

1 Introduction to surface processes

In this opening chapter, section 1.1 introduces some of the thermodynamic ideas which are used to discuss small systems. In section 1.2 these ideas are developed in more detail for small crystals, both within the terrace–ledge–kink (TLK) model, and with examples taken from real materials. Section 1.3 explores important differences between thermodynamics and kinetics; the examples given are the vapor pressure (an equilibrium thermodynamic phenomenon) and ideas about crystal growth (a non-equilibrium phenomenon approachable via kinetic arguments); both discussions include the role of atomic vibrations.

Finally, in section 1.4 the ideas behind reconstruction of crystal surfaces are discussed, and section 1.5 introduces some concepts related to surface electronics. These sections provide groundwork for the chapters which follow. You may wish to come back to individual topics later; for example, although the thermodynamics of small crystals is studied here, we will not have covered many experimental examples, nor more than the simplest models. The reason is that not everyone will want to study this topic in detail. In addition to the material in the text, some topics which may be generally useful are covered in appendices.

1.1 Elementary thermodynamic ideas of surfaces

1.1.1 *Thermodynamic potentials and the dividing surface*

The idea that thermodynamic reasoning can be applied to surfaces was pioneered by the American scientist J.W. Gibbs in the 1870s and 1880s. This work has been assembled in his collected works (Gibbs 1928, 1961) and has been summarized in several books, listed in the further reading at the end of the chapter and in Appendix A. These references given are for further exploration, but I am not expecting you to charge off and look all of them up! However, if your thermodynamics is rusty you might read Appendix E.1 before proceeding.

Gibbs' central idea was that of the 'dividing surface'. At a boundary between phases 1 and 2, the concentration profile of any elemental or molecular species changes (continuously) from one level c_1 to another c_2 , as sketched in figure 1.1. Then the extensive thermodynamic potentials (e.g. the internal energy U , the Helmholtz free energy F , or the Gibbs free energy G) can be written as a contribution from phases 1, 2 plus a surface

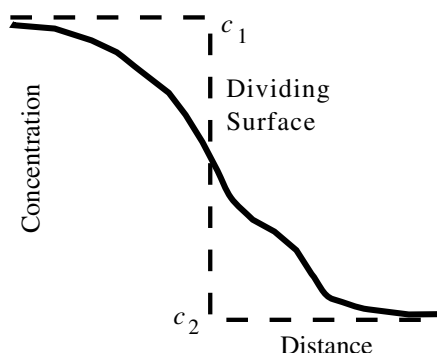


Figure 1.1. Schematic view of the ‘dividing surface’ in terms of macroscopic concentrations. See text for discussion.

term. In the thermodynamics of bulk matter, we have the bulk Helmholtz free energy $F_b = F(N_1, N_2)$ and we know that

$$dF_b = -SdT - pdV + \mu dN = 0, \quad (1.1)$$

at constant temperature T , volume V and particle number N . In this equation, S is the (bulk) entropy, p is the pressure and μ the chemical potential. Similar relationships exist for the other thermodynamic potentials; commonly used thermodynamic relations are given in Appendix E.1.

We are now interested in how the thermodynamic relations change when the system is characterized by a surface area A in addition to the volume. With the surface present the total free energy $F_t = F(N_1, N_2, A)$ and

$$dF_t = dF_b(N_1, N_2) + f_s dA. \quad (1.2)$$

This f_s is the extra Helmholtz free energy per unit area due to the presence of the surface, where we have implicitly assumed that the total number of atomic/molecular entities in the two phases, N_1 and N_2 remain constant. Gibbs’ idea of the ‘dividing surface’ was the following. Although the concentrations may vary in the neighborhood of the surface, we consider the system as uniform up to this ideal interface: f_s is then the surface excess free energy.

