vacancy concentration that is orders of magnitude greater than the equilibrium value. In Al–Cu alloys, for example, GP zones can form by ageing at room temperature, which would not be feasible without assistance from excess vacancies.

There is other evidence for the role of quenched-in vacancies in enhancing diffusion rates. If the alloy is quenched from different solution treatment temperatures and aged at the same temperature, the initial rate of zone formation is highest in the specimens quenched from the highest temperatures. Also, if the quench is interrupted at an intermediate temperature, so that a new equilibrium concentration can be established, the rate of transformation is reduced. Reducing the rate of cooling from the solution treatment temperature produces a similar effect by allowing more time for vacancies to be lost during the quench. This is important when large parts are to be heat treated as the cooling rate varies greatly from the surface to the centre when the specimen is water-quenched for example.

Apart from dislocations, the main sinks for excess vacancies are the grain boundaries and other interfaces within the specimen. Since vacancies have such a high diffusivity it is difficult to avoid losing vacancies in the vicinity of grain boundaries and interfaces. This has important effects on the distribution of precipitates that form in the vicinity of grain boundaries on subsequent ageing. Figure 5.35a shows the vacancy concentration profiles that should be produced by vacancy diffusion to grain boundaries during quenching. Close to the boundary the vacancy concentration will be the equilibrium value for the ageing temperature, while away from the boundary it will be that for the solution treatment temperature. On ageing these alloys it is found that a precipitate-free zone (PFZ) is formed as shown in Fig. 5.35b. The solute concentration within the zone is largely unchanged, but no nucleation has occurred. The reason for this is that a *critical vacancy supersaturation* must be exceeded for nucieation to occur. The width of the PFZ is determined by the vacancy concentration as shown in Fig. 5.35c. At low temperatures, where the driving force for precipitation is high, the critical vacancy supersaturation is lower and narrower PFZs are formed. High quench rates will also produce narrow PFZs by reducing the width of the vacancy concentration profile. Similar PFZs can also form at inclusions and dislocations.

Finally, it should be mentioned that another cause of PFZs can be the nucleation and growth of grain boundary precipitates during cooling from the solution treatment temperature. This causes solute to be drained from the surrounding matrix and a PFZ results. An example of this type of PFZ is shown in Fig. 5.36.

### 5.5.4 *Age Hardening*

The reason for the interest in alloy systems that show transition phase precipitation is that great improvements in the mechanical properties of these alloys can be achieved by suitable solution treatment and ageing operations.




Fig. 5.35 A PFZ due to vacancy diffusion to a grain boundary during quenching. (a) Vacancy concentration profile. (b) A PFZ in an Al–Ge alloy (× 20 000). (c) Dependence of PFZ width on critical vacancy concentration $X_v^c$ and rate of quenching. [(b) After G. Lorimer in *Precipitation in Solids*, K.C. Russell and H.I. Aaronson (Eds.). The Metallurgical Society of AIME, 1978.]

