Title
PHASE TRANSFORMATIONS, CRYSTALLOGRAPHIC ASPECTS

Author
Dahmen, U.

Publication Date
1986-06-01

LBL-21423
Reprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Materials & Chemical Sciences Division

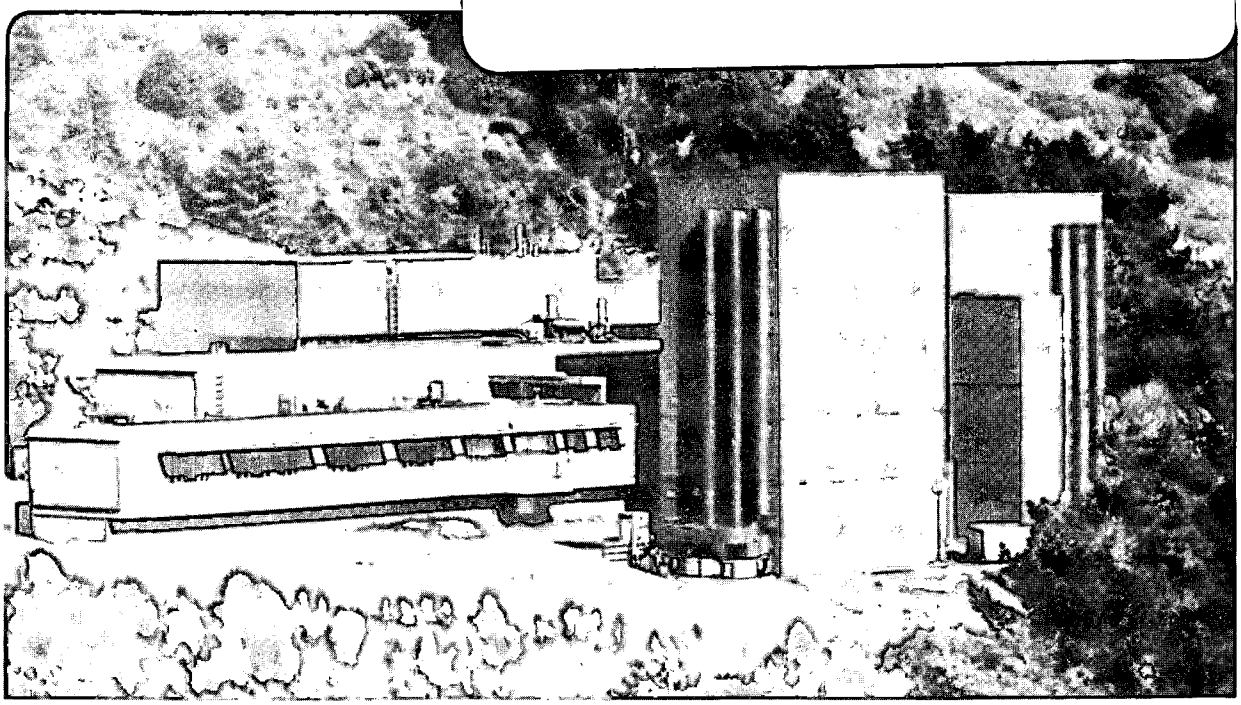
National Center for Electron Microscopy

Published as a chapter in **Encyclopedia of Physical Science and Technology**, Vol. 10, Academic Press, San Diego, CA

Phase Transformations, Crystallographic Aspects

U. Dahmen

June 1986



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

FROM: Encyclopedia of Physical Science and Technology
Academic Press (1988)

PHASE TRANSFORMATIONS, CRYSTALLOGRAPHIC ASPECTS

U. Dahmen *Lawrence Berkeley Laboratory*

GLOSSARY

Anisotropy: Directional dependence of properties or behavior.

Coincidence site lattice (CSL): Lattice formed by the translations common to two misoriented grains.

Domains: Symmetry-related crystalline regions. As a result of a symmetry-breaking transition (such as ordering) a single crystal transforms to an assembly of domains. Any two domains are related by one of the broken symmetries.

Habit plane: Major plane of contact between two phases.

Homogeneous: When referring to elastic inclusions: having the same elastic constants as the matrix; when referring to nucleation: random nucleation without the aid of defects. When referring to strain: everywhere the same, as in a linear transformation; opposite of heterogeneous.

Invariant plane (line) strain: Form of strain that leaves a plane (line) unstretched and unrotated.

Motif: Group of atoms associated with each lattice point in a crystal lattice; also called basis.

Orientation relationship: Relative orientation of two crystals, specified by pairs of parallel planes and directions, by a rotation tensor, or by the angle and axis of rotation.

Parent/product: Phase before/after transformation. Also referred to as matrix/precipitate, matrix/inclusion, high-temperature phase/low-temperature phase, group/subgroup, austenite/martensite, disordered phase/ordered phase.

Space group: Set of symmetry operations that leaves a crystal structure invariant.

Special point: Point of high symmetry lying at intersection of symmetry elements in a lattice.

Strain accommodation: Process of elastic or plastic deformation around an inclusion allowing a change in shape, orientation, or volume during a transformation.

Variant: One of a set of crystallographically equivalent inclusions; for example [1 0 0], [0 1 0], and [0 0 1] needles are variants of the same type of ⟨1 0 0⟩ needle precipitates. The term is also used to describe the four (usually nonequivalent) solutions of the martensite problem.

A phase is a structurally and chemically homogeneous volume of material. A phase transformation (sometimes called transition) entails a change in either structure or composition, or both. The structural aspects of phase transformations in crystalline solids include the diffusionless martensitic transformation which is a purely structural change and proceeds by the athermal movement of a glissile interface. However, structure is also important in transformations involving short-range diffusion such as order–disorder reactions, polymorphic changes, or recrystallization and grain growth. Even in those transformations that require long-range diffusion such as precipitation, eutectoid, or discontinuous reactions, the structure of the inter-

face between parent and product phase has important implications for the morphology or the growth mechanism of the new phase.

Two types of structural change may be distinguished, a purely crystallographic change due, for example, to a rearrangement of atoms within a rigid lattice, and a dimensional change due to a distortion of the lattice. The former type is described by the concepts of pure crystallography and group theory and forms the basis of the theory of group–subgroup transitions. The latter type focuses on the change in lattice dimensions, largely ignoring the atomic arrangement in the lattice; it forms the basis of the theory of elastic inclusions and of martensite theory. Most transformations induce both types of structural change, crystallographic and dimensional, but usually one or the other dominates.

A general homogeneous distortion can be written as the distortion tensor **A**, which may be separated into a pure distortion **D** and a rigid body rotation **R**:

$$\mathbf{A} = \mathbf{RD}$$

The pure distortion may be decomposed further into a pure shear **S** and a pure volume expansion **V**:

$$\mathbf{A} = \mathbf{RSV}$$

This decomposition is useful because it separates the changes in orientation (**R**), shape (**S**), and volume (**V**). Grain boundaries are described by a change in orientation, martensite transformations are dominated by the shape change, and precipitation reactions often have large orientation, shape, and volume changes, but each component has a different effect on the final morphology.

I. Crystallography Principles

In most solids the atoms are arranged in a periodic network such that their mutual coordination is optimized. The basic repeat unit of this network or crystal lattice is the unit cell. The atomic structure remains invariant under translation through any vector **t** that is the sum of integral multiples of three basic vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. The set of all such translation vectors **t** forms the translational group T. When a mathematical point in space is repeated through the vectors of the translation group the Bravais lattice is obtained (Fig. 1a). For certain angles and ratios between the basis vectors additional symmetries such as reflection, inversion, or rotation

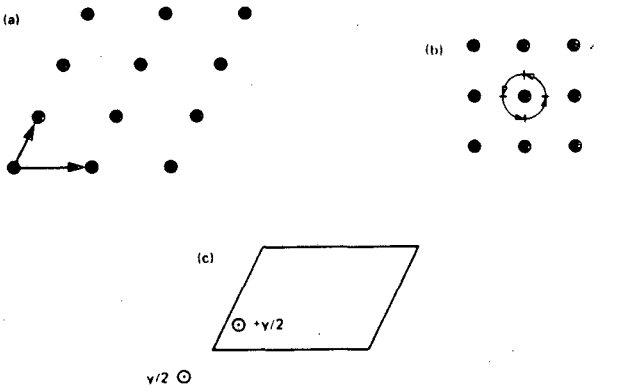


FIG. 1. Illustration of three basic symmetry elements: translation symmetry leads to the Bravais lattice (a), rotation symmetry exists for special angles (here 90°) between basis vectors of the Bravias lattice (b), and screw axis symmetry results from special distance (here $y = \frac{1}{2}$) between atoms in the motif (c).

arise, for example, the 4-fold rotation axis in Fig. 1b. The set of all such point symmetry elements **R** that leave a given Bravais lattice invariant is called its point group G. [*See* CRYSTALLOGRAPHY.]

If all lattice points of the Bravais lattice are occupied by identical atoms we have a simple (primitive) crystal with all the translation and point symmetry elements of the Bravais lattice. Not all crystals are this simple, however, and most solids contain more than one atom per lattice site. The group of atoms that occupies a lattice point is called the basis or motif. The position of atoms in the basis and the arrangement of identical bases in the Bravais lattice are dictated by bonding requirements. For certain bond angles and distances in the motif two new symmetry elements arise (screw axes and glide mirrors) that are a combination of a point symmetry element **R** with a fractional translation τ in the unit cell (e.g., $\frac{1}{2}$ 0 0). An example of a motif giving rise to either a 2-fold axis or a 2-fold screw axis, or no symmetry, depending on the bond distance y is shown in Fig. 1c. A Bravais lattice is projected along its 2-fold b axis. The two identical atoms in the motif are a vertical distance y apart. If y is irrational the crystal has no rotational symmetry, if $y = 0$ it has a 2-fold rotation axis and if $y = \frac{1}{2}$ it has a 2-fold screw axis, that is, an identical position to any point in the cell is reached by rotation through 180° followed by a translation of $\tau = \frac{1}{2}$ along the rotation axis. Applying this operation twice is the same as a simple translation along the axis. The set of all such symmetry elements $(\mathbf{R}|\tau)$ (including sim-

ple point symmetry elements for which $\tau = 0$) forms the space group.

Thus translation group T, point group G, and space group S are characterized by their specific symmetry operations **t**, **R**, and $(\mathbf{R}|\tau)$. Only a relatively small number of different crystallographic translation, point, and space groups exists, the 14 Bravais lattices, the 32 point groups, and the 230 space groups. Any crystal belongs to one of the 230 space groups. Each space group has a characteristic point group and Bravais lattice. Since the translations τ in the space group elements are on the order of unit cell dimensions, they are considered microscopic symmetry elements, in contrast to point symmetry operations that concern directions only and therefore represent the macroscopic symmetry of a crystal. A complete listing of crystallographic symmetry groups is compiled in the International Tables for Crystallography. A brief account of the concepts and notations necessary to represent and understand symmetry groups will be given below.

A. Translation Group

Many of the physical properties of crystals result from their invariance under translations **t** of the translation group. It is this periodicity that allows investigation of crystal structures and phase transformations by diffraction techniques. Some aspects of phase transformations are best described in terms of Fourier series as the most natural way to express periodic functions. Crystallography makes use of the fact that in order to describe a crystal structure it is sufficient to specify the dimensions and content of a single unit cell. The strain tensor of a transformation from one crystal structure to another is derived from the correspondence between unit cells. These and other aspects of translational symmetry are utilized in the theoretical and experimental investigation of phase transformations.

B. Point Group

For certain dimensions of the unit cell a Bravais lattice is invariant under rotation as well as translation. The set of rotations, reflection, or inversion operations **R** that leave the lattice points of such unit cells invariant is called the point group. Point symmetry in general characterizes the set of equivalent directions, hence it describes the macroscopic symmetry of an object. For example, many flowers have 5-fold rotational symmetry, a cylinder has infinite rotational symmetry ($\mathbf{R} = \infty$) about its axis, and a sphere has infinite rotational symmetry about any axis. However, only 1-, 2-, 3-, 4-, and 6-fold rotational symmetries are compatible with translation symmetry and therefore only those rotations occur in the 32 crystallographic point groups. To illustrate this consider the row of atoms in Fig. 2, which is a section of a larger space lattice with translation period **t**. If the whole assembly is rotated about a lattice point by $\pm\alpha$ degrees the new lattice points must again conform to translational symmetry, that is, they must be an integral multiple N of horizontal translations **t** apart. Thus

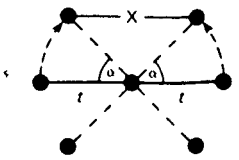


FIG. 2. Construction illustrating the rotation angles α that are consistent with translation symmetry **t**.

$$\mathbf{X} = N\mathbf{t} = 2\mathbf{t}\cos\alpha, \quad \text{or } \cos\alpha = N/2$$

The five possible rotation symmetries and their notation are listed in Table I. Each type of rotation axis (or rotor) is characteristic for a particular crystal system. In addition there are two more crystal systems that arise from combinations of these elements: the orthorhombic systems with two diads at right angles and the cubic system with three tetrads and four triads. Each of these seven crystal system has a characteristic space lattice. The angles and edge lengths of the unit cell of each space lattice must be compatible with the symmetry of the crystal system. The unit cells of the seven space lattices are shown in Fig. 3. It can be seen that each cell has more symmetry elements than just translation and the unique n-fold rotor. For example the monoclinic unit cell which has a diad through each lattice point automatically has other diads going through the face and edge centers [see Fig.

TABLE I. The Five Rotation Angles Compatible with Translation

N	α	n	Notation	Crystal system
2	360°	1	Monad	Triclinic
−2	180°	2	Diad	Monoclinic
−1	120°	3	Triad	Rhombohedral
0	90°	4	Tetrad	Tetragonal
1	60°	6	Hexad	Hexagonal

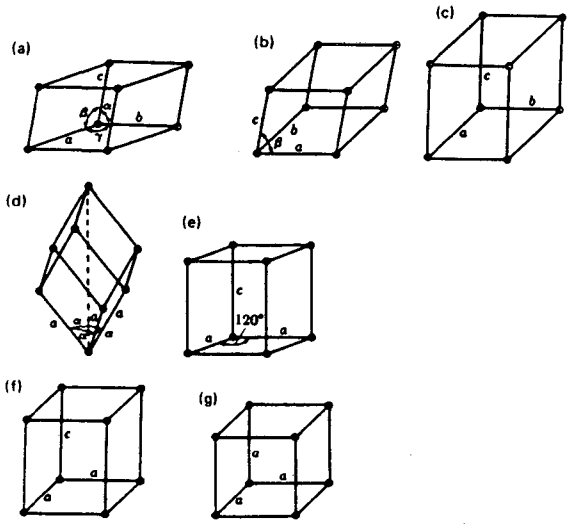


FIG. 3. Unit cells of the seven primitive space lattices: triclinic (a), monoclinic (b), orthorhombic (c), rhombohedral (d), hexagonal (e), tetragonal (f), and cubic (g).

4a). In addition to mirror planes on the faces normal to the unique axis it also has a mirror halfway between them. Any of these symmetry elements, when applied to the monoclinic unit cell, will reproduce its lattice points in identical positions, that is, leave the lattice invariant. These symmetry elements are shown in two projections (Fig. 4b and c) using diad symbols and mirror notation to indicate their locations. A complete list of standard symbols used for rotation axes, mirror planes, centering translations, screw axes, and glide mirrors is given in Table II.

C. Group Properties

The properties of the point group are easily visualized using the point symmetry of the monoclinic crystal system, characterized by a diad. The lowest symmetry monoclinic point group contains only two elements, the identity (1) and the diad (2). This set is written as $G = \{1, 2\}$, in shorthand $G = 2$, and it represents a group since (1) it contains the identity operation (1), (2) it is closed, that is, the product of any two successive operations is equivalent to a single operation which is itself part of the group (e.g., 2 followed by 2 is the identity 1), and (3) each element $\mathbf{R}$ has an inverse $\mathbf{R}^{-1}$ such that $\mathbf{R}\mathbf{R}^{-1} = \mathbf{R}^{-1}\mathbf{R} = 1$.

There are two other monoclinic point groups, $G = m$ and $G = 2/m$, again characterized by the diad but containing additional symmetry elements. The holohedral (highest symmetry) monoclinic group contains four elements: $G = \{1, \bar{1}, 2, m\}$. In shorthand this is referred to as

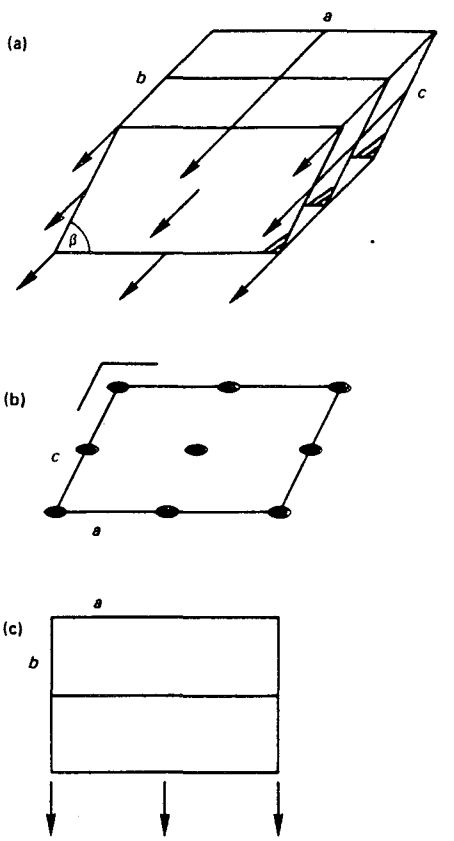


FIG. 4. Monoclinic space lattice with location of diads and mirrors: perspective view (a), projection along b axis (b), and perpendicular to b axis (c) in notation used in International Tables for Crystallography (see Table II).

$G = 2/m$ with the oblique line indicating that the diad and the mirror refer to the same axis. The operation of such a group is best illustrated in a multiplication table (Table III) which lists the products (successive application) of any two elements. It is clear from this table that any two successive operations are equivalent to a single operation which is itself an element of the group. In our example, each element happens to be its own inverse, that is, yields the identity when applied twice in succession. This need not always be the case, for example, a tetrad needs to be applied four times to produce the identity. This is known as the order of the symmetry element. Thus a hexad is of order six and a mirror of order two. The inverse of a 60° rotation (6^1) is a 300° rotation (6^5).

D. Stereographic Representation

A convenient graphic representation of the point group symmetry is the stereographic projection. Consider a general direction, indicated by a pole in the stereogram (Fig. 5a) in a crystal with $2/m$ monoclinic symmetry, that is, with point group $G = \{1, \bar{1}, 2, m\}$. Operation on this

TABLE II. Crystallographic Symmetry Elements and Their Notation

Type of symmetry element	Written symbol	Graphical symbol
Center of symmetry	$\bar{1}$	o
		Perpendicular to paper / In plane of paper
Mirror plane	m	——
Glide planes	$a\ b\ c$	-----
		Glide in plane of paper / Arrow shows glide direction
		..... Glide out of plane of paper
	n	-·-·-·-
Rotation	2	●
	3	▲
	4	◆
	6	⬢
Screw axes	2_1	●
	$3_1, 3_2$	▲ ▲
	$4_1, 4_2, 4_3$	◆ ◆ ◆
	$6_1, 6_2, 6_3, 6_4, 6_5$	⬢ ⬢ ⬢ ⬢ ⬢
Inversion axes	$\bar{3}$	△
	$\bar{4}$	◈
	$\bar{6}$	⬡

crystal by all four symmetry operations will take this pole to the four positions shown in Fig. 5b (open and closed circles represent poles in different hemispheres). The number of crystallographically equivalent poles in a stereogram is equal to the number of symmetry elements in its points group. This number is known as the order of the group. Stereograms characterizing all 32 crystallographic point groups are shown in Fig. 6. Note that they are subdivided into the seven crystal systems and that each crystal system has more than one possible point group. The holohedral group describes the symmetry of the space lattice of that crystal system. Thus the monoclinic space lattice has $2/m$ symmetry.