To make matters concrete, we might think of a one-component solid–vapor interface, where c_1 is high, and c_2 is very low; the exact concentration profile in the vicinity of the interface is typically unknown. Indeed, as we shall discuss later, it depends on the forces between the constituent atoms or molecules, and the temperature, via the statistical mechanics of the system. But we can define an imaginary dividing surface, such that the system behaves as if it comprised a uniform solid and a uniform vapor up to this dividing surface, and that the surface itself has thermodynamic properties which scale with the surface area; this is the meaning of (1.2). In many cases described in this book, the concentration changes from one phase to another can be sharp at the atomic level. This does not invalidate thermodynamic reasoning, but it leads to an interesting

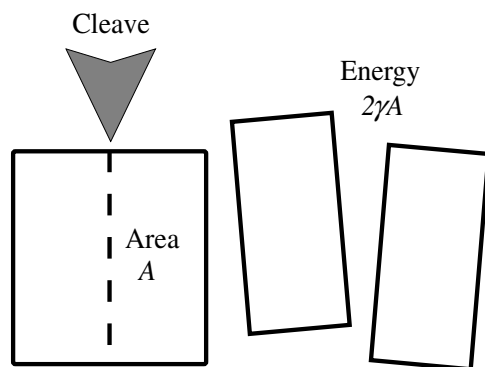


Figure 1.2. Schematic illustration of how to create new surface by cleavage. If this can be done reversibly, in the thermodynamic sense, then the work done is $2\gamma A$.

dialogue between macroscopic and atomistic views of surface processes, which will be discussed at many points in this book.

1.1.2 Surface tension and surface energy

The surface tension, γ , is defined as the reversible work done in creating unit area of new surface, i.e.

$$\gamma = \lim (dA \rightarrow 0) dW/dA = (dF_t/dA)_{T,V}. \quad (1.3)$$

In the simple illustration of figure 1.2, $\Delta F = F_1 - F_0 = 2\gamma A$; $dF_t = \gamma dA$. At const T and V ,

$$dF_t = -SdT - pdV + \sum \mu_i dN_i + f_s dA = f_s dA + \sum \mu_i dN_i. \quad (1.4)$$

Therefore,

$$\gamma dA = f_s dA + \sum \mu_i dN_i. \quad (1.5)$$

In a one-component system, e.g. metal–vapor, we can choose the dividing surface such that $dN_i = 0$, and then γ and f_s are the same. This is the sense that most physics-oriented books and articles use the term. In more complex systems, the introduction of a surface can cause changes in N_i , i.e. we have $N_1 + N_2$ in the bulk, and $dN_i \rightarrow$ surface, so that dN_i , the change in the bulk number of atoms in phase i , is negative. We then write

$$dN = -\Gamma dA \text{ and } \gamma = f_s - \sum \Gamma_i \mu_i, \quad (1.6)$$

where the second term is the free energy contribution of atoms going from the bulk to the surface; γ is the surface density of $(F - G)$ (Blakely 1973, p. 5). An equivalent view is that γ is the surface excess density of Kramers' grand potential $\Omega = -p(V_1 + V_2) + \gamma A$, which is minimized at constant T , V and μ (Desjonquères & Spanjaard 1996, p. 5). You might think about this – it is related to statistical mechanics of open systems using the grand canonical ensemble . . . ! Realistic models at $T > 0$ K need to map onto the

relevant statistical distribution to make good predictions at the atomic or molecular level; such points will be explored as we proceed through the book.

The simple example leading to (1.6) shows that care is needed: if a surface is created, the atoms or molecules can migrate *to* (or sometimes *from*) the surface. The most common phenomena of this type are as follows.

- (1) A soap film lowers the surface tension of water. Why? Because the soap molecules come out of solution and form (mono-molecular) layers at the water surface (with their ‘hydrophobic’ ends pointing outwards). Soapy water (or beer) doesn’t mind being agitated into a foam with a large surface area; these are examples one can ponder every day!
- (2) A clean surface in ultra-high vacuum has a higher free energy than an oxidized (or contaminated) surface. Why? Because if it didn’t, there would be no ‘driving force’ for oxygen to adsorb, and the reaction wouldn’t occur. It is not so clear whether there are exceptions to this rather cavalier statement, but it is generally true that the surface energy of metal oxides are much lower than the surface energy of the corresponding metal.

