vibrations of adsorbate bonds are in the infrared. For a typical value of $\lambda = 30$ μm or $1/\lambda$ (wavenumbers) = 333 cm^{-1}, we have $\nu_k = c_o/\lambda = 10^{13}$ s^{-1}, which means that at typical film deposition T, $\mathbf{h}\nu_k/k_B T > 1$, and Z_v is near unity. In other words, the vibrational modes are mostly in their ground states, because the excited states are just beginning to be accessible at ordinary T.

In all reactions, the one component of vibration that is aligned with the reaction coordinate (x, here) is transformed into a translational component crossing the barrier. The loss of this transformed vibrational component from $Z^\dagger_v$ in the numerator of Eq. (5.12) makes no difference in our calculation, since it is near unity anyway. Conversely, the newly created $Z^\dagger_t$ will increase the value of the numerator, since translational-energy quantum levels are much more closely spaced and therefore much more accessible at ordinary T. The translational-energy partition function is

$$Z^\dagger_{tx} = a\frac{\sqrt{2\pi m k_B T}}{\mathbf{h}} \quad (5.14)$$

Note here that the value of $Z^\dagger_{tx}$ is proportional to the linear dimension *a*. *If* the x component is the only translational component that increases in moving to the transition state, then Eq. (5.12) becomes

$$n^\dagger_s = n_s Z^\dagger_{tx} e^{-E_s/RT} \quad (5.15)$$

Upon inserting these two equations into Eq. (5.10), we have, finally

$$R_s = n_s\left(\frac{k_B T}{\mathbf{h}}\right)e^{-E_s/RT} = n_s \nu_{os} e^{-E_s/RT} = n_s k_s \quad (5.16)$$

which is the main result of absolute-reaction-rate theory. Note that in the second equality we have arrived at the Eq. (5.2) Arrhenius expression for the rate constant, k_s (s^{-1}), of a chemical reaction. For this particular reaction, the pre-exponential factor is $k_B T/\mathbf{h}$, which is 2×10^{13} s^{-1} at 960 K, for example. The rate constant here represents the frequency with which an individual adsorbate molecule "hops" to an adjacent site.

The above important result gives us some insight into the meaning of ν_{os}. In particular, note that ν_{os} is *not* the frequency of any vibrational component, ν_k, of the adsorbate, although such an implication is often made in the literature. That would be the case only in the event that $k_B T >> \mathbf{h}\nu_k$ in Eq. (5.13), in which case Z_v would become equal to $k_B T/\mathbf{h}\nu_k$. Then, since there is one more vibrational component in the adsorbed state than in the transition state, this $k_B T/\mathbf{h}\nu_k$ would end up in the denominator of Eq. (5.12), cancelling $k_B T/\mathbf{h}$ and leaving the pre-exponential factor equal to ν_k. However, at any reasonable T we do *not* have $k_B T >> \mathbf{h}\nu_k$, as pointed out after Eq. (5.13). Note also that Eq. (5.16) assumed that the only translational component acquired upon entering the transition state was the one in the direction of the reaction coordinate. This is a good assumption if sites 1 and 2 are identical, as in Fig. 5.4*a*. However, if the adsorbate is moving from a chemisorption site to a physisorption site in the course of diffusing, there may be an increase of as much as one additional translational component, which would raise ν_{os} by as much as 10^3 by evaluation of Eq. (5.14). Thus, ν_{os} may vary from 10^{13} to 10^{16} and may *not* simply be assumed to be ~10^{13} s^{-1} as is so often done.

The rate of surface diffusion also increases exponentially with T and with decreasing E_s, as seen in Eq. (5.16). E_s, the activation energy for surface diffusion from Fig. 5.4*a*, is always considerably lower than the desorption activation energy, which is E_c or E_d, depending on whether the species in question is chemisorbed or physisorbed, respectively (see Fig. 5.2). E_s is lower because the bonds are being only partially

broken in diffusion, whereas they are completely broken in re-evaporation. Therefore, at film deposition T approaching the onset of re-evaporation, where $\exp(-E_c/RT)$ is becoming significant, one expects a high rate of surface diffusion. This is one of the principal ways in which substrate T affects film structure, as we will see in Sec. 5.4. The ratio E_s/E_c is sometimes known as the "corrugation ratio," and it is lower for metals than for semiconductors because of the absence of bond directionality in metals.

We must now relate the molecular hopping rate, k_s, from Eq. (5.16) to the distance which an adsorbate molecule travels during film deposition. This is an adaptation of the classic random-walk problem. Since each hop is equally likely to be forward or backward in any given direction on the surface, there is no net motion in any one direction. However, as time passes, the molecule is more likely to be found further from its starting point. This is equivalent to saying that if one carries out a large number of trials of starting a molecule diffusing from a single point at t = 0, then with increasing t the final locations of these molecules become more widely dispersed from the starting point. In fact, for t corresponding to a large number of hops N_o ($N_o = k_s t >> 1$), the molecules will be dispersed in a Gaussian (normal) distribution whose median is at the starting point. The width of a Gaussian distribution is characterized by its standard deviation, σ, which is the root-mean-square (rms) deviation from the median and is also the half-width of the bell-shaped curve at its inflection point. For the diffusing molecule, if r is the per-hop rms change in distance from the starting point, then it turns out that $\sigma = r\sqrt{N_o}$. For diffusion on a given two-dimensional surface *lattice*, r is related to the hop distance, *a*, by a geometric factor, β, which depends upon the angles between the possible hopping directions on the lattice, but for present purposes we will assume $\beta \approx 1$. If we now consider σ as a measure of the diffusion length, Λ, of the molecule in time t, we can write

$$\Lambda = r\sqrt{N_o} = \beta a\sqrt{N_o} \approx a\sqrt{N_o} = a\sqrt{k_s t} \tag{5.17}$$

To get a feeling for the magnitudes involved here, assume that Eq. (5.16) holds and that the substrate T is 960 K, so that $\nu_{os} = k_B T/h = 2\times10^{13}\ s^{-1}$. If the molecule is physisorbed and thus has a relatively low E_s of, say, 20 kJ/mol, then $k_s = 1.6\times10^{12}\ s^{-1}$. For a typical *a* value of 0.3 nm and a t of 1 s, Λ = 380 μm, which is very large compared to the typical dimensions of thin-film topography. Conversely, for a *chemisorbed* molecule having an E_s of, say, 200 kJ/mol, $k_s = 2.6\times10^{2}\ s^{-1}$, and Λ = 4.8 nm, which is only tens of atomic diameters. Clearly, Λ varies enormously with bonding conditions at the surface.

We have arbitrarily chosen t = 1 s in the above example, but t will vary considerably with deposition conditions. Two regimes need to be considered separately: one in which t is the time between adsorption and burial by the next depositing monolayer, and another in which the adsorbate is more likely to desorb than to be buried. For the burial case,

$$t = \frac{n_o}{J_r} \tag{5.18}$$

where n_o = adsorption sites/cm²

J_r = deposition flux, mc/cm²·s

Inserting this and the expression for k_s [Eq. (5.16)] into Eq. (5.17), we have

$$\boxed{\Lambda = a\sqrt{\frac{\nu_{os} n_o}{J_r}}\, e^{-E_s/2RT}} \tag{5.19}$$

That is, Λ increases exponentially with T. This behavior appears as a straight line with a negative slope of $-E_s/2R$ on the Arrhenius plot of ln Λ versus 1/T in Fig. 5.5. Conversely, when T is high enough that film re-evaporation is becoming significant, diffusing species are likely to desorb before they become buried. Then, t is instead the adsorption

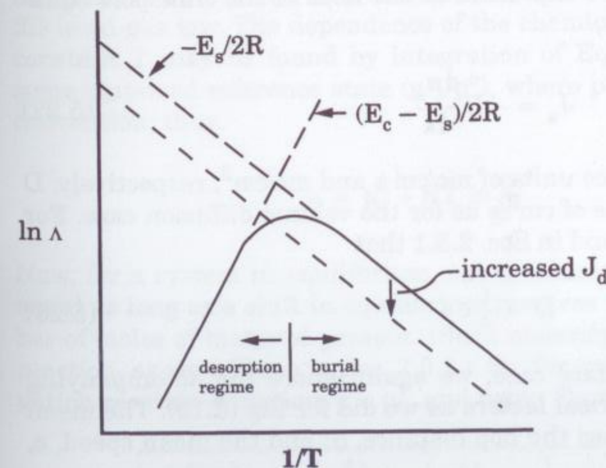


Figure 5.5 Behavior of surface diffusion length, Λ, with substrate T.

lifetime. To simplify its estimation, we neglect the precursor state, whose concentration is likely to be negligible at this high a T anyway, and we consider only desorption from the chemisorbed state. In this case,

$$t = \frac{1}{k_c} = \frac{1}{\nu_{oc}} e^{E_c/RT} \qquad (5.20)$$

where subscript c denotes the chemisorbed state as in Fig. 5.2. Inserting this and the expression for k_s [Eq. 5.16)] into Eq. (5.17), we have

$$\Lambda = a\sqrt{\frac{\nu_{os}}{\nu_{oc}}} e^{(E_c - E_s)/2RT} \qquad (5.21)$$

Because $(E_c - E_s)$ is always positive, as pointed out earlier, the slope of the Arrhenius plot is *positive* in this regime, as shown in Fig. 5.5. Note that this leads to a *maximum* in Λ at a T just below the onset of significant film re-evaporation. We will see in Sec. 5.4 that long Λ leads to films which are smoother, more homogeneous, and lower in crystallographic defects. Therefore, it is found empirically that the highest film quality is often obtained just below the re-evaporation point.

The above derivation of surface diffusion quantities has been based on examining the motion of individual adsorbed molecules. It is useful now to see how these quantities relate to familiar *macroscopic* quantities encountered in transport theory and thermodynamics. First of all, surface diffusion may be expressed in the form of the transport equation [Eq. (2.26)]; that is,

$$J_s = -D\frac{dn_s}{dx} \qquad (5.22)$$

With J_s and n_s in surface units of mc/cm·s and mc/cm^2, respectively, D here has the same units of cm^2/s as for the volume diffusion case. For the volume case, we found in Sec. 2.8.1 that

$$D = \tfrac{1}{4}\bar{c}\, l \qquad (5.23)$$

In adapting to the surface case, we again ignore the accompanying small changes in numerical factors as we did for Eq. (5.10). The mean free path, l, here becomes the hop distance, a; and the mean speed, c, will be $k_s a$. Thus,

$$D = \tfrac{1}{4}k_s a^2 \qquad (5.24)$$

and

$$\boxed{\Lambda = a\sqrt{k_s t} = 2\sqrt{Dt}} \qquad (5.25)$$

Now, using Eq. (5.16) for k_s, we can express D in Arrhenius form:

$$D = \tfrac{1}{4}\nu_{os}a^2 e^{-E_s/RT} = D_o e^{-E_s/RT} \qquad (5.26)$$

D is usually the quantity which appears in adsorbate diffusion experiments, while Λ is the quantity of interest for thin-film deposition.

The partition-function ratio of Eq. (5.12), which determines the pre-exponential factor for surface diffusion, can be related to the macroscopic thermodynamic concept of entropy which was discussed in Sec. 4.1. Indeed, this important relationship exists for any chemical reaction, and we will now develop it for the case of surface diffusion. Assuming thermal equilibrium as we did above, Eq. (4.9) for dG holds, and from it we may write for our pure adsorbate that

$$\left(\frac{\partial G_m}{\partial p}\right)_T = \left(\frac{\partial \mu}{\partial p}\right)_T = V_m = \frac{RT}{p} \qquad (5.27)$$

where the last equality also assumes the ideal-gas law, Eq. (2.10). A *mobile* adsorbate may in fact be thought of as a two-dimensional (2D) gas. We have essentially postulated a 2D gas already in writing Eqs. (5.10) and (5.24) above, and we will assume here that this gas obeys a 2D ideal-gas law. The dependence of the chemical potential, μ, on p at constant T may be found by integration of Eq. (5.27) starting from some standard reference state (μ°, p°), where $p^\circ = 10^5$ Pa $\approx$ 1 atm by convention; thus,

$$\mu = \mu^\circ + RT\ln\frac{p}{p^\circ} \qquad (5.28)$$

Now, for a system in equilibrium, the μ values of all components are equal as long as a shift in equilibrium involves no change in the number of moles of material present. (Such molarity changes can occur by reaction, as we will see in Sec. 7.3.1.). So, for our simple diffusion-activation reaction situation, $\mu = \mu^\dagger$, and using Eq. (5.28) we have

$$\boxed{\ln\left(\frac{p^\dagger}{p}\right) = \ln\left(\frac{n_s^\dagger}{n_s}\right) = \frac{-\left(\mu^{\circ\dagger} - \mu^\circ\right)}{RT} = -\frac{\Delta_r G^\circ}{RT}} \qquad (5.29)$$

In the first equality, we have used the ideal-gas law to switch from p to n_s units. The last equality defines the Gibbs free energy of reaction per mole at standard pressure, $\Delta_r G°$. This important equation relates the ratio of the concentrations of species in *any* two bonding configurations to their reference-pressure chemical potentials (free energies per mole) in those bonding configurations. The second and third equalities apply also to adsorbates of limited mobility, which are not behaving as 2D gases, provided that $n_s << 1$ ML as assumed previously, so that $n_s \propto p$ (more on this in Sec. 7.3.3).

Also, at equilibrium, forward rate equals reverse rate, so for this first-order reaction [Eq. (5.1)] we have

$$R_s = R_{-s} \quad \text{or} \quad n_s k_s = n^{\dagger}{}_s k_{-s} \tag{5.30}$$

where –s denotes the reverse reaction from the transition state back to adsorption site 1 in Fig. 5.4*a*. Rearranging this equation gives the definition of the equilibrium constant:

$$K = \frac{k_s}{k_{-s}} = \frac{n_s^{\dagger}}{n_s} = e^{-\Delta_r G°/RT} \tag{5.31}$$

where we have used Eq. (5.29) to obtain the last equality. Again, this expression is valid for any first-order reaction, not just for surface diffusion. Except for the $n_s{}^{\dagger}/n_s$ equality, the expression is valid for reactions of any order.

Now, using the definition of the Gibbs free energy in Eq. (4.5), we can write that

$$\Delta_r G° = U° + p°\Delta_r V° - T\Delta_r S° \tag{5.32}$$

The first term on the right is essentially the height of the potential-energy barrier, E_S. The second term is negligible for condensed-phase reactions. Upon inserting these results into Eq. (5.31), we have

$$\frac{n_s^{\dagger}}{n_s} = e^{-\Delta_r S°/R} \cdot e^{-E_s/RT} \tag{5.33}$$

Comparing this with Eq. (5.12) shows that

$$\Delta_r S° = R \ln\left(\frac{Z_r^{\dagger} Z_v^{\dagger} Z_t^{\dagger}}{Z_r Z_v Z_t}\right) \tag{5.34}$$

This result is equivalent to the expression for S presented in Eq. (4.3) and gives further insight into the meaning of entropy. We see from this that the pre-exponential term in the Arrhenius equation [Eq. (5.2)] is an entropy term; that is, it increases with the degree of delocalization or the degree of access to different quantum states in the numerator. It is useful to think of the Arrhenius pre-exponential term from this perspective. Note, however, that there may be other factors in the pre-exponential term as well. For example, if E_s has some T dependence, say $E_s = E_s^{\circ} - BT$, then a multiplier of $e^{B/R}$ will enter the pre-exponential [2].

Having now developed some ideas about the mechanisms and behavior of surface diffusion, we can proceed to discuss this key phenomenon in the context of thin-film nucleation and structure development.

5.3 Nucleation

5.3.1 Surface energy

To understand nucleation, we need to introduce the concept of surface energy. The familiar experiment of drawing a liquid membrane out of soapy water on a wire ring is illustrated in Fig. 5.6. The force required to support the membrane per unit width of membrane surface is known as the surface tension, γ, expressed as N/m in SI units, or more commonly as dynes/cm in cgs units. For a wire of circumference b, the width of surface is 2b, since the membrane has both an inner and an outer surface. Thus, the total force required to support the membrane is $F = 2b\gamma$. As the membrane is extended upward in the x direction, work $F\Delta x$ (N·m or J) is done to create the new surface, and the surface area created is $A = 2b\Delta x$, assuming for simplicity a constant membrane circumference. The work is stored as surface *energy* (as in stretching a spring), so the surface energy per unit area of surface is

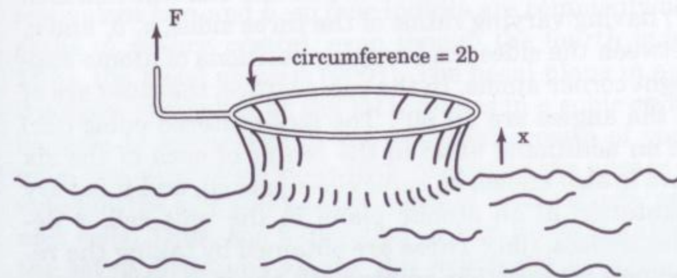


Figure 5.6 Surface tension of a liquid membrane.

$$F\Delta x/A\ (N{\cdot}m/m^2) = (2b\gamma)\Delta x/2b\Delta x = \gamma\ (N/m\ \text{or}\ J/m^2). \quad (5.35)$$

Thus, surface tension (N/m) and surface energy per unit area (J/m^2) are identical, at least for liquids. For solids at T > 0 K, surface Gibbs free energy is reduced by an entropy factor [Eq. (4.5)] which depends on the degree of surface disorder (Williams, 1994). For solids, there is also a quantity called surface stress, which differs from surface energy by a surface elastic-strain term (more in Sec. 5.6.1). Liquids cannot support such strain, because the atoms just rearrange to relax it.

Surface energy exists because the molecules of a condensed phase are attracted to each other, which is what causes the condensation. The creation of surface involves the removal of molecular contact from above that surface (bond-breaking), and thus involves energy input. Consequently, to the extent that motion within the condensed phase can occur, such motion will proceed so as to minimize the total surface energy, γA. In the liquid-membrane case, where γ is fixed, this means minimizing A. Thus, when the wire is lifted far enough, the membrane snaps taut into the plane of the ring; and when a bubble is blown, it becomes spherical. In the case of solids, surface energy proceeds to minimize itself by surface diffusion, and this process is fundamental to the development of structure in thin films. In thin-film growth, both A and γ are varying. Area, A, depends on surface topography, and γ depends on many properties of the exposed surface, including chemical composition, crystallographic orientation, atomic reconstruction, and atomic-scale roughness. In materials that have no orientation dependence of a particular property, that property is said to be isotropic. However, in most crystalline solids, γ is not isotropic, but is *an*isotropic.

Examining the anisotropy of γ calls for a brief digression to review crystallographic nomenclature. Some readers may be able to skip this paragraph, and others may need refer to a more thorough description in a solid-state physics or crystallography text. Figure 5.7 shows a crystal unit cell. There are fourteen variations of this cell (the fourteen "Bravais lattices") having varying ratios of the three sides, *a, b,* and *c,* varying angles between the sides, and varying positions of atoms supplementing the eight corner atoms. In the cubic lattice, the sides are of equal length and the angles are all 90°. The face-centered cubic (fcc) lattice shown has an additional atom in the center of each of the six faces; this structure is also known as cubic close-packed, and it is very common. The orientation of an atomic plane in the unit cell is described by its Miller indices, (jkl). These are obtained by taking the reciprocals of the plane's intersection points with the x, y, and z axes, measured in integral numbers of unit cells. Thus, the plane containing the atoms labeled 1, 2, 3, and 4, which intersects the axes at 1, ∞, and

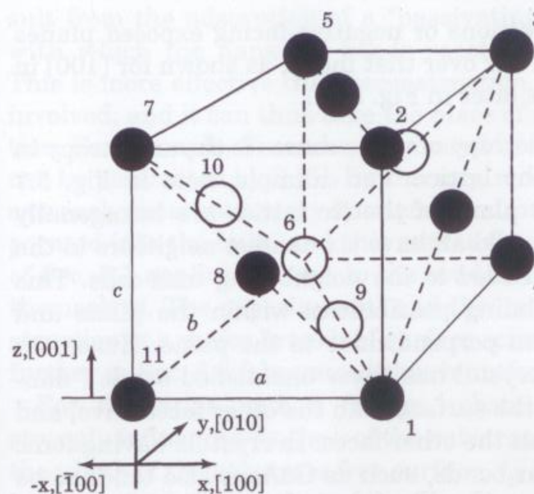


Figure 5.7 Geometry of a face-centered cubic crystal.

∞, has indices (100). Other planes with the same (100) symmetry in this crystal are (010) and (001), and the set of them is denoted using different brackets: that is, {001}. Similarly, the plane containing atoms 1, 2, 5, and 6, which intersects the axes at 1, 1, and ∞, has indices (110), and the set of planes with this symmetry is denoted as {011}. Finally, the set of planes like the one containing atoms 1, 6, and 7 is the {111} set. These are the three primary sets of planes in a cubic lattice.

We will see below that a crystal tends to expose surfaces having low Miller indices, and these exposed surfaces are known as "facets." The use of reciprocals in the indices comes from x-ray diffraction work, where the spacing of diffraction spots is reciprocally related to the spacing of atomic planes (more in Sec. 6.4.2). In the hexagonally close-packed or "hcp" lattice, *a* and *b* of Fig. 5.7 are at 120° in the so-called "basal" *a-b* plane; also, $c \perp a\text{-}b$, and $c \neq a = b$. There is a third direction in the basal plane, oriented at 120° to *a* and *b* and symmetrically equivalent to *a* and *b*, so *four* indices are commonly used to describe a plane in an hcp crystal, even though the fourth index is redundant. Thus, the basal plane is (0001). The basal plane in an hcp crystal has the same symmetry as the {111} planes in a cubic crystal, and this can be used to advantage in the epitaxial growth of one type of crystal upon the other. *Directions* with respect to the axes of a crystal are denoted by the x, y, and z components of their vectors, again measured in unit cells. Thus, the x direction in Fig. 5.7 is the [100] direction, and the set of directions having this symmetry is ⟨001⟩, where again the type of bracket identifies the meaning. When the axis angles are all 90°, a direction is always perpendicular to a plane having the same in-

dices. Negative-pointing directions or negative-facing exposed planes of a crystal are denoted by a bar over that index, as shown for $[\bar{1}00]$ in Fig. 5.7 and for various faces later in Fig. 5.11.