TABLE III. Multiplication Table for $2/m$ Monoclinic Point Group

	1	$\bar{1}$	2	m
1	1	$\bar{1}$	2	m
$\bar{1}$	$\bar{1}$	1	m	2
2	2	m	1	1
m	m	2	$\bar{1}$	1

Only when the lattice points are occupied by atoms or groups of atoms does the space lattice become a crystal. And only when the group of atoms at each lattice site, the motif, has lower symmetry than $2/m$ does the crystal have a lower than holosymmetric point group. An example is shown in Fig. 7 where the motif is made up of two dissimilar atoms aligned with the diad. This motif is not invariant under inversion or reflection perpendicular to the rotation axis. The crystal therefore cannot have this symmetry either. The point symmetry of this crystal is the set of elements common to the lattice G_0 =

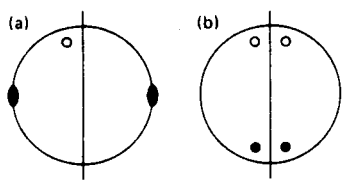


FIG. 5. Stereographic representation of point symmetry. The general point in (a) is taken to all its equivalent positions by the symmetry elements of the point group $2/m$ (b). Open circles are in upper hemisphere, solid dots in lower hemisphere.

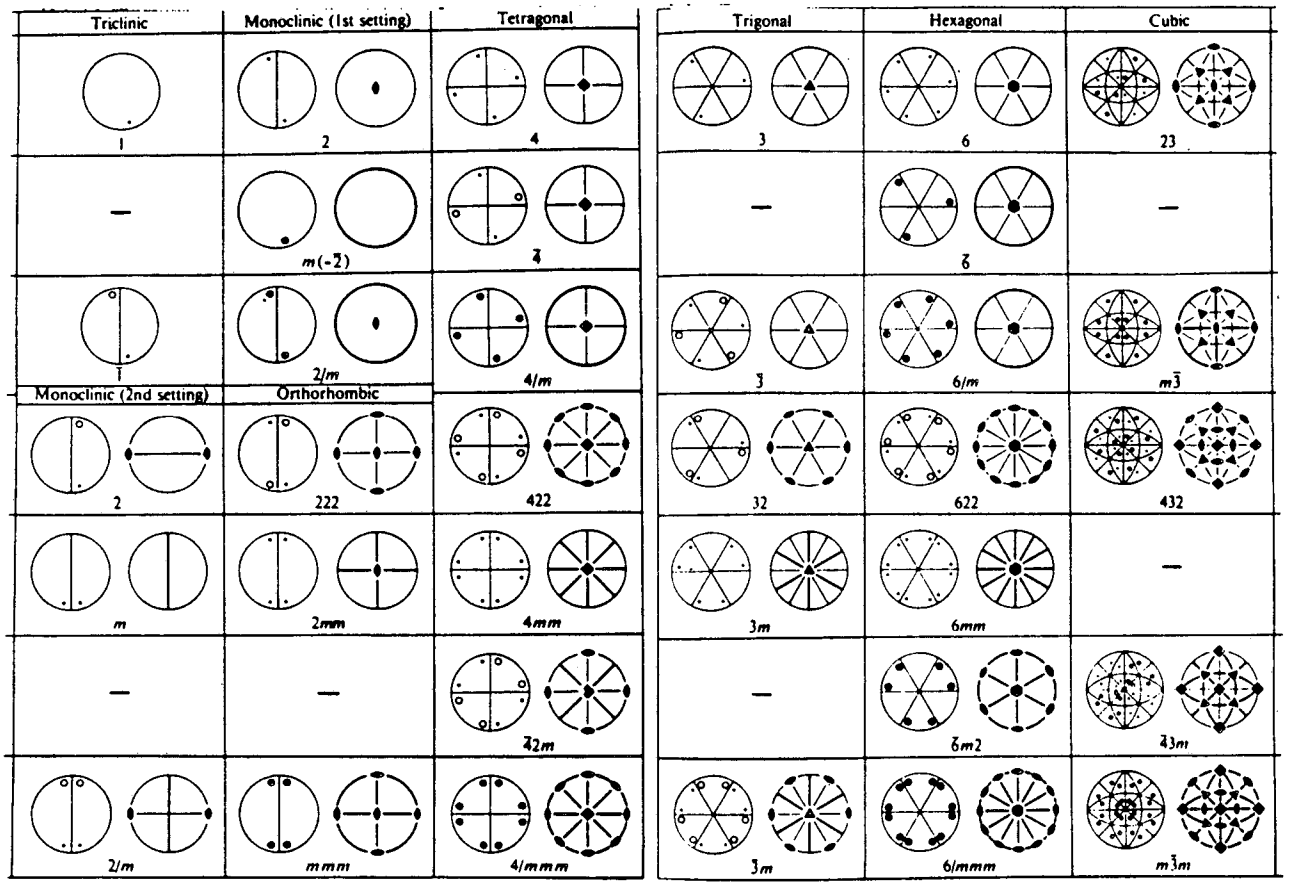


FIG. 6. Stereographic representation of the 32 crystallographic point groups. The order of a group is the multiplicity of a general point.

$\{1, \bar{1}, 2, m\}$ and the motif $G_1 = \{1, \infty, m\}$, also known as the intersection group $H = G_0 \cap G_1 = \{1, 2\}$ which is the lowest symmetry monoclinic group: 2. Note that the mirrors m of the lattice and the motif are not parallel and are therefore not part of the intersection group H. This is an important point: the orientation of the symmetry elements is essential in forming the intersection group. For example two cubic point groups share only the identity $\{1\}$ or inversion $\{1, \bar{1}\}$ if their symmetry axes are not aligned. The point symmetry of a crystal depends therefore on the space lattice, that is, the dimensions and angles of the unit cell, and on the motif, that is, the atomic arrangement about each lattice point.

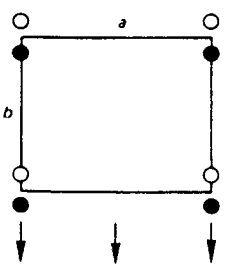


FIG. 7. Projection of monoclinic unit cell as in Fig. 4c illustrating absence of mirror due to two dissimilar atoms in the motif.

E. Tensor Representation

The point symmetry elements, 2, m, $\bar{1}$ etc. can be represented more explicitly by transformation tensors. Consecutive application of two symmetry operations is equivalent to matrix multiplication of the corresponding tensors. As an example a reflection in a plane perpendicular to the z axis of an orthogonal coordinate system is simply written as the tensor

$$m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \bar{1} \end{pmatrix}$$

Such symmetry elements $\mathbf{R}$ represent proper (det $\mathbf{R} = 1$) or improper (det $\mathbf{R} = -1$) rotations. Improper rotations change the handedness of a crystal. They are unitary operators, hence $\mathbf{R}^T = \mathbf{R}^{-1}$, (the transpose $\mathbf{R}^T$ is equal to the inverse $\mathbf{R}^{-1}$). In addition, as shown before, only certain

rotations are compatible with translational symmetry. Some simple point symmetry elements are listed in Table IV. Under a symmetry operation **R**, vectors **r** and tensors **M** transform as

$$\mathbf{r}' = \mathbf{Rr} \quad \text{and} \quad \mathbf{M}' = \mathbf{RMR}^{-1}$$

F. Special Points

The intersections of symmetry elements, for example the line where two mirror planes meet, are special positions in a lattice. When symmetry elements intersect in a point this becomes a special point. As shown later the special points of a lattice are of fundamental importance in the theory of phase transitions. The lattice points at the corners of the monoclinic cell shown in Fig. 4 are special points since they lie at the intersection of a diad with the mirror plane. The coordinates of these special points are characterized by the coordinates of the origin, (0 0 0), since the others can be derived from it by simple lattice translations **t**. However, there are other special points, for example at $(\frac{1}{2}\ 0\ 0)$, $(\frac{1}{2}\ \frac{1}{2}\ 0)$, and $(\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2})$, that do not coincide with lattice points of the primitive Bravais lattice. It is possible to place lattice points on special points without altering the point symmetry of the lattice. A lattice point at $(\frac{1}{2}\ \frac{1}{2}\ 0)$, for example, will change the primitive (symbol P) monoclinic Bravais lattice into a base-centered one (symbol C) without changing the number or location of point symmetry operations. The only change is an additional translation symmetry, the $(\frac{1}{2}\ \frac{1}{2}\ 0)$ base centering translation. Centering translations are denoted A, B, C for base, F for face, and I for body centering. If all such centering translations are added to the 7 primitive space lattices, a total of 14 different Bravais lattices is formed.

TABLE IV. Tensors **R** for Some Point Symmetry Elements

$1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Identity
$\bar{1} = \begin{pmatrix} \bar{1} & 0 & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & \bar{1} \end{pmatrix}$	Inversion
$m_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \bar{1} \end{pmatrix}$	Mirror perpendicular to z axis
$2_z = \begin{pmatrix} \bar{1} & 0 & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Diad along z axis
$3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$	Triad along cube body diagonal
$4 = \begin{pmatrix} 0 & 1 & 0 \\ \bar{1} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Tetrad along z axis

G. Space Group

As we have seen, the point group is concerned with macroscopic symmetry, that is, the equivalence of directions in an object. Directions are not changed by translations, although translations do limit the number of crystallographic point groups to 32 through the requirement of compatibility with translational symmetry. The space group is concerned with microscopic symmetry, including translations as well as directions. If a motif of lower symmetry is placed on a lattice of higher symmetry the space group is the set of operations that leaves this pattern (or crystal) invariant. In a unit cell of the monoclinic lattice, outlined in Fig. 8, it is apparent that the motif or group of atoms at the center of the cell cannot be obtained by a centering translation, but by reflection across the dashed mirror line followed by a fractional ($\tau = \frac{1}{2}$) translation parallel to the line. This is the operation of a glide mirror. Alternatively the same arrangement can be obtained by the 2-fold screw axes indicated by half arrows. However, this is no longer the case if the symmetry of the motif is lowered further by differentiating between its two atoms. The crystal structure then either has a screw

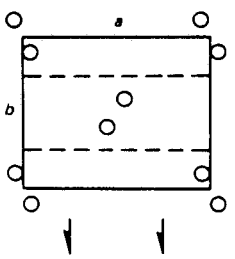


FIG. 8. Projection of monoclinic unit cell perpendicular to b axis with four-atom motif illustrating a-glide mirror perpendicular to b axis (dashed lines) and screw axis parallel to b axis (half arrows).

axis or a glide mirror but not both. By taking into account these small translations τ important in microscopic symmetry, the 32 point groups are further differentiated into 230 space groups.

In a manner similar to point symmetry elements, space group operations such as 2_1, n, 4_3 etc. (see Table II) may be represented by an operator $(\mathbf{R}/\tau)$ that denotes a point symmetry operation $\mathbf{R}$ followed by a translation τ. For example a 2-fold screw axis along the z direction would be

$$2_1 = \left(\begin{array}{ccc|c} \bar{1} & 0 & 0 & 0 \\ 0 & \bar{1} & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \end{array}\right)$$

that is, a 2-fold rotation followed by a translation of $\frac{1}{2}$ along the rotation axis. Some transformation and multiplication rules for such tensors are

$$(\mathbf{R}|\tau)\mathbf{r} = \mathbf{R}\mathbf{r} + \tau$$

$$(\mathbf{R}|\tau)(\mathbf{Q}|\varepsilon) = \mathbf{R}\mathbf{Q} + \mathbf{R}\varepsilon + \tau$$

$$(\mathbf{R}|\tau)^{-1} = (\mathbf{R}^{-1}|-\mathbf{R}^{-1}\tau)$$

H. Notation

The International (Hermann–Mauguin) notation used to describe space groups consists of two parts: (1) a letter (A, B, C, F, I, P) indicating the centering type of the unit cell, and (2) a set of characters giving symmetry elements along one, two, or three principal symmetry directions in the crystal. Table II lists the space group symbols along with their graphic representation. For example, the space group symbol $P2_1/c$ represents the primitive monoclinic crystal structure in Fig. 8 with the characters 2_1 (2-fold screw axis) and c (c-glide mirror) referring to the same principal symmetry direction in the crystal. The two symbols are therefore separated by an oblique line. On the other hand, the symbol $Im\bar{3}m$, representing a body-centered cubic structure, has three characters referring to the cube edge, cube diagonal, and face diagonal, respectively. By convention the first character after the letter indicating the centering type describes the characteristic symmetry direction in each crystal system, for example, [0 0 1] in the tetragonal and hexagonal systems and [1 1 1] in the rhombohedral system.

To summarize: crystals are a periodic arrangement of atoms on a space lattice that has translational and rotational symmetries. Only five rotational symmetries are compatible with translation, the 1-, 2-, 3-, 4-, and 6-fold axes. When combined in space, these give rise to 7 crystal systems, triclinic, monoclinic, orthorhombic, tetragonal, rhombohedral, hexagonal, and cubic, each with a characteristic shape of its unit cell as prescribed by symmetry and shown in the 7 primitive Bravais lattices. Without being occupied by atoms, these 7 simple space lattices have a number of point symmetry operations (rotation, inversion, mirror) in addition to the ones required for the crystal system. The set of these operations is called the point group of the system. Each of these seven holosymmetric point groups has a limited number of subgroups with lower symmetry but still compatible with the shape of its characteristic unit cell. This leads to a total of 32 possible point groups.

In addition to the primitive lattice translations, centering translations within the unit cell such as A, B, C, F, I are possible without changing the point group symmetry. When these centering translations are combined with the 7 primitive space lattices, the 14 different Bravais lattices result. Translations within the unit cell, which are either pure or combined with mirror or rotation symmetry, are registered in the space group. Due to these microscopic symmetries the 32 point groups are further differentiated into the 230 space groups.

Translation groups are characterized by symmetry operations $\mathbf{t}$, point groups by $\mathbf{R}$, and space groups by $(\mathbf{R}|\tau)$.

I. Symmetry Principles

The physical properties of crystals depend on their crystal structure and are often anisotropic, that is, different for different directions of the crystal. Any macroscopic physical property of a crystal must have at least the point symmetry of the crystal itself. Stated differently, this is known as Neumann's principle: "The symmetry elements of any physical property of a crystal must include the symmetry elements of its point group." Thus some physical property, for example, thermal expansion, of a crystal can be isotropic while the crystal itself is not. However, the converse is not true: if a physical property is anisotropic, the crystal structure must have the same anisotropy. For example, the spontaneous polarization in ferroelectric crystals has polar symmetry (∞m). According to Neumann's principle, the crystal point group must share the same polar symmetry, that is, it must be a sub-

group of (∞m). This allows only point groups with a unique rotor and any mirrors parallel to it. Of the 32 crystallographic point groups only 10 (1, 2, 3, 4, 6, m, $2mm$, $3m$, $4m$, $6mm$) are compatible with ferroelectricity.

As a result of Neumann's principle a proper description of the physical properties of a crystal must be invariant under the operation of its symmetry group. Physical properties can be described by matter tensors that give the response of the crystal to an external stimulus such as temperature or a stress field. The type and rank of the tensor depend on the property it describes. For example the tensor of thermal expansion is of rank two. It describes the strain field of a crystal when the stimulus temperature is applied. Other matter tensors of order two relate a vector stimulus to a vector response, for example, the tensors of electrical or thermal conductivity, diffusivity, permittivity, etc. Others relate a pair of vectors to a third as in the third-order tensor of the Hall constants. The elastic behavior of crystals is described by fourth-order tensors relating two second-order tensors, the stress field, and the strain field. Any of these matter tensors must remain invariant under all the symmetry operations $\mathbf{R}$ of the point group of the crystal. A tensor of rank two transforms as $\mathbf{M}' = \mathbf{RMR}^{-1}$. Since a matter tensor $\mathbf{M}$ must be invariant under the symmetry operations $\mathbf{R}$ of the point group, $\mathbf{M}' = \mathbf{M}$ for all symmetry operations $\mathbf{R}$ of the crystal. Using the rules of matrix multiplication, it can be shown, for example, that any second-order matter tensor in a tetragonal system must have the form

$$\mathbf{M} = \begin{pmatrix} M_1 & & \\ & M_1 & \\ & & M_3 \end{pmatrix}$$

and for a cubic system must be isotropic

$$\mathbf{M} = \begin{pmatrix} M_1 & & \\ & M_1 & \\ & & M_1 \end{pmatrix}$$

This means that diffusion, thermal expansion, or any physical property described by a second-order tensor is isotropic in cubic crystals. These results can be derived directly from Neumann's principle. A more general statement of this principle was given earlier by Curie: "When definite causes produce definite effects the elements of symmetry of the causes should be apparent in the effects." In this form the principle is applicable not just to crystallography but to any physical phenomena. If the cause is the crystal structure and the effect the physical properties we recover Neumann's principle. If, in the case of a phase transformation, the cause is a decrease in temperature and the effect is an arrangement of ordered domains due to the loss of symmetry in the phase transition, then the symmetry of the cause (isotropic) should be apparent in the symmetry of the effect, that is, all possible domains should appear at random. On the other hand, if the cause has a lower symmetry, such as a temperature gradient, uniaxial stress, or magnetic field then this symmetry should be apparent in the distribution of domains formed under the influence of this field. In order to predict and detect such effects we must answer the question of how many different domains can form in a given phase transition.

II. Modulated Structures

A. Characteristics

A large class of phase transformations can be described by a modulation of some physical quantity (e.g., composition, magnetization, displacement) associated with a crystal. In such transformations, the crystal lattice is modified with the periodicity of a static plane wave. For example, a sinusoidal composition modulation with wave length λ can be written as the (static) concentration wave

$$n(\mathbf{r}) = c + Q \sin 2\pi\mathbf{kr} \qquad (1)$$

where $\mathbf{k}$ is the wave vector ($|\mathbf{k}| = 1/\lambda$), and $n(\mathbf{r})$ the probability that the lattice site at position $\mathbf{r}$ is occupied by an atom whose concentration in the alloy is c. The amplitude Q of the modulation is proportional to the order parameter η which varies between zero for the disordered structure and unity for the fully ordered structure. An example is seen in Fig. 9 which illustrates the disordered and ordered state of a binary alloy with a simple cubic lattice. The ordered structure is tetragonal with the c axis parallel to the wave vector $\mathbf{k}$. When applied to an fcc lattice this describes the type of ordering found in equiatomic Cu–Au alloys. The wavelength $\lambda = na$ is an integer multiple n of the lattice periodicity a (in this particular case $\lambda = a$) and the wave is commensurate with the crystal lattice. The correspond-

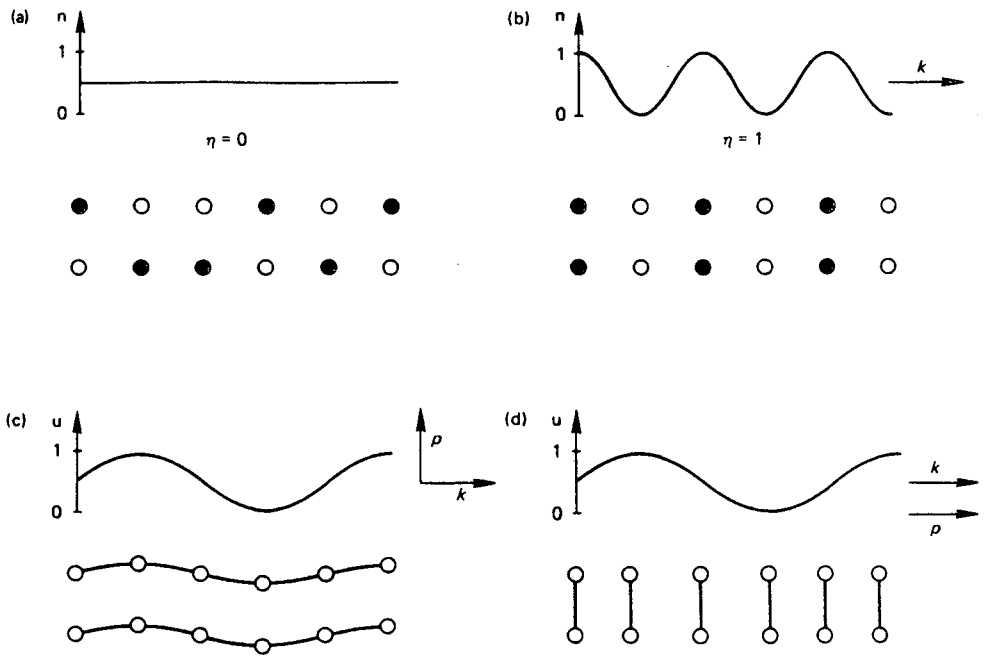


FIG. 9. Illustration of modulated structures: disordered solid solution (a), composition modulation (b), transverse displacement modulation (c), and longitudinal displacement modulation (d).

ing wave vector $\mathbf{k} = \langle 1\ 0\ 0 \rangle$ is on the Brillouin zone boundary, that is, the wavelength λ is on the order of the lattice parameter, a characteristic of an ordering modulation. A concentration wave with wave vector near the origin ($\lambda \gg a$) describes a clustering modulation. Continuous phase transitions that proceed by the gradual amplification of certain composition fluctuations are called spinodal ordering ($\lambda \simeq a$) and spinodal clustering or spinodal decomposition ($\lambda \gg a$).