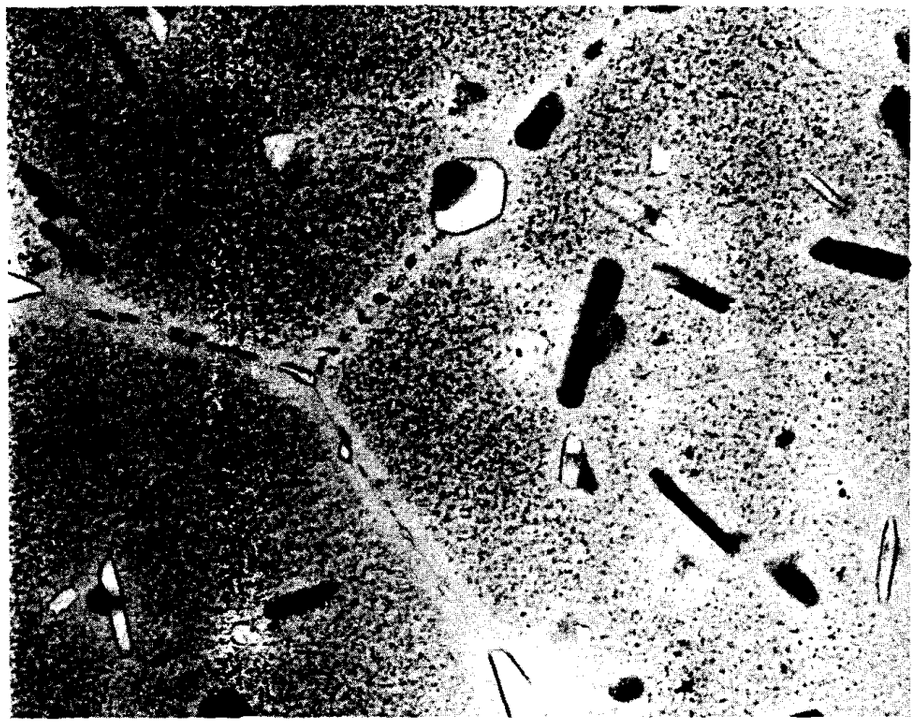

Fig. 5.36 PFZs around grain boundaries in a high-strength commercial Al–Zn–Mg–Cu alloy. Precipitates on grain boundaries have extracted solute from surrounding matrix. (×59 200)

This is illustrated for various Al–Cu alloys in Fig. 5.37. The alloys were solution treated in the single-phase α region of the phase diagram, quenched to room temperature and aged at either 130 °C (Fig. 5.37a) or 190 °C (Fig. 5.37b). The curves show how the hardness of the specimens varies as a function of time and the range of time over which GP zones, θ'' and θ' appear in the microstructure. Immediately after quenching the main resistance to dislocation movement is solid solution hardening. The specimen is relatively easily deformed at this stage and the hardness is low. As GP zones form the hardness increases due to the extra stress required to force dislocations through the coherent zones.

The hardness continues to increase with the formation of the coherent θ'' precipitates because now the dislocations must also be forced through the highly strained matrix that results from the misfit perpendicular to the θ'' plates (see Fig. 5.30b). Eventually, with the formation of θ' the spacing between the precipitates becomes so large that the dislocations are able to bow between the precipitates and the hardness begins to decrease. Maximum hardness is associated with a combination of θ'' and θ'. Further ageing increases the distance between the precipitates making dislocation bowing easier and the hardness decreases. Specimens aged beyond peak hardness are referred to as *overaged*.