If you need more details of multi-component thermodynamics, see Blakely (1973, section 2.3) Adamson (1990, section 3.5) or Hudson (1992, chapter 5). For now, we don’t, and thus $\gamma = f_s$ for one-component systems. We can therefore go on to define surface excess internal energy, e_s ; entropy s_s , using the usual thermodynamic relationships:

$$e_s = f_s + Ts_s = \gamma - T(d\gamma/dT)_V; \quad s_s = -(df_s/dT)_V. \quad (1.7)$$

The entropy s_s is typically positive, and has a value of a few Boltzmann’s constant (k) per atom. One reason, not the only one, is that surface atoms are less strongly bound, and thus vibrate with lower frequency and larger amplitude than bulk atoms; another reason is that the positions of steps on the surface are not fixed. Hence $e_s > f_s$ at $T > 0$ K. The first reason is illustrated later in figure 1.17 and table 1.2.

1.1.3 Surface energy and surface stress

You may note that we have not taken the trouble to distinguish surface energy and surface stress at this stage, because of the complexity of the ideas behind surface stress. Both quantities have the same units, but surface stress is a second rank tensor, whereas surface energy is a scalar quantity. The two are the same for fluids, but can be substantially different for solids. We return to this topic in chapter 7; at this stage we should note that surface stresses, and stresses in thin films, are not identical, and may not have the same causes; thus it is reasonable to consider such effects later.

1.2 Surface energies and the Wulff theorem

In this section, the forms of small crystals are discussed in thermodynamic terms, and an over-simplified model of a crystal surface is worked through in some detail. When

this model is confronted with experimental data, it shows us that real crystal surfaces have richer structures which depend upon the details of atomic bonding and temperature; in special cases, true thermodynamic information about surfaces has been obtained by observing the shape of small crystals at high temperatures.

1.2.1 General considerations

At equilibrium, a small crystal has a specific shape at a particular temperature T . Since $dF=0$ at constant T and volume V , we obtain from the previous section that

$$\gamma dA = 0, \text{ or } \int \gamma dA \text{ is a minimum,} \quad (1.8a)$$

where the integral is over the entire surface area A . A typical *non-equilibrium* situation is a thin film with a very flat shape, or a series of small crystallites, perhaps distributed on a substrate. The *equilibrium* situation corresponds to *one* crystal with $\{hkl\}$ faces exposed such that

$$\int \gamma(hkl) dA(hkl) \text{ is minimal,} \quad (1.8b)$$

where the surface energies $\gamma(hkl)$ depend on the crystal orientation. This statement, known as the Wulff theorem, was first enunciated in 1901 (Herring 1951, 1953). If γ is isotropic, the form is a sphere in the absence of gravity, as wonderful pictures of water droplets from space missions have shown us. The sphere is simply the unique geometrical form which minimizes the surface area for a given volume. With gravity, for larger and more massive drops, the shape is no longer spherical, and the ‘sessile drop’ method is one way of measuring the surface tension of a liquid (Adamson 1990, section 2.9, Hudson 1992, chapter 3); before we all respected the dangers of mercury poisoning, this was an instructive high school experiment. For a solid, there are also several methods of measuring surface tension, most obviously using the zero creep method, in which a ball of material, weight mg , is held up by a fine wire, radius r , in equilibrium via the surface tension force $2\pi r\gamma$ (Martin & Doherty, 1976, chapter 4). But in fact, it isn’t easy to measure surface tension or surface energy accurately: we need to be aware of the likelihood of impurity segregation to the surface (think soap or oxidation again), and as we shall see in section 1.3, not all surfaces are in true equilibrium.

The net result is that one needs to know $\gamma(hkl)$ to deduce the equilibrium shape of a small crystal; conversely, if you know the shape, you *might* be able to say something about $\gamma(hkl)$. We explore this in the next section within a simple model.