Returning now to the anisotropy of γ, one source is the anisotropy in bonding directions within the lattice. For example, note in Fig. 5.7 that the atoms in the (111) plane of the fcc lattice are hexagonally close-packed. That is, they would all have six nearest neighbors in the plane if the lattice were extended to the neighboring unit cells. This arrangement maximizes bonding possibilities within the plane and consequently minimizes them perpendicular to the plane. Hence, an exposed {111} face of an fcc crystal has fewer unsatisfied bonds ("dangling" bonds) sticking out of the surface than the other faces have, and it therefore has a lower γ than the other faces. In crystals having ionic bonds, such as CaF_2, or polar bonds, such as GaAs, γ also tends to be lower for faces containing equal numbers of cations (Ca, Ga) and anions (F, As), since this results in charge neutrality at the surface. Consequently, these are the *nonpolar* faces. In materials having a "layered" structure, such as graphite and MoS_2, there are *no* chemical bonds between the atomic layers of the basal plane. The sliding of these planes past each other accounts for the performance of layered materials as dry lubricants. The low-energy facets for various crystal structures are listed in Table 5.1. Aside from hcp, all of those listed are cubic and are distinguished from each other only by the positions of the atoms within the unit cell.

For most faces of crystals in general, γ is actually lower than one would predict from the dangling-bond density of the separated bulk lattice. This is because the dangling bonds and their atoms at the surface become distorted from their bulk lattice angles and positions, respectively, to cross-bond with each other and thereby reduce surface energy. The result is a "reconstructed" surface having patterns of atom positions and surface bonds that are different from those in the bulk and having γ reduced by as much as half. Further γ reduction may re-

TABLE 5.1 Facets of Lowest Surface Energy for Various Crystal Structures

Structure	Examples	Low-γ facets
Body-centered cubic (bcc)	Cr, Fe	{110}
Face-centered cubic (fcc)	Au, Al	{111}
Hexagonal close-packed (hcp)	Zn, Mg	{0001}
Diamond	Si, Ge	{111}
Zinc blende	GaAs, ZnSe	{110}
Fluorite	MgF_2, CaF_2	{111}
Rock salt	NaCl, PbTe	{100}

sult from the adsorption of a "passivating" monolayer of an element with which the dangling bonds react to become terminated bonds. This is more effective than reconstruction, because less bond strain is involved, and it can thus take the place of and prevent the reconstruction. For example, Si can be passivated by H [3] or by As [4], and GaAs can be passivated by S [5]. By the way, reconstruction and passivation can each generate energy barriers to adsorbing vapor becoming incorporated into the bulk, and this makes the precursor-adsorption model of Sec. 5.1 applicable even to the condensation of pure materials onto themselves. The experimental and theoretical study of surface reconstruction is a major branch of surface science, and it will be discussed further in Sec. 6.5.3 because it has a profound effect on epitaxy.

For deposition onto a foreign substrate, nucleation behavior is strongly influenced by the γ of the substrate. Here, we need to consider the γ both of the substrate free surface, γ_s, and of the substrate-film interface, γ_i, relative to that of the film free surface, γ_f. All three γ values will in general depend on crystallographic orientation, passivation, and sometimes other factors. Assume that there is enough surface diffusion so that the depositing material can rearrange itself to minimize γ; that is, assume that the nucleation is not kinetically limited and can approach equilibrium. For this, we must have $\Lambda >> a$ [see Eq. (5.17)]. In the opposite case, $\Lambda < a$, every atom sticks where it lands, and the growth behavior is "quenched." Quenched growth develops its own characteristic film structure which we will examine in Sec. 5.4. With our assumption of $\Lambda >> a$, there are two nucleation situations on the bare substrate, as shown in Figs. 5.8*a* and *b*. In *a*, the film spreads across or "wets" the substrate because

$$\gamma_f + \gamma_i < \gamma_s \tag{5.36}$$

That is, total surface energy is lower for the wetted substrate than for the bare one. This leads to smooth growth, atomic layer by layer, which is sometimes referred to as the Frank-van der Merwe growth mode. For this mode to occur, there must be strong enough bonding between film and substrate to reduce γ_i to where Eq. (5.36) holds. If there were no such bonding at all, we would have $\gamma_i = \gamma_f + \gamma_s$, so spreading the film across the substrate would always increase total surface energy by $2\gamma_f$, as in the case of the free-standing liquid membrane of Fig. 5.6. Thus, with insufficient substrate bonding, Eq. (5.36) ceases to hold, and the film does not wet the substrate but instead forms three-dimensional (3D) islands, shown in Fig. 5.8*b* and sometimes referred to as the Volmer-Weber growth mode. There is a third growth mode, Stranski-Krastanov, shown in Fig. 5.8*c*, in which the growth changes from layer to island after a monolayer or two due to a change in the

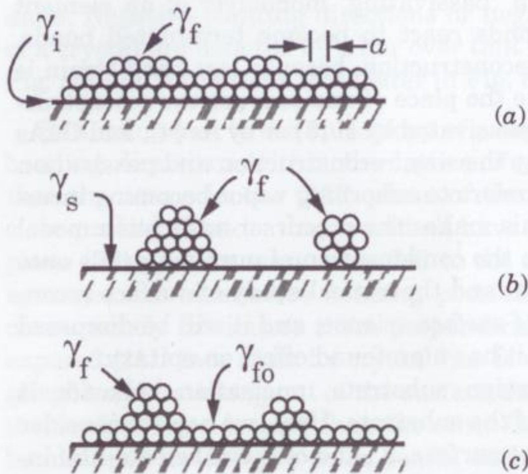


Figure 5.8 Film growth modes: (*a*) Frank-Van der Merwe (layer), (*b*) Volmer-Weber (island), and (*c*) Stranski-Krastanov.

energy situation with successive monolayers. This mode is associated with epitaxy and will be discussed in Sec. 6.7. For liquids contacting solids, the degree of wetting is most easily observed by the rise or depression of a liquid column in a narrow tube (a capillary). Thus, film nucleation analysis in terms of degree of wetting is known as the "capillarity" model. Nucleation and the three growth modes of Fig. 5.8 are further discussed by Venables (1984).

Three-dimensional nucleation is usually undesirable, since it leads to rough, nonuniform films, an extreme example being the diamond nuclei of Fig. 5.3. Often, it can be changed to 2D growth by manipulating one of the γ terms in Eq. (5.36) to satisfy the inequality. For example, γ_i decreases with increased film-substrate bonding. This bonding depends on the chemical reactivity of the film and substrate and on the similarity of the bonding within the two materials. There are three types of chemical bonding: (1) covalent, in which valence electrons are shared between atoms, as in Si and GaAs; (2) ionic, in which valence electrons are completely transferred from the cation to the anion, as in CaF_2 and NaCl; and (3) metallic, in which the valence electrons wander through the crystal, and the metal ions exist as an array of positive charges in a sea of negative charge (the "jellium" model). In general, interfacial bonding is stronger between materials having the same type of bonding. Bonding will still occur across types if there is enough chemical reactivity. For example, the chemically-active metal, Cr, will bond to glass by breaking an O-Si bond and forming a Cr-O or Cr-Si bond. Conversely, the noble metal, Au, cannot do this and does

not bond well to glass. Yet Au does form a strong metallic bond to clean (unoxidized) Cr. Thus, by the use of an intermediate "glue" layer which bonds well to both the film and the substrate, such as Cr between Au and glass, γ_i can be reduced and wetting accomplished. Another good bonding material is Ti. Only enough bonding material is needed to ensure a continuous layer—typically 10 nm. Alternatively, energy-enhanced techniques, particularly ion bombardment and sputtering (Chaps. 8 and 9), can provide the activation energy for bonding between film and substrate and thereby reduce γ_i. Ion bombardment can also break up the nuclei of 3D islands, thus *counteracting* the equilibrium tendency (more in Sec. 6.7.2). Energy-enhanced techniques are very powerful in modifying film structure and are being employed more and more routinely in a variety of applications.

Equation (5.36) can also be satisfied in principle by reducing γ_f or increasing γ_s. The former is the familiar "surfactant" effect known in liquids such as soap. However, I do not know of a case where this effect has been used to assist wetting in thin-film deposition. There has been a so-called surfactant effect reported in the suppression of 3D nuclei in epitaxy (Sec. 6.7.2), but it is not the result of γ_f reduction. In another approach to wetting, γ_s can be increased by activating the substrate surface through energetic irradiation. For example, polymers are treated with plasma exposure to increase bonding to metal films. Presumably, the irradiation creates dangling bonds at the surface.

5.3.2 Three-dimensional (3D) nucleation

Despite the above techniques to encourage smooth film growth, 3D nucleation is often a problem, so it is important to understand its behavior to deal with it effectively. There are two ways in which 3D nuclei can form. The more common one is when bonding initiates at active surface sites such as atomic steps, crystal defects, or impurities. At these nucleation sites, the activation energy for bonding is less than elsewhere, as discussed in Sec. 5.1; or equivalently, γ_i is lower in Eq. (5.36) because of the interfacial bonding which develops. The vapor may arrive at these sites either by surface diffusion or by direct impingement. The latter is the dominant mode when the vapor has a very low trapping probability, δ, elsewhere on the surface or when the desorption rate of the adsorbed precursor, R_d [Eq. (5.5)], is so high that most of it never reaches a nucleation site. However, it is more common for nuclei to accumulate by surface diffusion.

5.3.2.1 Classical nucleation. Even if there are no active nucleation sites on the surface, 3D nuclei can still form at random surface locations by

the spontaneous accumulation of mobile adsorbed atoms plus arriving vapor into "critical" nuclei which are big enough to be stable. This is the classic nucleation problem (Lewis, 1978). We will present below an elementary treatment of this difficult problem, using the "capillarity" approach, which is based on surface-energy minimization. This will serve to give a qualitative idea of the important factors which determine nucleation behavior. Quantitative evaluation of nucleation behavior in a particular deposition situation is more reliably done by direct microscopic observation of the nuclei using transmission electron microscopy (TEM) or scanning tunneling microscopy (STM).

Figure 5.9 illustrates the geometry of nucleus formation. In this example, wetting to the substrate is poor, so that the contact area, A_i, and its effect on total surface energy can be neglected. If anisotropy of the surface energy, γ_f, can also be neglected, the nucleus will be spherical as shown. Vapor arrives at the surface of the nucleus directly from the gas phase as well as by surface diffusion of adsorbed "monomers" (single atoms or molecules). The nucleus can also lose material by evaporation back into the gas phase or into the 2D gas of adsorbed monomers. Recall from the Sec. 4.1 discussion of evaporation that at vapor-liquid or vapor-solid equilibrium, the chemical potentials of the two phases are equal [Eq. (4.8): $\mu_v = \mu_c$] and the vapor is at its saturation vapor pressure, p_v. If the p of the vapor is increased above p_v, μ_v increases in accordance with Eq. (5.28) (for ideal gas), which we applied there to the 2D gas situation. Thus,

$$\mu_v - \mu_c = RT \ln\frac{p}{p_v} = RT \ln\frac{J_c}{J_v} \tag{5.37}$$

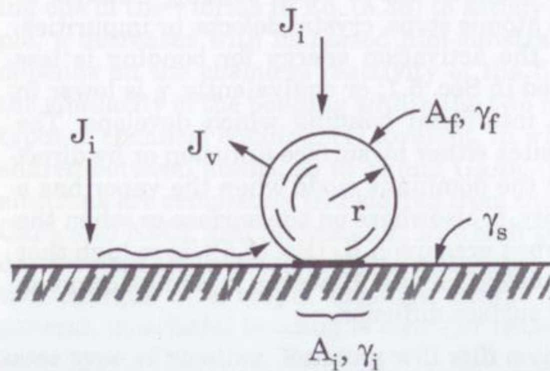


Figure 5.9 Geometry of condensate nucleus formation on nonwetting substrate.

Here, equality between the ratio of p values and the ratio of condensing and evaporating molecular fluxes, J_c and J_v, has been obtained by using Eqs. (2.18), (4.17), and (4.18) and by assuming that we are close enough to equilibrium so that $\alpha_c = \alpha_v$. This ratio is known as the "supersaturation" ratio. If the ratio were <1, the system would be called "undersaturated." For single-component systems, as we have here, $\mu = G_m$, the Gibbs free energy per mole. The difference in μ or G_m between the two phases is the driving force for condensation. Processes always proceed in the direction of decreasing G_m, which is toward condensation in the supersaturated case.

The quantity γ was not involved in the Sec. 4.1 treatment of vapor-condensate equilibrium, because there the condensate area, A, and therefore the total surface energy, γA, were constant. In that case, an infinitesimal amount of supersaturation was enough to drive the condensation. Conversely, in the nucleation situation of Fig. 5.9, nucleus surface area, A_f, is *increasing* with condensation, so there is a $\gamma_f A_f$ contribution to the total G of the nucleus. Consequently, the G change *per nucleus* (not per mole) for forming a nucleus of radius r and volume V from the vapor becomes

$$\Delta G = -(\mu_v - \mu_c)\frac{V}{V_{mc}} + \gamma_f A_f = -\left(RT \ln\frac{p}{p_v}\right)\cdot\frac{(4/3)\pi r^3}{V_{mc}} + \gamma_f \cdot 4\pi r^2 \tag{5.38}$$

where V_{mc} is the molar volume of the condensate. The γ_f term is positive and causes ΔG to *rise* with increasing r at small r. The behavior of this function is shown in Fig. 5.10 for water at two p/p_v ratios. For water, the surface tension γ_f = 73 dynes/cm = 0.073 J/m^2 at 20° C. nuclei having r less than that at which ΔG goes through a maximum will spontaneously decompose, since that is the direction of decreasing ΔG. Larger nuclei are stable and will spontaneously grow. The critical nucleus radius, r*, at which the slope of d(ΔG)/dr changes sign is found by setting d(ΔG)/dr = 0. Operating on Eq. (5.38) in this way, we obtain

$$r^* = \frac{2\gamma_f}{\left(\frac{RT}{V_{mc}}\right)\ln\left(\frac{p}{p_v}\right)} \tag{5.39}$$

and

$$\Delta G^* = \frac{(16/3)\pi\gamma_f^3}{\left[\left(\frac{RT}{V_{mc}}\right)\ln\left(\frac{p}{p_v}\right)\right]^2} \tag{5.40}$$

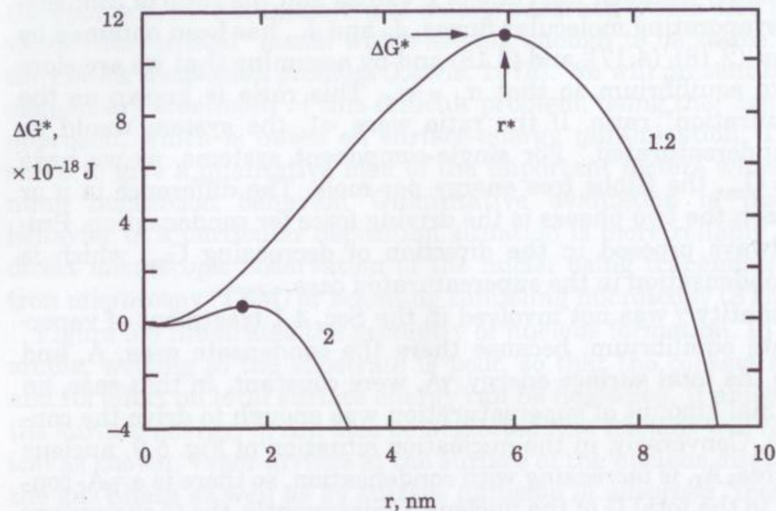


Figure 5.10 Classical nucleation behavior of water for two values of the supersaturation ratio (• denotes critical-cluster condition).

Here, ΔG^* is the energy barrier to nucleation and is analogous to the energy barrier for a chemical reaction.

The accumulation of individual adatom monomers into a nucleus may indeed be considered as a chemical reaction. Since nuclei with $r > r^*$ are stable and proceed to grow spontaneously, the reaction of adatoms with such nuclei is an irreversible one and leads rapidly to coalescence of the nuclei with each other to form a continuous film. Conversely, for $r < r^*$, the reaction is reversible because such nuclei spontaneously disintegrate. We may therefore view the population of nuclei with $r < r^*$ as being in a state of equilibrium with p, albeit a metastable one because of the ever-expanding "sink" provided by nuclei having $r > r^*$. Equation (5.31) relating reactant concentrations to ΔG now applies, with one modification. There, the reaction to the excited state was first-order, whereas here the reaction to a critical nucleus containing j atoms is j^{th} order. Thus, the equilibrium condition becomes

$$\mu_j = j\mu_1 \tag{5.41}$$

where μ_1 denotes the chemical potential of the adsorbed monomers. This leads to

$$\frac{n_j}{n_1^j} = \frac{n^*}{n_1^j} = e^{-\Delta G_m^*/RT} = e^{-\Delta G^*/k_B T} \tag{5.42}$$

where n_j is the surface concentration (cm^{-2}) of nuclei containing j monomers, and n_1 is the surface concentration of adsorbed monomers. Note that here G_m denotes a mole of *nuclei*, not a mole of atoms or molecules, and thus we obtain the last equality, where ΔG^* is per nucleus and is given by Eq. (5.40). Here, ΔG^* is equivalent to the standard-state ΔG° from Eq. (5.31), because the concentration chosen for the standard state is arbitrary and will not affect the difference of μ° terms or the ratio of n values significantly.

The concentration of critical nuclei, n^*, will be determined by both n_1 and ΔG^* in Eq. (5.42). The behavior of n_1 depends on whether the adatom desorption rate is appreciable. If it is, then n_1 will reach a nearly steady-state fractional surface coverage given by Eq. (5.4). It will not be precisely steady-state, because there will always be some finite rate of critical-nucleus generation, and these nuclei will ultimately take over the surface. Nevertheless, before this happens, n_1 is largely determined by the vapor arrival rate, J_i, which is proportional to p by the Knudsen equation [Eq. (2.18)]. Since increasing p also raises n^* by reducing ΔG^* [Eq. (5.40)], n^* becomes a very strong function of p, such that nucleation appears to initiate suddenly as p is raised. The dependence of n^* on T is more complicated. With increasing T at fixed p, p_v increases exponentially by Eq. (4.14), but this effect is more than cancelled by the RT factor in Eq. (5.40), so that ΔG^* decreases. But meanwhile, n_1 is decreasing exponentially by Eq. (5.4), and thus it is experimentally observed that n^* decreases with increasing T. For example, it is the difficulty of nucleating Zn and Cd on glass before monomer desorption which accounts for the "splotchy" appearance of thin deposits of these metals despite high supersaturation during deposition, as discussed in Sec. 5.1. The situation can be remedied by cooling the substrate to inhibit desorption.

If precursor desorption is negligible, then n_1 does not approach steady state, but increases steadily with time at a constant arrival rate (constant supersaturation) of vapor, J_i. In this case, n^* rises much more rapidly than when precursor desorption is appreciable. In either case, the rate at which n^* rises determines the "coarseness" of the nucleation. When the process objective is to produce a smooth, uniform film, a large n^* and small r^* are preferred; that is, less coarse nucleation. One way to achieve this is by using a very high vapor arrival rate (high supersaturation), at least until the nucleation phase is over and the film is continuous. At sufficient supersaturation, the number of atoms in the critical nucleus approaches unity, and ΔG^* becomes negligible. For example, this is believed to be the situation for the classic nucleation study case of Au on NaCl, which is characterized by poor adherence. Coarsening will still occur even with one-atom critical nuclei, to the extent that the atoms and nuclei are mobile on the sur-

face. Mobility allows the nuclei to migrate and coalesce with each other, becoming larger and fewer with time. It also allows transfer of atoms from smaller nuclei to larger ones, a process driven by the higher γ_f of the former which results from the curvature effect to be discussed in Sec. 5.4.2. These processes are analogous to the "Ostwald ripening" process of bulk metallurgy, in which similar crystallite (grain) growth occurs with annealing. Coarsening will continue to occur even if deposition is stopped, which is why the arrival rate needs to remain high until the film is continuous, if smoothness is the objective. Sometimes, it is instead desired to obtain polycrystalline films with large grain size and a correspondingly low concentration of grain boundaries. Grain boundaries can degrade film performance by acting as electrical-carrier traps in semiconductors or as channels for diffusion through chemical-barrier layers. In such cases, coarser nucleation may be preferable.

5.3.2.2 Kinetics vs. thermodynamics. Another approach to achieving smooth growth is to lower the substrate T to inhibit surface diffusion and thus "freeze out" the nucleation and coalescence processes. If the arriving species do not have enough thermal energy to either desorb or diffuse, they remain where they land, which leads to the quenched growth mentioned earlier. In this case, the nucleation process is *kinetically inhibited* by the surface-diffusion activation-energy barrier, E_s, in Eq. (5.16). This is also the case for ion-bombardment dissipation of 3D nuclei, as mentioned at the end of Sec. 5.3.1: the nuclei do not have time to reassemble themselves by surface diffusion before they are buried by depositing material.

The question of whether a process is approaching equilibrium or is instead limited by kinetics is an important one, and it arises often in thin-film deposition. Process behavior and film properties are profoundly affected by the degree to which one or the other situation dominates. The answer is not always apparent in a given process, and this often leads to confusion and to misinterpretation of observed phenomena. Therefore, to elaborate briefly, the generalized mathematical representation of this dichotomy is embodied in Eq. (5.30) describing the rate balance of a reversible reaction and Eq. (5.31) defining its equilibrium constant. Approach to equilibrium requires the forward and reverse rates to be fast enough so that they become balanced *within the applicable time scale*, which may be the time for deposition of one monolayer, for example. Then, the concentrations of reactant and product species are related by the difference in their free energies, $\Delta_r G°$. If, on the other hand, the forward rate is so slow that the product concentration does not have time to build up to its equilibrium

level within this time scale, then the product concentration is determined not by $\Delta_r G°$ but, instead, by the forward rate. This rate is governed by Eqs. (5.1) and (5.2), in which E_k/T plays the dominant role. So it is, that reactions can be frozen out and equilibration avoided if so desired, by lowering the T.

The difficulty of answering the question of kinetics versus thermodynamics arises from the fact that the applicable rate constants, k_k, are often unknown or not known accurately enough. The measurement of k_k is much more difficult than just measuring equilibrium concentration, both because it is a dynamic measurement and because it must be made in the absence of the reverse reaction. The problem is perhaps most troublesome in CVD, where many reactions are involved, as we will see in Sec. 7.3.

5.3.2.3 Other complications. The assumption of sufficient surface diffusion for approach to equilibrium is inherent in the above treatment of nucleation. Other simplifying assumptions have also been made, and it is important to be aware of these in order to recognize the limitations of the model. As mentioned earlier, active sites for nucleation have deliberately not been included in order to examine spontaneous nucleation on a homogeneous surface. Nevertheless, in almost all practical situations, nucleation will occur *predominantly* at active sites. The extent to which this mechanism increases nucleation density will depend on the activity (γ_s) and concentration of such sites on the particular substrate involved, so there is no way to construct a general model. Remaining simplifying assumptions relate to surface energy, γ. We have neglected any reduction of total surface energy arising from reduced γ_i over the contact area, A_i, in Fig. 5.9. Such reduction will reduce ΔG^* rapidly through the γ^3 factor in Eq. (5.40). The concept of γ also implies that the surface is a continuum, whereas in fact when the nuclei become smaller than perhaps 10^3 atoms (radius r = 6 atoms), the surface is more accurately treated as an array of discrete atoms. Also, when r is small, γ_f becomes larger than it is on a plane surface, because the atoms on a convex surface are more exposed and less well connected to the bulk. This convexity effect is very important to structural development, as we will see in Sec. 5.4.2. Finally, we have ignored the crystallographic anisotropy of γ_f, which we will now discuss.