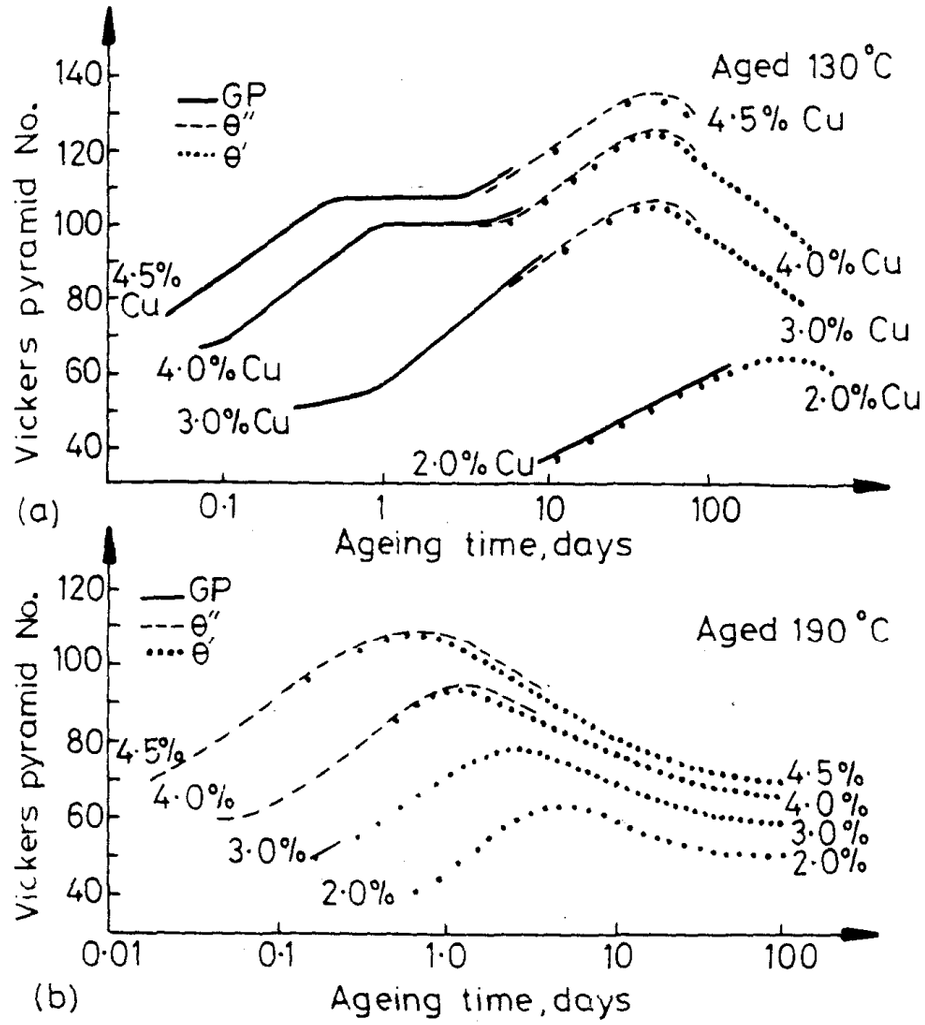


Fig. 5.37 Hardness v. time for various Al–Cu alloys at (a) 130 °C (b) 190 °C. (After J.M. Silcock, T.J. Heal and H.K. Hardy, *Journal of the Institute of Metals* **82** (1953–1954) 239.

If Al–4.5 wt% Cu is aged at 190 °C, GP zones are unstable and the first precipitate to form is θ''. The volume fraction of θ'' increases with time causing the hardness to increase as shown in Fig. 5.37b. However, at 190 °C the θ'' nucleates under the influence of a smaller driving force than at 130 °C and the resultant precipitate dispersion is therefore coarser. Also the maximum volume fraction of θ'' is reduced. Both of these factors contribute to a lower peak hardness on ageing at the higher temperature (compare Fig. 5.37a and b). However, diffusion rates are faster at higher temperatures and peak hardness is therefore achieved after shorter ageing times.

It can be seen that at 130 °C peak hardness in the Al–4.5 wt% Cu alloy is not reached for several tens of days. The temperatures that can be used in the

heat treatment of commercial alloys are limited by economic considerations to those which produce the desired properties within a reasonable period of time, usually up to ~24 h. In some high-strength alloys use is, therefore, made of a double ageing treatment whereby ageing is carried out in two steps: first at a relatively low temperature below the GP zone solvus, and then at a higher temperature. In this way a fine dispersion of GP zones obtained during the first stage can act as heterogeneous nucleation sites for precipitation at the higher temperature. This type of treatment can lead to a finer precipitate distribution than would be obtained from a single ageing treatment at the higher temperature.

Another treatment used commercially is to give the alloy a controlled deformation either before a single-stage age or between the two stages of a double-ageing treatment. The strength of the alloy after this treatment can be increased by a higher precipitate density, resulting from a higher nucleation rate, and by the retained dislocation networks which also act as a barrier to further deformation. However, deformation prior to ageing does not always result in an improvement in properties. In some cases deformation can lead to a coarser precipitate distribution.

Precipitation hardening is common to many alloy systems. Some of the more important systems are listed in Table 5.2. Some commercial alloys are listed in Table 5.3, along with their mechanical properties. In many of these systems it is possible to come very close to the maximum theoretical strength of the matrix, i.e. about $\sim\mu/30$. However, engineering alloys are not heat treated for maximum strength alone. Consideration must also be given to toughness, stress corrosion resistance, fatigue, etc., when deciding on the best heat treatment in practice.

### 5.5.5 *Spinodal Decomposition*

It was mentioned at the beginning of this chapter that there are certain transformations where there is no barrier to nucleation. One of these is the spinodal mode of transformation. Consider a phase diagram with a miscibility gap as shown in Fig. 5.38a. If an alloy with composition $X_0$ is solution treated at a high temperature $T_1$ and then quenched to a lower temperature $T_2$ the composition will initially be the same everywhere and its free energy will be $G_0$ on the $G$ curve in Fig. 5.38b. However, the alloy will be immediately *unstable* because small fluctuations in composition that produce A-rich and B-rich regions will cause the total free energy to decrease. Therefore 'up-hill' diffusion takes place as shown in Fig. 5.39 until the equilibrium compositions $X_1$ and $X_2$ are reached.