When the wavelength is not an integral multiple of a lattice spacing the modulation is incommensurate. In the direction of the wave vector, the crystal then loses its true translational symmetry. Most ordering modulations are commensurate. Although incommensurate structures exist incommensurations are often found to be a mixture of two commensurate phases. If the quantity that is modulated is not a scalar such as concentration but a vector such as displacement or magnetization, the wave description of the position $\mathbf{u}(\mathbf{r})$ of an atom originally at $\mathbf{r}$ becomes

$$\mathbf{u}(\mathbf{r}) = \mathbf{r} + \mathbf{p} \sin 2\pi\mathbf{kr} \qquad (2)$$

where $\mathbf{p}$ is the polarization vector and $\mathbf{k}$ the wave vector. Since $\mathbf{p}$ is a vector quantity we now have to distinguish between transverse ($\mathbf{p} \perp \mathbf{k}$) and longitudinal ($\mathbf{p} \parallel \mathbf{k}$) modulations, illustrated in Fig. 9c and d. Short wavelength ($\lambda \simeq a$) commensurate displacement waves describe shuffles whereas long wavelength ($\lambda \gg a$) modulations describe lattice distortions such as the premartensitic "tweed" effect. A simple example of a short wavelength longitudinal displacement wave is the ω-transformation found in Ti and Zr base alloys. It can be described as the local collapse of every other pair of $\{1\ 1\ 1\}$ planes of the bcc lattice, or a wave with $\mathbf{k} = \frac{2}{3}\langle 1\ 1\ 1 \rangle$ and $\mathbf{p} = \frac{1}{6}\langle 1\ 1\ 1 \rangle$.

Due to their periodic nature modulated structures give rise to diffraction peaks. Long wavelength modulations lead to satellite reflections near the main Bragg peaks while their short wavelength counterparts cause extra reflections at rational positions between Bragg peaks of the unmodulated structure. As an example, the ordered structure of Fig. 9b would give rise to extra reflections at the positions of the wave vector ($\mathbf{k} = \langle \frac{1}{2}\ 0\ 0 \rangle$) halfway between the Bragg peaks of the disordered structure. The wave vector $\mathbf{k}$ can thus be read directly from a diffraction pattern as the position of the extra reflections in the first Brillouin zone (superlattice reflections). This is shown in Fig. 10a for a short wavelength composition modulation (ordering). The equivalent for a long wavelength (clustering) modulation is shown in Fig. 10b where satellites are seen around each Bragg peak as well as the origin.

In a displacement modulation not all satellite reflections are allowed due to the directional nature of the polarization $\mathbf{p}$. Satellite reflections are forbidden near all Bragg peaks $\mathbf{g}$ for which

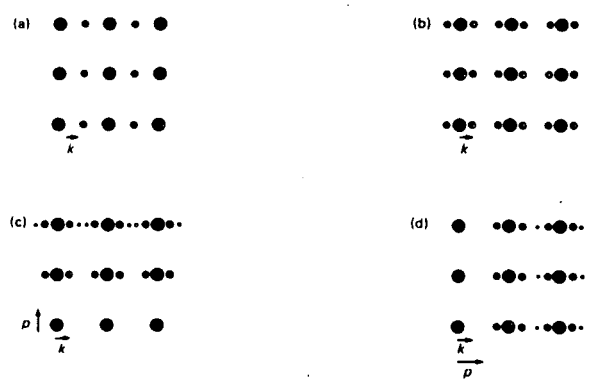


FIG. 10. Schematic diffraction patterns from modulated structures: short wavelength composition modulation (ordering) (a), long wavelength composition modulation (clustering) (b), and long wavelength transverse (c) and longitudinal (d) displacement modulation.

$\mathbf{gp} = 0$. Long wavelength modulated transverse and longitudinal displacements would thus give rise to the diffraction patterns shown in Fig. 10c and d. The short wavelength equivalent can lead to either extra reflections or extinctions depending on the magnitude and direction $\mathbf{p}$ of the modulation, that is, the nature of the shuffle.

The structures considered so far were simple modulations by a single plane wave. More complex structures are obtained from combinations of several waves. For composition modulations this is expressed as a sum of concentration waves

$$n(\mathbf{r}) = c + \sum_j Q_j \exp(2\pi i \mathbf{k}_j \mathbf{r}) + c.c. \quad (3)$$

where Q_j is the amplitude of the jth wave with wave vector $\mathbf{k}_j$ defined in the first Brillouin zone, $c.c.$ is the complex conjugate, and $n(\mathbf{r})$ is the probability of site $\mathbf{r}$ to be occupied by a particular species whose concentration in the alloy is c.

Due to crystal symmetry some of the wave vectors $\mathbf{k}_j$ are crystallographically equivalent, for example [0 0 1] and [1 0 0] in cubic systems. Such vectors are said to belong to the same star. The star of a wave vector $\mathbf{k}$ is the set of vectors obtained by applying all the symmetry elements of the point group of the crystal. If the summation is rewritten as the sum over different stars s it becomes

$$n(\mathbf{r}) = c + \sum_s \eta_s \sum_j \gamma_s(j) \exp(2\pi i \mathbf{k}_j \mathbf{r}) + c.c. \quad (4)$$

Here the summation j is carried out over the equivalent vectors in a star and the summation s runs over the different stars. The $\gamma_s(j)$ are coefficients determining the relative contribution of each of the equivalent waves in a star and the η_s are the long-range order parameters.

B. Special Point Ordering

Whether or not an ordered structure is thermodynamically favorable depends on the specific interaction between the different atoms. Generally the wave vector $\mathbf{k}$ of a stable ordered phase will change continuously with the form of this interaction. However, there are some special wave vectors, determined by the crystal symmetry alone, whose positions are independent of atomic interactions. It can be shown that at such special points any function with the periodicity of the lattice must have a symmetry-dictated extremum, that is a minimum, saddle point, or maximum. Not surprisingly, ordered structures with special point wave vectors are often found to be the most stable phases over a wide range of composition and temperature. Special point wave vectors are the reciprocal space equivalent of special points in a unit cell (see Section I,F) and are thus easily determined. For each disordered structure it is therefore possible to enumerate all possible special point ordered structures simply by summing over all compatible combinations or special wave vectors. This leads to an elegant derivation of the most commonly found ordered structures. Examples for phases based on the fcc lattice (special points are the stars of $\langle 0\ 0\ 0\rangle$, $\langle 1\ 0\ 0\rangle$, $\langle \frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}\rangle$, $\langle 1\ \frac{1}{2}\ 0\rangle$) and the bcc lattice (special points are the stars of $\langle 0\ 0\ 0\rangle$, $\langle 1\ 1\ 1\rangle$, $\langle \frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}\rangle$, $\langle \frac{1}{2}\ \frac{1}{2}\ 0\rangle$) together with some alloys forming such phases and the operating special point wave vectors are listed in Table V. Some of these structures are shown in Fig. 11. Note that the method applies to substitutional and interstitial phases alike since it is based entirely on symmetry and not on the kind of atomic interaction.

C. First- and Second-Order Transformations

First-order transformations are accompanied by drastic changes in macroscopic properties of the material and the coexistence of two phases at the transformation temperature. Typical examples of such transitions are melting, evaporation, precipitation, or polymorphic changes. In second-order transformations the thermodynamic properties of a material change continuously but their second derivatives such as spe-

TABLE V. Wave Vectors of Some Common Special-Point Ordered Structures

k	Substitutional	Interstitial
fcc based		
[0 0 1]	CuAuI	$(Fe,Ni)_2N$
[1 0 0], [0 1 0], [0 0 1]	Cu_3Au	Fe_4N
[1/2 1/2 1/2]	CuPt	
[0 0 0], [1/2 1/2 1/2], [1/2 1/2 $\bar{1}$/2]	$CuPt_3$	
[0 0 1], [1 0 1/2]	Al_3Ti	Ni_4NII
[1 0 0], [0 1 0], [0 0 1], [$\bar{1}$/2 1/2 1/2], [1/2 $\bar{1}$/2 1/2] [1/2 1/2 $\bar{1}$/2], [1/2 1/2 1/2]	$CuPt_7$	Fe_8N
bcc based		
[1 1 1]	CuZn	
[1/2 1/2 0]		Ta_2O
[1/2 1/2 1/2]	NaTl	
[1 1 1], [1/2 1/2 1/2]	Fe_3Al	
[1 1 1], [1/2 1/2 0], [1/2 $\bar{1}$/2 0]		Ta_4O
[1 1 1], [1/2 1/2 0], [0 1/2 $\bar{1}$/2], [1/2 0 1/2], [1/2 0 $\bar{1}$/2], [1/2 $\bar{1}$/2 0]		Fe_8N

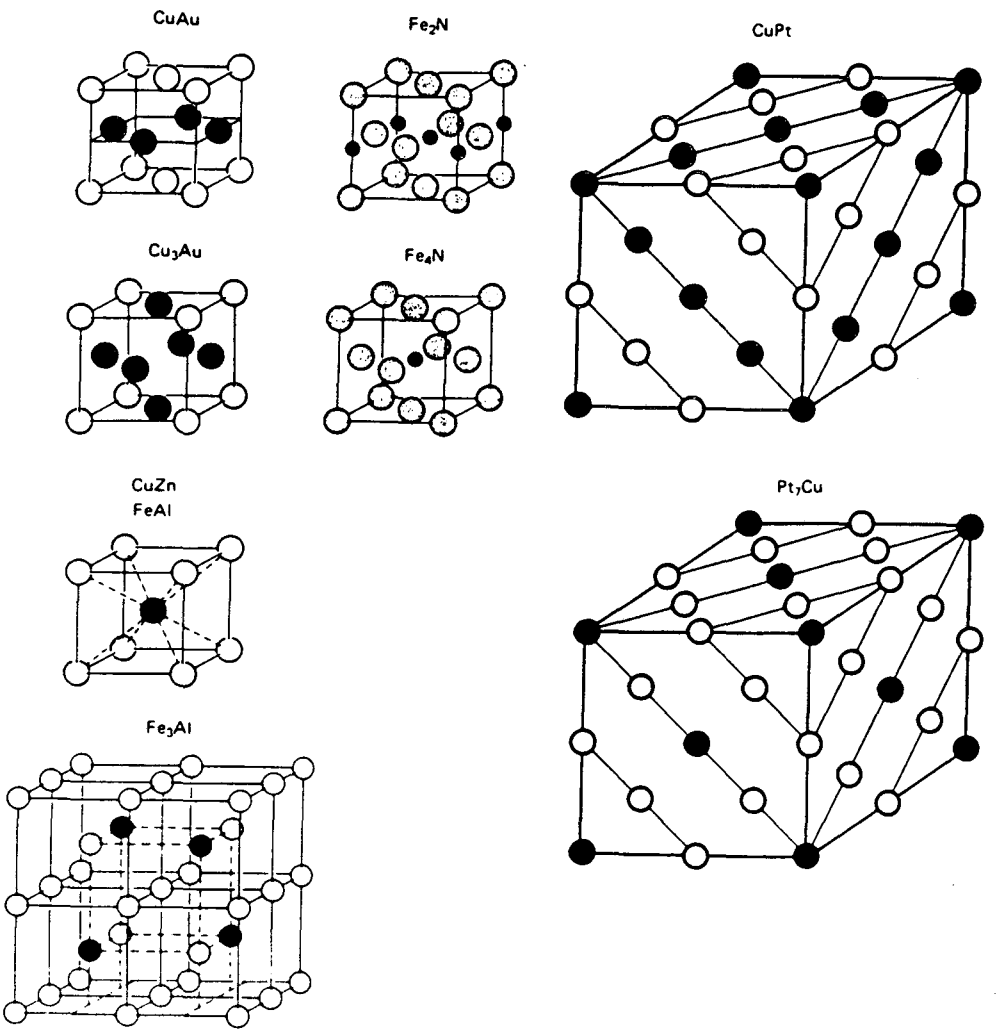


FIG. 11. Unit cells of some special point ordered structures.

cific heat undergo a discontinuous change. Second-order transformations arise from instability to small fluctuations such as, for example, static concentration waves in order–disorder reactions. The two phases involved in the transition cannot coexist at any temperature. [*See* CHEMICAL THERMODYNAMICS.]

Landau and Lifshitz in a symmetry-based theory have derived criteria necessary for a second-order transition:

1. The symmetry of the product phase must be a subgroup of that of the parent phase.
2. The transition must be generated by special point wave vectors.
3. The sum of any three wave vectors **k** of the star generating the transition must not be equal to a reciprocal lattice vector **g** of the parent structure, that is

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 \neq \mathbf{g}$$

This third criterion is a simplified version of the original Landau criterion.

All group–subgroup transitions fulfill the first condition. This includes, for example, order–disorder, ferroelectric, ferroelastic, or magnetic transitions. CuAu and Cu_3Au for instance (see Fig. 11) are both subgroups of the fcc structure and are generated by the wave vectors of the $\langle 1\ 0\ 0 \rangle$ star. The second criterion predicts this transition to be of first order since $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = \langle 1\ 1\ 1 \rangle = \mathbf{g}$. On the other hand for structures generated by the $\langle \frac{1}{2}\ \frac{1}{2}\ \frac{1}{2} \rangle$ star, such as CuPt, a second-order transition is allowed by symmetry.

D. Long Period Superlattices

In addition to special point structures many other ordered structures can form and have been observed. These can be enumerated by considering all the possibilities of distributing different atomic species on a given lattice, and their stability can be evaluated theoretically. A large class of such structures arises from a long period modulation of the stacking order in the basic structure, that is, periodic stacking faults. In the basic structure of SiC many periodic arrangements of stacking faults, termed polytypes, have been found. If the change in stacking alters only the modulated, or ordered, structure the stacking faults are called antiphase boundaries and the resulting structure is a long period superlattice. An example of a one-dimensional long period superlattice is that found in Cu–Au alloys. As illustrated in Fig. 12, an antiphase boundary in every fifth unit cell changes the tetragonal

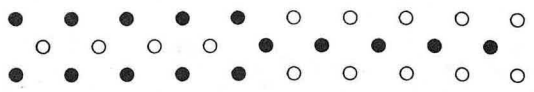


FIG. 12. CuAu II long period superlattice due to periodic stacking shifts.

structure of CuAu I to the orthorhombic structure of CuAu II. Similar one-dimensional long period superlattices are found in Cu–Al alloys. Au–Mn alloys exhibit both one- and two-dimensional examples of these structures. In semiconductors long period superlattices have been produced artificially to form multiple quantum well structures with unique properties.

III. Domain Structures

A. Group–Subgroup Transitions

In an order–disorder reaction a compositionally disordered crystal transforms to an ordered one on crossing the critical ordering temperature during cooling. The low-temperature product phase has lower symmetry than the high-temperature parent phase. A single crystal of the parent phase can therefore transform to several orientations of the product phase and the resulting regions of different orientation are called domains. The simplest case, that of antiphase domains, is illustrated below using the β-brass order–disorder transition based on the body-centered cubic (bcc) lattice. In a solid solution of Cu and Zn in a ratio of 1 : 1 each lattice site is occupied at random by Cu and Zn atoms with a probability of $n(\mathbf{r}) = 0.5$ [see Eq. (1)]. Upon crossing the critical ordering temperature the ordered structure shown in Fig. 11 forms. Due to the ordering the corner sites are no longer equivalent to the body-centered sites and thus two different domains are possible depending on whether the Cu atoms occupy one or the other sites. Both possibilities are equally likely and do in fact occur. Where two such domains meet they form an antiphase domain boundary such as the ones shown schematically in Fig. 13a. These can be made visible in electron micrographs and an example is shown in Fig. 13b. The boundaries mark the impingement of domains nucleated out of phase in different regions (thermal boundaries).

The order–disorder reaction in Cu–Zn is a simple example of the more general class of group–subgroup transitions. It illustrates an im-

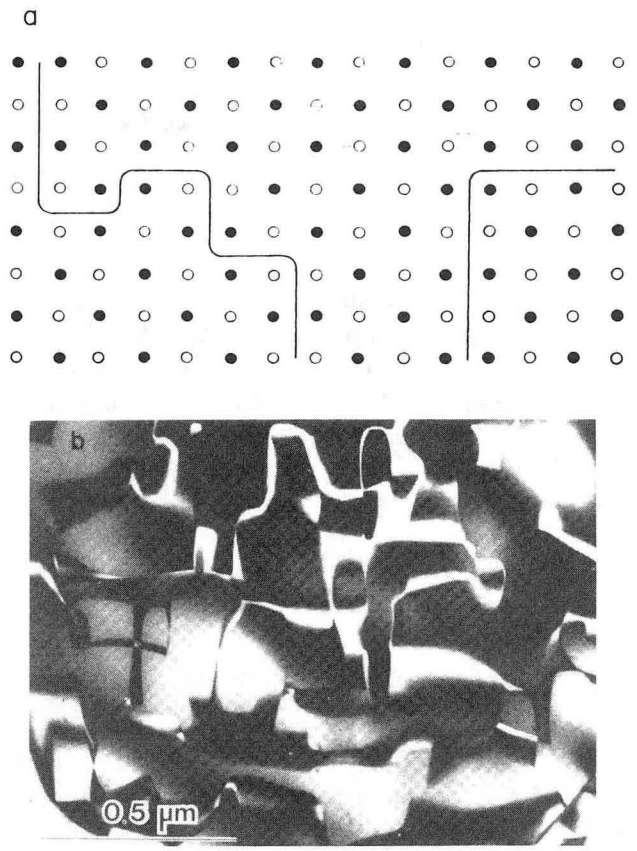


FIG. 13. Antiphase domain boundaries: schematic representation (a) and TEM micrograph of antiphase boundaries in Ni–Al alloy (b). (Courtesy K. H. Westmacott.)

portant point: domains are due to broken symmetries. Generally, every symmetry element lost in a transition will give rise to a possible domain boundary. For CuZn, only one symmetry element is lost, the $\frac{1}{2}\langle 1\ 1\ 1\rangle$ body-centering translation. Hence only one type of domain boundary can be formed (either thermally or by glide). In general the loss of translation elements gives rise to antiphase domains, the loss of inversion symmetry causes inversion (enantiomorphic) domains, the loss of a mirror symmetry results in twin domains, and the loss of a rotation axis leads to orientation domains.

It was shown earlier that the symmetry lowering in a group–subgroup transition may be described as a commensurate modulation of a physical quantity associated with the lattice, such as concentration, magnetization, or displacement. Extra spots (superlattice reflections) will appear in a diffraction pattern at the positions of the wave vectors $\mathbf{k}$ generating the ordered structure, (such as $\mathbf{k} = \langle 1\ 1\ 1\rangle$ in CuZn). Domain boundaries can be made visible in electron microscopy when one of these wave vectors fulfills the Bragg condition (see Fig. 13b). Thus a (1 1 1) superlattice reflection could be used to image antiphase domains in CuZn and a (1 0 0) reflection in CuAu.

Since each symmetry element lost in a transition gives rise to a domain, the number of possible boundaries is equal to the index n of the space group symmetry G_1 of the product crystal in the group G_0 of the parent crystal, that is, the ratio of the orders of the two groups:

$$n = \text{order of } G_0/\text{order of } G_1 \qquad (5)$$

When changes in the unit cell size are taken into account, this number is multiplied by the ratio of the primitive unit cell volumes: V_1 and V_0:

$$n = (V_1/V_0)(\text{order of } G_0/\text{order of } G_1) \qquad (6)$$

For example in the Cu_3Au ordering transformation (see Fig. 11) the parent phase is $Fm\bar{3}m$ and the product phase $Pm\bar{3}m$, $n = 192/48 = 4$. These are the four possible antiphase domains, separated by the three possible domain boundaries, each of which is characterized by a lost translation of the type $\langle \frac{1}{2}\ \frac{1}{2}\ 0\rangle$. Note that for n domains there are only $n - 1$ boundaries.