The shape of a nucleus will try to adjust itself to minimize total surface energy; that is,

$$\boxed{\sum_k \gamma_k A_k = \text{minimum}} \tag{5.43}$$

Subscript k denotes terms corresponding to the nucleus free surface, the interface to the substrate, and the substrate free surface. In the case of liquid or amorphous nuclei, which have no γ anisotropy, there is only one term, $\gamma_f A_f$, for the nucleus free surface. In the more common case of crystalline nuclei, these surface terms include all of the various exposed atomic planes or facets, and the nucleus shape-adjustment process is called facetting. Consider first the case of poor wetting ($A_i \rightarrow 0$ in Fig. 5.9), so that the nucleus free-surface terms are the only significant terms in Eq. (5.43). The "Wulff theorem" states that when total surface energy of a crystal of constant volume is minimized per Eq. (5.43), then it will also hold that

$$\gamma_k / r_k = \text{constant} \tag{5.44}$$

where r_k is the perpendicular distance from the center of the crystal to the kth facet. This situation is easier to visualize in two dimensions, so it is illustrated in Fig. 5.11 for the case of a needle crystallite oriented along the *z* axis, perpendicular to the paper. For this hypothetical crystallite, we have specified that γ on the {110} facets (γ_1) is 20 percent higher than γ on the {100} facets (γ_0), so $r_1 = 1.2 r_0$ as shown. The facets of lowest γ for various crystal structures are listed in Table 5.1. These are the preferred faces of exposure and therefore the largest ones on an equilibrated crystallite. Facetted crystals are abundant in Nature. Even grains of common table salt show up as perfect little cubes under the microscope. Covalent and ionic crystals have much larger γ anisotropy than metallic crystals, because they have strongly

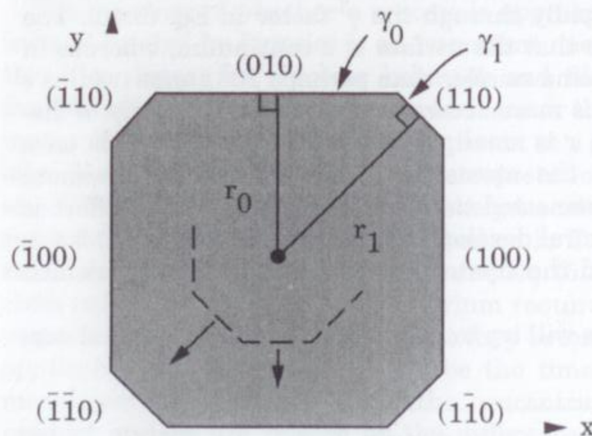


Figure 5.11 Wulff construction for a needle crystallite oriented along the z axis (perpendicular to the paper).

preferred bond directionality and bond polarity, respectively, while metals have neither. Polar covalent crystals and ionic crystals have, in addition, a preference for exposing charge-neutral (nonpolar) facets. In metals, γ anisotropy arises from anisotropy in the surface packing density of atoms and is typically only a few percent, but often this is still enough to result in preferred planes of exposure.

There are two ways in which a growing crystallite evolves in shape: deposition anisotropy and surface diffusion. Deposition rate from the vapor is often higher on facets of high γ because bonding is stronger and therefore δ or S_c is higher (see Fig. 5.1). It can be seen from the arrows in Fig. 5.11 leading from the nonequilibrated dashed-line facets that such deposition-rate anisotropy will cause the facets of high γ to become proportionately smaller as growth proceeds. Meanwhile, surface diffusion will redistribute mobile atoms so as to minimize total surface energy whether growth is occurring or not. Without surface diffusion, crystallites would evolve toward a "growth shape" which is in general different from the equilibrium "Wulff shape." Not all nuclei are single crystals, however, and this leads to shapes other than the Wulff shape even with surface diffusion occurring. In particular, there are multiple-twinned particles (MTPs). The twin of a crystal is its mirror image. For example, in the crystal of Fig. 5.7, if the pyramidal portion enclosed by the origin and atoms 1, 6, and 7 in the (111) plane were to form a twin on the (111) plane, the result would be two pyramids with a common base on (111) and six exposed {100} facets. MTPs consist of many single-crystal facetted particles nested together along twin planes to form one particle. In summary, we have seen that γ anisotropy in thin-film nuclei will cause them to develop facets and shapes that are governed by the anisotropy, by the dynamics of the deposition process, and by the extent of twinning.

In the case of liquid or amorphous nuclei, Eq. (5.43) still applies and in general has three terms—γ_f, γ_i, and γ_s—where γ_f is now isotropic and therefore constitutes only one term. At substrate T above 2/3 of the absolute melting T of the film material in bulk form, small nuclei are likely to be molten [6], so the liquid case is encountered more often than one might expect. In previously discussing 3D nucleation, we neglected for simplicity the effect of γ_i and γ_s on total surface energy. Accounting for them now in Eq. (5.43) leads to the spherical-cap nucleus shown in Fig. 5.12, and leads to the "wetting angle," θ, with the substrate (Exercise 5.5). Angle θ may also be obtained by considering the γ values as surface-tension vectors as shown and by writing the force balance known as Young's equation:

$$\gamma_i + \gamma_f \cos\theta = \gamma_s \tag{5.45}$$

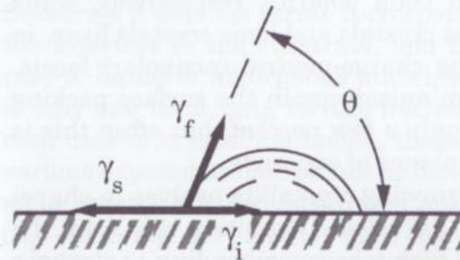


Figure 5.12 Wetting angle of a liquid nucleus.

The poorer the wetting, the larger θ becomes. When there is no wetting, θ = 180°, and γ_i is just $\gamma_s + \gamma_f$. When wetting is complete, θ = 0°, and Eq. (5.45) turns into the inequality of Eq. (5.36).

5.3.3 Two-dimensional nucleation

When wetting is complete and Eq. (5.36) holds, the adsorbing atoms do not accumulate into 3D islands but, instead, spread out on the surface in a partial monolayer as shown in Fig. 5.8*a*. Because total surface energy is *reduced* rather than increased by this process, there is no nucleation barrier in going from the vapor state to the adsorbed state; that is, the γ term in Eq. (5.38) is negative when the interfacial area is included. This means that deposition can proceed even in *under*saturated conditions. A familiar example of this situation is the oxidation of a metal, which involves O_2 adsorbing into a chemisorbed state (that is, reacting with the metal) at well below its saturation vapor pressure, p_v. Also in Example 2 of Sec. 5.1 on ZnSe deposition from Zn and Se vapor, each element will chemisorb onto the other at well below the element p_v values.

Assuming, as we did for 3D nucleation, that there is sufficient surface diffusion for equilibration, the partial monolayer of adsorbed atoms will behave as a 2D gas. By analogy to a 3D gas condensing into 3D nuclei, the 2D gas then condenses into 2D nuclei as illustrated in Fig. 5.13. Here, only the top monolayer of atoms is drawn. The "atomic terrace" to the left represents a monolayer which is one atomic step (*a*) higher than the surface to the right. But unlike the 3D nucleation case, 2D nucleation from a 2D gas involves no change in any of the γ values, so one might expect there to be no nucleation barrier. However, the chemical potential, μ, of a 2D nucleus is higher than that of a continuous monolayer because of the exposed *edge*. This situation may be viewed in terms of an excess edge energy, β (J/m), which is analogous to the surface energy, γ, of the 3D case. The surface concentration of the 2D gas for which its μ is the same as that on the straight terrace

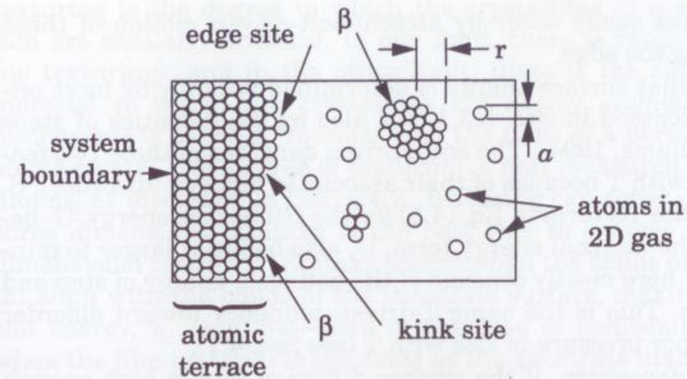


Figure 5.13 Geometry of 3D nucleation, looking down at the surface.

edge of a continuous monolayer (μ_c) may be thought of as the 2D saturation vapor concentration, n_v (mc/m^2). If n_s is the actual concentration of the 2D gas, then (n_s/n_v) becomes the 2D supersaturation ratio. By the same procedures as in the 3D case, we may then derive expressions for the critical nucleus which are analogous to Eqs. (5.39) and (5.40):

$$r^* = \frac{\beta}{a\left(\frac{RT}{V_{mc}}\right)\ln\left(\frac{n_s}{n_v}\right)} \tag{5.46}$$

and

$$\Delta G^* = \frac{\pi\beta^2}{a\left(\frac{RT}{V_{mc}}\right)\ln\left(\frac{n_s}{n_v}\right)} \tag{5.47}$$

Here, *a* is the monolayer thickness.

Once supercritical nuclei form, the 2D gas continues to attach to their edges until coalescence occurs and the monolayer is complete. Meanwhile, the next monolayer is beginning to form, and the film continues to build up in this way, atomic layer by layer. In the special case of single-crystal film deposition (epitaxy), the surface may contain many atomic terraces with straight edges as shown in Fig. 5.13. Their importance for epitaxy will be discussed in Sec. 6.7. The "kink" sites shown in Fig. 5.13 are also important surface features. Attachment of a 2D gas atom to a random site on the *straight* edge involves an increase in total edge energy, because it increases the length of the edge. Conversely, attachment to the kink site makes no change in the length of the edge; this is therefore an energetically preferred site, and edge

growth can most easily occur by attachment-driven motion of these kink sites along the edge.

We now see that surface energy is determined not only by facet orientation as discussed in Sec. 5.3.1, but also by the densities of steps and kinks (Williams, 1994). The equilibrium densities of these two features increase with T because of their associated entropy (disorder), S. That is, when the TS term in Eq. (4.5) for the Gibbs free energy, G, becomes larger, the internal energy term, U, also becomes larger to minimize G, and U here mostly consists of the potential energy of step and kink formation. This is the same T-driven tendency toward disorder that causes vapor pressure to rise with T (see Sec. 4.1).

During film deposition, if the surface diffusion rate is high enough and n_s is low enough so that the 2D gas atoms are more likely to attach to an edge than to form a critical nucleus within an atomic terrace, then edge attachment becomes the dominant growth mode; that is, we have $\Lambda > L$, where is the surface diffusion length from Eq. (5.25) and L is the distance between terraces. This is called the "continuous" growth mode, as opposed to the nucleated mode. The continuous mode of 2D growth is analogous to the type of 3D nucleation in which nucleation is more likely to occur at active surface sites than by spontaneous nucleation elsewhere on the surface. Active sites and step edges, especially kinked edges, break the nucleation barrier by providing wetting at those sites.

Two-dimensional nucleation is usually preferred to 3D because it leads to smooth growth. In nonepitaxial growth, large grain size (coarse nucleation) may be desired in addition to smoothness. Unlike in the 3D nucleation case, here large grain size and smoothness are not incompatible. That is, if adatom mobility on the substrate is sufficient, large 2D nuclei will form before the first monolayer coalesces, and then subsequent monolayers will grow epitaxially on those nuclei. But there is another problem. High adatom mobility requires a low surface-diffusion activation energy, E_s, in accordance with Eq. (5.16), but E_s tends to increase with the strength of the adsorption, E_d or E_c, as suggested in Fig. 5.2*b*. At the same time, good wetting requires low γ_i and therefore requires strong adsorption. As a result, it will not always be possible to achieve strong enough adsorption for wetting without immobilizing the adsorbate and preventing grain growth. Even so, small-diameter grains can become wider as the film grows thicker, as we will see in Sec. 5.4.2.

5.3.4 Texturing

Here, we are referring to *crystallographic* texturing rather than to surface topography, although they are often correlated. The degree of

texturing is the degree to which the crystallites in a polycrystalline film are similarly oriented. In one limit, there is random orientation (no texturing), and in the other limit, there is the single crystal. A material in which the crystallites are *nearly* aligned in all three dimensions is called a "mosaic," and the limit of a perfect mosaic is a single crystal. The degree of texturing is best measured by x-ray techniques, as discussed in Sec. 6.4.2. Texturing can occur in one, two, or three dimensions. Epitaxy is the best way to achieve perfect three-dimensional texturing. Epitaxy occurs when the bonds of the film crystal align with the bonds of the substrate surface, making the interfacial energy, γ_i, very low—zero in the case of homoepitaxy, which is when the film material is the same as the substrate material. In other cases, when no such alignment is operative, the most common form of thin-film texturing is a two-dimensional one in which the crystallite planes are aligned with respect to rotation about the two axes which lie in the plane of the substrate. This means that the film has a preferred growth plane parallel to the substrate but has random orientation with respect to rotation about the axis perpendicular to the plane of the substrate—the "azimuthal" axis.

One frequently wants to deposit a film onto a substrate to which no crystallographic alignment is possible, such as an amorphous substrate (glass, for example) or one which has a crystal symmetry or lattice dimension very different from that of the film. The achievement of 2D texturing in such cases can be very desirable when the desired film property is also crystallographically anisotropic. For example, ZnO exhibits its largest piezoelectric effect along the [0001] axis, which also happens to be its preferred growth plane, so texturing is very beneficial in that case. The magnetic saturation properties of ferromagnetic polycrystalline films used in memory discs are also highly anisotropic. The *thermodynamically* driven orientation of nuclei into a texture in nonepitaxial situations requires two conditions: (1) surface energy anisotropy, so that there will be a preferred facet; and (2) adatom mobility, so that the adatoms can arrange themselves to minimize surface energy. Since these requirements are often met, some degree of texturing is commonly observed, but the degree varies greatly with materials and deposition conditions. It is also possible that texturing can be *kinetically* driven, if the deposition rate on certain facets is higher than that on others due to a difference in S_c [Eq. (5.6)]. The kinetically favored texture is likely to be different from the thermodynamically favored one, as is the case for shape evolution in bulk crystals (Sec. 5.3.2.3).

Even with good 2D texturing, grain boundaries can be a problem. In the electronics industry, thin-film polycrystalline semiconductor materials are often used to obtain electrical isolation of transistors built on

an insulating sublayer or in order to avoid the cost of single-crystal bulk material. The latter is particularly important in large-area applications such as solar cells and flat-panel display screens. The performance of these polycrystalline semiconductors is degraded from that of single-crystal ones, because the dangling or strained bonds present at grain boundaries act as traps and recombination centers for the charge carriers. Consequently, the achievement of single-crystal semiconductor growth on inexpensive substrates has been a long-standing goal. More recently, work on high-T superconductors has shown that grain boundaries lower the current-carrying capacity by orders of magnitude [7]. Avoidance of grain boundaries requires not only a preferred growth plane (two-axis rotational alignment) but also rotational alignment with respect to the azimuthal axis. Attempts have been made for about two decades to encourage rotational alignment in this dimension by generating on an amorphous substrate a pattern matching the film's crystallographic symmetry—a template [8]. Presumably, nuclei should preferentially align to the edges of the template to minimize their edge energy, β, thus achieving what is called "graphoepitaxy." These attempts have had varying degrees of success, but a reliable, large-scale process has yet to emerge. Some reasons are discussed below.

For a nucleus to diffuse to and align itself with an edge of the template pattern, it must be mobile on the surface. The mobility of a nucleus decreases rapidly as its radius increases, for several reasons: (1) increased contact area and bonding to the substrate; (2) increased mass and therefore lower velocity for the same thermal energy; and (3) reduction of the translational-energy fluctuations that result from impact by adatoms of the 2D gas (2D Brownian motion). Not much is known about nucleus mobility, and it is often ignored in models of nucleus growth. In graphoepitaxy attempts, it is likely that a principal cause of failure relates to the scale of the template relative to the scale of the largest mobile nuclei. This situation is illustrated in Fig. 5.14, in which we are looking down onto a patterned substrate containing various square nuclei. There are two matters of scale involved here. First, if nucleation is occurring on terraces between adjacent pattern edges, as in the case of the large nucleus in (*a*), we must have $\Lambda > L$ for the largest nuclei, where Λ is the nucleus diffusion length and L is the pattern spacing. If instead, the large nucleus in (*a*) becomes immobilized where it is shown, graphoepitaxy will fail. The second matter of scale involves the radius, r, of nuclei forming on or diffusing to the pattern edge relative to the spatial wavelength, λ, of the edge roughness. For $r < \lambda$, nuclei will misalign at the edge as shown in (*a*); for $r > \lambda$, they will align well as shown in (*b*). In a relatively successful recent attempt [9], 85 percent alignment of ZnS was achieved in a pattern of

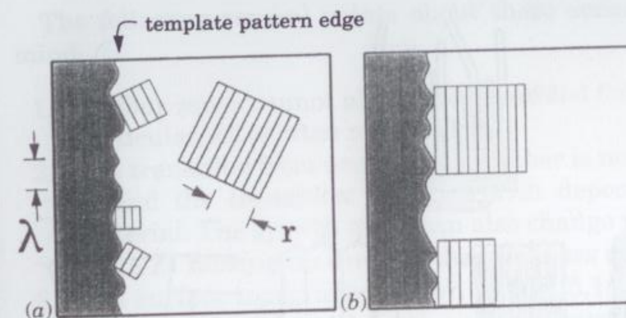


Figure 5.14 Graphoepitaxy on an irregular template, when the larger nuclei are (*a*) immobile and (*b*) mobile on the surface.

pyramidal pits transferred to polyimide from Si(100) which had been etched into {111} pyramids. Success was attributed to several factors: (1) strongly preferred {111} planes of exposure for ZnS; (2) smoothness of the template pattern; (3) nonwetting of the nuclei to the substrate, which increased their freedom to rotate; and (4) selective nucleation in the bottoms of the pyramidal pits.

5.4 Structure Development

Upon coalescence of the surface nuclei to form a continuous film, the nucleation step of film deposition is complete, and the fourth step begins: development of the bulk film structure. The form of this structure changes dramatically with the amount of thermal motion taking place during film growth, which scales with the ratio of the substrate T to the melting point of the film, T_s/T_m (in K, not °C), known as the "homologous" or "reduced" T. Structural form also changes with the amount of additional energy being delivered to the growth surface. Three structural zones (Z1, 2, and 3) were initially identified in an evaporative deposition study which included both metals and ceramics [10], and these zones have since been observed in a wide variety of film materials deposited by all of the vapor-phase processes. A fourth "transitional" zone (ZT) between Z1 and Z2 was identified in sputter deposition [11] and has since been found prominent in other energy-enhanced processes. Occasionally, anomalous structural forms occur, in particular the whiskers, illustrated in Fig. 5.15, that have been seen in Ti, for example [12]. The occurrence of whiskers implies an extreme preference for growth along the vertical direction. The characteristic structures of the four basic zones are also illustrated in Fig. 5.15. They are described briefly below and are analyzed in more detail in the next two subsections.

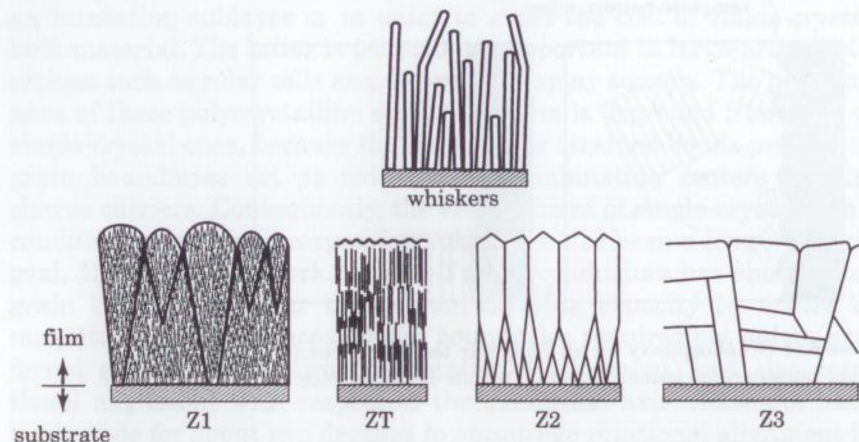


Figure 5.15 Characteristics of the four basic structural zones and of whiskers, in cross section. The ratio of substrate T to film melting T (T_s/T_m) increases in the direction Z1→ZT→Z2→Z3.

- *Z1* occurs at T_s/T_m so low that surface diffusion is negligible; that is, $\Lambda < a$ in Eq. (5.17). Z1 consists of columns typically tens of nm in diameter separated by voids a few nm across. The columns have poor crystallinity (many defects) or are amorphous. In thicker films, there becomes superimposed upon this structure an array of cones with wider voids between them. The cones terminate in domes at the surface, and the size of the domes increases with film thickness.
- *ZT* also occurs when $\Lambda < a$. It contains defected columns similar to those of Z1, but the voids and domes are absent. ZT is usually associated with energy-enhanced processes.
- *Z2* occurs at $T_s/T_m > 0.3$ or so, high enough so that surface diffusion is becoming significant. It consists of columns having tight grain boundaries between them and having a characteristic diameter which increases with T_s/T_m. Crystalline columns are less defected than in Z1 and ZT and are often facetted at the surface. The Z2 structure can also occur in amorphous films; there, the column boundaries are planes of reduced bonding rather than planes of crystallographic discontinuity.
- A transition to *Z3* occurs in certain instances at $T_s/T_m > 0.5$ or so, high enough so that considerable bulk annealing of the film is taking place during deposition. Z3 is characterized by more isotropic or equiaxed crystallite shapes. Film surfaces are often smoother at $T_s/T_m > 0.5$ for either Z2 or Z3; however, the grain boundaries can develop grooves.

The following several points about these zones need to be kept in mind.

1. All four zones cannot always be identified for a given material. In particular, Z3 is often not observed.
2. The transition from one zone to another is not always abrupt with T, and the transition T varies with deposition conditions and material. The growth mode can also change from Z3 to Z2 or from ZT to Z1 moving up through the thickness of a film.
3. The surface topography shown in Fig. 5.15 is typical, but it can vary considerably with factors such as surface-energy anisotropy and incident angle of depositing vapor.
4. Epitaxial films exhibit none of this bulk structure—at least when they are free of mosaic texture (see Sec. 5.3.4). However, they can still have highly facetted and therefore very rough surfaces when grown on planes of high surface energy, due to restructuring of the surface to minimize energy in accordance with Eq. (5.43).
5. Amorphous films exhibit bulk structure only when they are inhomogeneous, because they have no crystallographic pattern by which to define the boundaries between grains.