The above process can occur for any alloy composition where the free energy curve has a negative curvature, i.e.

$$\frac{d^2G}{dX^2} < 0 \qquad (5.42)$$



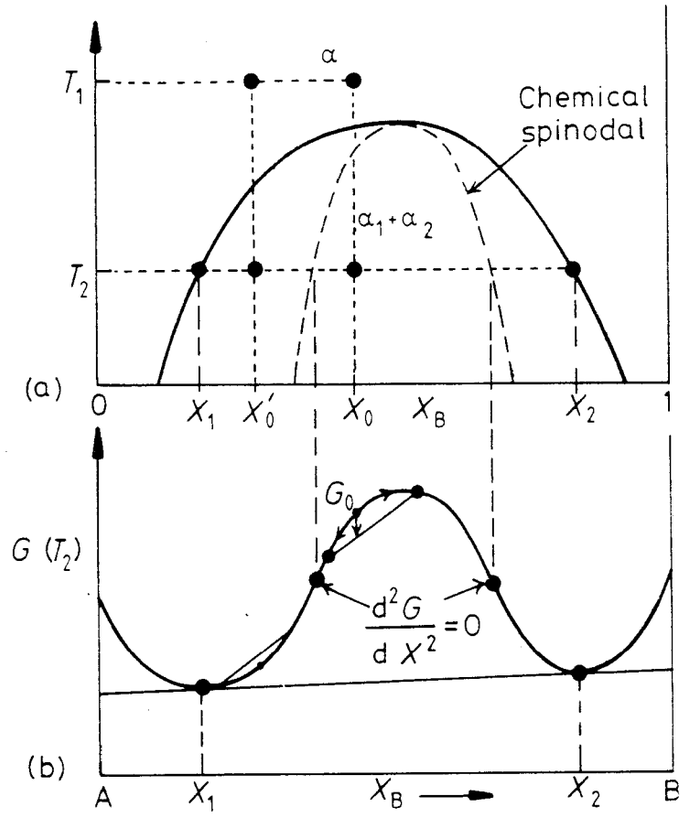


Fig. 5.38 Alloys between the spinodal points are unstable and can decompose into two coherent phases $\alpha_1$ and $\alpha_2$ without overcoming an activation energy barrier. Alloys between the coherent miscibility gaps and the spinodal are metastable and can decompose only after nucleation of the other phase.

Therefore the alloy must lie between the two points of inflection on the free energy curve. The locus of the points on the phase diagram, Fig. 5.32a, is known as the chemical spinodal.

If the alloy lies outside the spinodal, small variations in composition lead to an increase in free energy and the alloy is therefore *metastable*. The free energy of the system can only be decreased in this case if nuclei are formed with a composition very different from the matrix. Therefore, outside the spinodal the transformation must proceed by a process of nucleation and growth. Normal down-hill diffusion occurs in this case as shown in Fig. 5.40.

The rate of spinodal transformation is controlled by the interdiffusion coefficient, $D$. Within the spinodal $D < 0$ and the composition fluctuations shown in Fig. 5.39 will therefore increase exponentially with time, with a