1.2.2 The terrace–ledge–kink model

Consider a simple cubic structure, lattice parameter a , with nearest neighbor (nn) bonds, where the surface is inclined at angle θ to a low index (001) plane; a two-dimensional (2D) cut of this model is shown in figure 1.3, but you should imagine that the 3D crystal also contains bonds which come out of, and go into, the page.

In this model, bulk atoms have six bonds of strength ϕ . The sublimation energy L , per unit volume, of the crystal is the $(6\phi/2)(1/a^3)$, where division by 2 is to avoid double

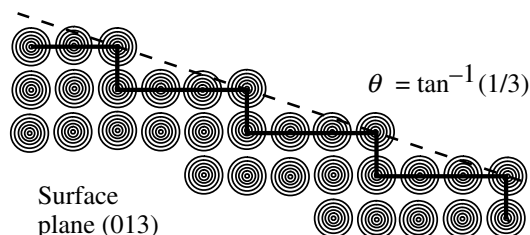


Figure 1.3. 2D cut of a simple cubic crystal, showing terrace and ledge atoms in profile. The tangent of the angle θ which the (013) surface plane makes with (001) is $1/3$. The steps, or ledges, continue into and out of the paper on the same lattice.

counting: 1 bond involves 2 atoms. Units are (say) eV/nm^3 , or many (chemical) equivalents, such as kcal/mole . Useful conversion factors are $1 \text{ eV} \equiv 11\,604 \text{ K} \equiv 23.06 \text{ kcal}/\text{mole}$; these and other factors are listed in Appendix C.

Terrace atoms have an extra energy e_t per unit area with respect to the bulk atoms, which is due to having five bonds instead of six, so there is one bond missing every a^2 . This means

$$e_t = (6 - 5) \phi / 2a^2 = \phi / 2a^2 = La/6 \text{ per unit area.} \quad (1.9a)$$

Ledge atoms have an extra energy e_l per unit length *over terrace atoms*: we have four bonds instead of five bonds, distributed every a . So

$$e_l = (5 - 4) \phi / 2a = La^2/6 \text{ per unit length.} \quad (1.9b)$$

Finally kink atoms have energy e_k *relative to the ledge atoms*, and the same argument gives

$$e_k = (4 - 3) \phi / 2 = La^3/6 \text{ per atom.} \quad (1.9c)$$

More interestingly a kink atom has 3ϕ *relative to bulk atoms*. This is the same as L/atom , so adding (or subtracting) an atom from a kink site is equivalent to condensing (or subliming) an atom from the bulk.

This last result may seem surprising, but it arises because moving a kink around on the surface leaves the number of T, L and K atoms, and the energy of the surface, unchanged. The kink site is thus a ‘repeatable step’ in the formation of the crystal. You can impress your friends by using the original German expression ‘wiederholbarer Schritt’. This schematic simple cubic crystal is referred to as a Kossel crystal, and the model as the TLK model, shown in perspective in figure 1.4. The original papers are by W. Kossel in 1927 and I.N. Stranski in 1928. Although these papers seem that they are from the distant past, my own memory of meeting Professor Stranski in the early 1970s, shortly after starting in this field, is alive and well. The scientific ‘school’ which he founded in Sofia, Bulgaria, also continues through social and political upheavals. This tradition is described in some detail by Markov (1995).

Within the TLK model, we can work out the surface energy as a function of (2D or 3D) orientation. For the 2D case shown in figure 1.3, we can show that

1.2 Surface energies and the Wulff theorem

7

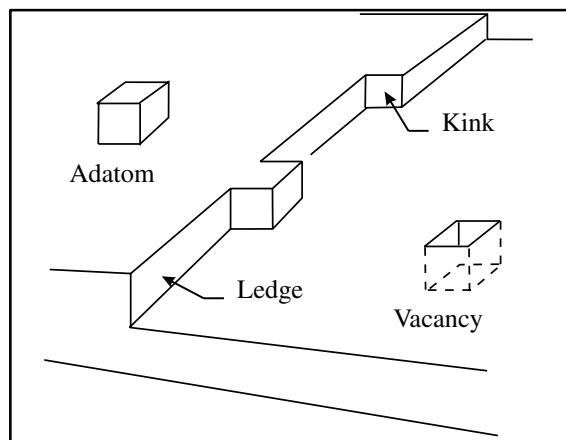


Figure 1.4. Perspective drawing of a Kossel crystal showing terraces, ledges (steps), kinks, adatoms and vacancies.