Homologous T is the main determinant of structure. Z1 and ZT films result from "quenched growth" processes in which thermal migration of the adsorbed material is negligible, whereas Z2 and Z3 films result from thermally activated rearrangement on or within the film, respectively. We will discuss these two regimes of structural development separately below, and we will see how the growth dynamics affect both the bulk structure and the surface topography.

5.4.1 Quenched growth

When initial bonding of the arriving vapor to the film surface is strong enough and film T is low enough, surface diffusion does not have time to occur before deposition of the next atomic layer, so that atoms become immobilized where they land. This is known as "ballistic" deposition, since the only motion of atoms which affects the deposition is the projectile direction of the arriving vapor. One might think that such deposition would result in smooth and homogeneous films, especially for vapor arriving uniformly and perpendicularly, as in high-vacuum deposition from a point source, and on very smooth substrates such as glass or chemically polished Si wafers. However, even under these conditions two destabilizing factors are always at work which generate roughness and voids. These factors are statistical roughening and self-shadowing, and we will examine them in turn below.

Statistical roughening arises because of statistical fluctuation in the vapor arrival flux. This effect can be illustrated with a simple model in which atoms of diameter a arrive perpendicularly at a steady rate but at positions randomly chosen from among the surface sites of a linear array. In this model, each atom is constrained to stick on the site it lands on even if it is on top of a pillar of atoms. The resulting surface topography after depositing an average number, $\bar{N}$, of 25 atoms per site across the array is shown in Fig. 5.16. At this point, the average column height—that is, the film thickness—is $\bar{h} = a\bar{N}$, but there is considerable variation in the heights, h, of atomic columns across the array. For large enough $\bar{N}$, as shown, this variation is mathematically described by a Gaussian distribution whose standard deviation is given by

$$\sigma = a\sqrt{\bar{N}} = \sqrt{a\bar{h}} \tag{5.48}$$

The σ of the distribution is a measure of the roughness of the film, or the "dispersion" of h about $\bar{h}$. Note that Eq. (5.48) is analogous to Eq. (5.17) for the dispersion in the *lateral* direction that arises from surface diffusion, because they are both random processes. Statistical roughening and surface diffusion are also *competing* processes, the first increasing the film roughness and the second smoothing it out. In Z1 growth, surface diffusion is completely quenched.

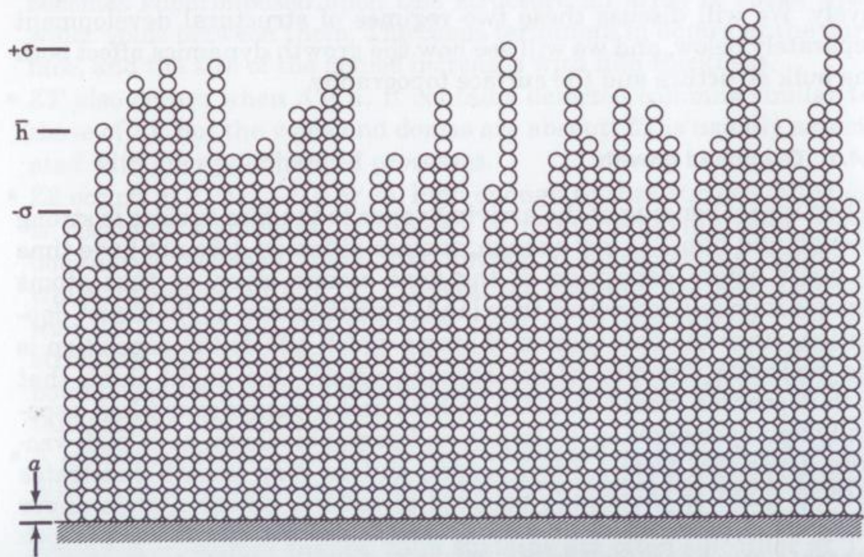


Figure 5.16 Statistical roughening in random ballistic deposition of a 25-atom-thick film. (Pascal solution courtesy of Jared Smith-Mickelson.)

A more realistic refinement of the Fig. 5.16 model avoids the single-atom pillars by constraining neighboring pillar height differences to be ±1 atom or zero. This changes the 1D model into a 2D model, because now the columns are interacting laterally, and it changes the predicted dependence of roughness on film thickness, h. In general, roughness is found to increase as h^{β}, and for the 1D model of Eq. (5.48), the roughness exponent, β = 1/2. One group [13] has calculated, using the ±1 atom constraint, that β = 1/3 for the 2D case and 1/4 for the 3D case, this last corresponding to a 2D substrate. However, that work also incorporated the "solid-on-solid" assumption which excludes voids from the structure. In Z1 growth, voids are a key feature, and observations of a wide range of actual Z1 films [14] has shown that their roughness increases with β = 3/4. That is, voids increase the rate of roughening considerably. We will see below that there are many factors which influence the Z1 microstructure. Because of this and because of the geometric complexity of a voided and rough morphology, the mathematical and computational challenges of modeling such growth are formidable, and an accurate comprehensive model has yet to emerge.

In the above discussion of roughness, we have only considered the overall variation in film thickness, h, with lateral position; this is known as the "interface width," Δh, and for the 1D case it was found to be 2σ of a Gaussian distribution. A more complete description of surface topography would give Δh as a function of the lateral spatial frequency or the spatial wavelength, λ_s, of the h fluctuation, which is essentially a spatial "noise" spectrum. Topography having a given Δh over short lateral distances (short λ_s) is going to have more effect on film structure and properties than that having long λ_s, because it will have a steeper slope, $\approx 2\Delta h/\lambda_s$. We will see in the next section that surface diffusion smooths out the short-λ_s roughness.

Various methods for measuring film surface roughness after deposition are discussed in Sec. 10.1.2. For monitoring roughness *during* deposition, laser light scattering is a convenient technique that is especially sensitive to Δh for λ_s on the order of the laser wavelength. The λ_s of maximum sensitivity depends on the scattering angle being monitored [15]: it occurs at longer λ_s for angles close to the specular angle, and at shorter λ_s for angles farther away. Thus, angle-resolved scattering patterns give a more complete description of the topography [16]. Even submonolayer roughness is detectable despite the fact that the typical (Ar-ion) laser wavelength of 488 nm is 10^3 larger!

Self-shadowing is the second factor that destabilizes surface smoothness during film deposition, and it is the cause of the characteristic Z1 voided columnar structure. To understand this effect, we

first make the Fig. 5.16 model more realistic by introducing the rule that arriving atoms cannot perch on top of each other, but rather settle sideways into the nearest "cradle" position in which they establish relaxed bond lengths to their nearest neighbors. This is known as "ballistic aggregation," and it is shown for 2D geometry in Fig. 5.17*a*. Because of the finite size of atoms ($a > 0$), aggregation can result in overhang structures such as that of Fig. 5.17*b*, which shadow the low areas from deposition. In effect, the area of vapor collection by the high areas increases from A_h to A_h', and that of the low areas decreases from A_l to A_l'; this is known as the "finite-size effect." Additional overhangs are produced by the atomic attraction of arriving vapor into the sidewalls of columns as shown in Fig. 5.17*c*. Thus, even

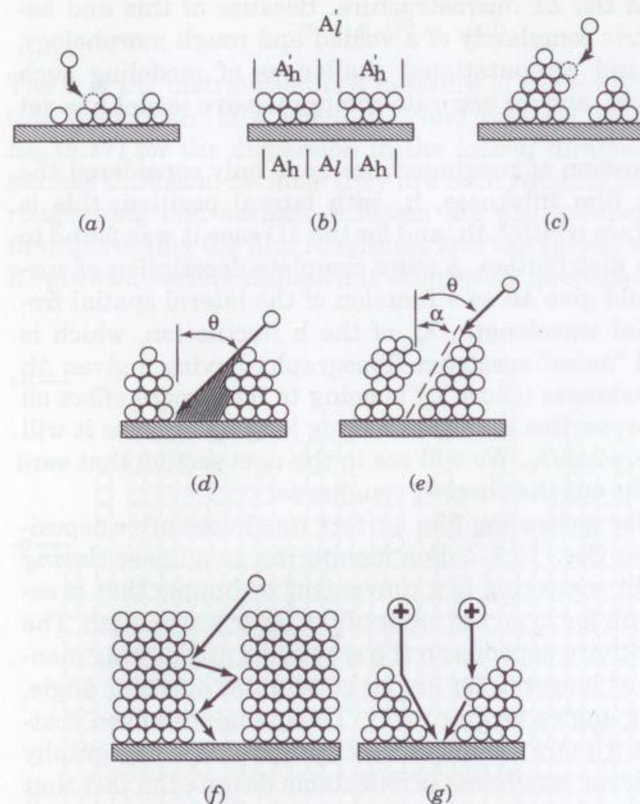


Figure 5.17 Atomistic processes in quenched-growth structure development: (*a*) ballistic aggregation, (*b*) finite-size effect, (*c*) sideways attraction, (*d*) oblique shadowing, (*e*) tilt effect, (*f*) low sticking coefficient, (*g*) void-filling by energetic particles due to enhanced mobility (left) and forward sputtering (right).

in perpendicular deposition, there develops a Z1 structure of columns with shadowed voids between them. One can readily see in Figs. 5.17*b* and 5.17*c* how self-shadowing increases the degree of roughening beyond what one would obtain just from statistical fluctuation in arrival rate, and this is why the roughness exponent, β, is so high for Z1 films.

When vapor is arriving obliquely at angle θ or over a range of angles, as shown in Fig. 5.17*d*, self-shadowing increases, thus increasing the T_s/T_m to which Z1 persists before surface diffusion can counteract it. Oblique incidence at a fixed *azimuthal* (rotational) angle also causes Z1 and Z2 columnar structures to tilt toward the incident direction at an angle α from the perpendicular (Fig. 5.17*e*). However, the so-called "tangent rule," which specifies that $\tan \alpha = (1/2)\tan \theta$, is not always obeyed. In addition to tilting, the columns become elliptical in cross section [17], with the long axis perpendicular to the plane of incidence except at high θ, when it can switch to parallel.

Incidence over a *range* of θ occurs when the vapor source subtends a large solid angle as viewed from the substrate, or when the process is operating in the fluid-flow regime (Kn << 1). The direction of vapor emanating even from a point source becomes randomized by collisions with background gas during fluid-flow transport to the substrate, so that the vapor can arrive with an angular distribution approaching the completely random cosine distribution. Thus, for processes such as sputter deposition and laser ablation whose operating pressure can span the range from molecular to fluid flow, raising pressure spreads the range of incident angles and increases self-shadowing, so that deposition pressure becomes an important determinant of film structure in these processes. At the even higher pressures of CVD, incidence is always random. Note that even at 10^5 Pa (1 atm), the vapor mean free path, $l \approx 100$ nm by Eq. (2.24), is still much longer than the Z1 void diameter, so that self-shadowing persists. That is, the "fluid" never penetrates the voids. Nevertheless, there are often compensating effects in CVD that reduce self-shadowing: namely, a low sticking coefficient, S_c, for the depositing vapor, and surface diffusion of the adsorbed precursor species even at low substrate T. Low S_c allows vapor to bounce down the void walls toward the bottom as shown in Fig. 5.17*f*. Indeed, in the limit as $S_c \rightarrow 0$, the self-shadowing effect vanishes, but of course then there is no deposition. For $S_c > 0$, the vapor flux decreases with depth into the void channel as it progressively deposits, an effect known as "nutrient depletion," which may be thought of as partial self-shadowing. Thus, Z1 growth can occur in a CVD process as well as in PVD, even when that CVD process is producing conformal coverage on *macroscopic* topography because of fluid-flow transport on that scale. CVD coverage in channels will be discussed further in Sec. 7.3.3.

In the above discussion, we have considered the initial roughness to be due to statistical roughening of a perfectly smooth substrate. At some thickness of the growing film, self-shadowing takes over from statistical fluctuation as the dominant roughening mechanism, and at that point, void formation begins. The crossover thickness depends on all of the deposition factors mentioned above. In addition, initial roughness either on the substrate or on the film nuclei upon coalescence hastens the crossover and also increases the final roughness of the Z1 film. One might think that under the quenched Z1 growth conditions, the critical nucleus would always be only one atom and would not represent a roughening factor. However, remember that the surface-diffusion activation energy of the adsorbing vapor [E_s from Eq. (5.19)] can be lower on the substrate than on the film surface, so that the initial nucleation can involve more surface diffusion than the subsequent deposition. This usually corresponds to a nonwetting situation in which the film material bonds more readily to itself than to the substrate and thus forms 3D nuclei, which would leave a rough surface after coalescence.

As one might expect, Z1 microstructure is undesirable in a thin film unless the application specifically benefits from porosity, as in the following examples. In gas-detector applications, a film property is changed by adsorption of the gas; the amount adsorbed is proportional to surface area and therefore to porosity. High surface area is also useful in catalytic applications such as fuel-cell electrodes. Columnar porosity is useful where lateral mechanical rigidity would lead to cracking or buckling due to thermal or other stresses (see Sec. 5.6), as in ZrO_2 coatings used as thermal barriers in rocket nozzles and other high-T parts. High porosity fractions such as those occurring in whisker growth cause films of opaque materials to look black as a result of light trapping, and this is useful in some optical applications. However, for most of the applications in Table 1.1, porosity is undesirable. In optical coatings, it causes light absorption and scattering. In electronics, it decreases the conductance of metal films, increases the leakage of insulators, and causes charge trapping in semiconductors. It compromises the effectiveness of chemical-barrier films and weakens films in mechanical applications. However, it is often necessary to restrict T_s to below the Z2 regime. In such cases, Z1 can usually be transformed into the void-free ZT structure by the use of energy-enhanced deposition.

The energy-enhanced processes that can reduce void formation are those in which the energy is translational kinetic energy carried by massive particles (ions and superthermal atoms), because the void reduction mechanisms involve *momentum* transfer to the growing surface. The ions can be those of the depositing vapor itself, as in

cathodic-arc and electron-beam evaporation (Chap. 8), or those of an inert gas such as Ar, as in glow-discharge plasma processes (Chap. 9). Superthermal atoms are present in supersonic beams (Sec. 4.5.4), laser ablation (Sec. 8.4), and sputter deposition (Sec. 8.5.4.3). Some of the typical energies involved are listed in Table 4.2.

Insight into the ways in which energy enhancement prevents voids has been obtained by "molecular-dynamics" computer modeling, in which all the atoms involved are followed in vibratory motion and in bond formation and breakage. In this way, four mechanisms have been identified [18]:

1. local heating due to the "thermal spike" produced upon impact, resulting in local surface diffusion
2. less trajectory curvature (Fig. 5.17*c*) due to the higher approach velocity
3. higher "impact mobility" of the adatoms so that they can move down into the voids
4. forward sputtering of other adatoms into the voids.

The last two effects are the dominant ones and are illustrated in Fig. 5.17*g*. Impact mobility amounts to a high surface diffusion rate for a few atomic distances until the excess kinetic energy upon impact becomes dissipated into the bulk. In the forward-sputtering effect, momentum transfer from the approaching particle to the adatom causes the latter to be knocked loose and scattered forward.

Figure 5.18 shows a 2D molecular-dynamics simulation of structure development versus the kinetic energy carried by atoms perpendicularly incident on a film held at $T_s = 0$ K. The lowest of the three ratios of incident energy, E_t, to the adatom potential-well depth, E_c, corresponds to thermal deposition and is dominated by ballistic aggregation, so that voids quickly develop. The effect of added E_t on void filling is quite dramatic in this simulation, as is the effect observed experimentally in energy-enhanced deposition. The highest ratio shown (1.5) would typically correspond to an E_t of 5 eV, which is characteristic of sputtered particles and of the lowest-energy ions emanating from plasmas. More recent molecular-dynamics simulations done in 3D, which uses much more computer time than 2D, have shown that the overhanging atoms predicted in 2D are actually unstable even in thermal deposition [20]. This emphasizes the importance of accurate modeling. Of course, some overhanging atoms *must* be stable in order for a Z1 structure to develop under perpendicular incidence. Collisions of energetic particles with surface atoms are discussed further in Sec. 8.5.

In practice, ZT is encouraged by depositing at low pressure so that incident particles neither become scattered into more oblique incident

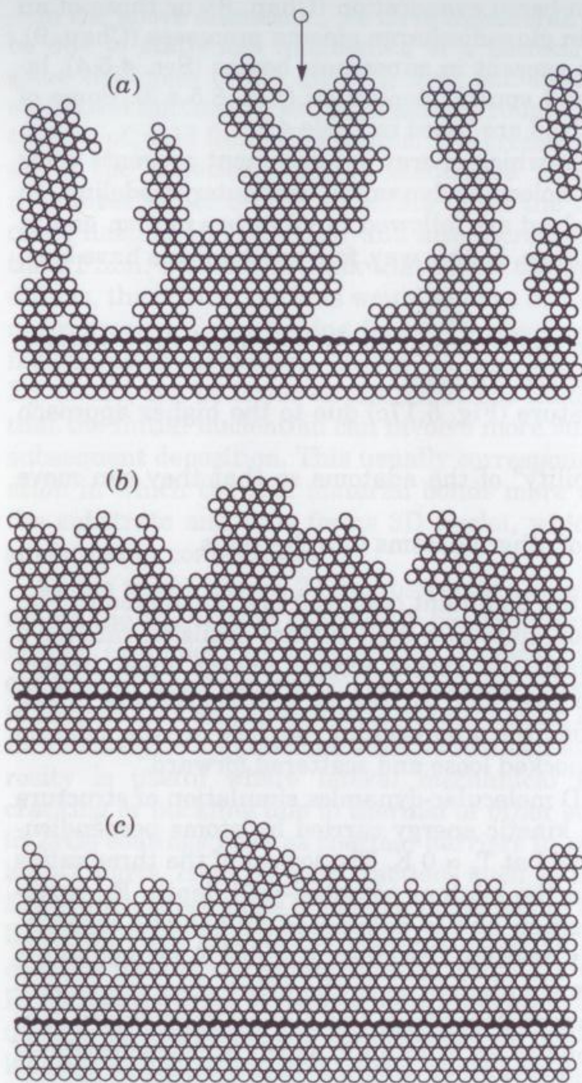


Figure 5.18 Two-dimensional molecular-dynamics simulation of the deposition of energetic atoms impinging perpendicularly onto a substrate held at 0 K. The horizontal line is the substrate interface. Normalized impinging energy E_i/E_c, as defined in text, is (*a*) 0.02, (*b*) 0.5, and (c) 1.5. (Source: Reprinted from Ref. 19 by permission.)

angles nor dissipate their kinetic energy in gas collisions. When the energy is being provided by a supplemental source of inert ions, void filling increases with ion flux and decreases with increasing deposition

rate. Microstructure can also vary with *position* on substrates having convoluted topography, which is common in microcircuit fabrication, for example. Figure 5.19 illustrates schematically a sputter-deposition situation discussed by Thornton (1986), in which growth is Z1 on the top surface due to a broad angle of incidence of the vapor flux, but is ZT at the bottom of the trench due to flux collimation by the sidewalls. The sidewalls are Z1 with even higher void fraction as well as tilted columns, due to the oblique flux. Since film properties vary with microstructure, this situation is clearly undesirable. The remedy is to increase particle collimation or energy input so that the whole film becomes ZT, or to increase T_s so that it becomes Z2.

The degree of collimation depends on the solid angle that the vapor source subtends at the substrate. When one wants to coat large areas uniformly, a row of evaporation sources (Fig. 4.13) or a long rectangular-magnetron sputter source (Sec. 9.3.4) is used in conjunction with substrate transport perpendicular to the row. If, in such cases, the distance from source to substrate is made small for efficient utilization of source material, the solid angle will be large in both directions. Thus, there is a trade-off between utilization and collimation. Collimation can be improved either by increasing transport distance or, in the magnetron case, by inserting a honeycomb baffle into the transport space to trap obliquely-directed vapor. This trapping decreases mate-

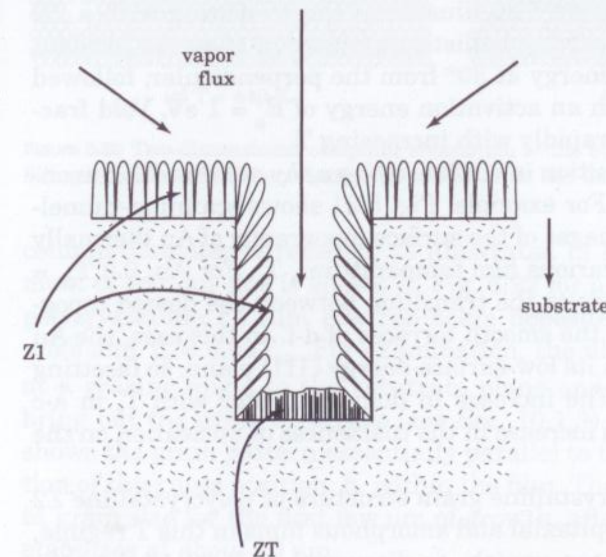


Figure 5.19 Microstructure variation with substrate geometry for sputter deposition into a trench. [Adapted from Thornton (1986).]

rial utilization by the same amount as does increasing transport distance for a given improvement in collimation, but it can also cause particulate contamination from deposits flaking off of the honeycomb.

5.4.2 Thermally activated growth

When film T is raised to where $\Lambda > a$ in Eq. (5.19), surface diffusion becomes a key factor in determining film morphology. Voids become filled by diffusing adatoms, and the film develops the characteristic Z2 columnar grain structure. Since Λ increases exponentially with T, the transition occurs over a narrow T range, and empirically it is found to occur at about 0.3 T_m for many materials, where T_m is the melting T in K. The correlation with T_m occurs because both diffusion and melting depend largely on the binding energy of atoms in the solid. Indeed, the activation energy for *bulk* solid-state diffusion has been shown to be proportional to T_m, at least for crystals of a given structure [21]. Assuming that this rule holds for surface diffusion, with a proportionality factor of B, then $E_s = BT_m$ in Eq. (5.19) and the exponential becomes proportional to T_m/T. Various factors can shift the transition T from the 0.3 T_m value, however. If background impurities are adsorbing strongly on the growing film, such as oxygen on metals, they can inhibit surface diffusion. The transition T also increases slowly with increasing deposition flux, J_r, due to the appearance of $1/\sqrt{J_r}$ in Eq. (5.19). Figure 5.20 illustrates the transition with a 2D computer simulation involving ballistic aggregation of atoms incident with negligible kinetic energy at 30° from the perpendicular, followed by surface diffusion with an activation energy of E_s = 1 eV. Void fraction is seen to decrease rapidly with increasing T.