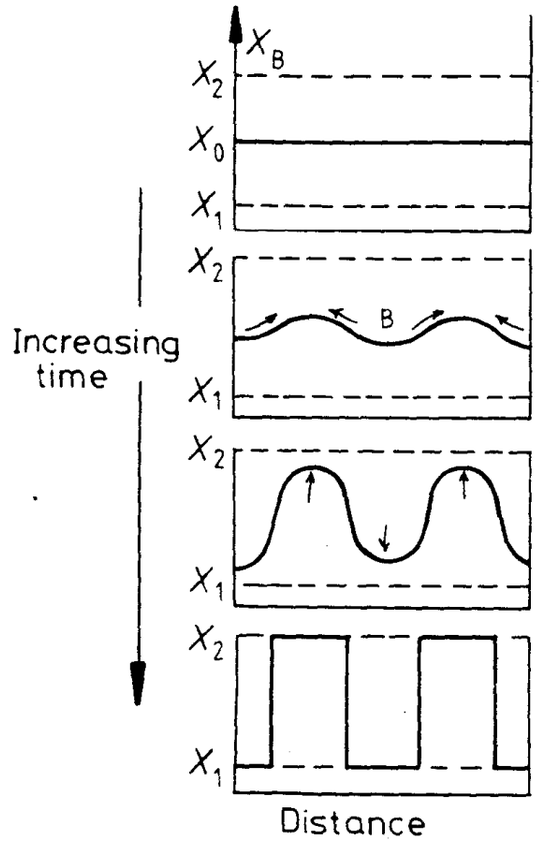


Fig. 5.39 Schematic composition profiles at increasing times in an alloy quenched into the spinodal region ($X_0$ in Fig. 5.38).

characteristic time constant $\tau = -\lambda^2/4\pi^2 D$, where $\lambda$ is the wavelength of the composition modulations (assumed one-dimensional). The rate of transformation can therefore become very high by making $\lambda$ as small as possible. However, as will be shown below, there is a minimum value of $\lambda$ below which spinodal decomposition cannot occur.

In order to be able to calculate the wavelength of the composition fluctuations that develop in practice it is necessary to consider two important factors that have been omitted from the above discussion: (1) interfacial energy effects, and (2) coherency strain energy effects.

If a homogeneous alloy of composition $X_0$ decomposes into two parts one with composition $X_0 + \Delta X$ and the other with composition $X_0 - \Delta X$, it can be shown that[7] the total chemical free energy will change by an amount $\Delta G_c$ given by

$$\Delta G_c = \frac{1}{2}\frac{d^2G}{dX^2}(\Delta X)^2 \qquad (5.43)$$

If, however, the two regions are finely dispersed and coherent with each



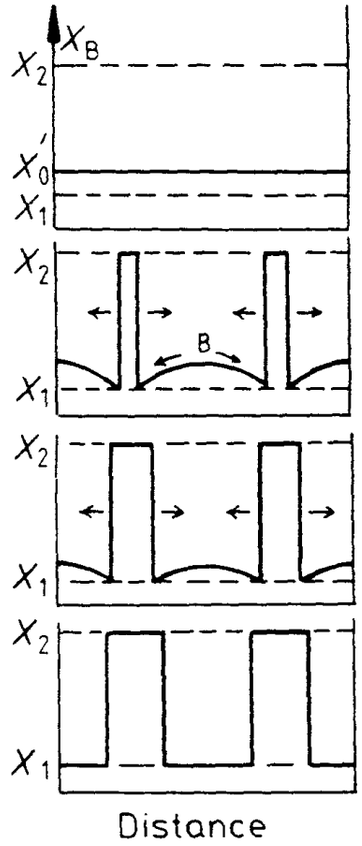


Fig. 5.40 Schematic composition profiles at increasing times in an alloy outside the spinodal points ($X_0'$ in Fig. 5.38).

other there will be an additional energy change due to interfacial energy effects. Although, during the early stages of spinodal decomposition, the interface between A-rich and B-rich regions is not sharp but very diffuse, there is still an effective interfacial energy contribution. The magnitude of this energy depends on the composition gradient across the interface, and for this reason it is known as a 'gradient energy'. In solid solutions which tend to cluster the energy of like atom-pairs is less than that of unlike pairs. Thus the origin of the gradient energy is the increased number of unlike nearest neighbours in a solution containing composition gradients compared to a homogeneous solution. For a sinusoidal composition modulation of wavelength $\lambda$ and amplitude $\Delta X$ the maximum composition gradient is proportional to $(\Delta X/\lambda)$ and the gradient energy term $\Delta G_\gamma$ is given by

$$\Delta G_\gamma = K\left(\frac{\Delta X}{\lambda}\right)^2 \qquad (5.44)$$

where $K$ is a proportionality constant dependent on the difference in the bond energies of like and unlike atom pairs.