$$e_s = (e_t + e_l/na) \cos \theta. \quad (1.10a)$$

But $1/n = \tan |\theta|$. Therefore, $e_s = e_t \cos |\theta| + e_l/a \sin |\theta|$, or, within the model

$$e_s = (La/6)(\cos |\theta| + \sin |\theta|). \quad (1.10b)$$

We can draw this function as a polar diagram, noting that it is symmetric about $\theta = 45^\circ$, and repeats when θ changes by $\pm 90^\circ$. This is sufficient to show that there are cusps in all the six $\langle 100 \rangle$ directions, i.e. along the six $\{100\}$ plane normals, four of them in, and two out, of the plane of the drawing. The $|\theta|$ form arises from the fact that θ changes sign as we go through the $\{100\}$ plane orientations, but $\tan |\theta|$ does not. In this model it does not matter whether the step train of figure 1.3 slopes to the right or to the left; if the surface had lower symmetry than the bulk, as we discuss in section 1.4, then the surface energy might depend on such details.

1.2.3 Wulff construction and the forms of small crystals

The Wulff construction is shown in figure 1.5. This is a polar diagram of $\gamma(\theta)$, the γ -plot, which is sometimes called the σ -plot. The Wulff theorem says that the minimum of $\int \gamma dA$ results when one draws the perpendicular through $\gamma(\theta)$ and takes the inner envelope: this is the equilibrium form. The simplest example is for the Kossel crystal of figure 1.3, for which the equilibrium form is a cube; a more realistic case is shown in figure 1.5.

The construction is easy to see qualitatively, but not so easy to prove mathematically. The deepest cusps (C in figure 1.5) in the γ -plot are always present in the equilibrium form: these are *singular* faces. Other higher energy faces, such as the cusps H in the figure, may or may not be present, depending in detail on $\gamma(\theta)$. Between the singular faces, there may be rounded regions R, where the faces are *rough*.

The mathematics of the Wulff construction is an example of the calculus of variations; the history, including the point that the original Wulff derivation was flawed, is

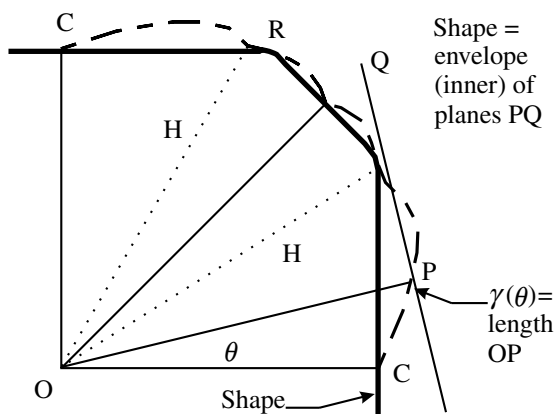


Figure 1.5. A 2D cut of a γ -plot, where the length OP is proportional to $\gamma(\theta)$, showing the cusps C and H , and the construction of the planes PQ perpendicular to OP through the points P . This particular plot leads to the existence of facets and rounded (rough) regions at R . See text for discussion

described by Herring (1953). There are various cases which can be worked out precisely, but somewhat laboriously, in order to decide by calculation whether a particular orientation is mechanically stable. Specific expressions exist for the case where γ is a function of one angular variable θ , or of the lattice parameter, a . In the former case, a face is mechanically stable or unstable depending on whether the surface stiffness

$$\gamma(\theta) + d^2\gamma(\theta)/d\theta^2 \text{ is } > \text{ or } < 0. \quad (1.11)$$

The case of negative stiffness is an unstable condition which leads to *faceting* (Nozières 1992, Desjonquères & Spanjaard 1996). This can occur at 2D internal interfaces as well as at the surface, or it can occur in 1D along steps on the surface, or along dislocations in elastically anisotropic media, both of which can have unstable directions. In other words, these phenomena occur widely in materials science, and have been extensively documented, for example by Martin & Doherty (1976) and more recently by Sutton & Balluffi (1995). These references could be consulted for more detailed insights, but are not necessary for the following arguments.