A similarly rapid transition in surface *topography* occurs with the onset of surface diffusion. For example, Fig. 5.21 shows scanning-tunneling-microscope (STM) images of the surface topography of Au thermally deposited onto mica at various film temperatures, T_s. For Au, 0.3 T_m = 128° C, which corresponds to the transition between the domed topography of images a-c and the smooth terraces of d-f. In this case, the Au is growing epitaxially on its low-surface-energy (111) plane, so facetting is not observed in d-f. The increase in dome diameter with T_s in a-c probably results from an increase in the coarseness of nucleation on the mica with increasing T_s.

In cross section, the crystalline grain structure of polycrystalline Z2 films is columnar. (For epitaxial and amorphous films in this T regime, the column boundaries can vanish, as discussed earlier.) At the substrate interface, the grains of a polycrystalline film begin as nuclei and grow upwards as columns after coalescence. With increasing thickness, some columns expand at the expense of others until a limiting

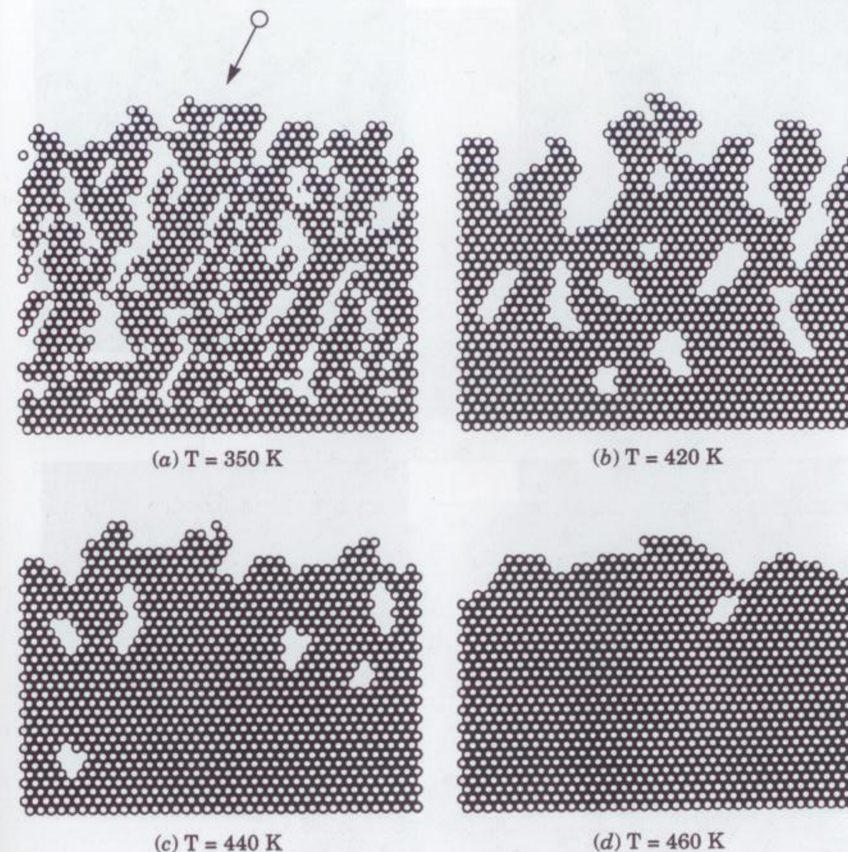
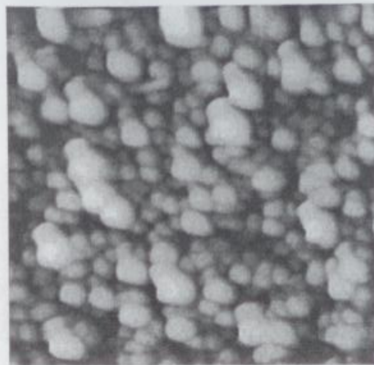


Figure 5.20 Two-dimensional computer simulation of the effect of substrate T on void filling by surface diffusion. (Source: Reprinted from Ref. 22 by permission.)

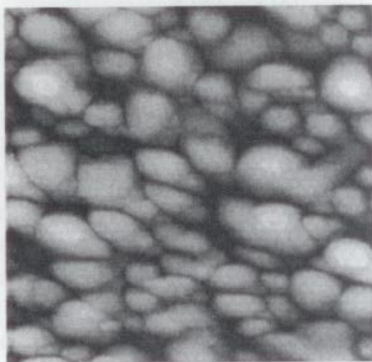
column diameter is reached, as illustrated in Fig. 5.15. The development of column size is shown in Fig. 5.22 for a 40-μm-thick columnar polycrystalline Fe film electron-beam deposited onto alumina at 80 nm/s and $T_s = 0.64T_m$. For examination, the film was lapped through at a grazing angle to the substrate plane and then etched lightly to bring out the grain structure (see Sec. 10.1.3). Thus, the micrograph shows the grain pattern essentially parallel to the substrate as a function of thickness position, h, within the film. There is a steady increase in grain size for the first few μm of growth, after which the grain size stabilizes at about 40 μm.

The driving force for columnar grain growth in Z2 is surface-energy minimization. Computer modeling has provided insight into this process and has predicted the increase in column diameter with thickness

(*a*) $T_s = -150°$ C

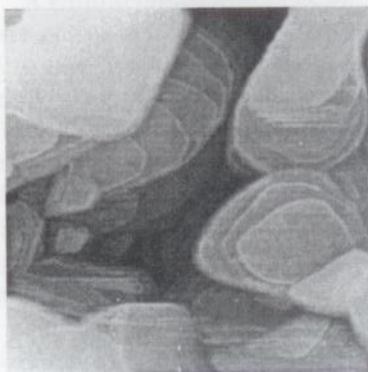


(*b*) $T_s = -50°$ C

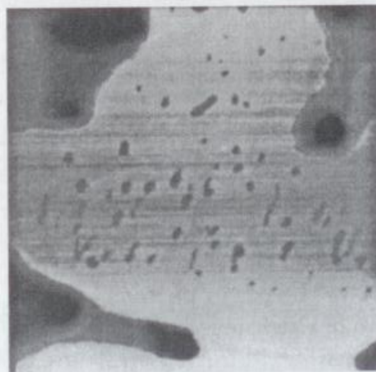


(*c*) $T_s = 50°$ C

(*d*) $T_s = 150°$ C



(*e*) $T_s = 250°$ C



(*f*) $T_s = 400°$ C

Figure 5.21 STM images of Au(111) grown on mica. Image size = 250 × 250 nm. (Source: Reprinted from Ref. 23 by permission.)

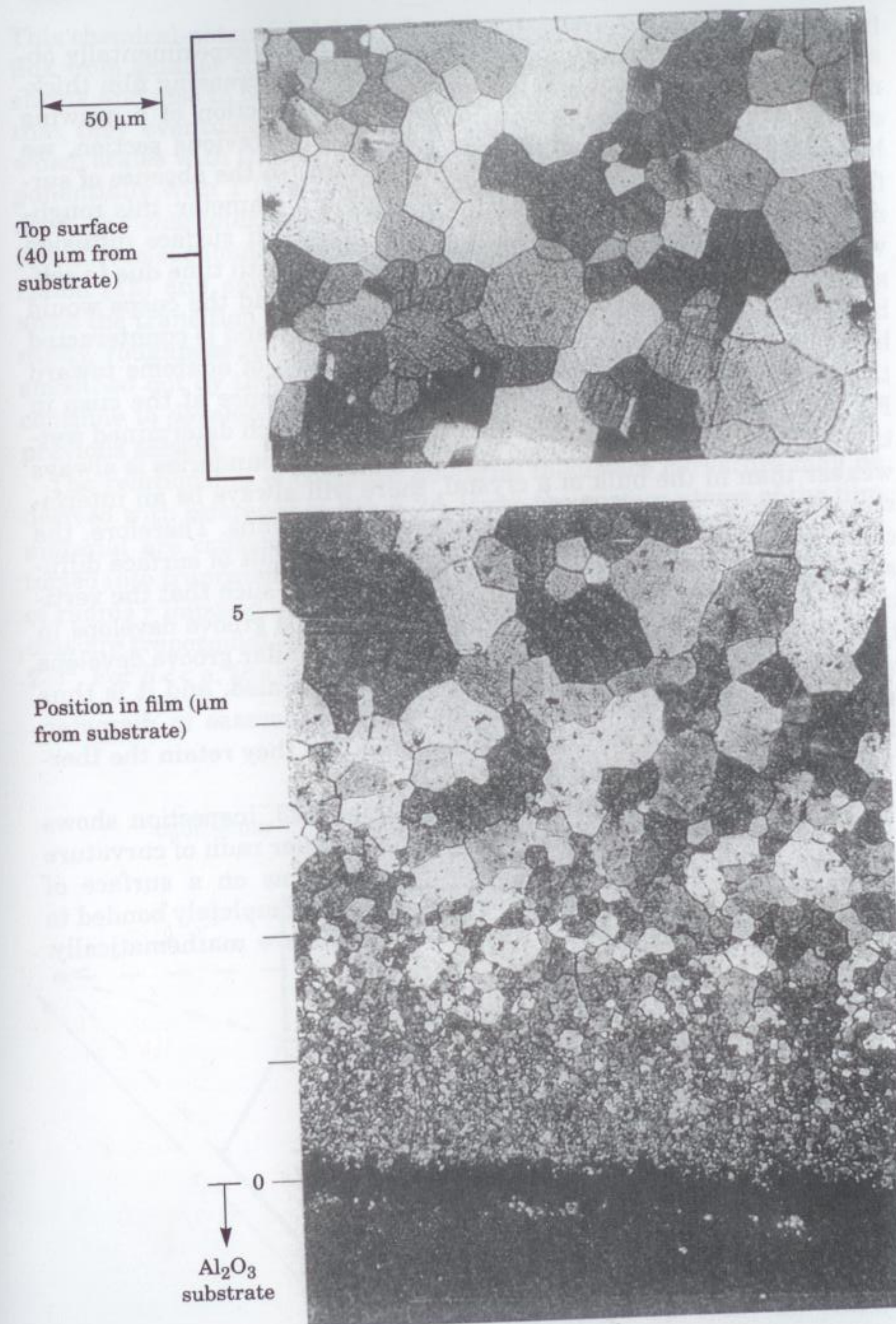


Figure 5.22 Development of columnar grain diameter with thickness in an evaporated Fe film. (Source: Author's unpublished work.)

[24], although the model cited fails to predict the experimentally observed [10] *saturation* of column diameter with increasing film thickness. The model is developed using the cross section of a growing columnar structure shown in Fig. 5.23. In the previous section, we have seen that films tend to develop roughness. In the absence of surface-energy (γ) anisotropy and for small column diameter, this roughness can be approximated as the domes shown. If surface diffusion were negligible, these domes would grow steeper with time due to self-shadowing of the cusped valleys between them, and the cusps would become the voids of the Z1 structure. In Z2, this trend is counteracted by surface diffusion, which results in net motion of adatoms toward the cusps in an attempt to satisfy the γ force balance at the cusp in Fig. 5.23. This balance is analogous to the one which determined wetting angle in Fig. 5.12. Because bonding at grain boundaries is always weaker than in the bulk of a crystal, there will always be an interfacial tension, γ_i, at the boundary between two columns. Therefore, the column surfaces do not completely flatten as a result of surface diffusion, but instead approach equilibrium dome radii such that the vertical components of γ_1 plus γ_2 balance γ_i, and thus a groove develops in the film surface along the column boundary. A similar groove develops when polished polycrystalline materials are annealed, and it is thus known as "thermal grooving." As the columns increase in diameter, they flatten on top to minimize surface area, but they retain the thermal grooves on the edges.

To satisfy the *lateral* force balance in Fig. 5.23, inspection shows that smaller-diameter columns will develop smaller radii of curvature (also see Exercise 5.10). Now, because the atoms on a surface of smaller convex radius are more exposed and less completely bonded to the bulk, they have higher γ, as we will soon show mathematically.

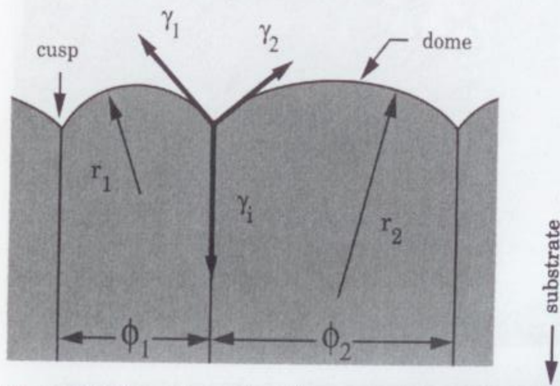


Figure 5.23 Geometry and surface energetics of columnar grain growth.

This chemical-potential driving force generates a net surface-diffusion flux from smaller columns to larger ones, causing the smaller ones to shrink and also to become more shadowed as deposition progresses, so that they eventually close out. This process continues to an extent which scales with the surface diffusion length, Λ. Now, Λ increases exponentially with T [Eq. (5.19)] or, more precisely, with T_s/T_m; and indeed, the characteristic column diameter, ϕ_c, behaves the same way [10]. Cross-sectional micrographs of metal films grown by both evaporation [12] and sputtering [11] clearly show this behavior and also show the transition from Z1 or ZT to Z2. With regard to surface *topography*, roughness features having spatial wavelength $\lambda_s < \Lambda$ become smoothed out by the surface diffusion, while those having $\lambda_s > \Lambda$ will continue to develop in accordance with the statistics discussed in the previous section.

The relationship of chemical potential to surface curvature may be derived with reference to Fig. 5.24. Here, the surface atoms of the film material are considered to be cubes of side a which have been distorted into truncated, square-based pyramids by a spherical curvature of radius r imposed on the surface. In the cross section shown, the atoms are trapezoidal, and the base exposed to the surface has area $(a + 2\delta)^2$. For $\delta \ll a$, the increase in surface area per atom versus that of a

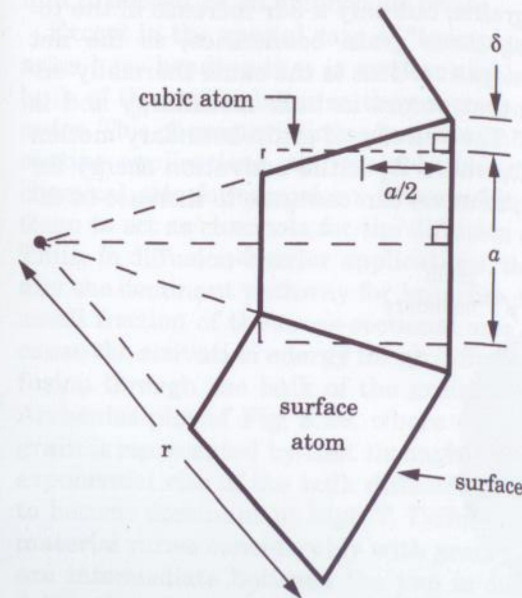


Figure 5.24 Geometry for calculating the effect of curvature on chemical potential.

flat surface (a^2) is $4a\delta$, and the corresponding increase in chemical potential is $\Delta\mu = 4a\delta\gamma_0$, where γ_0 is for a flat surface. Now, δ is related to r by similar right triangles: $(a/2 + \delta)/r \approx (a/2)/r = \delta/(a/2)$; and thus,

$$\Delta\mu = a^3\gamma_0/r = \Omega\gamma_0 K \tag{5.49}$$

where Ω = atomic volume
K = 1/r = curvature (positive is convex; negative is concave)

This is the classic expression for the curvature effect on μ.

At still higher T_s, grain-boundary migration can also occur within the bulk of the film. This is essentially bulk annealing proceeding during deposition. Figure 5.25 shows a section through the bulk of a film and parallel to the surface. A small cylindrical grain is surrounded by three larger grains. Since the curvature effect causes the μ of the atoms on the convex edge of the small grain to be larger than that of the atoms on the concave edges of the surrounding grains, atoms will transfer across the grain boundary from the small grain to the surrounding ones and eventually annihilate the small grain. Note that this process is also favorable in terms of minimization of total surface energy for the Fig. 5.25 geometry, because a decrease of dr in the radius of the small grain results in a 2πdr decrease in its boundary length with the surrounding grains, but only a 3dr increase in the total length of the surrounding three grain boundaries, so the net change in boundary length is negative. This is the same thermally-activated grain growth process that occurs in bulk metallurgy and is known as "Ostwald ripening." The velocity of grain-boundary motion is proportional to $\exp(-E_a/RT)$, where E_a is the activation energy for grain-boundary motion. Thus, columns can continue to increase in diameter beneath the growth surface at elevated T_s, so that their diameter approaches the film thickness to give the equiaxed grain structure characteristic of Z3.

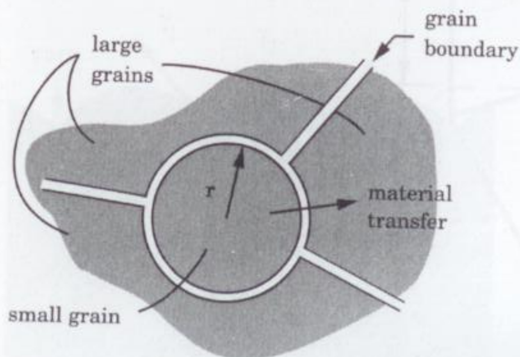


Figure 5.25 Film section parallel to the substrate showing how lateral grain growth occurs.

In the above discussion of Z2, we have assumed that γ is independent of crystallographic plane. In fact, γ anisotropy is common, and it provides an additional and often dominant driving force for surface diffusion and column selection during growth. Consequently, even a film which nucleates with completely random orientation can develop its preferred crystallographic texture with increasing thickness, as those columns oriented so that they have high-γ surfaces lose material and get closed out. Moreover, a small-diameter column can be favored over a large one despite the curvature effect if the surface γ of the small column is lower; this situation is known as "abnormal" grain growth. Column selection might also be *kinetically* determined due to varying deposition rates, as pointed out in Sec. 5.3.4.

5.4.3 Amorphous films

Grain boundaries are often undesirable—even the void-free ones of the Z2 structure. In some cases, a grain boundary (gb) can be eliminated by growing epitaxially, as we will see in the next chapter, but in the more general case, the only way to eliminate them is to cause the film to deposit as an amorphous phase.

Except in the special case of "twin-plane" boundaries, grain boundaries have bonding that is weaker and less complete than within the bulk of the grains because they represent a disruption of crystalline order. This characteristic leads to a number of problems. In protective coating applications, the grain boundaries are sites for preferential chemical attack (corrosion). Their relative openness can also cause them to act as channels for the diffusion of corrosives and impurities. Thus, in diffusion-barrier applications, the grain boundaries are usually the dominant pathway for impurity diffusion, even though only a small fraction of the cross-sectional area is represented by them, because the activation energy for gb diffusion is much lower than for diffusion through the bulk of the grain. This situation is shown in the Arrhenius plot of Fig. 5.26, where diffusion through the bulk of the grain is represented by that through a single-crystal film. The steeper exponential rise of the bulk diffusion component sometimes causes it to become dominant at high T. Diffusion behavior for polycrystalline material varies considerably with grain morphology. Amorphous films are intermediate between the two in diffusion behavior: diffusion is faster than through a single crystal but can still be orders of magnitude slower than through a polycrystalline film of the same composition (Kattelus, 1988). Diffusion of the film's own atoms also occurs

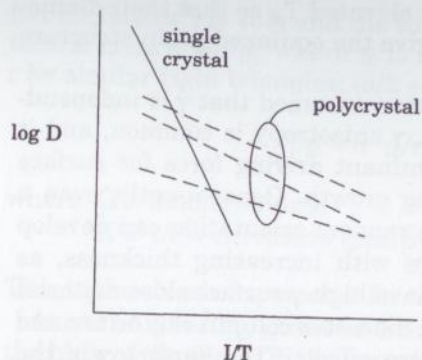


Figure 5.26 Typical diffusion behavior in a polycrystalline solid. D = diffusivity.

more rapidly along grain boundaries than through the grains, and this can cause bulk changes in film structure with time after deposition, such as hillock growth in compressively stressed films (Sec. 5.6.2) and electromigration failure in microcircuit metal lines carrying high currents. Diffusion along grain boundaries in metals appears to be inhibited by "stuffing" the boundaries (Kattelus, 1988) with a low-solubility and reactive impurity that tends to segregate to the grain boundaries and form strong bonds there (e.g., O or N in Ti). This effect may also explain the effect of Cu in inhibiting electromigration in Al microcircuit lines.

The grain boundaries cause further problems unrelated to diffusion. Mechanically-stressed films slip and fracture more easily at the boundaries because of the weak bonding. Thus, amorphous metal alloys (metallic glasses) tend to be harder and stronger than their polycrystalline counterparts. In polycrystalline semiconductors, the dangling bonds along the grain boundaries act as charge-carrier traps and thus degrade electrical behavior, as was discussed in Sec. 5.3.4. Of course, amorphous semiconductors contain dangling bonds, too. In the case of Si, most of the dangling bonds in either the amorphous or the polycrystalline phase can be terminated with H so that satisfactory electrical properties are obtained.

Because Nature likes to grow crystals, the growth of amorphous materials requires deliberately blocking this tendency, either by quenching the crystallization process (as in metastable metallic glasses) or by using a composition which does not easily crystallize. Thus, pure glassy SiO_2 crystallizes (devitrifies) as quartz; however, ordinary glass does not crystallize, because it is a solid solution of SiO_2 with various metal oxides that do not fit the quartz crystal lattice. In general, solid

solutions are less likely to crystallize when their component materials have different crystal symmetries or large differences in lattice constant. *Metallic* solid solutions (alloys) usually require the above conditions as well as quenched growth ($T_s << T_m$) to form amorphous films, because the nondirectionality of metallic bonds makes it particularly easy for the atoms to settle into a crystalline arrangement. Likely candidates for amorphous metals are alloys having the "eutectic" (minimum T_m) composition, because low T_m is an indicator of difficulty in crystallization. *Elemental* metals crystallize even at cryogenic T! Elements with covalent bonds, which are directional, crystallize less easily. Thus, the amorphous phase of Si can be stabilized at T_s up to ≈800 K by the incorporation of ≈10 percent H to terminate dangling bonds and thus inhibit the development of tetrahedral crystalline bonding. This is done most effectively by deposition from a silane (SiH_4) plasma (more in Sec. 9.6.4.1).

5.4.4 Composites

So far, we have considered only films having uniform composition from grain to grain. Composite films are made by codepositing two or more immiscible materials which then accumulate into grains of different phases. Composites can provide properties unachievable in uniform materials. Structurally, these materials tend to be equiaxed rather than columnar, because column propagation is periodically interrupted by the nucleation of another phase.

Ceramic-metal composites ("cermets") are used as resistors, wear coatings, and optical absorbers. Resistivity is determined by the ratio of ceramic to metal; it can be varied over many orders of magnitude and controlled accurately within the high-value range between that of insulators and that of metals [25]. For example, the resistivity of the popular cermet Cr-SiO varies with composition from 10^{-4} to 10^{-2} Ω-cm. In wear-coating applications of cermets, the ceramic provides hardness while the metal provides toughness (resistance to cracking). Thus, cermets such as TiC-Mo are very hard but much less brittle than the pure ceramic material.