If the sizes of the atoms making up the solid solution are different, the generation of composition differences will introduce a coherency strain energy term, $\Delta G_s$. If the misfit between the A-rich and B-rich regions is $\delta$, $\Delta G_s \propto E\delta^2$ where $E$ is Young's modulus. For a total composition difference $\Delta X$, $\delta$ will be given by $(da/dX)\Delta X/a$, where $a$ is the lattice parameter. An exact treatment of the elastic strain energy shows that

$$\Delta G_s = \eta^2(\Delta X)^2 E' V_m \tag{5.45}$$

where

$$\eta = \frac{1}{a}\left(\frac{da}{dX}\right) \tag{5.46}$$

i.e. $\eta$ is the fractional change in lattice parameter per unit composition change. $E' = E/(1-\upsilon)$, where $\upsilon$ is Poisson's ratio, and $V_m$ is the molar volume. Note that $\Delta G_s$ is independent of $\lambda$.

If all of the above contributions to the total free energy change accompanying the formation of a composition fluctuation are summed we have

$$\Delta G = \left\{\frac{d^2G}{dX^2} + \frac{2K}{\lambda^2} + 2\eta^2 E' V_m\right\}\frac{(\Delta X)^2}{2} \tag{5.47}$$

It can be seen therefore that the condition for a homogeneous solid solution to be unstable and decompose spinodally is that

$$-\frac{d^2G}{dX^2} > \frac{2K}{\lambda^2} + 2\eta^2 E' V_m \tag{5.48}$$

Thus the limits of temperature and composition within which spinodal decomposition is possible are given by the conditions $\lambda = \infty$ and

$$\frac{d^2G}{dX^2} = -2\eta^2 E' V_m \tag{5.49}$$

The line in the phase diagram defined by this condition is known as the *coherent spinodal* and it lies entirely within the chemical spinodal ($d^2G/dX^2 = 0$) as shown in Fig. 5.41. It can be seen from Equation 5.48 that the wavelength of the composition modulations that can develop inside the coherent spinodal must satisfy the condition

$$\lambda^2 > -2K\Big/\left(\frac{d^2G}{dX^2} + 2\eta^2 E' V_m\right) \tag{5.50}$$

Thus the minimum possible wavelength decreases with increasing undercooling below the coherent spinodal.

Figure 5.41 also shows the coherent miscibility gap. This is the line defining the equilibrium compositions of the coherent phases that result from spinodal decomposition ($X_1$ and $X_2$ in Fig. 5.39). The miscibility gap that normally



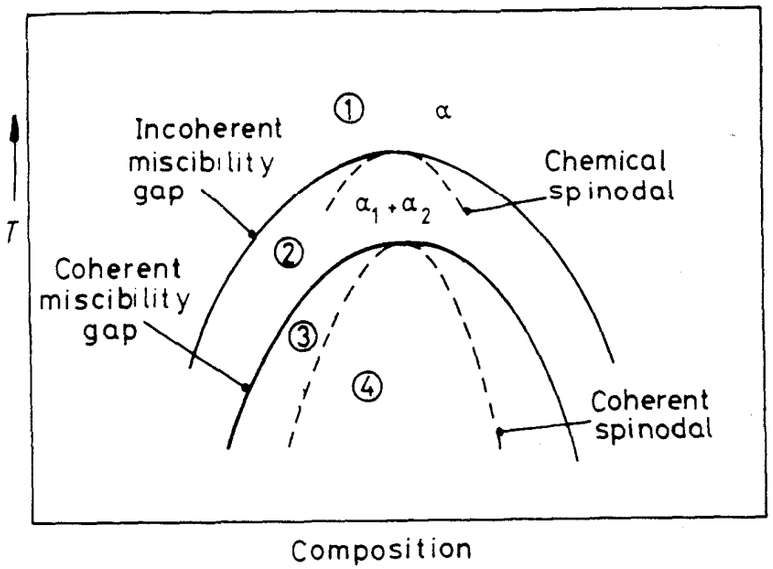


Fig. 5.41 Schematic phase diagram for a clustering system. Region 1: homogeneous α stable. Region 2: homogeneous α metastable, only incoherent phases can nucleate. Region 3: homogeneous α metastable, coherent phases can nucleate. Region 4: homogeneous α *unstable* , no nucleation barrier, spinodal decomposition occurs.

appears on an equilibrium phase diagram is the incoherent (or equilibrium) miscibility gap. This corresponds to the equilibrium compositions of *incoherent* phases, i.e. in the absence of strain fields. The chemical spinodal is also shown in Fig. 5.41 for comparison, but it is of no practical importance.