Each boundary is characterized not only by one but by a whole set of symmetry elements, the coset of the boundary. For example if a symmetry element $(\mathbf{R}|\boldsymbol{\tau})$ generates a boundary between domains, then $(\mathbf{R}|\boldsymbol{\tau})g_1$ describes the same boundary if g_1 is an element of the symmetry group G_1 of the product phase. This is true for all elements of G_1, and the set $(\mathbf{R}|\boldsymbol{\tau})G_1$ contains all the symmetry elements generating this boundary. Each boundary has its own unique coset. Two cosets have either no elements in common or are identical. If two symmetry elements belong to the same coset they generate the same boundary. The parent symmetry group G_0 can be decomposed into a sum of cosets characteristic for the domains formed in the transition to G_1:

$$G_0 = G_1 + \Sigma_i(\mathbf{R}_i/\boldsymbol{\tau}_i)G_1$$

where $(\mathbf{R}_i/\boldsymbol{\tau}_i)$ are the lost symmetry elements. As an example, in Cu_3Au ordering, the coset decomposition becomes

$$Fm\bar{3}m = Pm\bar{3}m + \{(1|\tfrac{1}{2}\ \tfrac{1}{2}\ 0) + (1|\tfrac{1}{2}\ 0\ \tfrac{1}{2}) + (1|0\ \tfrac{1}{2}\ \tfrac{1}{2})\}Pm\bar{3}m$$

B. The Effect of Strain

The change of symmetry in a transition is usually associated with a distortion of the unit cell. In most group–subgroup transitions this trans-

formation strain is small ($\lesssim 1\%$) and it is usually neglected in determining the possible domain boundaries. However, larger transformation strains have a significant effect on the possible domain configurations and the orientation of the domains themselves. Even for small strains, domain boundaries are not arbitrary in their orientation but tend to assume positions of minimum strain. The location $\mathbf{u}$ of such boundaries is given by

$$\mathbf{u}(\mathbf{s} - \mathbf{s}')\mathbf{u} = 0 \tag{7}$$

where $\mathbf{s}$ and $\mathbf{s}'$ are the small strain transformation tensors for the two domains meeting at the boundary. Since $\det|\mathbf{s} - \mathbf{s}'| = 0$, the solutions $\mathbf{u}$ describe invariant planes and Eq. (7) is the small strain equivalent of the invariant plane condition in martensite theory, as described later. Consequently domains usually meet along planar interfaces, a condition that becomes more stringent for larger strains.

In the case of the cubic to tetragonal transformation the strain tensors have the form

$$\mathbf{s} = \begin{pmatrix} e & & \\ & e & \\ & & -2e \end{pmatrix} \quad \text{and} \quad \mathbf{s}' = \begin{pmatrix} e & & \\ & -2e & \\ & & e \end{pmatrix}$$

Equation (7) has solutions $\mathbf{u} = \langle \mathrm{h\ k\ k} \rangle$, hence the invariant plane interfaces are of the {0 1 1} type. Figure 14 illustrates that two domains meeting at a boundary undergo a slight rotation toward each other. The rotation angle $\alpha = (3/2)e$ is on the order of one degree if the strain e is on the order of 1%. Since the domains will rotate in opposite directions when they meet along two different {0 1 1} planes (see Fig. 14a and b) they strictly become four different domains. The total transformation that leads to only 3 orientation domains if the small strains are neglected will form 12 orientation domains if the rotations resulting from the strain are taken into account. The number of domains (also called variants) is thus no longer given by the index n of G_1 in G_0, since due to the small rotation the 4-fold axes of the parent and product phase are no longer exactly parallel.

C. Variants

Even though it does not change the symmetry of the low-temperature phase itself, the rotation due to strain causes the loss of symmetry elements common to parent and product phases. The set of remaining common symmetry elements is given by the intersection H of the parent group G_0 and product group G_1 in its slightly rotated orientation, $H = G_1 \cap G_0$. The number of domains is then the index of H in G_0:

$$n = \text{order of } G_0/\text{order of } H \tag{8}$$

If the product phase coexists with the parent phase as in any first-order transformation, isolated domains or inclusions form in the matrix crystal. Different crystallographically equivalent inclusions are called variants. Their rotation from a symmetrical orientation determines the orientation relationship and their number is given by Eq. (8) with G_0 and G_1 being the point groups rather than space groups since the transformation strain usually destroys all common translations. The generation of variants is again due to those symmetry elements of the parent phase that are not shared by the product phase, that is, the broken symmetries. Figure 15 illustrates this process in a hypothetical cutting and welding operation similar to the ones used to find the strain energy of an inclusion (see Fig. 32) or the dislocation network in a grain boundary (see Fig. 20). A spherical volume of the matrix containing the inclusion is cut out, rotated, inverted, or reflected, and rewelded in the ma-

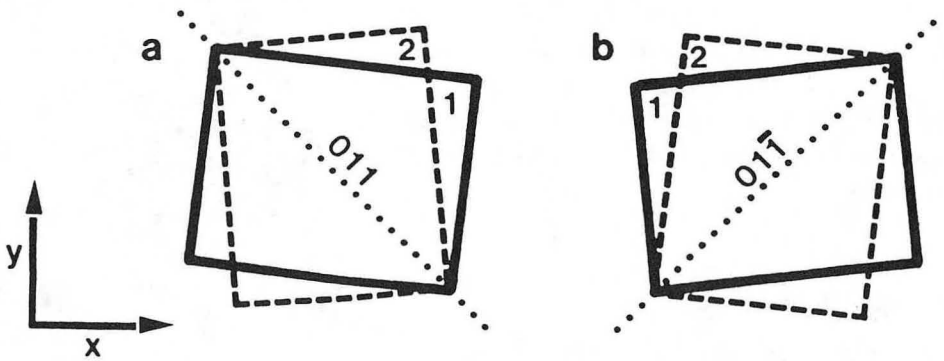


FIG. 14. Two tetragonal domains 1 and 2 meeting on (0 1 1) plane (a) and (0 1 $\bar{1}$) plane (b). A slight rotation of the domains toward their common boundary leads to relative changes in orientation.

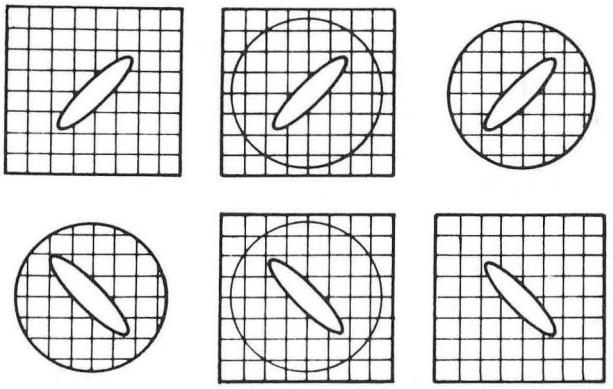


FIG. 15. Hypothetical sequence of operations illustrating the generation of precipitate variants.

trix. The matrix crystal is undisturbed by this operation because by definition it is invariant under a symmetry operation. However, the inclusion is in a new orientation, that is, it has become a different variant. Symmetry elements that are common to matrix and inclusion will not generate new variants. These elements form the intersection group $H = G_1 \cap G_0$ and an example of one such element is the 2-fold rotation about the plane normal in Fig. 15.

The electron micrograph in Fig. 16 shows a distribution of many variants of Cr precipitate needles along ⟨6 5 1⟩ directions in a Cu matrix. The intersection group of the matrix $G_0 = m\bar{3}m$ and the precipitate $G_1 = m\bar{3}m$ is only $H = \bar{1}$ of order 2 because of the low-symmetry orientation relationship adopted by the precipitate lattice. The number of variants is therefore $n = 48/2$. The distribution of the 24 variants can be visualized as poles on a stereogram showing a general direction (see Fig. 6) and can be enumerated by permutation of the indices ⟨6 5 1⟩. Notice that both the product and parent lattices are cubic but are misoriented so that they have only an inversion center in common. The orientation relationship is one of low symmetry since it is dictated by the invariant line or invariant plane criterion discussed in Sections V and VI. The symmetry of the orientation relationship is given by the intersection group H.

D. Equilibrium Inclusion Shape

It is easy to see that H is also the symmetry group of the equilibrium inclusion shape: the structure and energy of a given interface plane are identical for two planes related by a symmetry operation of H and hence all equivalent faces

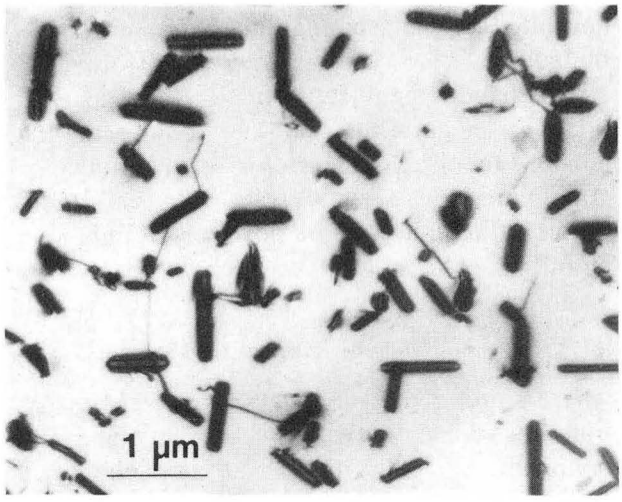


FIG. 16. TEM micrograph with ⟨0 0 1⟩ beam direction showing needle-shaped precipitates along ⟨6 5 1⟩ directions in Cu–Cr alloy.

are expected to be found bounding a particle. The needles shown in Fig. 16 have a low-symmetry equilibrium shape, that of a pinacoid ($H = \bar{1}$). Particles with higher intersection symmetries H are shown in cross section in Fig. 17a–c. In all three cases the germanium particle and the aluminum matrix share a mirror parallel to the image plane and a 2-fold axis perpendicular to it. When there are no other common elements (see Fig. 17a) the intersection group is monoclinic, $H = 2/m$. Higher symmetry intersection groups are seen in Fig. 17b ($H = mmm$, orthorhombic) and

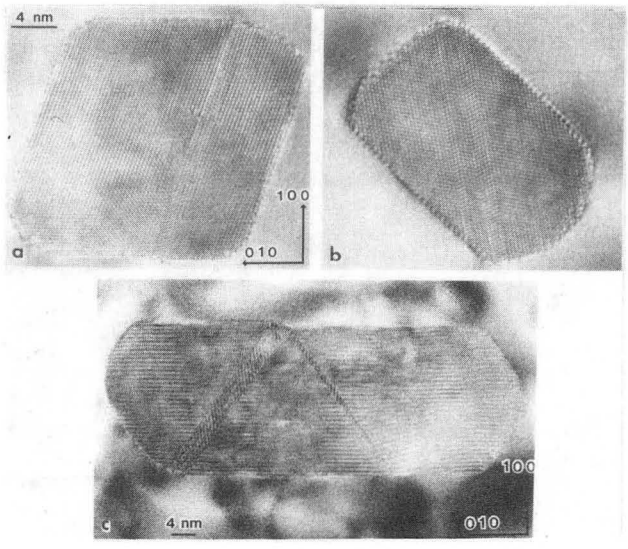


FIG. 17. TEM micrographs of Ge needle precipitates in Al with different common symmetries, seen in cross-section; monoclinic (a), orthorhombic (b), and tetragonal (c). Note corresponding shapes.

Fig. 17c ($H = 4/mmm$, tetragonal). Note that the overall particle shapes conform to the expected symmetries, that is, they are close to their equilibrium shape.

Usually the more symmetrical groups H (higher order) are preferred over those of low symmetry. When H is such that any misorientation would lower its symmetry it is said to be at a symmetry-dictated extremum.

IV. Grain Boundaries

Grain boundaries are interfaces separating two crystals identical in structure but different in orientation. The misorientation between two grains is described by the axis **l** and angle θ of misorientation (axis/angle pair) or alternatively by a rotation tensor **R**. If in addition to the misorientation there is also a shift τ between the grains the interface operation is $(\mathbf{R}|\tau)$. Most geometrical properties can be derived from this operation. However, a boundary is completely described only if the boundary plane is specified as well. Thus nine parameters are necessary for a complete description of a grain boundary, three for the axis and angle of rotation, three for the direction and magnitude of the translation, and three for the position of the boundary plane.

Existing grain boundary models are based either on geometric or energetic criteria. The latter usually require large computer programs and the results depend strongly on the interatomic potential chosen. The former generally have simpler, analytical solutions but are not directly related to the energy. However, experience has shown boundaries with optimum geometry are usually low-energy boundaries. Some geometrical models of grain boundaries are described below.

A. Low-Angle Boundaries

If the misorientation between two grains is less than ~10° their boundary is usually considered a low-angle boundary. The dislocation array shown in Fig. 18a is a simple example of a low-angle boundary and a high-resolution image of such a boundary in molybdenum is shown in Fig. 18b. The misorientation θ is related to the dislocation spacing d as

$$\tan \theta = b/d \tag{9}$$

where b is the Burgers vector of the dislocation. At 10° misorientation the dislocations would thus be spaced $d \cong 6b$ apart, causing considerable overlap between their strain fields. It would therefore be physically meaningless to describe larger misorientations as arrays of individual dislocations, and other constructions become important. These will be discussed under high-angle boundaries. Some further principles can be illustrated using low-angle boundaries. The example shown in Fig. 18 is a symmetrical tilt

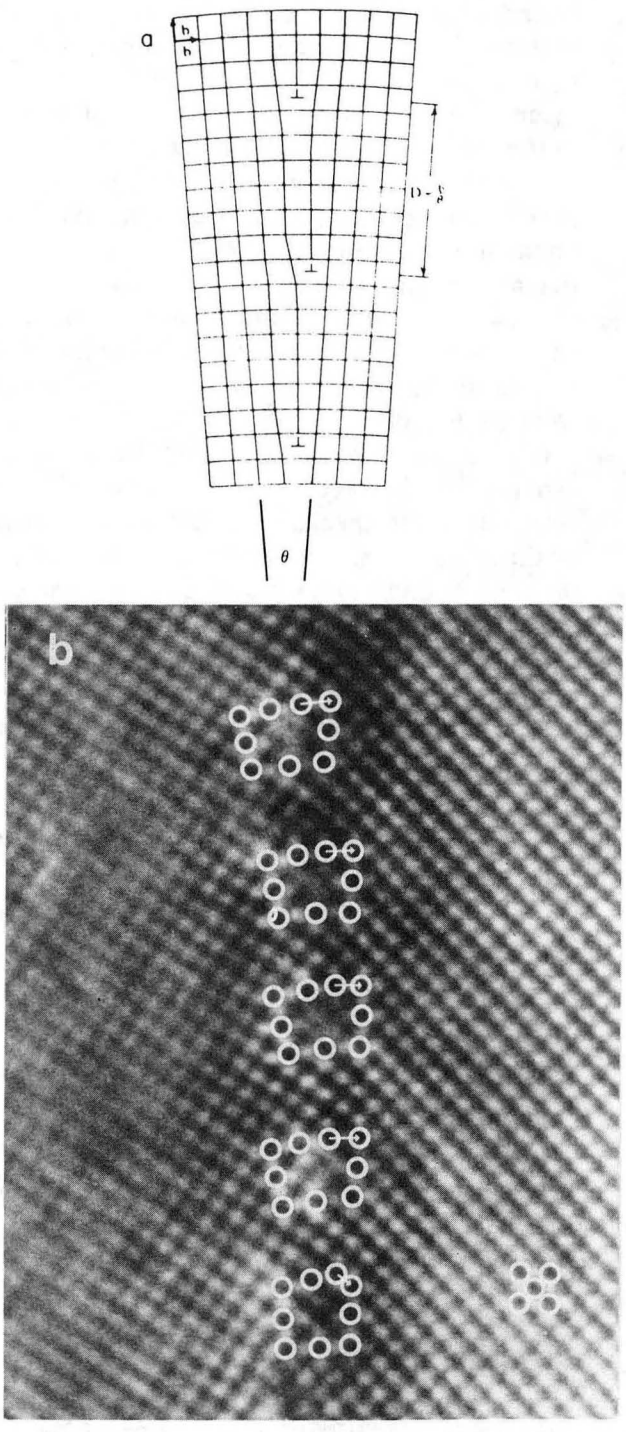


FIG. 18. Low-angle tilt boundaries: schematic of symmetrical boundary (a) and TEM micrograph of nonsymmetrical boundary in Mo (b). (Courtesy of J. M. Pénisson and R. Gronsky.)

boundary; it is located on the plane of symmetry between the two crystals and it contains the axis of misorientation. Because of their simplicity, symmetrical tilt boundaries figure prominently in the modeling of grain boundaries.

Another high-symmetry configuration is the pure twist boundary shown in Fig. 19. Such boundaries are normal to the axis of misorientation and require not a single but at least two sets (a network) of dislocations. As for a symmetrical tilt boundary, the angle of misorientation is related to the spacing of each array in the network by Eq. (9).

For the same axis/angle pair, the boundary can take arbitrary orientations which then of course lack the special symmetry of pure twist or tilt boundaries. One way to visualize all the possible boundary orientations is to treat one grain as an inclusion in the other. The grain boundary is then a closed surface and the dislocation network a set of closed loops on that surface. To obtain their spacing and orientation, consider the misorientation of the enclosed grain to be the result of a phase transformation whose sole effect is a small change in orientation of the lattice without changing the volume, shape, or crystal structure of the "new phase." In order to accommodate this change in orientation it is necessary to deform the grain by slip, resulting in a dislocation network in the interface. This is best seen in the schematic sequence in Fig. 20. From a single crystal cut a region (a), remove and rotate it (b), deform it plastically such that it will fit into its previous space (c) and (d), and insert and reweld it in the original crystal (e). The slip steps formed in (c) and (d) now become dislocation loops as shown in the perspective drawing (f). By extension it is easy to see from this description how a small-angle grain boundary of any axis/angle pair and on any plane could be constructed. In particular, notice that the same misorientation could be achieved with a different set of dislocations, depending on the slip systems operating in steps (c) and (d). Also note that in the particular example shown the planes of the dislocation loops intersect in a line parallel to the rotation axis. The rotation axis is an invariant line of the "transformation" and must be unaffected by the dislocations. Not every rotation axis conveniently lies at the intersection of two crystallographic slip planes. More than two sets of dislocation loops are then necessary to accommodate a rotation around an irrational axis but again the total set must leave the rotation axis unchanged.

Since the rotations involved are small, the rotation matrix can be written as

$$\mathbf{R} = \mathbf{I} + \boldsymbol{\omega} \tag{10}$$

where $\mathbf{I}$ is the identity matrix and $\boldsymbol{\omega}$ is an antisymmetrical tensor, that is, $\omega_{ij} = -\omega_{ji}$. This "rotational strain" tensor $\boldsymbol{\omega}$ must be countered by an equal and opposite plastic strain $-\boldsymbol{\omega}$, made up of dislocation arrays. Each array of dislocation loops of given orientation, spacing, and Burgers vector has a characteristic strain tensor. For example, an array of shear loops on

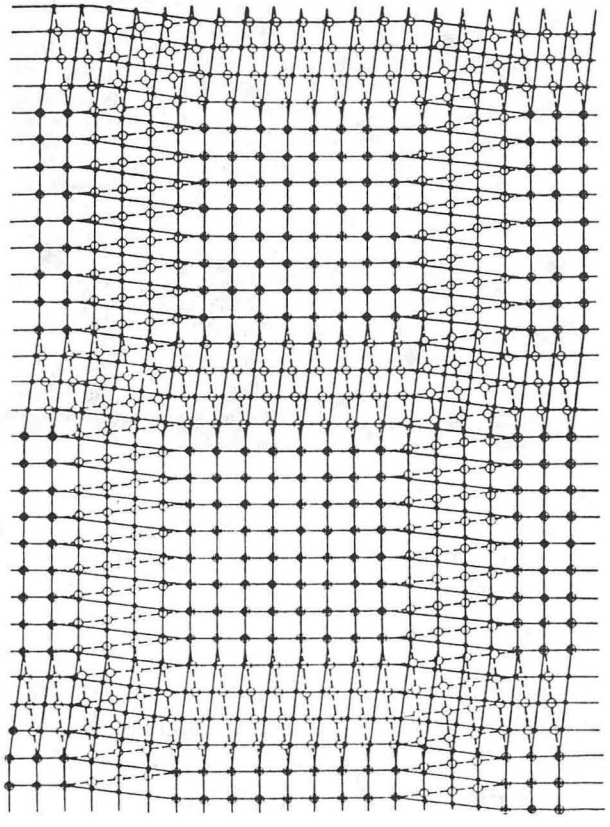


FIG. 19. Low-angle twist boundary seen face-on. [From W. T. Read, Jr., "Dislocations in Crystals." Copyright 1953, with permission of McGraw-Hill.]