High optical absorption is needed in films used in solar heaters and in calorimeters for optical-power measurement ("bolometers"). This can be accomplished by light trapping in cermets, which occurs as follows. Many pure ceramic materials are transparent to visible light in single-crystal or amorphous form, but in polycrystalline form, multiple scattering of light off the grain boundaries causes them to appear white whenever there is a refractive-index (ñ) discontinuity at the grain boundaries. This happens when ñ is orientation-dependent (birefringence) or when there are voids. Thus, fused quartz (amorphous

SiO_2) is clear, and devitrification (crystallization) turns it white due to the birefringence of crystalline SiO_2. However, in cermets, each scattering event off of a metal grain is accompanied by partial absorption, so that multiple scattering results in efficient light trapping. By using a metal of low reflectivity and by optimizing metal ratio and grain size, >99 percent absorption can be obtained.

Absorptive light scattering also occurs in films consisting of whiskers (Fig. 5.15) or other highly porous structures, so such films often appear black. A very effective way to obtain a highly porous film structure is to run a PVD process at a pressure so high that homogeneous nucleation occurs during the transport step; that is, "macroparticles" condense in the gas phase. This is a variation of the gas-condensation technique of cluster formation which was discussed in Sec. 4.5.4. Homogeneous nucleation is discussed further in Secs. 7.3.2 and 9.6.2. The macroparticles settle on the substrate and agglomerate into a highly porous film. Such films are very weak mechanically, however, and cermets are far preferable from that standpoint.

5.5 Interfaces

Thin films inevitably involve interfaces, both with the bulk substrate and between layers of different film materials. The thinner the film, the larger is the fraction of material occupied by interfaces. One often wants the interfaces to be abrupt, but this is seldom the equilibrium situation. Most materials dissolve in each other to some extent, even in the solid state. This process is driven by the reduction of free energy, G, which arises from the randomization of order (configurational entropy increase) and from any heat of mixing which is released [see Eq. (4.5)]. Many elements and some compounds also *react* with each other at the interface to form new compounds having lower G of formation, $\Delta_f G$, than the starting materials. These various interfacial processes are illustrated in Fig. 5.27 and will be discussed in turn below. Interfacial reaction is particularly troublesome when eross sectionpitaxy is required, because the new compound is unlikely to match

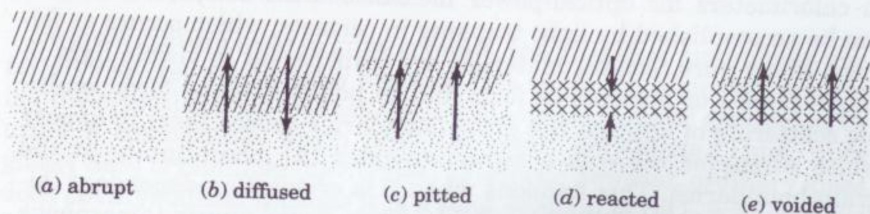


Figure 5.27 Interface types. (Arrows indicate direction of diffusion.)

the other layers in crystalline symmetry or lattice constant, and it will thus disrupt the epitaxy (more in Sec. 6.3).

Both dissolution and reaction proceed by solid-state interdiffusion across the interface, and the extent to which this occurs is determined by time and T. That is, the degree of approach to equilibrium is limited by the kinetics of interdiffusion. Eq. (5.25) for the surface diffusion length also applies here to the bulk diffusion length; that is, $\Lambda = 2\sqrt{Dt}$. When reactions occur, those which involve an activation-energy barrier, E_k, have rate-limiting constants, k_k, given by Eq. (5.2). Those which have $E_k \rightarrow 0$ will have rates limited instead by diffusion of the reactants to the interface. Since D and k_k both increase exponentially with T, it is important to keep substrate temperature, T_s, as low as possible during deposition whenever interface abruptness or avoidance of interfacial compounds is important. This is one reason why films must often be deposited in the Z1 or ZT structural regime (Sec. 5.4) even when the bulk film properties of Z2 would be preferred. When high T_s is needed to activate a CVD film-forming reaction at the surface, time at T_s can be minimized by using *rapid thermal processing* (RTP), in which the substrate is heated and cooled very rapidly by maximizing thermal coupling and minimizing thermal load in accordance with the principles of Sec. 4.5.3. Another approach to the CVD problem is to use energy enhancement to activate film-forming surface reactions while keeping the bulk film T low. This energy is often supplied from a plasma (Sec. 9.3.2).

The interdiffusion rate between any two particular materials can vary considerably with their structure. In polycrystalline materials, the rate varies with grain structure because of the dominance of grain-boundary diffusion (Fig. 5.26). In epitaxial materials, diffusion occurs through the lattice either "interstitially" (that is, between atomic sites) in the case of small diffusing atoms, or substitutionally (on atomic sites) in the case of larger ones. In the latter case, the diffusion rate increases with atomic vacancy concentration, since substitutional diffusion proceeds by the diffusing atoms swapping places with neighboring vacancies. Vacancy concentration can vary considerably with film deposition conditions; in particular, it increases with ion bombardment (Sec. 8.5.3). Interdiffusion rate also increases in the presence of an electric field (electromigration) or mechanical strain.

Thus, the extent of film interaction at interfaces is hard to predict, because it is determined by rates of solid-state diffusion and reaction which are widely variable. To understand the interaction, the interface needs to be examined for the particular deposition process conditions, using the analytical techniques of Chap. 10. Nevertheless, valuable information as to what *kind* of interaction is likely to occur is available in the phase diagrams [26], which give the compositions of

phases that have reached equilibrium with each other. The simplest situation consists of two elements that are completely miscible with each other and do not form any compounds. In this case, a planar interdiffused region will develop as shown in Fig. 5.27*b*.

Another type of binary phase diagram is represented by Al + Si in Fig. 5.28. The Al-Si interface has been studied extensively because of its importance in the metallization of integrated-circuit chips. No compounds appear between these elements, but there is a "miscibility gap"; that is, the elements have limited solubility in each other. This is indicated by the presence of two solid phases, labeled (Al) and (Si) on the phase diagram. The boundary of the (Al) phase is the solubility of Si in Al, which has a maximum value of 1.5 at.% at 577° C. No corresponding boundary is discernible on the Si-rich side, which means that Al has <0.1 at.% solubility in Si. Thus, most of the solid-state region between pure Al and pure Si consists of a two-phase mixture of Si-saturated Al and pure Si, which is labeled "(Al) + (Si)." Also shown is an intermediate composition having a minimum melting point—a "eutectic" point—at 12.2 at.% Si and 577° C. To avoid melting, T_s must never be allowed to exceed the eutectic T during deposition or subsequent processing.

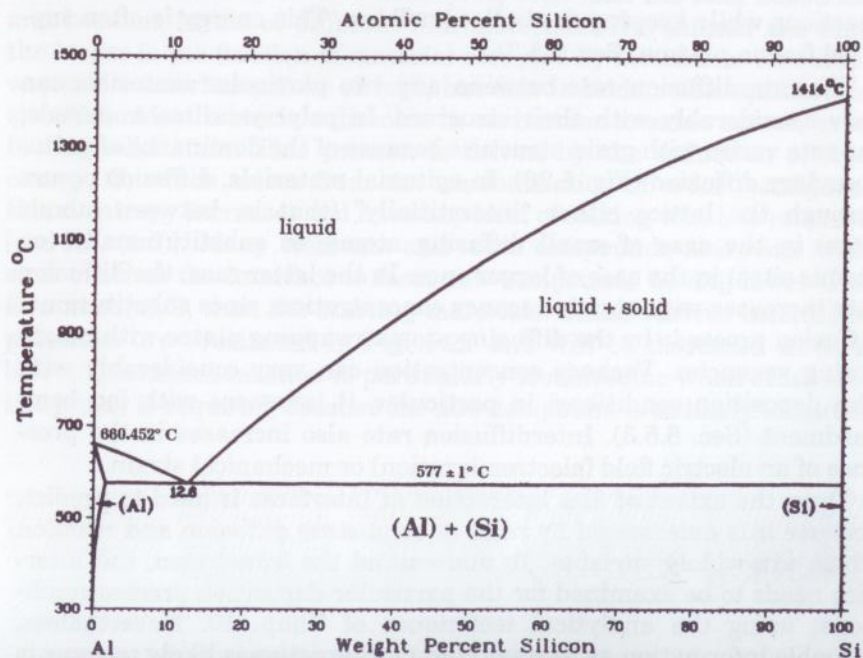


Figure 5.28 Phase diagram of Al-Si. (Source: Reprinted from Ref. 26 by permission.)

The solubility asymmetry of Al and Si means that Si will diffuse into Al, but Al will not diffuse into Si. This situation is shown in Fig. 5.27*c*, and it leads to the following problem. The space vacated by the out-diffusing Si becomes filled with the (Al) phase in an uneven manner, leading to the development of pits or "spikes" in the Si. When an Al film is etched off of a Si substrate, these pits are revealed quite dramatically, as seen in Fig. 5.29. Al spiking can short out shallow p/n junctions in the Si, so it is counteracted in the chip industry by inserting 100 nm or so of some film material which acts as a diffusion barrier to Al and Si, such as Ti-W alloy (Kattelus, 1988). The Al spiking of Fig. 5.29 actually resulted from breakthrough of an amorphous W-Re barrier layer due to Al reaction with the barrier. Structural considerations in the behavior of films as diffusion barriers were discussed in Sec. 5.4.3. Chemical considerations involve the degree to which the barrier film reacts with the adjoining films.

Many pairs of elements do react to form compounds rather than just interdiffusing. In addition, compound films can react at the interface to form new compounds there. In another example of semiconductor metallization problems, it is found that most metals (M) react with GaAs to form intermetallics (M_mGa_n) and arsenides (M_mAs_n). Known compounds appear as vertical lines on the binary phase diagrams, as was shown for the generic compound MY in Fig. 4.6*a*. When more than one compound appears on the diagram, all or only some of these may form at a film interface depending on the thermodynamic favorability

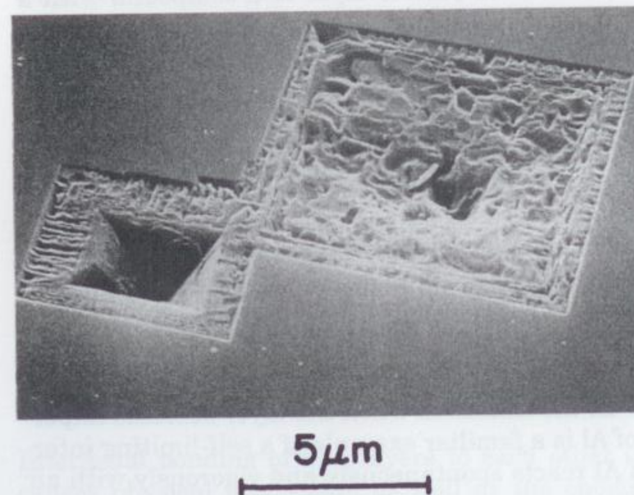
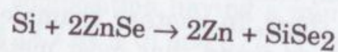


Figure 5.29 Etch pits in Si(001) substrate after removing an Al film. (Source: Reprinted from Ref. 27 by permission.)

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-

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

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

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

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

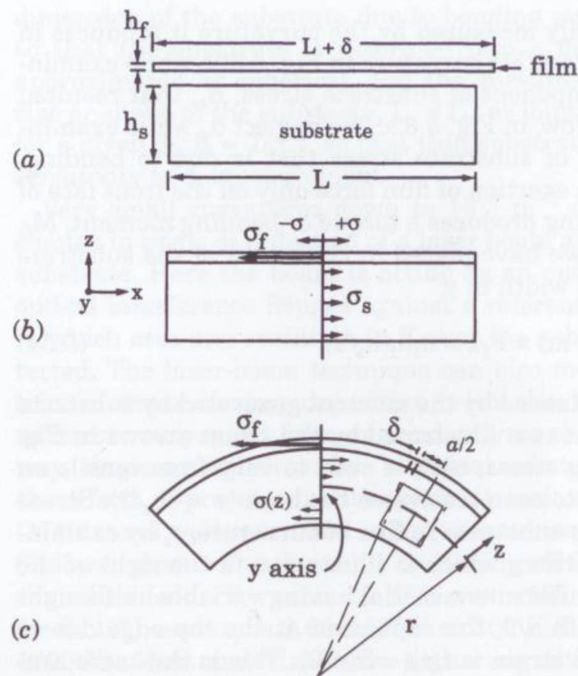


Figure 5.32 Stresses between film and substrate: (*a*) film unbonded to remove stress, (*b*) expansion component of substrate stress, and (*c*) bending component.

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:

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

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

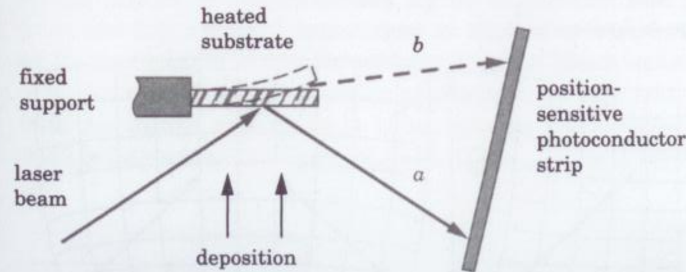


Figure 5.33 Technique for measuring intrinsic-stress build-up during deposition: (*a*) light beam before deposition and (*b*) after deposition of a compressive film.

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.

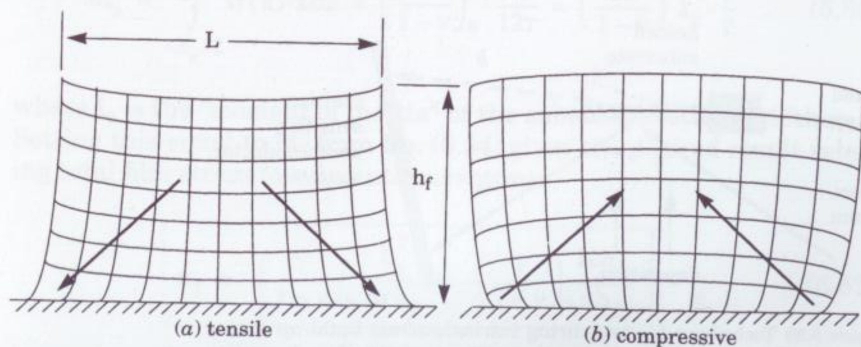


Figure 5.34 Qualitative cross-sectional strain distributions near the edges of patterned film lines that are under stress.

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

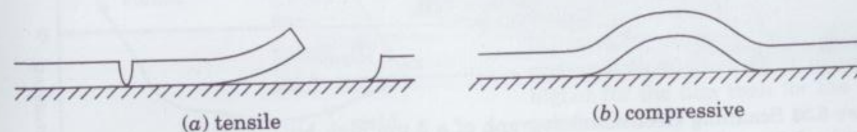


Figure 5.35 Catastrophic failure from film stress.

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

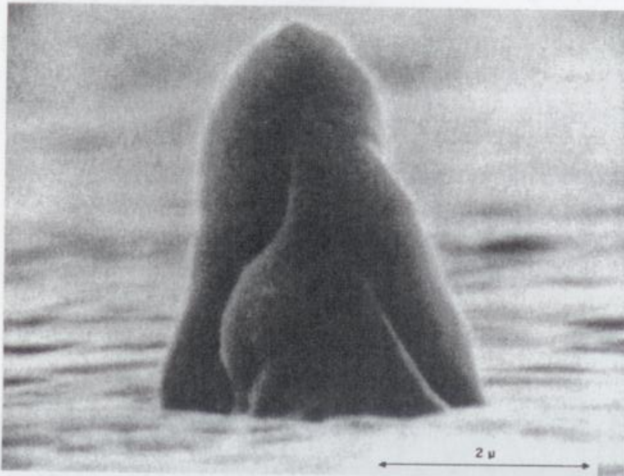


Figure 5.36 Scanning electron micrograph of a 3 μm high hillock in a 0.7 μm thick Al film. (Source: Reprinted from Ref. 35 by permission.)

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-

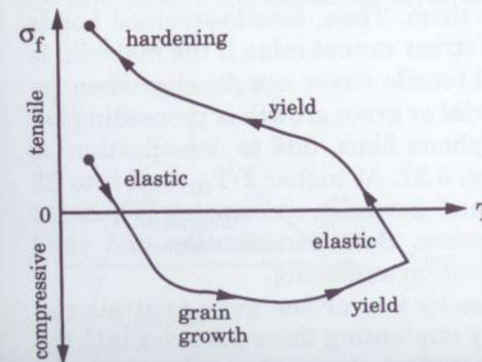


Figure 5.37 Typical hysteresis behavior during thermal cycling of metal films when α_T is higher for the film than for the substrate.

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

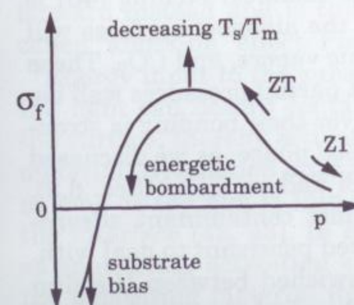


Figure 5.38 Behavior of film stress with sputter-deposition conditions.

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

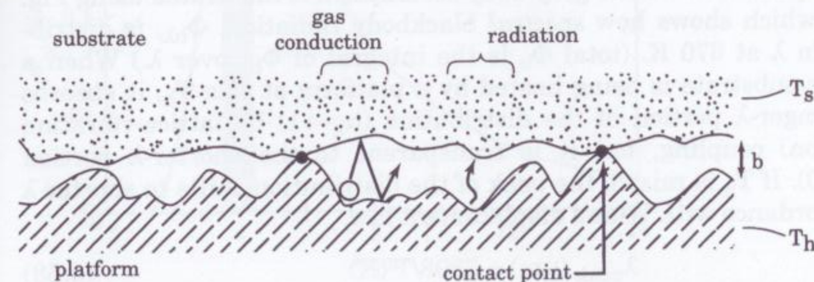


Figure 5.39 Microscopic schematic of the typical interface between a substrate and its heated platform.

heat conduction occurs in parallel to the radiation. In the following subsections, we will analyze these two mechanisms and then discuss T measurement techniques. Usually, T_s is lower than T_h because of heat loss from the substrate to the surroundings. However, if sufficient energy is arriving at the film surface during deposition, T_s will increase with deposition time and can actually become higher than T_h. This effect is especially important in energy-enhanced processes; but even in thermal evaporation, radiation is arriving from the hot evaporation source. In sputter deposition, energy is arriving both as kinetic energy of the depositing atoms and as ion bombardment.

5.8.1 Radiation

Heat transfer can occur in a vacuum only by radiation. We will analyze this process in some detail here for common substrate-heating situations in order to illustrate the large T differences which can arise and how to predict and minimize them.

The maximum amount of radiation flux which can be emitted by a surface, Φ_b, is given by the Stefan-Boltzmann blackbody radiation law [Eq. (4.19)], which is also plotted in Fig. 5.40. Real surfaces emit a fraction ε_λ of that amount, where ε_λ is the emissivity and is a function of the radiation wavelength, λ. Real surfaces also absorb only a fraction α_λ of the radiation incident on them. To allow analytical solutions to heat-transfer problems, one must make the "gray-body" assumption of constant ε_λ and α_λ over the λ involved; that is, $\varepsilon_\lambda = \varepsilon$ and $\alpha_\lambda = \alpha$. Since this is not always realistic, and since ε and α usually are not known accurately anyway, radiative-heat-transfer calculations of T_s are best backed up by calibration under actual deposition-process conditions. Nevertheless, the analytical solutions give important insight into radiative behavior as well as useful estimates of T_s. The gray-body assumption also means that $\varepsilon = \alpha$ for a given surface, a thermodynamic fact known as Kirchoff's law.

The problem with the gray-body assumption is illustrated using Fig. 5.41, which shows how *spectral* blackbody radiation, $\Phi_{b\lambda}$, is distributed in λ at 670 K. (total Φ_b is the integral of $\Phi_{b\lambda}$ over λ.) When a quartz substrate is being heated by a platform at this T_h, it absorbs the longer-λ portion of the distribution ($\alpha_\lambda \rightarrow 1$) by lattice-vibration (phonon) coupling, but it is transparent to the shorter-λ portion ($\alpha_\lambda \rightarrow 0$). If T_h is raised, the peak of the distribution shifts to shorter λ in accordance with Wien's displacement law,

$$\lambda_{peak}\ (\mu m) = 2898/T\ (K) \tag{5.58}$$

This causes the λ-averaged α of quartz and other visibly transparent or white (high-band-gap) materials to decrease with increasing T_h.

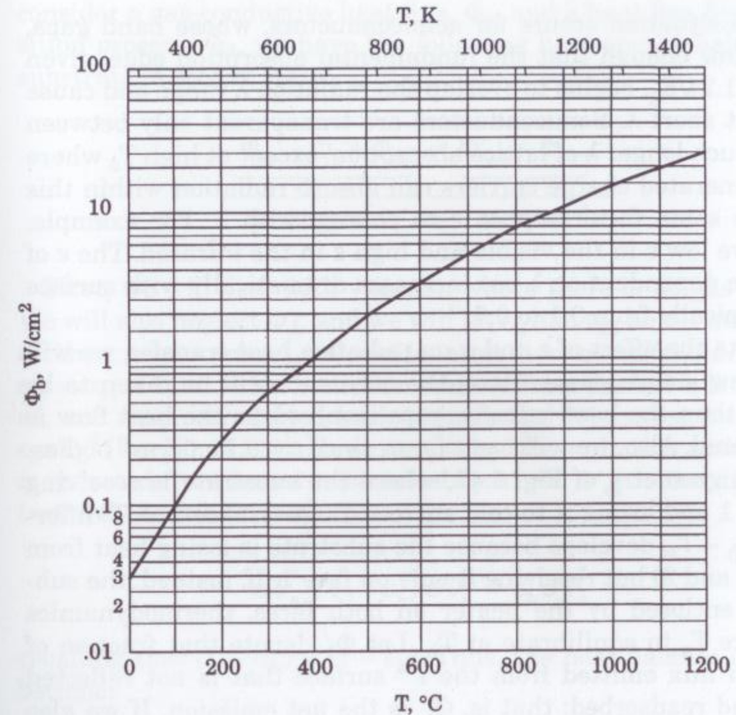


Figure 5.40 Blackbody radiation (ε = 1) vs. T, into a medium of unity refractive index.