Spinodal decomposition is not only limited to systems containing a stable miscibility gap. All systems in which GP zones form, for example, contain a *metastable* coherent miscibility gap, i.e. the GP zone solvus (see the Al–Ag system in Fig. 5.34 for example). Thus it is possible that at high supersaturations GP zones are able to form by the spinodal mechanism. If ageing is carried out below the coherent solvus but outside the spinodal, GP zones can only form by a process of nucleation and growth, Fig. 5.40. Between the incoherent and coherent miscibility gap, Fig. 5.41, $\Delta G_v - \Delta G_s < 0$ and only incoherent strain-free nuclei can form.

The difference in temperature between the coherent and incoherent miscibility gaps, or the chemical and coherent spinodals in Fig. 5.41, is dependent on the magnitude of $|\eta|$. When there is a large atomic size difference $|\eta|$ is large and a large undercooling is required to overcome the strain energy effects. As discussed earlier large values of $|\eta|$ in cubic metals can be mitigated if the misfit strains are accommodated in the elastically soft ⟨100⟩ directions. This is achieved by the composition modulations building up parallel to {100}.

Figure 5.42 shows a spinodal structure in a specimen of Al–22.5 Zn–0.1 Mg (atomic %) solution treated at 400 °C and aged 20 h at 100 °C. The wavelength in the structure is 25 nm, but this is greater than the initial microstructure due to coarsening which occurs on holding long times at high temperatures.

### 5.5.6 *Particle Coarsening*[8]

The microstructure of a two-phase alloy is always unstable if the total interfacial free energy is not a minimum. Therefore a high density of small precipitates will tend to *coarsen* into a lower density of larger particles with a smaller total interfacial area. However, such coarsening often produces an undesirable degradation of properties such as a loss of strength or the disappearance of grain-boundary pinning effects (see Section 3.3.5). As with grain growth, the rate of coarsening increases with temperature and is of particular concern in the design of materials for high temperature applications.