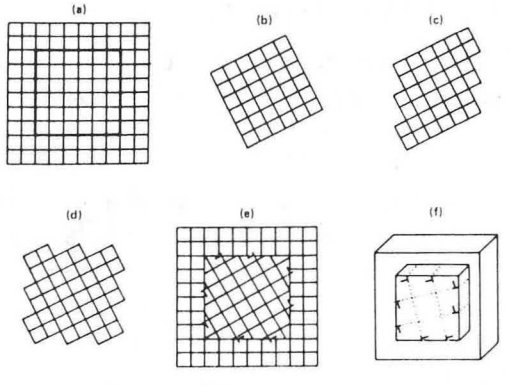


FIG. 20. Sequence of hypothetical operations illustrating the formation of an arbitrary grain boundary.

(0 0 1) planes with Burgers vector **b** in the [1 0 0] direction is written as the simple shear strain:

$$\mathbf{s} = \begin{pmatrix} 0 & 0 & b/d \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

By adding appropriate arrays to form an antisymmetrical tensor ω any small angle boundary can be described. This addition applies only for small strains since the superposition principle (strains are additive) holds only when products of strains are negligible.

B. O-Lattice

Another method to determine the Burgers vector content **B** along a vector **p** in a boundary is by Frank's formula:

$$\theta(\mathbf{l} \times \mathbf{p}) = \mathbf{B}$$

where $(\mathbf{l}, \theta)$ is the axis/angle pair. Frank's formula is valid only for small rotation angles θ. The concept of measuring the misfit **B** in an arbitrary direction **p** in the boundary is very useful and is central to the more general theory of surface dislocations. By decomposing the misfit **B** into a sum of lattice translations it becomes possible to devise a corresponding set of dislocation lines that must be traversed by the vector **p**. If **p** is selected so that the misfit along it is equal to a Burgers vector **b** then **p** becomes an O-lattice vector $\mathbf{x}^0$, a basic concept in O-lattice theory. For low-angle boundaries

$$\omega\mathbf{p} = \Sigma_i \mathbf{b}_i \qquad (11a)$$

or

$$\omega\mathbf{x}^0 = \mathbf{b} \qquad (11b)$$

where ω is the antisymmetrical tensor describing the "rotational strain." For misorientations greater than ~10°, the small-angle approximation is no longer valid and Eqs. (11a,b) become

$$(\mathbf{I} - \mathbf{R}^{-1})\mathbf{p} = \Sigma_i \mathbf{b}_i \qquad (12a)$$

$$(\mathbf{I} - \mathbf{R}^{-1})\mathbf{x}^0 = \mathbf{b} \qquad (12b)$$

A dislocation is necessary in the boundary every time the rotational mismatch $(\mathbf{I} - \mathbf{R})$ measured along a vector $\mathbf{x}^0$ equals a Burgers vector **b**. Perhaps this is most apparent for coincidence site lattices (CSL) for which both **b** and $\mathbf{x}^0$ are lattice translation vectors.

C. Coincidence Site Lattice

When two identical interpenetrating lattices are rotated from initial coincidence around a lattice point, there are certain discrete rotation angles for which lattice points other than the origin coincide. An example is shown in Fig. 21a where a black lattice and a white lattice are outlined by a dashed and a solid square. The coincident points (solid dots) form a lattice themselves, termed the CSL, shown in Fig. 21b. The CSL is characterized by the inverse density of coincidence sites, Σ. For the example illustrated in Fig. 21, $\Sigma = 5$. The rotation angle θ is 36.9° ($\tan \theta/2 = 1/3$) due to the coincidence of the [3 1 0] vector in one lattice with the [3 $\bar{1}$ 0] vector in the other lattice. Any rational $\langle h\ k\ l \rangle$ lattice vector can be used to generate a CSL characterized by $\Sigma = h^2 + k^2 + l^2$. If $h^2 + k^2 + l^2$ is even, Σ is its largest odd divisor. For example a $\langle 3\ 1\ 0 \rangle$ vector $\Sigma = 10$ and a $\langle 2\ 1\ 0 \rangle$ vector $\Sigma = 5$ generate the same $\Sigma 5$ CSL shown in Fig. 21. In

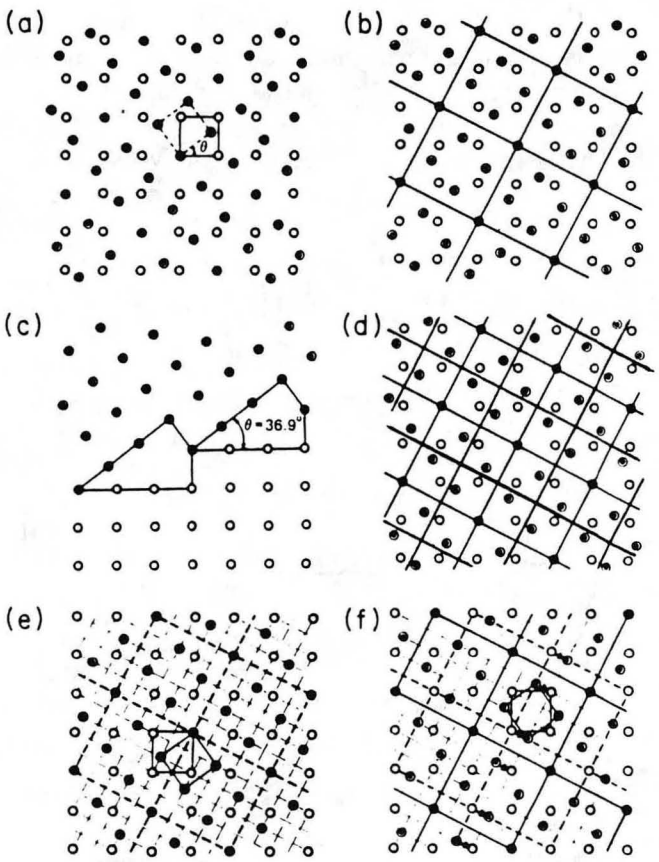


FIG. 21. Σ5 coincidence site lattice (CSL); dichromatic pattern of two interpenetrating square lattices of black and white atoms with solid dots showing coincidence points (a), CSL outlined (b), structural units along possible boundary plane (c), O-lattice outlined (d), DSC lattice outlined (e), and effect of DSC shift on CSL (f). [From H. F. Fischmeister, *J. Phys. Colloque* C4-3 (1983).]

Fig. 21c a boundary plane has been chosen and the black (white) lattice discarded from below (above) the plane of the boundary. Structural units typical of this boundary are outlined. The density Γ of coincidence sites in a boundary plane depends on its position. Special grain boundaries have high values of Γ and low values of Σ.

Any lattice vectors of the CSL are solutions $\mathbf{x}^0$ of the O-lattice equation (12b) since at each coincidence site the rotational displacements between the two lattices amount to a lattice translation **b**. However, not all O-lattice vectors $\mathbf{x}^0$ are CSL vectors. The complete O-lattice for the $\Sigma 5$ boundary is shown in Fig. 21d. When the misorientation between the lattices is not exactly equal to one necessary for a CSL, all the coincidence sites (except the center of rotation) are lost. However the O-lattice is still maintained with lattice vectors $\mathbf{x}^0$ which are now irrational although close to the rational CSL vectors. Thus the O-lattice changes continuously between the discrete CSLs.

In order to find the geometrically necessary network of dislocations in an interface from the solution of Eq. (12b) one constructs cell walls midway between any two O-lattice points. These represent the positions of worst match since the O-lattice points are the positions of best match. An example is shown in Fig. 22 in which the O-lattice is a hexagonal set of parallel lines and the cell walls are shown as hexagonal prisms. The vectors $\mathbf{x}^0$ corresponding to three coplanar Burgers vectors **b** lie in the basal plane. The predicted dislocation network lies at the intersection of the chosen boundary plane with the cell walls (heavy outline). Notice that the boundary could also be curved or even a closed surface as considered earlier. The same dislocation network as shown for example in Fig. 20 would be predicted. However, in this view, the physical origin of the dislocations is not considered since they arise only as a geometrical necessity.

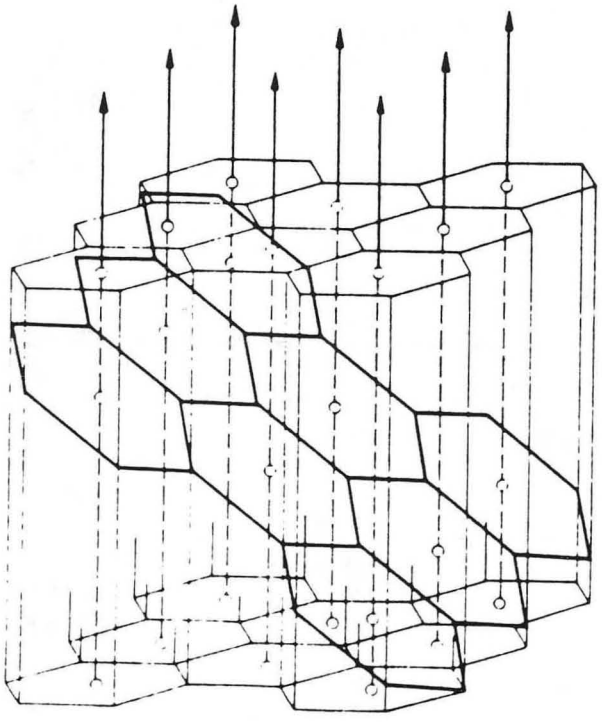


FIG. 22. Hexagonal O-lattice due to misorientation of two grains around a common 3-fold axis. Heavy outlines show predicted dislocation network where the boundary plane intersects the O-lattice cell walls.

Even though the equations are valid for any angle of misorientation this construction is physically meaningful only for low-angle boundaries since, as shown earlier, for misorientations $>10°$, their cores are too close to each other.

D. High-Angle Boundaries

High-angle boundaries are best treated as small deviations from the nearest CSL. They are then similar to low-angle boundaries that are a small deviation from the $\Sigma 1$ CSL, that is, a single crystal. Each CSL, in particular $\Sigma 1$, may be considered a low-energy configuration. Usually the lower Σ is, the lower the energy. Any deviation from a CSL is accommodated by dislocations, lines of high local distortion in favor of relaxed or undistorted regions of the CSL in between. In the case of low-angle boundaries the dislocations are primary dislocations and the boundary between them perfect crystal ($\Sigma 1$). This is shown in Fig. 18. In the case of high-angle boundaries the dislocations are secondary dislocations and the boundary between them perfect CSL. These secondary dislocations have small Burgers vectors given by the smallest difference vectors between lattices 1 and 2 in the exact CSL orientation. A dislocation with such a Burgers vector will translate the complete CSL and is part of the so-called DSC (displacement shift complete) lattice. The example of the $\Sigma 5$ DSC lattice is shown in Fig. 21e and the effect of a DSC lattice translation is illustrated in Fig. 21f. Any misorientation from the perfect $\Sigma 5$ orientation can be accommodated by these dislocations with areas of perfect $\Sigma 5$ boundary in between.

When a boundary is between two adjacent low-energy CSL orientations it could approximate one or the other with an appropriate set of dislocations. In fact, experience has shown that

it does both. By mixing well-matched structural units characteristic for each of the neighboring CSLs in the right proportion the boundary energy is minimized. Any degree of misorientation can be accommodated in this manner and a change in misorientation is similar to a change in the structural "composition" of the boundary. Rearrangement of structural units in a boundary, by separation or mixing, may in a sense be considered a phase transformation. This must be accompanied by a local change in boundary orientation. Structural units are a refinement of the purely geometrical CSL concept since local atomic relaxations and relative translations between the two lattices are taken into account. This can be done by computer simulation assuming an interatomic potential, or more simply by geometrical criteria based on rigid sphere packing. The latter model is successful, at least in fcc metals where the hard sphere model is a good approximation.

E. Grain Boundary Crystallography

In the wider context of interphase boundaries the concepts of CSL, DSC, and O-lattice can be derived alternatively as follows. A general interphase boundary described by a transformation **A** followed by a translation τ is called a heterophase boundary, characterized by the operation $(\mathbf{A}|\tau)$. If **A** leaves the crystal structure invariant (**A** = **R**) it describes a homophase boundary. This can be a domain boundary if $(\mathbf{R}|\tau)$ is a symmetry operation of the parent phase, or a grain boundary if $(\mathbf{R}|\tau)$ is a general rotation/translation. The translation may be irrational, for example, if due to relaxation at a grain boundary, or a rational lattice translation **t**, for example, at an antiphase domain boundary.

Due to crystal symmetry (crystal space group G) the description of a boundary by a single operator $(\mathbf{R}|\tau)$ is ambiguous because any symmetrically equivalent orientation of either crystal will give a different but equivalent boundary operation. The boundary is completely described by the set of all such operations $(\mathbf{R}|\tau)G$, called the coset of the boundary. Of this set, often the operation with the smallest rotation angle is used to characterize the boundary.

If $(\mathbf{R}|\tau)$ is a symmetry operation of the crystal, that is, an element of the space group G, it leaves the entire lattice invariant. Each lattice point is brought to an equivalent position. For a general boundary, $(\mathbf{R}|\tau)$ is not a symmetry element of the crystal. Thus formation of a boundary is a symmetry-breaking transition and the boundary operation $(\mathbf{R}|\tau)$ brings lattice points to nonequivalent positions. However, there may still be geometrical locations **r** in the lattice which remain invariant:

$$(\mathbf{R}|\tau)\mathbf{r} = \mathbf{r} \tag{13}$$

Such locations **r** exist if $(\mathbf{R}|\tau)$ is a reducible operator. For these positions, the effect of the rotation (**Rr**) equals the effect of the translation, $(\mathbf{r} - \tau)$. This is shown schematically in Fig. 23a.

If τ (but not **r**) is a lattice translation **t**, one of the elements of the translation group T_1 of lattice 1 (for example a Burgers vector $\mathbf{b} = -\mathbf{t}$), then **Rr** becomes an O-lattice vector $\mathbf{x}^0$ (see Fig. 23b) and Eq. (13) becomes

$$(\mathbf{R}^{-1}|\mathbf{b})\mathbf{x}^0 = \mathbf{x}^0$$

This is identical to Eq. (12b) and the solutions $\mathbf{x}^0$ to these equations describe points, lines, or planes that remain invariant in the transition $(\mathbf{R}|\mathbf{t})$. Finally, if both **r** and τ in Eq. (13) are members $\mathbf{t}_1$ and **t** of the translation group T, then **r** becomes a CSL vector $\mathbf{t}_1$:

$$(\mathbf{R}|\mathbf{t})\mathbf{t}_1 = \mathbf{t}_1$$

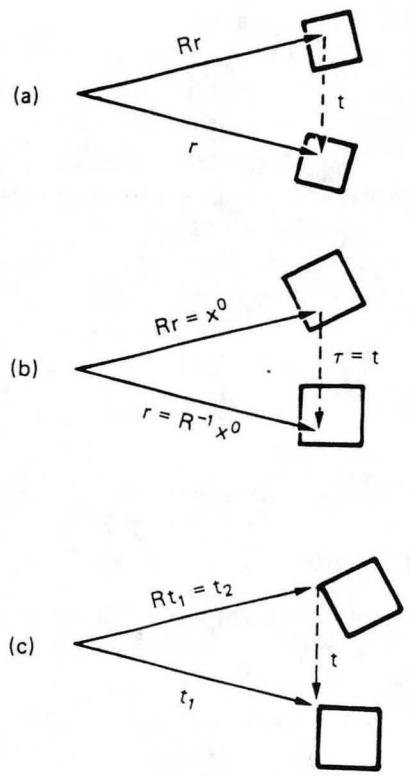


FIG. 23. Schematic illustration of rotation/translation operation: general operation (a), generation of O-lattice (b), and generation of CSL (c).

as shown schematically in Fig. 23c. If this is rewritten as

$$\mathbf{R}\mathbf{t}_1 = \mathbf{t}_1 - \mathbf{t}$$

the left-hand side is a translation in crystal 2 while the right-hand side is a translation in crystal 1. The CSL is thus the group of those vectors $\mathbf{t}_1$ which are translations in both crystals, that is, the intersection of the translation groups

$$\text{CSL} = T_1 \cap \mathbf{R}T_1 \tag{14}$$

The DSC lattice is simply the corresponding union of the translation groups

$$\text{DSC} = T_1 \cup \mathbf{R}T_1 \tag{15}$$

It was shown that the operation $(\mathbf{R}|\tau)$, if it is reducible, leaves a set of locations $\mathbf{r}$ invariant, that is $(\mathbf{R}|\tau)$ is a symmetry operation of the bicrystal. The entire set of bicrystal symmetries is the union of the symmetry elements common to both crystals $G_1 \cap G_2$, with those operations $(\mathbf{R}|\tau)G_1$ that transform crystal 1 into crystal 2 while simultaneously transforming crystal 2 into crystal 1: $G_1(\mathbf{R}|\tau)^{-1}$. Thus the bicrystal symmetry is

$$(\mathbf{G}_1 \cap \mathbf{G}_2) \cup [(\mathbf{R}|\tau)G_1 \cap G_1(\mathbf{R}|\tau)^{-1}]$$

Operations that interchange the two crystals are also called colored symmetry operations, referring to the colors, that is, black and white, assigned to the two identical but misoriented crystals shown for example in Fig. 21. The corresponding pattern is called the dichromatic pattern. Its symmetry is described by the Shubnikov groups, and it is interesting to note that it can have symmetries not present in either lattice. If, for example, $(\mathbf{R}|\tau)$ is the twinning operation in an fcc crystal where $\mathbf{R}$ is a {1 1 1} mirror m and $\tau = 0$, the bicrystal symmetry of the resulting $\Sigma 3$ CSL is

$$(\bar{3}m) \cup (m/o) = 6/mmm$$

V. Martensite

A. Characteristics

Martensite transformations are characterized by a surface relief that indicates that a shape change is associated with the transformation. The surface relief can be measured from the displacement of scratches placed on a flat polished surface before transformation (see Fig. 24). The interface plane between the martensite inclusion and the matrix (habit plane) remains macroscopically undistorted (invariant plane) and the total deformation is an invariant plane strain (IPS). Due to the constraint from the solid matrix, martensite inclusions take the form of plates, laths, or needles, with an internal structure resulting from slip or twinning. The transformation is diffusionless and fast, proceeding at rates near the speed of sound. This implies that large numbers of atoms transfer rapidly and in an orderly fashion from the high-temperature structure (austenite) to the low-temperature structure (martensite). Nearest neighbors remain nearest neighbors.

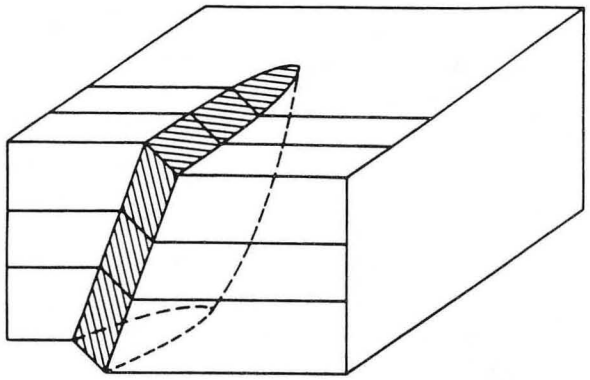


FIG. 24. Schematic of surface relief where a martensite plate intersects the surface.