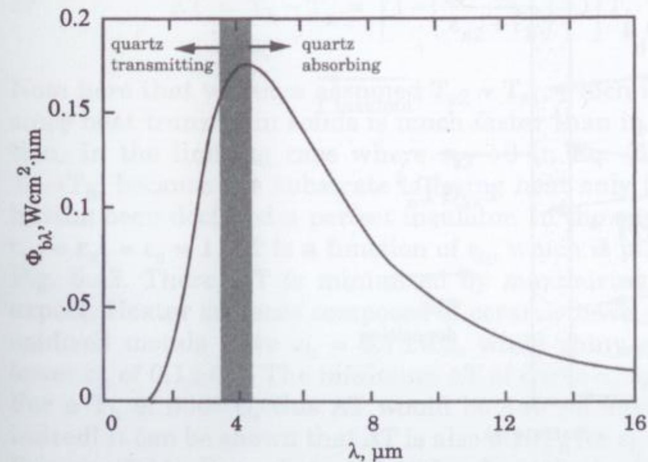


Figure 5.41 Planck's-law emissive-power function for blackbody radiation at 670 K.

The opposite situation occurs for semiconductors, whose band gaps, E_g(eV), are low enough that the fundamental absorption edge, given by λ_e (μm) ≈ 1.24/E_g, begins to overlap the radiation λ range and cause absorption at short λ. Semiconductors are transparent only between λ_e and the much longer λ of lattice absorption, except at high T_s where thermally generated charge carriers can absorb radiation within this window. The same factors cause ε to change with λ. For example, ceramics have low ε in the visible and high ε in the infrared. The ε of metals is less dependent on λ but increases dramatically with surface oxidation, typically from 0.1 to 0.7.

To illustrate the effect of ε and α on radiative heat transfer, we will examine a few simple cases. Here, the substrate will be taken to be much wider than the heat-transfer gaps involved, so the heat flow is one-dimensional. Also, we will use ε for α, since ε = α for "gray" bodies. Consider the geometry of Fig. 5.42, where the substrate is receiving heat on face 1 and losing it to cold surroundings on face 2. A T difference, $\Delta T = T_h - T_s$, develops because the substrate is losing heat from both faces (1 and 2) but receiving it only on face 1. If, instead, the substrate were enclosed by the heater on both faces, thermodynamics would require T_s to equilibrate at T_h. Let Φ_i' denote that fraction of the radiation flux emitted from the ith surface that is not reflected back to it and readsorbed; that is, Φ_i' is the net emission. If we also

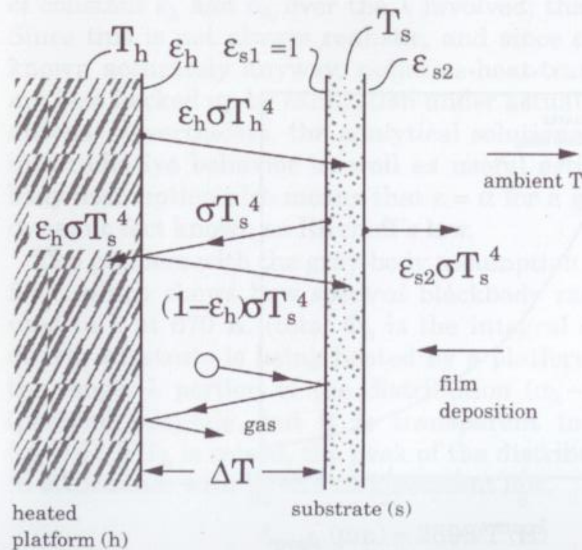


Figure 5.42 Geometry for radiative and gas-conductive heating of a substrate.

consider a gas-conductive heat flux, Φ_c, and a heat flux from the deposition process, Φ_d, we have the following heat-energy balance for the substrate in steady state:

$$\underbrace{\Phi_h' + \Phi_c + \Phi_d}_{\text{(input)}} = \underbrace{\Phi'_{s1} + \Phi'_{s2}}_{\text{(output)}} \tag{5.59}$$

First, we will deal with vacuum, where there is no gas conduction. We will also neglect Φ_d, and we will let ε_h and ε_{s2} be < 1 and $\varepsilon_{s1} = 1$. (If all the ε values are < 1, calculations get more complicated.) Given that $\varepsilon_{s1} = \alpha_{s1} = 1$, all of the radiation emitted from surface h is absorbed by s1, and none is reflected back to h, so the net emission is just $\Phi_h' = \varepsilon_h \sigma T_h^4$ [see Eq. (4.20)]. However, part of the radiation emitted from s1 *is* reflected from h and readsorbed by s1, so Φ'_{s1} is the difference between these two fluxes:

$$\sigma T_s^4 - (1-\varepsilon_h)\sigma T_s^4 = \varepsilon_h \sigma T_s^4$$

recalling that $(1-\alpha_h) = (1-\varepsilon_h)$. Thus, the heat balance of Eq. (5.59) becomes

$$\varepsilon_h \sigma T_h^4 = \varepsilon_h \sigma T_s^4 + \varepsilon_{s2} \sigma T_s^4$$

or
$$\Delta T = T_h - T_s = \left[1 - \left(\frac{\varepsilon_h}{\varepsilon_{s2} + \varepsilon_h}\right)^{1/4}\right] T_h \tag{5.60}$$

Note here that we have assumed $T_{s2} = T_{s1}$, which is quite reasonable since heat transfer in solids is much faster than in gases or by radiation. In the limiting case where $\varepsilon_{s2} \to 0$ in Eq. (5.60), we find that $T_s \to T_h$, because the substrate is losing heat only from face 1, face 2 having been declared a perfect insulator. In the opposite limit, where $\varepsilon_{s2} = \varepsilon_{s1} = \varepsilon_s = 1$, ΔT is a function of ε_h, which is plotted as curve *a* in Fig. 5.43. There, ΔT is minimized by maximizing ε_h, as one would expect. Heater surfaces composed of ceramic have $\varepsilon_h \approx 1$, and heavily oxidized metals have $\varepsilon_h = 0.7 \pm 0.2$, while shiny metals have much lower ε_h of 0.1 ± 0.1. The minimum ΔT of curve *a*, for $\varepsilon_h = 1$, is $0.16T_h$. For a T_h of 600° C, this ΔT would be 140°, a substantial difference indeed! It can be shown that ΔT is also $0.16T_h$ for $\varepsilon_h = 1$ and any ε_s (see Exercise 5.14). For substrate heating from the back by radiation only, as in Fig. 5.42, ΔT can be reduced further only by reducing ε_{s2}, but of course ε_{s2} is fixed by the film being deposited. In fact, a difference in

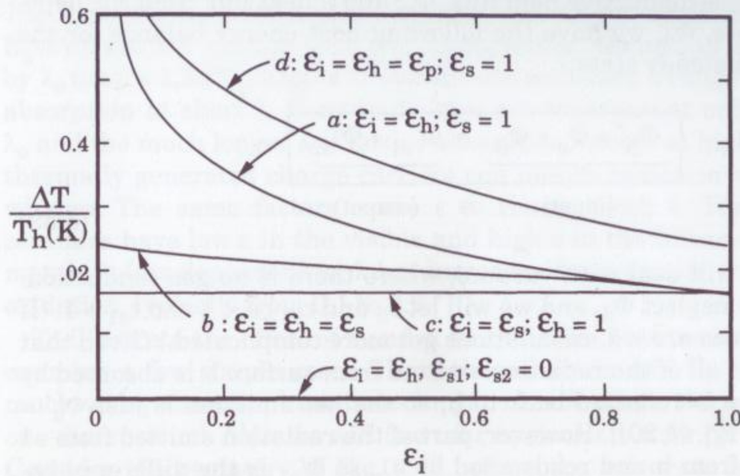


Figure 5.43 Behavior of ΔT with various ε values. Curves *a*, *b*, and *c* are for the geometry of Fig. 5.42, and solutions given are as follows: curve *a*, Eq. (5.60); curve *b*, Eq. (5.63); curve *c*, Exercise 5.15. Curve *d* is for the geometry of Fig. 5.45 and Eq. (5.66). Note that $\Delta T = 0$ whenever $\varepsilon_{s2} = 0$.

ε_{s2} from substrate to film will cause T_s to begin rising or falling toward a new steady-state value once deposition begins.

Consider next the case where all the ε terms are <1, and let them all be equal for simplicity. This leads to the multiple reflections shown in Fig. 5.44 between facing surfaces h and s1. Surface h emits a fraction ε of the blackbody radiation, and $(1-\alpha)$ [$= (1-\varepsilon)$] of this fraction is reflected from s1. A fraction α ($= \varepsilon$) of the reflection is reabsorbed by h, and this is now a fraction $\varepsilon^2(1-\varepsilon)$ of the blackbody radiation. Following this progression, we see that the sum of the reabsorptions is an infinite series having a closed-form solution:

$$\varepsilon^2(1-\varepsilon)\sum_{n=0}^{\infty}(1-\varepsilon)^{2n} = \frac{\varepsilon^2(1-\varepsilon)}{[1-(1-\varepsilon)^2]} = f(\varepsilon) \qquad (5.61)$$

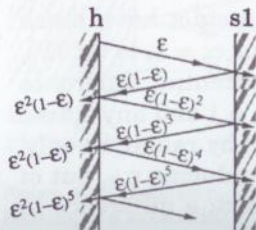


Figure 5.44 Multiple reflections between surfaces.

Thus, the net emission from h (and also from s1) is

$$\Phi_i' = [\varepsilon - f(\varepsilon)]\sigma T_i^4 = \left(\frac{\varepsilon}{2-\varepsilon}\right)\sigma T_i^4 \qquad (5.62)$$

and the heat balance for the substrate in Fig. 5.42 becomes

$$\left(\frac{\varepsilon}{2-\varepsilon}\right)\sigma T_h^4 = \left(\frac{\varepsilon}{2-\varepsilon}\right)\sigma T_s^4 + \varepsilon\sigma T_s^4 \quad \text{or} \quad \Delta T = \left[1 - \left(\frac{1}{3-\varepsilon}\right)^{1/4}\right]T_h \qquad (5.63)$$

This function is plotted as curve *b* in Fig. 5.43.

Another common substrate-heating situation is illustrated in Fig. 5.45, where the substrate is held by a pallet or carrier plate that is placed in front of the heater or transported past it during film deposition. This introduces two more surfaces, p1 and p2, between the heater and the substrate, which further increases the ΔT between them. For the case where $\varepsilon_h = \varepsilon_{p1} = \varepsilon_{p2} = \varepsilon$ and $\varepsilon_{s1} = \varepsilon_{s2} = 1$, the heat balance for the pallet is essentially that of Eq. (5.62) with a second input term added for back-radiation from the substrate:

$$\left(\frac{\varepsilon}{2-\varepsilon}\right)\sigma T_h^4 + \varepsilon\sigma T_s^4 = \left(\frac{\varepsilon}{2-\varepsilon}\right)\sigma T_p^4 + \varepsilon\sigma T_p^4 \qquad (5.64)$$

The heat balance for the substrate is equivalent to that of Eq. (5.60):

$$\varepsilon\sigma T_p^4 = (1+\varepsilon)\sigma T_s^4 \qquad (5.65)$$

Combining these two heat balances to eliminate T_p yields the remarkably simple result that

$$\Delta T = \left[1 - \left(\frac{\varepsilon}{3}\right)^{1/4}\right]T_h \qquad (5.66)$$

which is plotted as curve *d* in Fig. 5.43. To minimize ΔT, it is important that ε_h, ε_{p1}, and ε_{p2} all be large. The geometry of Fig. 5.45 also applies

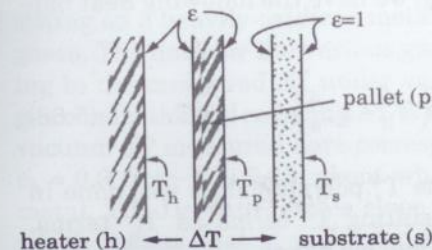


Figure 5.45 Geometry for radiative heating of a substrate held on a pallet.

to the situation where the substrate is directly facing the heater (no pallet) but is covered by a shutter on the deposition face to block deposition during vapor-source start-up. In this case, the pallet of Fig. 5.45 would become the substrate, and the substrate would become the shutter. Clearly, the substrate will be hotter when the shutter is in place, due to radiation and reflection from the shutter. When the shutter is swung away to commence deposition, the substrate will begin to re-equilibrate at a lower T (see Exercise 5.16).

Because a large ΔT always develops between the heater and the substrate whenever radiation is the only mode of heat transfer, it is best to monitor T_s directly during deposition. When this is not possible, T_s can be calibrated against T_h by using a "dummy" substrate that has a T sensor firmly attached. However, the accuracy of such a calibration will be limited by variation in the ε values with time or by variation in the heat-balance conditions such as the removal of a shutter or the input of heat from the deposition process.

5.8.2 Gas conduction

When the pressure during film deposition is more than a few Pa, gas-conductive heat transfer begins to provide a significant parallel path to radiation for thermal coupling between the heater platform and a substrate placed on it, as illustrated in Fig. 5.39. That is, the Φ_c term in the basic heat balance [Eq. (5.59)] becomes significant. From the kinetic theory of gas heat conduction (Sec. 2.8.3), we know that

$$\Phi_c = h_c \Delta T \qquad (5.67)$$

The gas heat transfer coefficient, h_c, is given by Eq. (2.32) when the gas mean free path, l, is larger that the gap, b, between the two surfaces. Then, h_c is independent of b, because the molecules are simply bouncing back and forth between the surfaces. On the other hand, when $l \ll b$, h_c is replaced by K_T/b in accordance with Eq. (2.31), where K_T is the bulk thermal conductivity of the gas.

Using Eq. (5.67) and the geometry of Fig. 5.42, and assuming for simplicity that ε_h is unity and $\varepsilon_{s1} = \varepsilon_{s2}$, we have the following heat balance on the substrate:

$$\varepsilon_s \sigma T_h^4 + h_c \Delta T = 2\varepsilon_s \sigma T_s^4 \quad \text{or} \quad \varepsilon_s \sigma\left(T_h^4 - 2T_s^4\right) = -h_c \Delta T \qquad (5.68)$$

This can be simplified by factoring the T^4 polynomial as was done in Eq. (4.27), and by taking all the resulting T^3 terms as T_h^3 terms, which is reasonable when $\Delta T \ll T_h$. Then we have

$$\left(T_h^4 - 2T_s^4\right) \approx \left(1 + 2^{1/4} + 2^{1/2} + 2^{3/4}\right) T_h^3 \left(T_h - 2^{1/4} T_s\right)$$

$$= T_h^3 (6.3\Delta T - T_h) \qquad (5.69)$$

or, upon inserting into Eq. (5.68),

$$\Delta T = \frac{\varepsilon_s \sigma T_h^4}{6.3\varepsilon_s \sigma T_h^3 + h_c} \quad \text{or} \quad \frac{1}{\Delta T} = \frac{6.3}{T_h} + \frac{h_c}{\varepsilon_s \sigma T_h^4} \qquad (5.70)$$

Note that for $h_c \to 0$ (radiation only), $\Delta T/T_h \to 0.16$, in agreement with Fig. 5.43 for $\varepsilon_h = 1$. Conversely, for $h_c \to \infty$, $\Delta T \to 0$. For intermediate h_c, if $\varepsilon_h < 1$, the 6.3 factor in Eq. (5.70) will change, but as long as $\Delta T \ll T_h$, the linear form will remain; that is, $1/\Delta T = A + Bh_c$, where A and B are functions only of T_h and the ε values. Since by Eq. (2.32), $h_c \propto \gamma' c_v p / (\sqrt{MT})$, we can also write, assuming $T \approx T_h$, that

$$\boxed{\frac{1}{\Delta T} = A + B\frac{\gamma' c_v p}{\sqrt{MT_h}}} \qquad (5.71)$$

The above formula is very useful for simplifying T_s calibration using a "dummy" substrate, which would typically have a thermocouple glued to the surface as discussed in the next subsection. When a deposition process is to be run at various conditions of p and gas composition, ΔT can vary significantly; but rather than calibrating for all process conditions, the following procedure can be used. For each T_h, the constants A and B are determined by making two measurements of T_s: one in vacuum and one at some typical process gas pressure. Then, interpolation can be used to determine ΔT for other gas compositions and pressures. If c_v is not known for a particular gas, it can be estimated by the procedures in Sec. 2.4. The thermal-accommodation factor, γ', can be taken as unity except for He, as discussed in Sec. 2.8.3. Figure 5.46 shows the Eq. (5.71) correlation for a Si substrate sitting on a heavily-oxidized metal heater at T_h = 730 K, using various gases. The data for the various gases fall on a straight line extrapolating to the measured ΔT under vacuum, as expected, and the effect of gas conduction on reducing ΔT is seen to be substantial at 140 Pa. The vacuum ΔT measured here corresponds to $\Delta T/T_h$ = 0.17, meaning that $\varepsilon_h \approx 0.9$ from Fig. 5.43, about what is expected for a heavily oxidized metal. The He data have a slope of 0.2 times that of the other gases, which means that γ = 0.2 for He on these particular surfaces. Al-

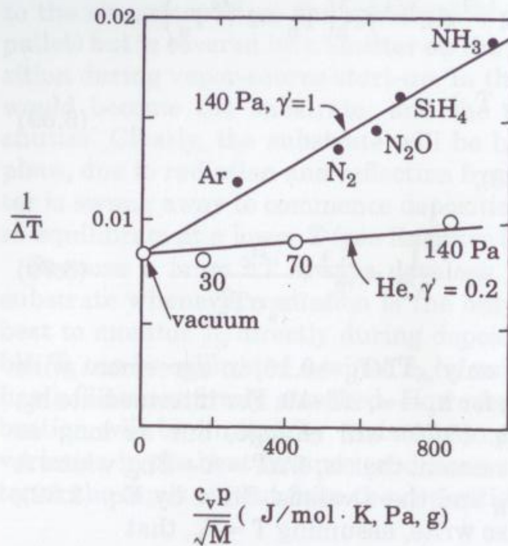


Figure 5.46 Gas-conduction correlation for T_h = 730 K [44].

though He is often thought of as a good heat-transfer gas because of its high atomic speed, Fig. 5.46 shows that it is actually poor because of the low γ. Conversely, NH_3 is quite good, because its molecular weight, M, is reasonably low, and yet it has high c_v from its many molecular degrees of freedom. For process gases other than He, where $\gamma \approx 1$, one ΔT measurement under vacuum and one at a typical process pressure are sufficient to construct the Fig. 5.46 plot from which ΔT at other process gas pressures and compositions can be interpolated.

The above discussion has assumed gas conduction with Kn > 1. However, when the gap, b, is larger or pressure,p, is higher so that Kn << 1, h_c must be replaced by K_T/b, so that the gas-conduction contribution to T_s now decreases with increasing b, does *not* increase with increasing p, and does not involve γ (see Fig. 2.9). At p approaching one atmosphere, however, "convective" gas heat transfer is added for b > 1 cm or so, due to gas circulation in T gradients (more in Sec. 7.2.3).

We have mentioned above several instances of a change in heat flux to the substrate occurring at the start of film deposition, including shutter opening, emissivity change due to the presence of the film, or process heat input. Heat flux also changes if T_h is changed between layers of a multilayer film structure. These flux changes, which we will call $\Delta\Phi$, cause a transient response in substrate T from its initial value, T_s, to a new steady-state value, T_s', as illustrated in Fig. 5.47. A similar transient response was analyzed in Sec. 4.5.3 for an evaporation crucible subjected to a change in set-point T. The heat balance given there, Eq. (4.26), can be rewritten for the present case as

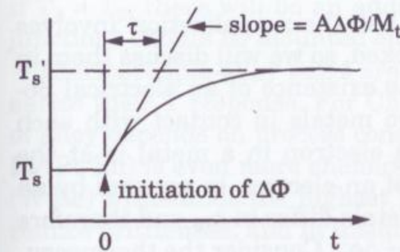


Figure 5.47 Transient response of substrate T to a step change, $\Delta\Phi$, in heat flux at time t = 0.

$$A\Delta\Phi = m_T\left(\frac{dT_s}{dt}\right)_o = c_{gs}m_s\left(\frac{dT_s}{dt}\right)_o \tag{5.72}$$

where A = substrate area
m_T = thermal load (J/K)
c_{gs} = heat capacity of the substrate, J/g·K

and where subscript o denotes the initial T_s slope. (When the transient of a substrate plus its carrier plate is being considered, the m_T of the carrier plate must be included too.) If the T_s slope is measured using a substrate having a T sensor attached, $\Delta\Phi$ can be found from Eq. (5.72). This procedure is especially useful for determining process heat input, Φ_d, as has been done for plasmas [45]. Here, the substrate itself is being used as a calorimeter. Equation (5.72) can also be used to estimate the time constant, τ, for T_s restabilization if $\Delta\Phi$ and the T change, $T_s' - T_s$, can be calculated from known ε values and h_c. Then, assuming exponential decay behavior,

$$\tau = \frac{T_s' - T_s}{(dT_s/dt)_o} \tag{5.73}$$

as shown in Fig. 5.47.

5.8.3 Measurement

Temperature (T) can be measured using any material property that changes with T, as most properties do. We will focus here on a few of the more common techniques. Because of the large ΔT associated with radiative heat transfer, T measurements at low pressure (where gas conduction is poor) are preferably made with the T sensor in solid contact with the object being measured. Alternatively, the emitted radiation itself can be measured with an optical or infrared pyrometer. We will first discuss contact sensors, the most common one being the thermocouple.

Thermocouples are simple devices, but their application involves several pitfalls which are easily overlooked, so we will discuss them in some detail. The device is based on the existence of an electrical potential difference, $\Delta\phi$, between any two metals in contact with each other. The chemical potential, μ, of an electron in a metal is at the Fermi level, which is lower than that of an electron in vacuum by an amount called the work function, ϕ_w. Metals differ in ϕ_w, and therefore they differ in electrical potential by $\Delta\phi = \Delta\phi_w$. Consider the thermocouple measurement circuit shown in Fig. 5.48*a*, which involves the common thermocouple-alloy pair Chromel and Alumel (labeled Ch and Al) plus the Cu wiring and terminals of the millivoltmeter. There are two Ch/Al junctions, a "hot" one at T_h and a "reference" one at T_r. However, *every* junction between dissimilar metals will involve a $\Delta\phi$, including the Ch/Cu junctions at T_1 and T_2. Now if the whole circuit is at the same T, then $\Delta\phi_h = -\Delta\phi_r$, since these are identical junctions in reverse; and similarly, $\Delta\phi_1 = -\Delta\phi_2$. Thus, the potentials around the circuit add up to zero, as they must to avoid generating energy out of nothing. But $\Delta\phi$ is a function of T, so if T_h is higher than T_r, a potential of $\Delta V = |\Delta\phi_h| - |\Delta\phi_r|$ is read on the meter (shown as mV). Similarly,

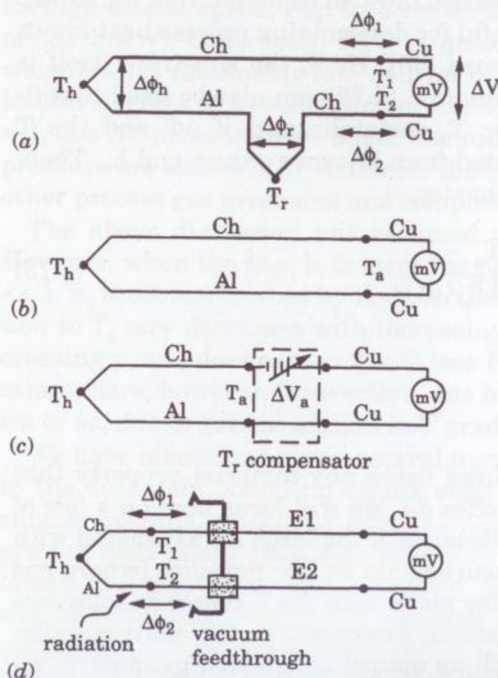


Figure 5.48 Typical thermocouple measurement circuits.

if $T_1 \neq T_2$, there will be an additional but unwanted ΔV, so these two junctions should be mounted on the same solid block.