A full set of 3D bond-counting calculations has been given in two papers by MacKenzie *et al.* (1962); these papers include general rules for nearest neighbor and next nearest neighbor interactions in face-centered (f.c.c.) and body-centered (b.c.c.) cubic crystals, based on the number of broken bond vectors $\langle uvw \rangle$ which intersect the surface planes $\{hkl\}$. There is also an atlas of 'ball and stick' models by Nicholas (1965); an excellent introduction to crystallographic notation is given by Kelly & Groves (1970). More recently, models of the crystal faces can be visualized using CD-ROM or on the web, so there is little excuse for having to duplicate such pictures from scratch. A list of these resources, current as this book goes to press, is given in Appendix D.

The experimental study of small crystals (on substrates) is a specialist topic, aspects of which are described later in chapters 5, 7 and 8. For now, we note that close-packed

faces tend to be present in the equilibrium form. For f.c.c. (metal) crystals, these are $\{111\}$, $\{100\}$, $\{110\}$. . . and for b.c.c. $\{110\}$, $\{100\}$. . .; this is shown in γ -plots and equilibrium forms, calculated for specific first and second nearest neighbor interactions in figure 1.6, where the relative surface energies are plotted on a stereogram (Sundquist 1964, Martin & Doherty 1976). For really small particles the discussion needs to take the discrete size of the faces into account. This extends up to particles containing $\sim 10^6$ atoms, and favors $\{111\}$ faces in f.c.c. crystals still further (Marks 1985, 1994). The properties of stereograms are given in a student project which can be found via Appendix D.

The effect of temperature is interesting. *Singular* faces have low energy and low entropy; *vicinal* (stepped) faces have higher energy and entropy. Thus for increasing temperature, we have lower free energy for non-singular faces, and the equilibrium form is more rounded. Realistic finite temperature calculations are relatively recent (Rottman & Wortis 1984), and there is still quite a lot of uncertainty in this field, because the results depend sensitively on models of interatomic forces and lattice vibrations. Some of these issues are discussed in later chapters.

Several studies have been done on the anisotropy of surface energy, and on its variation with temperature. These experiments require low vapor pressure materials, and have used Pb, Sn and In, which melt at a relatively low temperature, by observing the profile of a small crystal, typically 3–5 μm diameter, in a specific orientation using scanning electron microscopy (SEM). An example is shown for Pb in figures 1.7 and 1.8, taken from the work of Heyraud and Métois; further examples, and a discussion of the role of roughening and melting transitions, are given by Pavlovskaya *et al.* (1989).

We notice that the anisotropy is quite small (much smaller than in the Kossel crystal calculation), and that it decreases, but not necessarily monotonically, as one approaches the melting point. This is due to three effects: (1) a nearest neighbor bond calculation with the realistic f.c.c. structure gives a smaller anisotropy than the Kossel crystal (see problem 1.1); (2) realistic interatomic forces may give still smaller effects; in particular, interatomic forces in many metals are less directional than implied by such bond-like models, as discussed in chapter 6; and (3) atomistic and layering effects at the monolayer level can affect the results in ways which are not intuitively obvious, such as the missing orientations close to (111) in the Pb crystals at 320°C, seen in figure 1.7(b). The main qualitative points about figure 1.8, however, are that the maximum surface energy is in an orientation close to $\{210\}$, as in the f.c.c. bond calculations of figure 1.6(b), and that entropy effects reduce the anisotropy as the melting point is approached. These data are still a challenge for models of metals, as discussed in chapter 6.