The best-known and technologically most important martensite transformation after which the whole class of transformations is named is that in steel. It occurs on rapid cooling below the transformation temperature M_s which depends on carbon concentration. The volume fraction of martensite formed is determined by an equilibrium between the temperature-dependent chemical driving force and the strain energy of accommodating the shape change. In the case of steel the parent austenite is fcc and the product martensite bct. A characteristic orientation relationship between austenite and martensite is usually observed which is generally nonrational and of low symmetry. This leads to a large number of crystallographically equivalent orientation variants and complex morphologies. The semicoherent habit plane, usually also irrational, must be glissile, a condition that imposes restrictions on the interface dislocations.

Not all of these characteristics are typical for martensite transformations alone. Orientation relationships, habit planes, surface relief, and geometrically glissile interfaces are often also observed in diffusion-controlled precipitation

reactions. For example the Bainite reaction in steel bears all of these marks but is not martensitic because interstitial carbon diffuses during the reaction, sweeping ahead of the interface and eventually stopping the reaction front. Other transformations that exhibit a surface relief while allowing some diffusion are sometimes classified as bainitic. Diffusionless rapid transformations that do not show surface relief are termed massive transformations.

B. IPS Geometry

The phenomenological theory of martensite transformations is based on the observation of an IPS deformation, and well-defined orientation relationship. Since the habit plane is undistorted and unrotated it can be modeled geometrically rather than on some energetic criterion but it can be shown that an IPS is favored also by strain energy considerations. This is similar to the structure of grain boundaries where optimized geometry usually coincides with minimum energy configurations.

An exceptionally simple example of a martensitic transformation is that from fcc to hcp cobalt. As shown in Fig. 25a the only strain is due to a change in stacking of the close-packed planes. The close packed plane is undistorted and naturally becomes the habit plane. In this case the transformation strain itself is an IPS of the form

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where s is the shear due to the change in stacking sequence. For the two-dimensional case illustrated in Fig. 25a this reduces to

$$\mathbf{B} = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

The habit plane and orientation relationship both have simple rational form. However, this is not generally the case, and usually there is no lattice plane that remains unstretched and unrotated in the transformation. If, for example, the close packed plane in Fig. 25a were stretched by a factor a in the transformation it could no longer be an undistorted habit plane, and a slightly rotated, irrational habit plane would

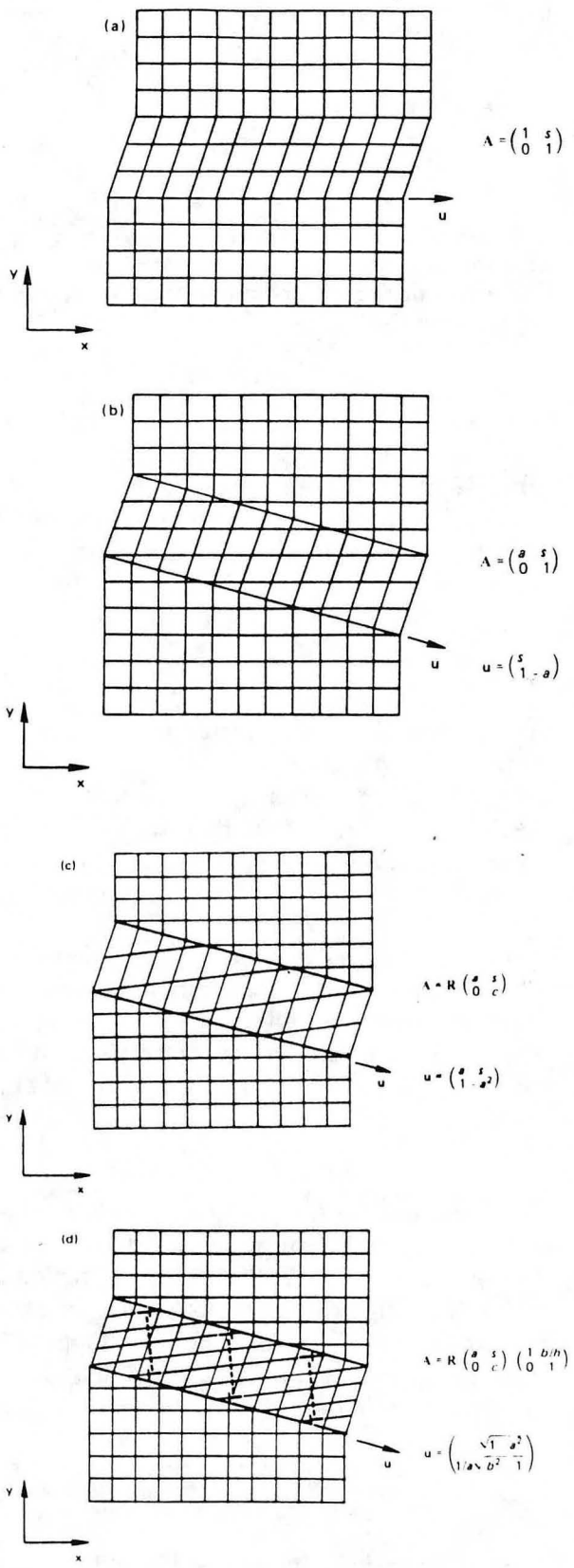


FIG. 25. Two-dimensional illustration of IPS geometry; simple shear on the habit plane (a), additional expansion a in the habit plane (b), additional expansion c normal to the habit plane (special case) (c), and additional lattice-invariant shear (general case) (d).

have to be adopted. The transformation tensor would be

$$\mathbf{B} = \begin{pmatrix} a & s \\ 0 & 1 \end{pmatrix}$$

and the invariant plane (or here in two dimensions the invariant line), given by the condition $\mathbf{Bu} = \mathbf{u}$, hence $\mathbf{u}\ (\frac{s}{1-a})$. This is shown in Fig. 25b where the perfect match along the direction $\mathbf{u}$ is apparent. At an angle

$$\tan\theta = (1 - a)/s \tag{16}$$

to the x axis, this habit plane is irrational but the orientation relationship is still rational with the close-packed planes exactly parallel.

If we further allow the spacing of the close-packed planes to change by a factor c the transformation

$$\mathbf{B} = \begin{pmatrix} a & s \\ 0 & c \end{pmatrix}$$

no longer has an invariant interface. However, if there were a direction that was undistorted but not unrotated, it could simply be rotated back by a rigid body rotation $\mathbf{R}$ so that $\mathbf{RBu} = \mathbf{u}$. In the present example, this requires that

$$\det|\mathbf{B}^2 - \mathbf{I}| = 0 \tag{17}$$

that is, the factors a, c, and s must follow the relations $(a^2 - 1)(c^2 - 1) = s^2$. If this relationship happens to be fulfilled by the transformation matrix, the coherent interface along the invariant line $\mathbf{u}$ shown in Fig. 25c is obtained at an angle

$$\tan\theta = (1 - a^2)/as \tag{18}$$

However, usually this relationship will not hold and must be fulfilled by adding a lattice-invariant deformation. This can be achieved by either twinning or slip in a number of ways. For example, an array of edge dislocation loops with Burgers vector b and spacing h would add a lattice-invariant deformation

$$\mathbf{S} = \begin{pmatrix} 1 + b/h & 0 \\ 0 & 1 \end{pmatrix}$$

or an array of shear loops with the same Burgers vector and spacing h would be

$$\mathbf{S} = \begin{pmatrix} 1 & b/h \\ 0 & 1 \end{pmatrix}$$

Considering the requirement of a glissile interface the array of shear loops is more realistic. The combined transformation

$$\mathbf{A} = \mathbf{RBS} \tag{19}$$

now has an invariant interface if the dislocation spacing h in the lattice-invariant shear $\mathbf{S}$ is adjusted so that $\mathbf{RBSu} = \mathbf{u}$. This requires that

$$\det|(\mathbf{BS})^2 - \mathbf{I}| = 0 \tag{20}$$

or explicitly $(a^2 - 1)(c^2 - 1) = (s + ab/h)^2$. This leads to the semicoherent interface shown schematically in Fig. 25d with the same angle of inclination as for the corresponding coherent interface in Fig. 25c, now given by

$$\tan\theta = \frac{1}{a}\sqrt{\frac{1 - a^2}{1 - c^2}} \tag{21}$$

Figure 25d is the two-dimensional analog of the crystallography of a general martensite transformation. It shows the shape strain $\mathbf{A}$, the lattice invariant shear $\mathbf{S}$ in the form of a shear dislocation array, and the lattice rotation $\mathbf{R}$.

To illustrate schematically the same situation in three dimensions becomes more difficult. However, the same principles apply and the phenomenological theory of martensite transformations theory provides an algorithm to determine the habit plane and orientation relationship for a given transformation and slip system.

C. The Phenomenological Theory

In order to find the transformation strain $\mathbf{B}$, a plausible correspondence between the austenite and martensite lattices must first be established. In some cases the choice is obvious, especially when the distortions are small and the relation between the two lattices is easily recognized. The fcc to bct transformation in steel is not so obvious because the transformation strains are large. The most widely used lattice correspondence for this case is the Bain correspondence shown in Fig. 26. The bct cell shown in heavy outline within the fcc lattice must be deformed along three orthogonal axes to the correct dimensions of the martensite ($c/a \simeq 1.08$). $\mathbf{B}$ is the matrix of distortions along these principal axes

$$\mathbf{B} = \begin{pmatrix} \eta_1 & & \\ & \eta_1 & \\ & & \eta_3 \end{pmatrix} \tag{22}$$

In most martensite transformations, the volume change (given by $\det|\mathbf{B}|$) is small compared to the shape change, and $\mathbf{B}$ causes almost a pure shear deformation. The transformation $\mathbf{B}$ alone

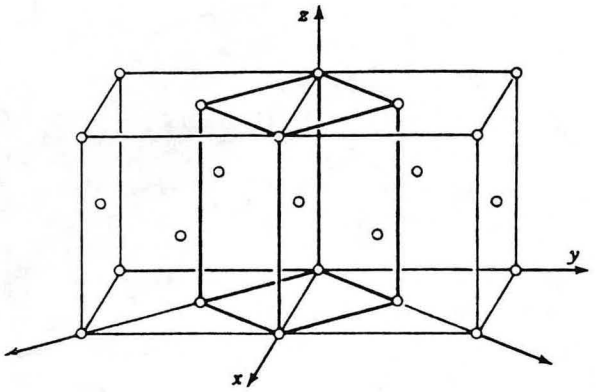


FIG. 26. Bain correspondence between fcc and bct lattices. The bct lattice can be deformed into a bcc lattice by three orthogonal strains.

leaves no direction or plane invariant although it leaves all directions **u** unextended for which

$$(\mathbf{Bu})^2 = \mathbf{u}^2 \tag{23}$$

In general,

$$\det|\mathbf{B}^2 - \mathbf{I}| \neq 0$$

and the solutions **u** of this equation lie on a cone of unextended lines. If, as shown above for the two-dimensional case, we add a lattice-invariant shear **S** of appropriate magnitude such that

$$\det|(\mathbf{BS})^2 - \mathbf{I}| = 0$$

the unextended lines **u** given by the equation

$$(\mathbf{BS}\,\mathbf{u})^2 = \mathbf{u}^2 \tag{24}$$

lie in a plane. By adding the appropriate rotation **R**, this undistorted plane becomes an invariant plane, and

$$\mathbf{A} = \mathbf{RBS}$$

is an invariant plane strain. The normal to the invariant or habit plane can be found from the eigenvectors **e** of the symmetrical transformation matrix $\mathbf{A}^2 = (\mathbf{BS})^2$, that is, the solutions of

$$\mathbf{A}^2\mathbf{e} = \lambda^2\mathbf{e} \tag{25}$$

Since $\mathbf{A}^2$ is a symmetrical matrix it has orthogonal eigenvectors, and since $\det|\mathbf{A}^2 - \mathbf{I}| = 0$ one eigenvalue, say λ_2^2 is equal to one, that is, the corresponding eigenvector is an invariant line. If the other two eigenvalues are $\lambda_1^2 < 1$ and $\lambda_3^2 > 1$, a second unextended line can be found between $\mathbf{e}_1$ and $\mathbf{e}_3$ at an angle θ to the eigenvector $\mathbf{e}_1$, given in a form similar to Eq. (21):

$$\tan\theta = \sqrt{\frac{1 - \lambda_1^2}{\lambda_3^2 - 1}} \tag{26}$$

An explicit normalized expression for the normal **n** to the habit plane can be written

$$\mathbf{n} = \sqrt{\frac{\lambda_3^2 - 1}{\lambda_3^2 - \lambda_1^2}}\,\mathbf{e}_3 + \sqrt{\frac{1 - \lambda_1^2}{\lambda_3^2 - \lambda_1^2}}\,\mathbf{e}_1 \tag{27}$$

D. Special Solutions

This equation still requires the solution of the eigenvector equation, Eq. (25), but for two special cases of high symmetry, analytical solutions are available for the fcc → bct transformation in terms of the principal strains η_1 and η_3. For fcc slip on (1 0 1) [1 0 $\bar{1}$], or bcc twinning, the solution is $\mathbf{n} = [h\ k\ l]$ with

$$h = \frac{1}{2\eta_1}(M - N)$$

$$k = \frac{1}{\eta_1}\sqrt{\frac{\eta_1^2 - 1}{1 - \eta_3^2}}$$

$$l = \frac{1}{2\eta_1}(M + N)$$

where

$$M = \sqrt{\frac{\eta_1^2 + \eta_3^2 - 2\eta_1^2\eta_3^2}{1 - \eta_3^2}}$$

and

$$N = \sqrt{\frac{2 - \eta_1^2 - \eta_3^2}{1 - \eta_3^2}}$$

and for simultaneous slip on the two conjugate bcc slip systems (0 1 1) [0 $\bar{1}$ 1] and (0 $\bar{1}$ 1) [0 1 1] the solution is $\mathbf{n} = [h\ k\ l]$ with

$$h = \sqrt{\frac{\eta_1^2(\eta_1^2 - 1)}{2C}}$$

$$k = \sqrt{\frac{\eta_1^2(\eta_1^2 - 1)}{2C}}$$

$$l = \sqrt{\frac{\eta_1^2 - \eta_3^2(2\eta_1 - 1)^2}{C}}$$

where $C = \eta_1^4 - \eta_3^2(2\eta_1 - 1)^2$.

E. Orientation Relationship

The orientation relationship **R** can be described by the rotation axis **w** and rotation angle α:

$$R_{ij} = \delta_{ij}\cos\alpha + w_i w_j(1 - \cos\alpha) - \varepsilon_{ijk}\,w_k\sin\alpha$$

where ε_{ijk} is the permutation symbol. Since **A**, **B**,

and **S** are known, **R** may be back-calculated as

$$\mathbf{R} = \mathbf{A}\mathbf{S}^{-1}\mathbf{B}^{-1}$$

A simpler, more direct way of finding the orientation relationship is based on the fact that the lattice-variant part of the transformation **RB** and its inverse $\mathbf{B}^{-1}\mathbf{R}^{T}$ are invariant line strains as is apparent from

$$\mathbf{RB} = \mathbf{A}\mathbf{S}^{-1} \qquad (28)$$

Both **A** and $\mathbf{S}^{-1}$ are IPS tensors and the two invariant planes must intersect along an invariant line. The invariant line **i** must therefore lie where the slip plane **q** of $\mathbf{S}^{-1}$ intersects the cone of unextended lines of **B**:

$$(\mathbf{Bi})^2 = (\mathbf{i})^2 \quad \text{and} \quad \mathbf{iq} = 0 \qquad (29)$$

A conjugated relation may be written for the inverse transformation valid for plane normals. The unextended normal **m** is given by

$$(\mathbf{B}^{-1}\mathbf{m})^2 = (\mathbf{m})^2 \quad \text{and} \quad \mathbf{mp} = 0 \qquad (30)$$

where **p** is the slip direction. Using Euler's theorem the rotation axis **w** is then simply the cross-product of the rotational strains

$$\mathbf{w} = (\mathbf{B} - \mathbf{I})\mathbf{i}(\mathbf{B}^{-1} - \mathbf{I})\mathbf{m}/D$$

where

$$D = (\mathbf{B} + \mathbf{I})\mathbf{i}(\mathbf{B}^{-1} - \mathbf{I})\mathbf{m}$$

The rotation angle α is

$$\alpha = 2 \arctan |\mathbf{w}|$$

This procedure is illustrated stereographically in Fig. 27 where the cone of unextended lines of **B** (initial cone), solution of Eq. (23) is shown as a solid circle. The transformation **B** moves directions **u** from the initial to the final cone and plane normals **h** in the opposite sense. The traces of the slip plane **q** and of the plane perpendicular to the slip direction **p** are shown as great circles. The possible solutions for the invariant lines **i** [Eq. (29)] and invariant normals **m** [Eq. (30)] are marked as solid and open circles. There are four combinations of invariant lines and normals, each giving a different orientation relationship R. The four corresponding habit planes and magnitudes of the lattice-invariant shear s and the shape shear g are also different. The interface energy is related to s through the density of dislocations necessary to produce an invariant plane whereas the strain energy is related to g, the shape change that remains to be accommodated by the matrix. This is how the geometrical solutions are related to energetic criteria, and the solutions with the smallest values of g and s are physically most realistic.

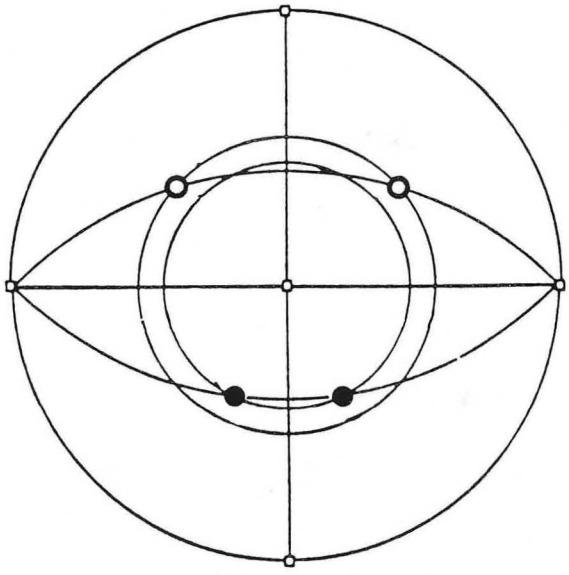


FIG. 27. Stereogram showing invariant lines (solid dots) at intersection of slip plane with initial cone of unextended lines and invariant normals (open circles) at intersection of plane normal to the slip direction with final cone of unextended lines.

Refinements of the basic theory have included a uniform dilatation δ of the habit plane and multiple lattice invariant shear systems

$$\mathbf{A} = \mathbf{RBS}_1\mathbf{S}_2\mathbf{S}_3 \cdots$$

These have improved the match with experimental results in some cases but have done so at the sacrifice of the clarity and relative simplicity of the basic theory.

Electron microscope observations of the interface structure can identify the slip plane(s) of the lattice-invariant shear S. This is easily seen by considering the lattice-invariant part **RB** of the transformation. As shown above this must be an invariant line strain since $\mathbf{RB} = \mathbf{A}\mathbf{S}^{-1}$, with the invariant line at the intersection of the habit plane (interface) with the slip plane. All the dislocations in the interface must lie along this direction so that the shear **S** does not interfere with the invariant line. The density of these dislocations can be measured provided they are not too closely spaced, a condition identical to low-angle grain boundaries.

F. Twinned Martensite

It has been assumed in the theory outlined above the lattice invariant shear S is accomplished by slip. Twinning is an alternative deformation mode. A schematic comparison of slip-

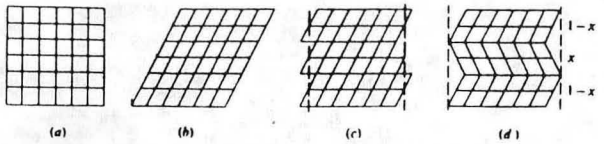


FIG. 28. Lattice-variant deformation (b) of crystal (a) with equal and opposite lattice-invariant deformation by slip (c) and twinning (d).

and twinning-produced lattice-invariant shear is shown in Fig. 28. If two twin-related regions of martensite form simultaneously, their relative volume fraction x determines the net amount of shear produced this way. Algebraically this is written as

$$\mathbf{A} = \mathbf{R}[x\mathbf{B}_1 + (1 - x)\mathbf{R}_2\mathbf{B}_2] \qquad (31)$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are different variants of the Bain strain. One eigenvector of $\mathbf{A}^2$ is made an invariant line by adjusting the volume fraction x appropriately. The remainder of the procedure is as before. Slip and twinning are equivalent modes of deformation and the resulting habit planes and orientation relationships are identical.