Thermocouple-alloy pairs are formulated to both maximize and linearize their T response. For Ch/Al, $d(\Delta V)/dT \approx 0.04$ mV/K. The choice of alloy depends on process conditions. Ch/Al is resistant to oxidation. Pt/(Pt-Rh) is even more chemically resistant but is expensive. (W-Re)/(W-Re) withstands the highest T. Here, Re is alloyed into both legs to reduce brittleness, and in unequal amounts to produce a $\Delta\phi_w$.

For accurate measurements, T_r must be known, because $\Delta V \propto T_h - T_r$. Since it used to be common to use a melting ice bath for T_r, many thermocouple calibration tables are calculated for $T_r = 0°$ C. If the reference junction is omitted as in Fig. 5.48*b*, the *effective* reference junction becomes the terminals of the meter, since a Ch/Cu junction and an Al/Cu one in series are equivalent to a Ch/Al junction. This equivalency can be seen by imagining the Cu leg shortening until it disappears. Thus, the reference T becomes the ambient T of the meter, T_a. To use 0° C reference tables with readings made this way, one must first add to the meter reading the ΔV which that thermocouple would generate with a T difference of $(T_a - 0)°$ C. If the meter reading is used in a feedback loop to control T_h, omitting the reference junction will cause T_h to drift if T_a drifts. In such applications, one can insert a "reference-junction compensator" as shown in Fig. 5.48*c*, which electronically generates the ΔV correction signal appropriate for its own T_a. Most thermocouple monitoring meters have built-in reference-junction compensators.

Two additional junctions are present in the circuit when "extension alloys" are used, as shown in Fig. 5.48*d* by E1 and E2. Extension alloys avoid long runs of expensive Pt-Rh and allow W-Re to be read through a vacuum feedthrough. W wire leaks gas because it is fibrous rather than solid, so it cannot be fed through a vacuum seal. Extension alloys are formulated so that $\Delta\phi_1 = -\Delta\phi_2$ *if* $T_1 = T_2$ in Fig. 5.48*d*, no matter what the T. However, if $T_1 \neq T_2$, large errors can result. Thus, these junctions should be close to each other and carefully shielded from heat sources such as radiation from evaporation sources and substrates. Also, the alloys must not be reversed!

Various thermocouple attachment techniques are shown in Fig. 5.49. The thermocouple is best attached solidly to the object to be measured, and it should be of fine wire so that it does not conduct too much heat away. When the object is metal, spot-welding can be used. If each leg is separately welded to the object (*a*), a broken weld is readily detected as an open circuit, whereas if the thermocouple junction were welded to the metal as a unit, a broken weld might go unnoticed and give bad T_h data because of the loss in thermal contact. Contact to nonmetallic substrates can be made with braze alloy or ce-

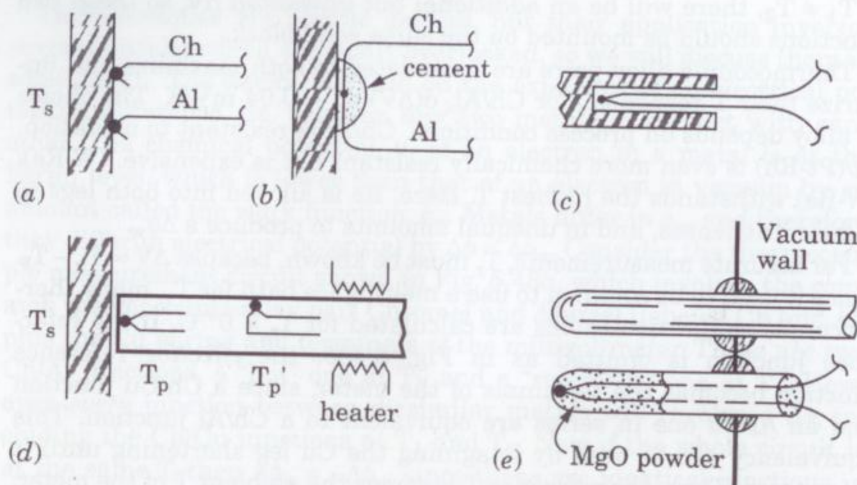


Figure 5.49 Various thermocouple attachment techniques: (*a*) separate spot-welding, (*b*) cementing, (*c*) burial in a well, (*d*) compensating probe [46], and (*e*) sheathed probe.

ramic cement for use at high T (*b*), and with polyimide at <400° C. Alternatively, the thermocouple can be buried in a deep well drilled into the substrate (*c*). In this case, it will reach the T of its surroundings as long as the heat lost down the wires is much less than the radiation it is receiving. Thermocouples simply pressed against a surface are often used at atmospheric pressure, but this can lead to a large error under vacuum because of the poor contact pointed out in Fig. 5.39. To cancel this error, a recently proposed compensating probe [46] uses two thermocouples, T_p and T_p', and adds heat as shown in (*d*) until the T drop along the probe shaft, $(T_p - T_p')$, is zero, at which point $(T_s - T_p)$ must also be zero; that is, there is no net heat flow from the substrate down the probe. For operation in electrically noisy environments such as plasma and for ease of accomplishing vacuum feedthrough, metal-sheathed thermocouples (*e*) are available with sheath diameters down to 0.25 mm.

There are other solid-contact T sensors which are not electrical and are therefore useful where substrate transport makes wiring awkward or where electrical interference is severe. One is an adhesive tape having dots which permanently turn black once a certain T has been reached. Another is the fluoroptic thermometer [47], which uses a dot of phosphor painted onto the substrate. The phosphor is activated by a pulse of UV light from a fiber-optic probe, and the decay time of the resulting phosphorescence is measured. Decay time decreases with increasing T in a well characterized and reproducible

manner. Both of the above devices are usable only up to 600 K or so, and they also outgas somewhat under vacuum.

When *no* contact to the substrate is acceptable, T_s can be found from its radiation flux at some λ using $\Phi_\lambda = \varepsilon_\lambda \Phi_{b\lambda}$, where $\Phi_{b\lambda}$ is given by Planck's law (Fig. 5.41). "Optical pyrometers" are used for this purpose, operating in the red for >900 K and in the infrared (IR) at lower T. The lower the T, the larger an area must be sampled to gather sufficient radiation for detectability. The viewing window must be big enough so as not to occlude the area being sampled, it must be transparent to the λ band being measured, and it must be kept free of deposits. The main problem with pyrometry is uncertainty in ε_λ, which is best determined directly using a thermocouple on a "dummy" substrate to calibrate the pyrometer. However, film deposition will change ε_λ, so this should be corrected for during deposition ("adaptive calibration" [48]). Also, spurious radiation sources must be avoided, such as heater radiation being transmitted *through* the substrate [49]. Above 900 K, the area sampled can be small enough to allow observation of the mouth of a well drilled into the substrate. Such a well behaves as a blackbody and therefore has an effective ε_λ of unity, as discussed in Sec. 4.2. When a well cannot be arranged, ε_λ must be determined as above.

5.9 Conclusion

We have examined here the factors determining film structure, topography, interfacial properties, and stress. There are two principal factors: the degree of interaction of the depositing vapor with the substrate and with itself, and the amount of energy input to the deposition surface. When the energy input is thermal, care must be taken to achieve good substrate-T control.

The next four chapters examine special techniques for modifying film characteristics, including the use of single-crystal substrates to achieve epitaxy, chemical vapors to improve film conformality or to obtain selective deposition, and energy-enhanced processes to improve structure or reduce substrate T.

5.10 Exercises

5.1 A molecule has a condensation coefficient of $\alpha_c = 0.2$ on a depositing film of its own solid phase. For this molecule, what are the minimum and maximum values of the thermal accommodation coefficient, γ; trapping probability, δ; sticking coefficient, S_c; and utilization factor, η?

5.2 Assuming that the pre-exponential factor for surface diffusion is given by $k_BT/\mathbf{h}$, (a) how low must the diffusion activation energy be to give a diffusion length, Λ, of roughly 100 nm between successive collisions with impinging vapor, for deposition of Si at 1 μm/h and T_s = 400° C? (b) How much of a decrease in T_s would correspond to a ×10 decrease in Λ?

5.3 For condensation onto nonwetting substrates at 300 K using a 10^{-4} Pa overpressure, what are the classical critical-nucleus radii for the following metals, neglecting substrate contact area and γ_f anisotropy? (a) Au, with γ_f = 1400 ergs/cm^2 (dynes/cm); (b) Pb, with γ_f = 560 ergs/cm^2.

5.4 Derive Young's equation [Eq. (5.45)] using surface-energy minimization.

5.5 Derive Eqs. (5.46) and (5.47) for the critical 2D nucleus.

5.6 What are the factors that influence whether adsorbing vapor will wet a substrate?

5.7 What factors inhibit nucleation?

5.8 Name at least five ways in which 3D nucleation may, in principle, be suppressed.

5.9 Explain the factors affecting the degree of void formation in Z1 growth.

5.10 Derive the relationship between dome radius and column diameter which satisfies the surface-energy force balance of Fig. 5.23, assuming isotropic γ and negligible curvature effect on γ.

5.11 100 nm of the elastic material CaF_2 is deposited with an intrinsic stress of +10 kpsi onto 300° C, 0.4-mm-thick Si(100) which is then cooled to 25° C. Assume 4×10^{10} Pa for the biaxial elastic modulus of the CaF_2. (a) What is the level of stress (MPa) and strain in the film? (b) What are the force-balance and bending-moment contributions to the strain in the substrate at the film interface? (c) What is the radius of curvature of the substrate (in m)?

5.12 A film is being deposited with an intrinsic stress of –500 MPa onto a 0.2-mm-thick Si(111) substrate mounted as shown in Fig. 5.33. If the laser beam is hitting the surface 2 cm from the edge of the support, at what film thickness (in nm) will the deflection of the reflected beam be 0.1°?

5.13 Explain at least three ways in which substrate T can vary during film deposition when its platform T is held constant.

5.14 If $\varepsilon_h = 1$ and $\varepsilon_{s1} = \varepsilon_{s2} < 1$ in the geometry of Fig. 5.42, show that $T_s = T_h/2^{1/4}$.

5.15 Show that Eq. (5.66) follows from the heat balances in Eqs. (5.64) and (5.65).

5.16 A substrate having $\varepsilon = 1$ reaches a steady state at T_s = 250° C in vacuum when placed in front of a platform having $\varepsilon = 1$ and heated to T_h and when covered on the deposition face by a shutter having $\varepsilon = 0.1$. (a) When the shutter is swung away at the start of deposition, what is the new steady state, T_s'? (b) If the substrate is a 0.4-mm-thick Si wafer (c_g = 0.18 cal/g, ρ_m = 2.34 g/cm^3), what is the time constant for T_s re-equilibration?

5.11 References

1. Feibelman, P.J. 1990. "Adsorption Energetics: First Principles Calculations of Adatom Interactions and Induced Local Lattice Relaxation." *J. Vac. Sci. Technol.* A8:2548.
2. Kang, H.C., T.A. Jachimowski, and W.H. Weinberg. 1990. "Role of Local Configurations in a Langmuir-Hinshelwood Surface Reaction: Kinetics and Compensation." *J. Chem. Phys.* 93:1418.
3. Grunthaner, P.J., F.J. Grunthaner, R.W. Fathauer, T.L. Lin, M.H. Hecht, L.D. Bell, W.J. Kaiser, F.D. Showengerdt, and J.H. Mazur. 1989. "Hydrogen-Terminated Silicon Substrates for Low-Temperature Molecular Beam Epitaxy." *Thin Solid Films* 183:197.
4. Uhrberg, R.I.G., R.D. Bringans, R.Z. Bachrach, and J.E. Northrup. 1986. "Symmetric Arsenic Dimers on the Si(100) Surface." *Phys. Rev. Lett.* 56:520.
5. Ueno, K., T. Shimada, K. Saiki, and A. Koma. 1990. "Heteroepitaxial Growth of Layered Transition Metal Chalcogenides on Sulfur-Terminated GaAs {111} Surfaces." *Appl. Phys. Lett.* 56:327.
6. Behrndt, K.H. 1966. "Phase and Order Transitions during and after Film Deposition." *J. Appl. Phys.* 37:3841.
7. Larbalestier, D.C. 1992. "High-Temperature Superconductors 1992: Bringing the Materials under Control." *MRS Bull.* August:15.
8. Givargizov, E.I. 1991. *Oriented Crystallization on Amorphous Substrates*. New York: Plenum.
9. Kanata, T., H. Takakura, H. Mizuhara, and Y. Hamakawa. 1988. "Graphoepitaxial Growth of ZnS on a Textured Natural Crystalline Surface Relief Foreign Substrate." *J. Appl. Phys.* 64:3492.
10. Movchan, B.A., and A.V. Demchishin. 1969. "Study of the Structure and Properties of Thick Vacuum Condensates of Nickel, Titanium, Tungsten, Aluminum Oxide, and Zirconium Dioxide." *Fiz. Metal. Metalloved.* 28:653.
11. Thornton, J.A. 1974. "Influence of Apparatus Geometry and Deposition Conditions on the Structure and Topography of Thick Sputtered Coatings." *J. Vac. Sci. Technol.* 11:666.
12. Bunshah, R.F., and R.S. Juntz. 1973. "Influence of Condensation Temperature on Microstructure and Tensile Properties of Ti Sheet Produced by High-Rate Physical Deposition Process." *Metallurg. Trans.* 4:21.
13. Kim, J.M., and J.M. Kosterlitz. 1989. "Growth in a Restricted Solid-on-Solid Model," *Phys. Rev. Lett.* 62:2289.
14. Messier, R., and J.E. Yehoda. 1985. "Geometry of Thin Film Morphology." *J. Appl. Phys.* 58:3739.
15. Bennett, J.M., and L. Mattsson. 1989. *Introduction to Surface Roughness and Scattering*. Washington, D.C.: Optical Society of America.
16. Pidduck, A.J., D.J. Robbins, D.B. Gasson, C. Pickering, and J.L. Glasper. 1989. "In Situ Laser Light Scattering." *J. Electrochem. Soc.* 136:3088.

17. Tait, R.N., T. Smy, and M.J. Brett. 1992. "Structural Anisotropy in Oblique Incidence Thin Metal Films." *J. Vac. Sci. Technol.* A10:1518.
18. Müller, K.-H. 1988. "Molecular Dynamics Studies of Thin Film Deposition." *J. Vac. Sci. Technol.* A6:1690.
19. Müller, K.-H. 1987. "Stress and Microstructure of Sputter-Deposited Thin Films: Molecular Dynamics Investigations." *J. Appl. Phys.* 62:1796.
20. Gilmore, C.M., and J.A. Sprague. 1991. "Molecular Dynamics Simulation of the Energetic Deposition of Ag Thin Films." *Phys. Rev. B* 44:8950.
21. Brown, A.M., and M.F. Ashby. 1980. "Correlations for Diffusion Constants." *Acta Metallurgica* 28:1085.
22. Müller, K.-H. 1987. "Models for Microstructure Evolution during Optical Thin Film Growth." *Proc. SPIE* 821:36.
23. Buchholz, S., H. Fuchs, and J.P. Rabe. 1991. "Surface Structure of Thin Metallic Films on Mica as Seen by Scanning Tunneling Microscopy, Scanning Electron Microscopy, and Low-Energy Electron Diffraction." *J. Vac. Sci. Technol.* B9:857.
24. Srolovitz, D.J., A. Mazor, and B.G. Bukiet. 1988. "Analytical and Numerical Modeling of Columnar Evolution in Thin Films." *J. Vac. Sci. Technol.* A6:2371.
25. Maissel, L. 1970. "Thin Film Resistors." Chap. 18 in *Handbook of Thin Film Technology*, ed. L.I. Maissel and R. Glang. New York: McGraw-Hill.
26. Massalski, T.B., (ed.). 1990. *Binary Alloy Phase Diagrams*, v. 1. Materials Park, Ohio: ASM International.
27. Harper, J.M.E., S.E. Hörnström, O. Thomas, A. Charai, and L. Krusin-Elbaum. 1989. "Mechanisms for Success or Failure of Diffusion Barriers between Aluminum and Silicon." *J. Vac. Sci. Technol.* A7:875.
28. Chase, M,W., et al. (eds.). 1985. *JANAF Thermochemical Tables*, 3rd ed. Washington, D.C.: American Chemical Society.
29. Wagman, D.D., et al. (eds.). 1982. "NBS Tables of Chemical Thermodynamic Properties." *J. Phys. Chem. Ref. Data* 11, suppl. no. 2. Washington, D.C.: American Chemical Society.
30. Beyers, R., K.B. Kim, and R. Sinclair. 1987. "Phase Equilibria in Metal-Gallium-Arsenic Systems: Thermodynamic Considerations for Metallization Materials." *J. Appl. Phys.* 61:2195.
31. Pretorius, R., A.M. Vredenberg, F.W. Saris, and R. de Reus. 1991. "Prediction of Phase Formation Sequence and Phase Stability in Binary Metal-Aluminum Thin-Film Systems using the Effective Heat of Formation Rule." *J. Appl. Phys.* 70:3636.
32. Andrä, W., and H. Danan. 1982. "Interface Energy as an Origin of Intrinsic Stress in Thin Films." *Phys. Stat. Sol. (a)* 70:K145.
33. Needs, R.J. 1987. "Calculations of the Surface Stress Tensor at Al(111) and (100) Surfaces." *Phys. Rev. Lett.* 58:53.
34. Murakami, M., T.-S. Kuan, and I.A. Blech. 1982. "Mechanical Properties of Thin Films on Substrates." Chap. 5 in *Treatise on Materials Science and Technology, v. 24.* New York: Academic Press.
35. Santoro, C.J. 1969. "Thermal Cycling and Surface Reconstruction in Aluminum Thin Films." *J. Electrochem. Soc.* 116:361.
36. Gardner, D.S., and P.A. Flinn. 1988. "Mechanical Stress as a Function of Temperature in Aluminum Films." *IEEE Trans. on Electron Devices* 35:2160.
37. Frost, H.J., and M.F. Ashby. 1982. *Deformation-Mechanisms Maps*. Oxford, U.K.: Pergamon Press.
38. d'Heurle, F.M. 1989. "Metallurgical Topics in Silicon Device Interconnections: Thin Film Stresses." *Internat. Mater. Reviews* 34:53.
39. Smith, D.L., A.S. Alimonda, C.-C. Chen, S.E. Ready, and B. Wacker. 1990. "Mechanism of SiN_xH_y Deposition from NH_3-SiH_4 Plasma." *J. Electrochem. Soc.* 137:614.

40. Walker, P., and W.H. Tarn. 1991. *CRC Handbook of Metal Etchants*. Boca Raton, Fla.: CRC Press (also covers many compounds).
41. Peden, C.H.F., K.B. Kidd, and N.D. Shinn. 1991. "Metal/Metal-Oxide Interfaces: A Surface Science Approach to the Study of Adhesion." *J. Vac. Sci. Technol.* A9:1518.
42. Sood, D.K., W.M. Skinner, and J.S. Williams. 1985. "Helium and Electron Beam Induced Enhancement in Adhesion of Al, Au and Pt Films on Glass." *Nucl. Instrum. and Methods in Physics Res.* B7/8:893.
43. Mitchell, I.V., G. Nyberg, and R.G. Elliman. 1984. "Enhancement of Thin Metallic Film Adhesion following Vacuum Ultraviolet Radiation." *Appl. Phys. Lett.* 45:137.
44. Smith, D.L., and A.S. Alimonda. Unpublished data.
45. Visser, R.J. 1989. "Determination of the Power and Current Densities in Argon and Oxygen Plasmas by *in situ* Temperature Measurements." *J. Vac. Sci. Technol.* A7:189.
46. Ekenstedt, M.J., and T.G. Andersson. 1991. "A Mechanical Probe for Accurate Substrate Temperature Measurements in Molecular Beam Epitaxy." *J. Vac. Sci. Technol.* B9:1605.
47. Luxtron Corporation, Mountain View, California.
48. Choi, B.I., M.I. Flik, and A.C. Anderson. 1992. "Adaptively Calibrated Pyrometry for Film Deposition Processes." In *Heat Transfer in Materials Processing*, ed. J.C. Khanpara and P. Bishop. New York: American Society of Mechanical Engineers 224:19.
49. Wright, S.L., R.F. Marks, and W.I. Wang. 1986. "Reproducible Temperature Measurement of GaAs Substrates during Molecular Beam Epitaxial Growth." *J. Vac. Sci. Technol.* B4:505.

5.12 Recommended Readings

Kattelus, H.P., and M.-A. Nicolet. 1988. "Diffusion Barriers in Semiconductor Contact Metallization." Chap. 8 in *Diffusion Phenomena in Thin Films and Microelectronic Materials*, ed. D. Gupta and P. S. Ho. Park Ridge, N.J.: Noyes Publications.

Lagally, M.G. 1993. "Atom Motion on Surfaces." *Physics Today* (November).

Lewis, B., and J.C. Anderson. 1978. *Nucleation and Growth of Thin Films*. Boston, Mass.: Academic Press.

Murakami, M. 1991. "Deformation in Thin Films by Thermal Strain." *J. Vac. Sci. Technol.* A9:2469.

Thompson, C.V. 1990. "Grain Growth in Thin Films." *Annual Rev. Mater. Sci.* 20:245.

Thornton, J.A. 1986. "The Microstructure of Sputter-Deposited Coatings." *J. Vac. Sci. Technol.* A4:3059.

Thornton, J.A., and D.W. Hoffman. 1989. "Stress-Related Effects in Thin Films." *Thin Solid Films* 171:5.

Weinberg, W.H. 1991. "Kinetics of Surface Reactions." Chap. 5 in *Dynamics of Gas-Surface Reactions*, ed. C.T. Rettner and M.N.R. Ashfold. Cambridge, U.K.: Royal Society of Chemistry.

Williams, E.D. 1994. "Surface Steps and Surface Morphology: Understanding Macroscopic Phenomena from Atomic Observations." *Surface Science* 299/300:502.

Venables, J.A., G.D.T. Spiller, and M. Hanbücken. 1984. "Nucleation and Growth of Thin Films," *Rep. Prog. Phys.*, 47:399.