1.3 Thermodynamics versus kinetics

Equilibrium phenomena are described by thermodynamics, and on a microscopic scale by statistical mechanics. However, much of materials science is concerned with kinetics, where the rate of change of metastable structures (or their inability to change) is

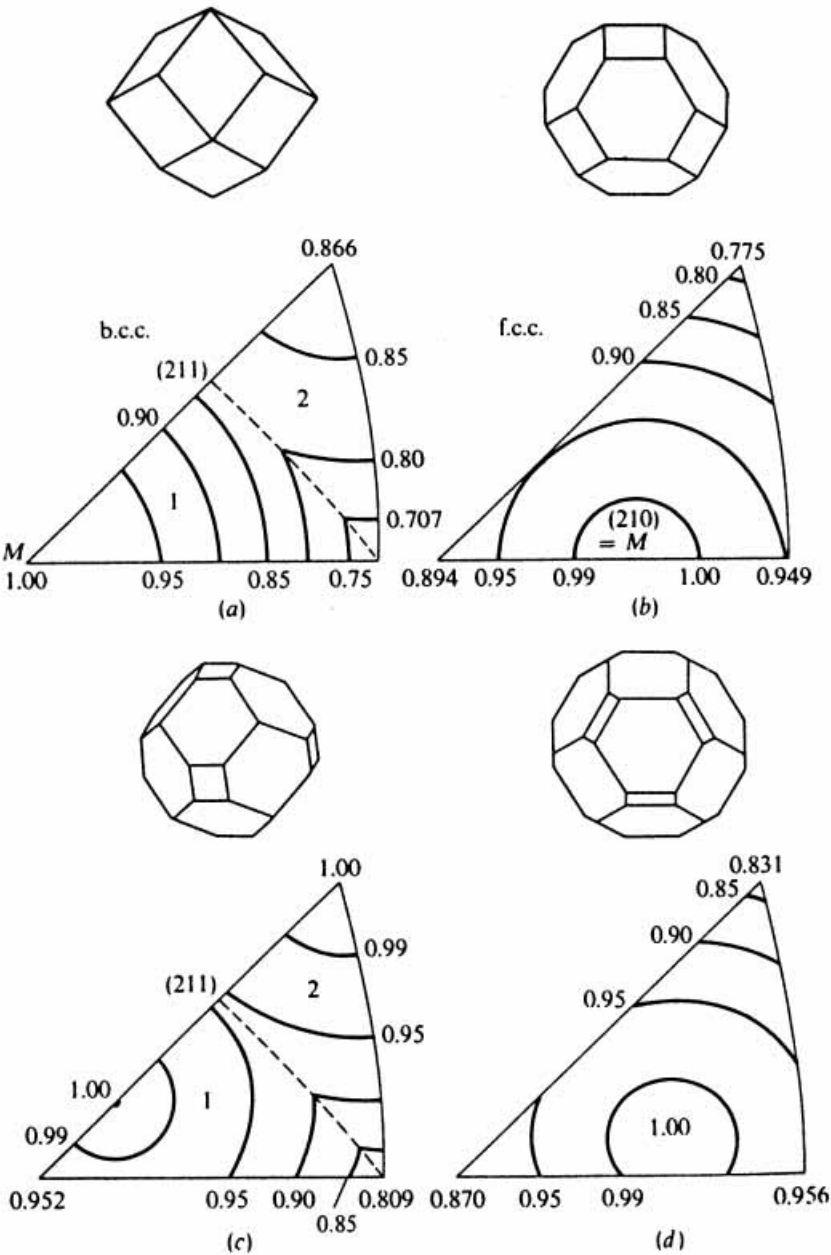


Figure 1.6. γ -Plots in a stereographic triangle (100, 110 and 111) and the corresponding equilibrium shapes for (a) b.c.c., (b) f.c.c., both with $\rho = 0$; (c) b.c.c. with $\rho = 0.5$, and (d) f.c.c. with $\rho = 0.1$; ρ is the relative energy of the second nearest bond to that of the nearest neighbor bond (from Sundquist 1964, via Martin & Doherty 1976, reproduced with permission).