The line of reasoning applied in the crystallographic theory of martensite transformations is most apparent for small strains. The decomposition of the macroscopic IPS $\mathbf{A} = \mathbf{RBS}$ where $\mathbf{R}$ is a rigid body rotation, $\mathbf{B}$ the Bain strain (approximately a pure shear), and $\mathbf{S}$ a simple shear becomes

$$\mathbf{A} = (\mathbf{I} + \boldsymbol{\omega})(\mathbf{I} + \boldsymbol{\varepsilon})(\mathbf{I} + \mathbf{s})$$

where $\boldsymbol{\omega}$ is the antisymmetrical matrix of a small rotational strain, $\boldsymbol{\varepsilon}$ the symmetrical matrix of a pure shear strain, and $\mathbf{s}$ the strain associated with a simple shear. Neglecting products of small strains we have

$$\mathbf{A} - \mathbf{I} = \boldsymbol{\omega} + \boldsymbol{\varepsilon} + \mathbf{s} \qquad (32)$$

the rotation, strain, and shear are additive and their order is unimportant (superposition principle). The shear strain $\mathbf{s}$ is chosen so as to reduce one principal strain to zero [equivalent to Eq. (20)], and a second invariant line is found between the other two principal strains by adjusting the orientation relationship $\boldsymbol{\omega}$ accordingly [equivalent to Eq. (28)]. If twinning is the deformation mode, Eq. (31) for small strains reduces to

$$\mathbf{A} - \mathbf{I} = \boldsymbol{\omega} + x\boldsymbol{\varepsilon}_1 + (1 - x)\boldsymbol{\varepsilon}_2 \qquad (33)$$

The procedure is as before: (1) choose the volume fraction x to make one invariant line, and (2) use $\boldsymbol{\omega}$ to produce a second invariant line.

An example of a small-strain martensite transformation with twinning as the lattice-invariant shear is the transformation in In–Tl alloys. The lattice correspondence is obvious since the transformation involves only small distortions. The transformation strain tensor is

$$\boldsymbol{\varepsilon}_1 = \mathbf{B}_1 - \mathbf{I} = \begin{pmatrix} -e & & \\ & -e & \\ & & 2e \end{pmatrix}$$

and for the twin

$$\boldsymbol{\varepsilon}_2 = \mathbf{B}_2 - \mathbf{I} = \begin{pmatrix} -e & & \\ & 2e & \\ & & -e \end{pmatrix}$$

and therefore

$$\mathbf{A} - \mathbf{I} = \boldsymbol{\omega} + x\boldsymbol{\varepsilon}_1 + (1 - x)\boldsymbol{\varepsilon}_2$$

$$= \boldsymbol{\omega} + \begin{pmatrix} -e & & \\ & (2 - 3x)e & \\ & & (3x - 1)e \end{pmatrix}$$

With $x = 2/3$, the transformation strain becomes

$$\mathbf{A} - \mathbf{I} = \boldsymbol{\omega} + \begin{pmatrix} -e & & \\ & 0 & \\ & & e \end{pmatrix}$$

It is apparent that another unextended line can be found midway between the remaining two strains at an angle $\tan \Theta = 1$ from the principal axes. This becomes an invariant line by addition of a small rotation

$$\boldsymbol{\omega} = \begin{pmatrix} 0 & 0 & -e \\ 0 & 0 & 0 \\ e & 0 & 0 \end{pmatrix}$$

which turns the pure shear into a simple shear and which determines the orientation relationship. This geometry is illustrated in Fig. 29a. Note that in this case the shape strain $\mathbf{A} - \mathbf{I}$ is a simple shear parallel to the habit plane. Another twinned martensite plate can be found with the same habit plane but opposite shear direction:

$$\mathbf{A}_1 - \mathbf{I} = \begin{pmatrix} -e & 0 & -e \\ 0 & 0 & 0 \\ e & 0 & e \end{pmatrix}$$

and

$$\mathbf{A}_2 - \mathbf{I} = \begin{pmatrix} e & 0 & e \\ 0 & 0 & 0 \\ -e & 0 & -e \end{pmatrix}$$

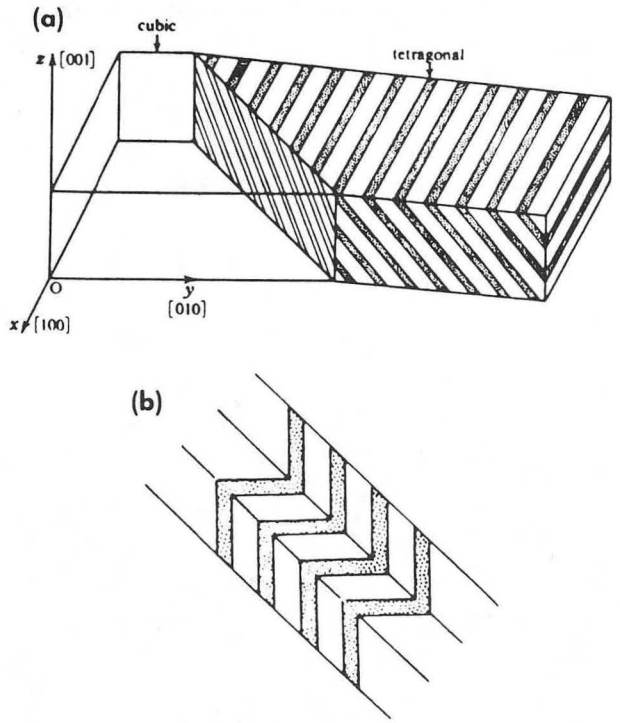


FIG. 29. Twinned martensite: habit plane of a single twinned plate in contact with matrix showing shape strain (a) and four self-accommodating twinned variants with no net shape strain (b). [From A. Kelly and G. W. Groves, "Crystallography and Crystal Defects." Copyright 1970, with permission of Addison-Wesley.]

Adding equal amounts of both variants leads to a vanishing total shape strain. Thus when formed in groups of four variants in the manner shown in Fig. 29b a self-accommodating configuration is achieved of the kind of that is commonly found in shape–memory alloys.

VI. Precipitation

A. Characteristics

Precipitation reactions are first-order phase transformations in which the parent and product phase (matrix and precipitate) coexist at the transformation temperature. Usually the precipitate is different from the matrix in composition, crystal structure, orientation, and atomic volume, causing a barrier to its nucleation and growth. Uniform distributions of fine precipitates are used in precipitation hardening alloys such as the Al–Cu alloy shown in the electron micrograph in Fig. 30. During nucleation and in the early stages of growth precipitates tend to be coherent. They lose coherency during continued growth, a process that is usually accompanied by a loss in alloy strength. The shape and distribution of precipitates depend strongly on such crystallographic factors as the relationship between parent and product lattices, elastic anisotropy, and the mode and mechanisms of accommodation of the new phase in the old. It is difficult to predict exact precipitate shapes with any accuracy and for most purposes it is sufficient to distinguish among three basic shapes: spheres, plates, and needles. The most widely encountered precipitate shape is that of a flat plate such as the particles seen in Fig. 30. The reasons why plates are preferred will become clear in the course of this section. Semicoherent precipitates are derived from coherent ones by introducing dislocations in the interface to relieve the elastic distortions. All three types of precipitates (schematic examples for each type are shown in Fig. 31), coherent, semicoherent, and incoherent, are subject to crystallographic constraints.

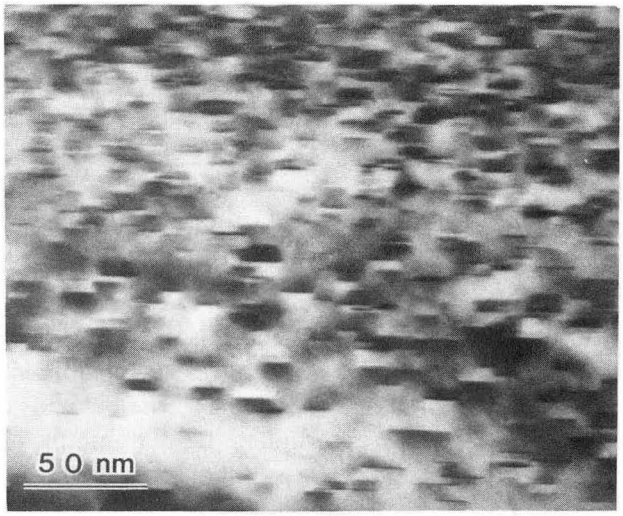


FIG. 30. TEM micrograph of homogeneous distribution of θ'' precipitates in Al–Cu alloy.

In this section we will first examine the factors that determine the optimum shape of a coherent precipitate (also called inclusion), and then describe a simple physical model for the loss of coherency in plates. It is possible to make a qualitative prediction of the shape of an inclusion and a quantitative prediction of its orientation. Examples of applications to real alloys will be given.

B. Elastic Inclusions

The theory of coherent elastic inclusions has been worked out in some detail and in this section some of its results are presented along with the underlying physical reasoning.

The strain energy of coherent inclusions in an

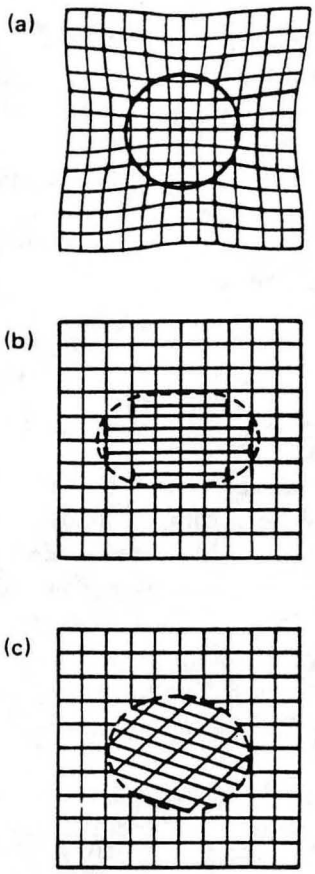


FIG. 31. Coherent (a), semicoherent (b), and incoherent (c) inclusion.

elastic solid is a function of inclusion shape and orientation. To find the optimum shape and orientation of an inclusion transforming under elastic constraint from a solid matrix is the main objective of the strain energy approach. Without constraint the inclusion would undergo the stress-free transformation strain ε_{ij}. On the other hand, if the matrix were infinitely rigid it would constrain the inclusion completely back to the shape it had before transforming, reversing the transformation strain elastically and setting up a stress σ_{ij}. The work W (per unit volume) done in this process is

$$W = \tfrac{1}{2}\sigma_{ij}\varepsilon_{ij}$$

(Throughout Einstein's summation convention is implied for repeated indices.) Using Hooke's law:

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl}$$

where c_{ijkl} are the elastic constants of the inclusion, the strain energy becomes

$$W = \tfrac{1}{2}c_{ijkl}\varepsilon_{ij}\varepsilon_{kl}$$

Note that at this stage the matrix is strain free and all the strain energy is contained in the inclusion. However, since in real materials the matrix is not rigid it will not constrain the inclusion completely. Instead, part of the transformation strain will be accommodated by elastic deformation of the matrix. As a result the strain energy is no longer contained entirely in the precipitate but is divided between the work done in constraining the precipitate partially and the work necessary to accommodate part of the transformation strain in the matrix. The hypothetical sequence of cutting, straining, and welding operations shown in Fig. 32 illustrates this point.

1. Make a cut around a volume within the matrix (a), remove it (b), and transform it stress free (c) (ε_{ij} is the transformation strain, $\sigma_{ij} = 0$).
2. Apply a stress σ_{ij} that results in an elastic strain $-\varepsilon_{ij}$, reversing the transformation strain elastically (d).
3. Replace and reweld the inclusion (e) and relax the applied stress (f). The first step corresponds to the extreme case of a transformation in an infinitely soft matrix; since $\sigma_{ij} = 0$ it causes no strain energy. The second step corresponds to the opposite extreme of transformation in an infinitely rigid matrix. The strain energy is $W = \frac{1}{2}\sigma_{ij}\varepsilon_{ij}$. In the third step the matrix is given more realistic elastic properties resulting in a total strain energy between zero and $\frac{1}{2}\sigma_{ij}\varepsilon_{ij}$. If the matrix is elastically identical to the precipitate, the inclusion is called homogeneous. If the elastic constants of matrix and precipitate are different, the inclusion is called heterogeneous.

The degree to which the inclusion is accommodated in the matrix depends on its shape and orientation. This is physically most apparent for an incoherent inclusion that, by definition, exerts only a hydrostatic pressure (no shear stresses), for example, an amorphous precipi-

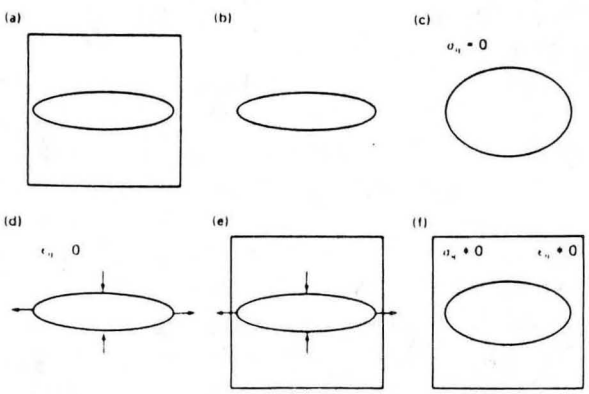


FIG. 32. Hypothetical sequence of operations in the formation of an elastic inclusion.

tate or a gas bubble. Hence, we will first consider the simplest case of an incoherent inclusion.

C. Incoherent Precipitates

For an incoherent inclusion no atomic correspondence is maintained across the interface. The strain energy is thus caused only by the volume difference between the inclusion and the hole it occupies. This energy depends on the shape of the hole. In an elastically isotropic material a spherical hole will yield uniformly in all directions. A disk-shaped hole on the other hand will yield unevenly. The flat faces are displaced easily under the internal pressure set up by the volume difference while the radial displacement is small. When the disk is very thin the entire volume change is accommodated by yielding of the flat faces. As a result a very thin incoherent precipitate is essentially stress free and therefore causes essentially no strain energy. To a good approximation the strain energy per unit volume of an incoherent plate is proportional to its aspect ratio c/a.

Thus of the three main possible shapes (spheres, plates, and needles), plates will always be preferred because they are easily accommodated. The smaller the aspect ratio the smaller is the strain energy and the flattening of a given incoherent precipitate is limited only by the increasing surface energy. A typical example of an incoherent inclusion is amorphous SiO_2 forming in Si.

D. Coherent Precipitates

Most applications of the strain energy approach deal with coherent rather than incoherent inclusions. Under constraint, incoherent inclusions are always in a state of hydrostatic stress whereas coherent inclusions can support shear stresses. This makes it necessary to distinguish between different types of stress-free transformation strains for coherent precipitates. These are best characterized by their principal strains e_1, e_2, and e_3. For many transformations this is self-evident. A cubic to tetragonal distortion, for example, is completely described by elongations or contractions along the three orthogonal cube axes. The strain tensor is thus automatically given in the diagonal form.

Other transformations are described by a general (nondiagonal) strain tensor e_{ij} which can always be decomposed into an antisymmetrical part $\omega_{ij} = \frac{1}{2}(e_{ij} - e_{ji})$ and a symmetrical part $\varepsilon_{ij} = \frac{1}{2}(e_{ij} + e_{ji})$. For small strains ω_{ij} represents a rigid body rotation and does not contribute to the strain energy. The symmetrical part ε_{ij} (the pure strain) can always be referred to three orthogonal principal axes (its eigenvectors). In this coordinate frame ε_{ij} is diagonal with the principal strains e_1, e_2, and e_3 (the eigenvalues) as the diagonal elements:

$$\varepsilon_{ij} = \begin{pmatrix} e_1 & & \\ & e_2 & \\ & & e_3 \end{pmatrix}$$

The three principal strains are all different in the general (orthorhombic) strain tensor. If two strains are equal, the strain tensor is cylindrically symmetric (sometimes called tetragonal), and if all three strains are equal the spherically symmetric strain tensor represents a pure dilatation. In addition, the strains can be all of the same sign or of mixed sign. This yields six cases that will be considered individually below.

The solution to the inclusion problem would be simple if the strain energy caused by a coherent inclusion were a simple analytical function of the principal strains and the variables "shape" and "orientation." Unfortunately no such function exists for the general case. However, for the special case of an oblate spheroidal inclusion (a flat ellipsoid of revolution) undergoing a stress-free strain (e_1, e_2, e_3) along its axes a, a, and c, an analytical expression that includes the shape effect explicitly through the aspect ratio c/a has been derived. The three shapes of plate, sphere, and needle can be approximated by $c/a < 1$, $c/a = 1$, and $c/a > 1$, respectively. The inclusion orientation in this case is fixed through the assumption that the axes of the spheroid are aligned with the principal axes of the strain. For this case, the strain energy takes the form $W = f(e_1, e_2) + g(e_1, e_2, e_3)c/a$.

It can be shown that for a pure dilatation ($e_1 = e_2 = e_3$) the total strain energy is independent of the shape: $g(e_1, e_2, e_3) = 0$. For a plate, the contribution of the strain e_3 normal to the plate disappears as the aspect ratio c/a approaches zero. Physically this means that the strain e_3 normal to the plane of the plate is accommodated well, as in the incoherent case, and a very thin plate is essentially free to expand in thickness without setting up a stress. An example of this is an interstitial dislocation loop that may be regarded as an inclusion plate undergoing an ex-

pansion e_3 with $e_1 = e_2 = 0$. The extra plane is incorporated with the same plane spacing as the equivalent planes in the matrix, that is, accommodation is essentially complete.

On the other hand, the strains e_1 and e_2 in the plane of the plate are constrained by the matrix, setting up large stresses and as the aspect ratio c/a goes to zero, the shape-independent term of the strain energy $f(e_1, e_2)$ remains finite.

These results may be summarized as follows: a plate-shaped inclusion with very small aspect ratio is stressed in its plane and essentially stress free normal to its plane (a state of plane stress). The strain energy in this situation is caused only by the principal strains e_1 and e_2 in the plane, as indicated by the 2×2 block in the upper left of the strain tensor

$$\varepsilon_{ij} = \begin{pmatrix} e_1 & & \\ & e_2 & \\ & & e_3 \end{pmatrix}$$

This corresponds to the case of the inclusion axes parallel to the principal strains.

Consider the more general situation with the inclusion not parallel to the principal strains. Referred to the inclusion frame the strain then takes the form of a general symmetrical tensor

$$\varepsilon_{ij} = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{pmatrix}$$

The same reasoning used before to illustrate easy accommodation of the expansion $\varepsilon_{33} = e_3$ for incoherent and coherent plates implies that the shears ε_{13} and ε_{23} are also easily accommodated by the matrix and thus contribute little to the strain energy of a plate. A shear dislocation loop is perhaps the best example to illustrate that almost all the strain (and hence strain energy) is located in the matrix outside the loop (or inclusion). Therefore the strain components ε_{11}, ε_{22}, and the shear ε_{12} in the plane of the plate (indicated by a block in the tensor ε_{ij}) dominate the strain energy. The approach used in the following sections seeks to minimize only the dominant part of the strain energy.

As seen above the strain energy of a homogeneous inclusion in the case of a pure dilatation is independent of the inclusion shape. For a spherical precipitate ($c/a = 1$) the total strain energy is only approximately one-third of the work that would be necessary to compress the particle elastically as in a rigid matrix. Thus the fact that the transformation strain in the particle is partially accommodated by elastic distortion of the matrix reduces the total strain energy by approximately two-thirds. Of this total strain energy one-third resides in the spherical particle and two-thirds in the matrix. This distribution of the strain energy between matrix and precipitate is reversed as the inclusion becomes flattened ($c/a \to 0$). Hence, as shown in the previous section, most of the strain energy resides in the inclusion, for a thin plate. Consider now an inclusion with elastic constants different from those of the matrix, that is, a heterogeneous inclusion. Having established that in the homogeneous case a spherical inclusion contains one-third of the strain energy, let one phase become slightly softer than the other. Clearly the energy savings are greater if the matrix softens since it carries two-thirds of the total strain energy. This is reversed for plates in which the strain energy reduction is larger if the particle softens rather than the matrix. Therefore the total strain energy of a heterogeneous inclusion is minimized if its major fraction in the corresponding homogeneous case is located in the softer phase: if the precipitate is elastically softer than the matrix it should be a plate, but if it is harder than the matrix a sphere would be preferred.