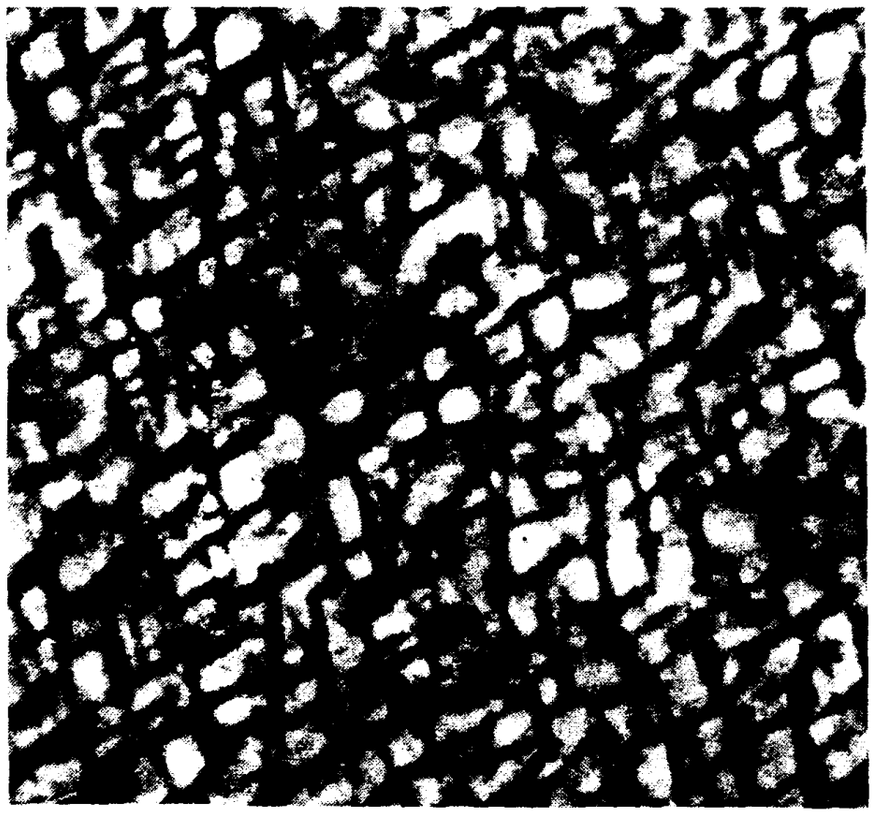

Fig. 5.42 A coarsened spinodal microstructure in Al–22.5 at% Zn–0.1 at% Mg solution treated 2 h at 400 °C and aged 20 h at 100 °C. Thin foil electron micrograph (× 314 000). (After K.B. Rundman, *Metals Handbook*, 8th edn., Vol. 8, American Society for Metals, 1973, p. 184.)



In any precipitation-hardened specimen there will be a range of particle sizes due to differences in the time of nucleation and rate of growth. Consider two adjacent spherical precipitates[9] with different diameters as shown in Fig. 5.43. Due to the Gibbs–Thomson effect, the solute concentration in the matrix adjacent to a particle will increase as the radius of curvature decreases, Fig. 5.43b. Therefore there will be concentration gradients in the matrix which will cause solute to diffuse in the direction of the largest particles away from the smallest, so that the small particles shrink and disappear while large particles grow. The overall result is that the total number of particles decreases and the mean radius ($\bar{r}$) increases with time. By assuming volume diffusion is the rate controlling factor it has been shown[10] that the following relationship should be obeyed:

$$(\bar{r})^3 - r_0^3 = kt \tag{5.51}$$

where

$$k \propto D\gamma X_e$$

$r_0$ is the mean radius at time $t = 0$, $D$ is the diffusion coefficient, $\gamma$ is the interfacial energy and $X_e$ is the equilibrium solubility of very large particles. Since $D$ and $X_e$ increase exponentially with temperature, the rate of coarsening will increase rapidly with increasing temperature. Fig. 5.44. Note that the rate of coarsening

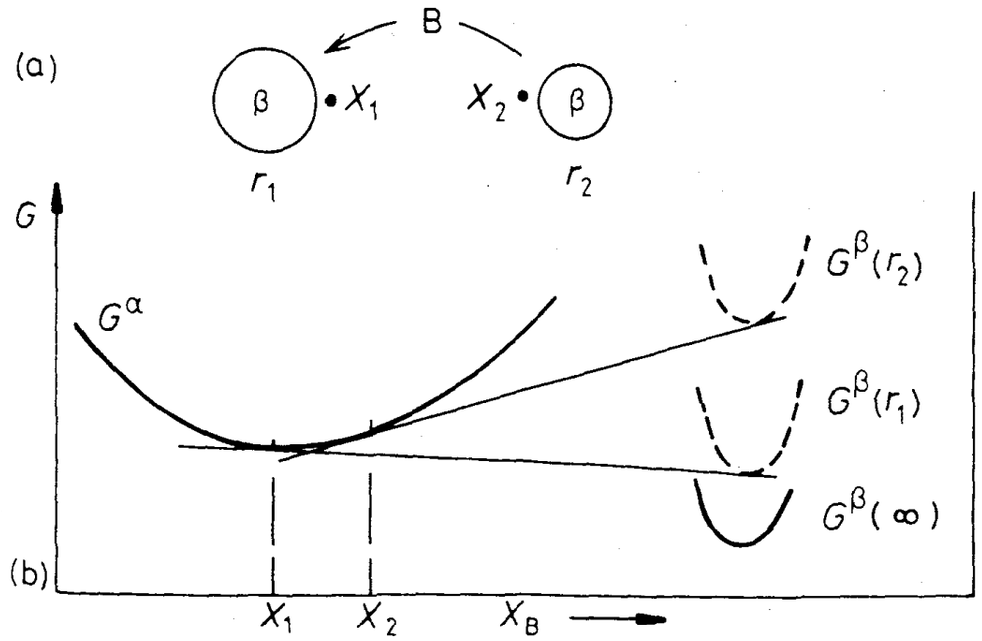


Fig. 5.43 The origin of particle coarsening. β with a small radius of curvature ($r_2$) has a higher molar free energy than β with a large radius of curvature ($r_1$). The concentration of solute is therefore highest outside the smallest particles.