Extending this analysis to the case of anisotropic elasticity is also straightforward. As shown above, a thin plate is in a state of plane stress which causes the major fraction of the total strain energy. If the plane of the stress is parallel to an "elastically soft plane" of the crystal the strain energy is smaller than for any other orientation. It can be shown that the strain energy of an elastically anisotropic plate inclusion on a plane $\{h\ k\ l\}$ depends on a term $(1 - A)(h^2k^2 + k^2l^2 + l^2h^2)$, where $A = 2c_{44}/(c_{11} - c_{12})$ is the Zener anisotropy ratio. The strain energy is then minimized for $\{1\ 0\ 0\}$ plates if A is larger than unity, as, for example, in all fcc metals. Conversely $\{1\ 1\ 1\}$ plates are preferred if $A < 1$. Since a very thin plate leaves the matrix essentially unstressed the strain energy is independent of the matrix elastic constants.

A heterogeneous anisotropic inclusion should form a plate if it is softer than the matrix. The only difference to the case of isotropic elasticity is that now two elastic shear moduli must be compared with that of the matrix. Thus an anisotropic inclusion with shear moduli $\frac{1}{2}(c_{11} - c_{12})$ and c_{44} in an isotropic matrix with shear modulus μ will form a plate if $\min[(\frac{1}{2}c_{11} - c_{12}), c_{44}] < \mu$. Otherwise a sphere is preferred.

In the previous section it was shown how the elastic strain energy of a coherent precipitate undergoing a pure dilatation can be minimized through an inclusion shape and orientation that concentrates the major part of the strain energy in the softer phase. The high symmetry of the dilatational transformation strain allows a degree of freedom in the orientation of the precipitate which is lost in the case of a cylindrical strain ($e_1 = e_2 \neq e_3$). The precipitate will be a plate in the plane containing the two equal strains e_1, if $e_1 < e_3$ and a needle if $e_1 > e_3$.

Consider a cylindral strain with mixed signs ($e_3/e_1 < 0$), for example, an expansion e_3 and a contraction e_1, in the plane normal to it. The strain energy of a thin plate on the plane normal to e_3 is dominated by the uniform strain e_1 in this plane. However, another orientation can be found in which the strain energy is still lower, if the plate is inclined to contain a direction in the interface in which the transformation strain is zero. Such a direction must exist since the strain changes sign between e_1 and e_3. The angle θ of the unextended line with the (0 0 1) plane is given by

$$\tan \theta = \sqrt{-e_1/e_3} \tag{34}$$

If a thin disk takes this orientation it is elastically distorted in its plane only by a unaxial extension $-e_1$ constraining the transformation strain e_1, while the direction normal to it is free of transformation strain (unextended line). This is the geometrical condition of an invariant (unextended and unrotated) line in the interface as it is used for instance in martensite theory [Eqs. (21) and (26)].

In an isotropic crystal the uniaxial tension constraining the plate could be in any direction in the (0 0 1) plane containing the uniform strain e_1. In an anisotropic cubic crystal this stress would be expected to lie in a soft direction in the (0 0 1) plane. Thus the elastic distortion of the plate causes minimum strain energy if it occurs in the ⟨1 0 0⟩ direction for $A > 1$ and the ⟨1 1 0⟩ direction for $A < 1$. The resulting habit plane will contain this "tensile axis" and lie at an angle θ to the (0 0 1) plane. This leads to $\{0\ k\ l\}$ habits for $A > 1$ and $\{h\ h\ l\}$ habits for $A < 1$.

Orthorhombic strains ($e_1 \neq e_2 \neq e_3$) with unmixed strains will lead to precipitates whose dimensions tend to be inverse to the transformation signs. The elastic behavior plays only a secondary role in determining the shape, and the orientation is entirely given by the principal strains.

In the case of orthorhombic strains with mixed signs the same principles as for the cylindrical strain apply. A plate should be inclined to have a stress-free line in the interface and most of the strain energy is caused by the uniaxial stress in the plane of the plate. Thus the smaller one of the two principal strains with equal sign will lie in the plate being elastically constrained and the stress-free line is formed between the remaining two principal strains. This determines the orientation uniquely and elastic anisotropy plays only a secondary role. The angle θ of inclination is given by Eq. (34).

E. Loss of Coherency

Beyond a critical size it becomes energetically favorable for a coherent particle to lose coherency. This size depends on the difference between the interface energy of a semicoherent inclusion (proportional to the surface area) and the strain energy of a coherent inclusion (proportional to the volume). At the critical size the accommodation of the transformation strains switches from an elastic to a plastic mode. A simple example of this transition is the punching out of prismatic dislocation loops. Interstitial loops are punched into the matrix leaving behind vacancy loops in the interface (or vice versa). A TEM micrograph of rows of punched loops around a He gas bubble in Nb–Zr alloy is shown in Fig. 33a. The corresponding schematic (Fig. 33b) emphasizes the fact that a spherical inclusion will punch out loops in all crystallographically equivalent directions, for example, in the case of Nb–Zr in all ⟨1 1 1⟩ directions. Such rows of punched prismatic loops will leave behind an equal number of opposite dislocation loops in the interface and render the particle semicoherent while accommodating the volume component of the transformation strain.

In order to accommodate the shear component of the transformation strain, shear loops are more efficient since they need not be generated in pairs. As a result the shape change is easier to accommodate than the volume change. Martensite plates and many semicoherent precipitates can be modeled successfully on the assumption that an array of shear dislocations accommodates part of the transformation strain. A schematic of a precipitate plate containing shear loops in its interface is shown in Fig. 34. The crystallographic constraints in this case are less apparent than in the case of a spherical inclusion under a pure dilatation (Fig. 33). One such con-

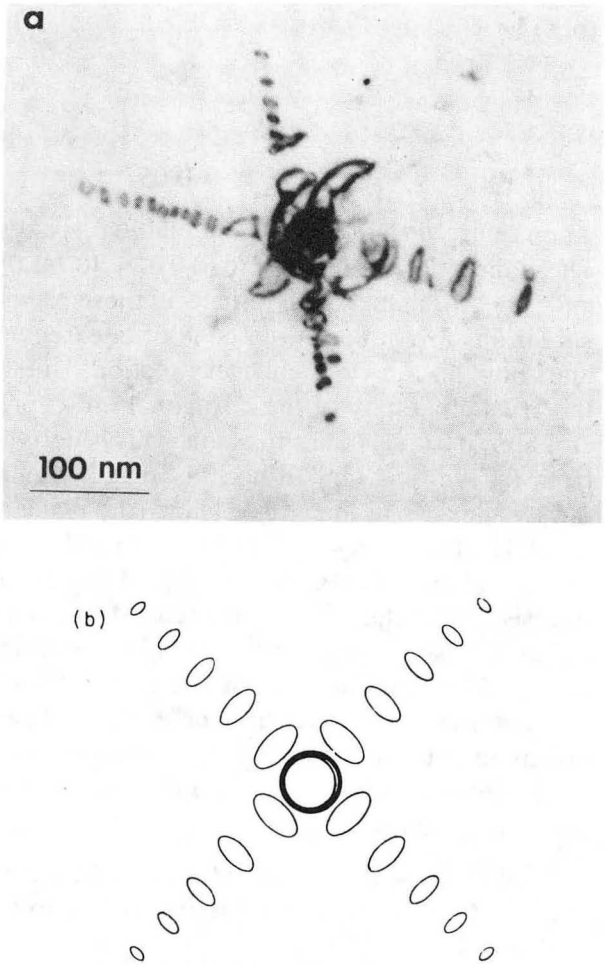


FIG. 33. Punched-out dislocation loops due to volume strain: TEM micrograph (a) and schematic (b) of punched loops around a helium bubble in Nb–Zr alloy. (Micrograph C. Echer.)

straint is that the loop plane must be a slip plane and contain the Burgers vector; but it is not immediately clear which of the crystallographically equivalent loop arrays are preferred for a given plate and what will be its habit plane.

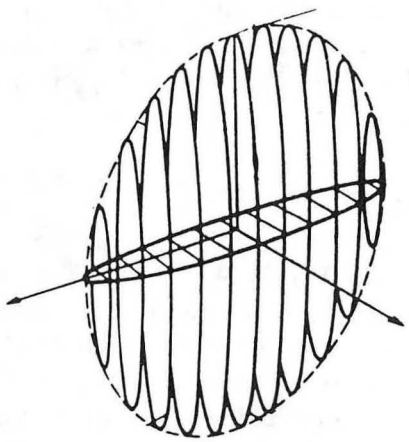


FIG. 34. Schematic of lens-shaped precipitate with array of shear loops in the interface.

An answer to these questions is provided by the theory of martensite transformations which gives an algebraic solution for the habit plane and orientation relationship if the slip plane and direction and of course the transformation strain are known. A more physical description on the loss of coherency and resulting interface structure based of the analogy with slip in a tensile test is summarized below.

Consider a thin coherent plate with mixed principal strains. As outlined above it will be in a state of uniaxial stress if it is oriented for minimum strain energy, that is, such that it contains an invariant line in the interface.

If it deforms plastically it is most likely to do so on the slip system with the highest resolved shear stress. This depends on the relative position of the tensile axis and the slip system; and the resolved shear stress is a maximum if both the slip plane and the slip direction make an angle of 45° with the tensile axis. The problem in the elastic case of finding a direction with a low Young's modulus in the plane of unmixed strains now translates to finding the tensile direction in this plane with the highest Schmid factor $R = \cos\phi \cos\lambda$ for a crystallographic slip system (where ϕ and λ are the angles of the tensile axis with the slip direction and the slip plane normal, respectively). As a first approximation it may be assumed that the Schmid factor is maximized when the tensile axis makes an angle of $\lambda = 45°$ with the slip plane. For slip on the (1 0 1) plane this direction is [1 0 0], as shown in the stereogram in Fig. 35a. The invariant line (open circle) must be located at the intersection of the cone of unextended lines with the slip plane so that it remains unaffected by the plastic deformation, a condition well known from martensite theory. The resulting habit plane (pole marked by star) contains the tensile axis and the invariant line. Applications of these principles to coherent and semicoherent precipitates are given below.

F. Applications

The transformation strain for α'' $Fe_{16}N_2$ (bct) in Fe (bcc) is cylindrically symmetric and

$$\varepsilon_{ij} = \begin{pmatrix} -0.0023 & & \\ & -0.0023 & \\ & & 0.0971 \end{pmatrix}$$

Since the anisotropy ratio A is greater than unity for Fe the habit plane predicted for α'' is (0 k l)

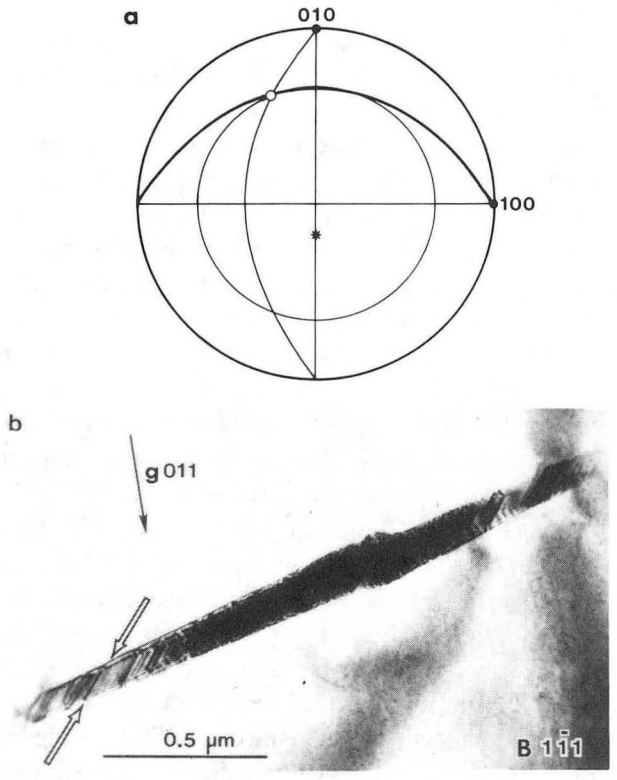


FIG. 35. γ' precipitate in Fe–N alloy: stereogram illustrating predicted habit plane (a) and corresponding TEM image showing inclined γ' plate with interface striations along invariant line direction (b).

with $k/l = \sqrt{0.0023/0.971}$, a plane that is inclined about 8.5° with respect to (0 0 1). This inclination is in agreement with experimental observations.

As for α'' $Fe_{16}N_2$, the transformation strain for γ' Fe_4N (fcc iron sublattice) in Fe (bcc) has cylindrical symmetry; with $a_{bcc} = 0.28678$ nm, $a_{fcc} = 0.3795$ nm, we obtain

$$\varepsilon_{ij} = \begin{pmatrix} -0.0643 & & \\ & -0.0643 & \\ & & 0.3233 \end{pmatrix}$$

These strains are large and the habit plane must be determined by the criterion of maximum Schmid factor. γ' slips on {1 1 1} fcc planes derived from the inclined {1 0 1} bcc slip planes as shown stereographically in Fig. 35a. The predicted (0 k l) habit plane defined by the [$\bar{4}$ 9 4] invariant line and the [1 0 0] tensile axis is (0 $\bar{4}$ 9), in agreement with the experimental observations. Figure 35b shows a TEM image of a typical γ' plate. The parallel striations are interfacial dislocations along the invariant line direction.

With the lattice parameters $a_{bcc} = 0.3147$ nm, $a_{hcp} = 0.3002$ nm, $c_{hcp} = 0.4724$ nm, the transformation strain tensor for Mo_2C (hcp) in Mo (bcc) is

$$\varepsilon_{ij} = \begin{pmatrix} 0.17 & & \\ & 0.06 & \\ & & -0.05 \end{pmatrix}$$

when referred to [1 1 0], [1 $\bar{1}$ 0], and [0 0 1] axes. These directions transform to [0 0 0 1], [0 1 $\bar{1}$ 0], and [2 $\bar{1}$ $\bar{1}$ 0], respectively. Since all these principal strains are different the cone of unextended lines now has an elliptical cross section. Due to the crystallography of the hexagonal Mo_2C precipitate, (0 0 0 1) is the slip plane. This is derived from the vertical (1 1 0) bcc slip plane. Again, the tensile axis in the (0 0 1) plane (the plane of unmixed strains) at 45° to the (1 1 0) bcc/(0 0 0 1) hcp slip plane is [1 0 0]. The invariant line at the intersection of the cone of unextended lines with the (1 1 0) slip plane is ~[$\bar{1}$ 1 3]. The resulting (0 k l) habit plane containing the [1 0 0] tensile axis and the [$\bar{1}$ 1 3] invariant line is (0 $\bar{3}$ 1), again in agreement with experimental observations.

The main results of this section can be summarized as follows:

1. For principal strains of unmixed sign the precipitate dimensions tend to be inverse to the strains.
2. For principal strains of mixed sign the interface is oriented so as to include an invariant line.
3. For coherent precipitates remaining degrees of freedom for shape and orientation are used to concentrate the major part of the equivalent homogeneous isotropic strain energy in the softer phase or orientation.
4. For plate-shaped inclusions the tensile axis must be a soft direction in the plane of unmixed strain [in our examples the (0 0 1) plane]. For coherent precipitates the tensile axis must be elastically soft, that is, have a low elastic modulus, and for semicoherent precipitates it must plastically soft, that is, be oriented for the maximum Schmid factor on the slip system.

G. Mechanisms

The mechanism of a phase transformation depends upon thermodynamic, kinetic, and structural parameters. If at a given temperature the parent phase is unstable with respect to small fluctuations of a physical quantity such as composition or atomic displacement then no barrier to nucleation exists and the transformation can proceed by the gradual amplification of some

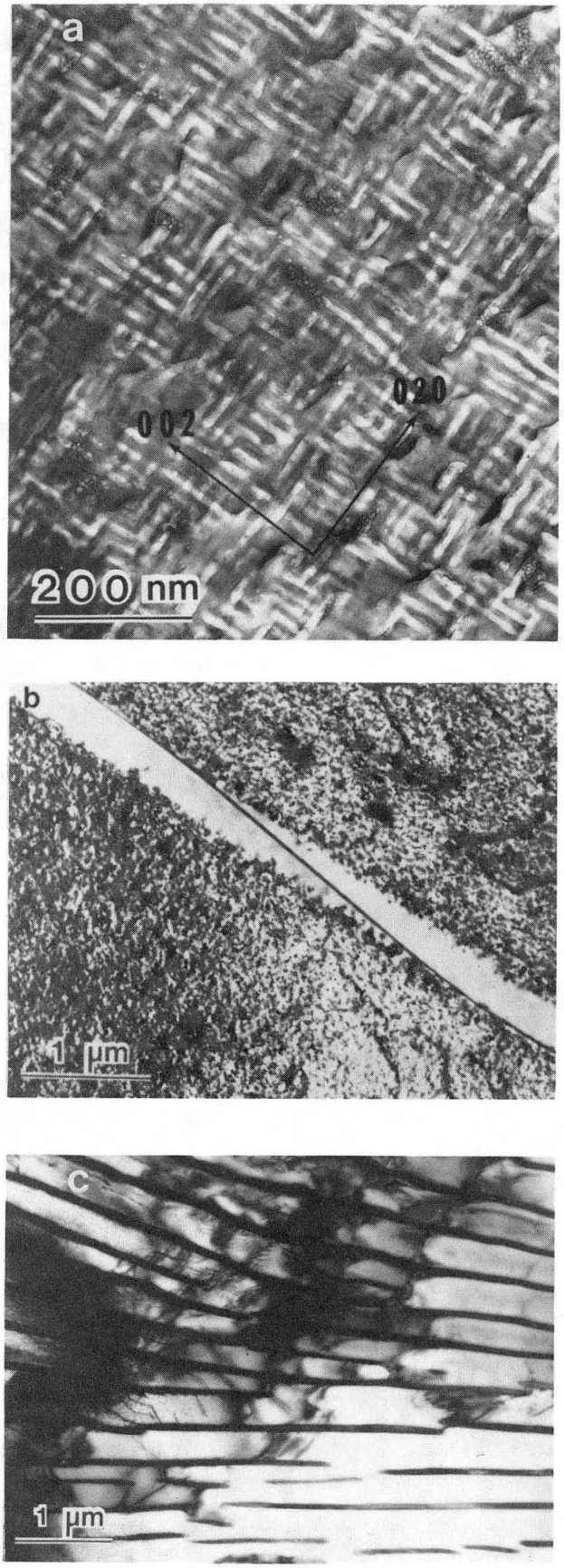


FIG. 36. TEM micrographs showing microstructures typical for different mechanisms of decomposition: spinodal decomposition in Cu–Ni–Fe alloy (a) (micrograph K. Kubarych), nucleation and growth in Al–Si alloy (b), and eutectoid reaction in Fe–C alloy (c).

fluctuations to form a modulated structure. Because of their barrier-free nucleation such structures are homogeneous and independent of preexisting microstructures and defects. This is illustrated in Fig. 36a with a Cu–Ni–Fe alloy which has undergone spinodal decomposition. On the other hand, when a barrier to nucleation exists the resulting two-phase microstructure depends on the type and distribution of defects available to help overcome the barrier. This situation is more commonly encountered than that of modulated structures and usually leads to less homogeneous microstructures such as the Al–Si alloy shown in Fig. 36b. Vacancies are necessary in this alloy to overcome the nucleation barrier due to the large volume increase on precipitation. As a result precipitate-free zones develop in regions depleted of vacancies such as grain boundaries, surfaces, or dislocations. Defects other than vacancies can also lead to preferred nucleation. Dislocations, stacking faults, grain boundaries, and other defects have all been found to provide heterogeneous nucleation sites. If the alloy decomposes only along an advancing interface, usually starting from a grain boundary, the reaction is called discontinuous, as opposed to the continuous reactions described above. Discontinuous reactions lead to characteristic cellular morphologies. Similar cellular morphologies are also observed in eutectoid decomposition (see Fig. 36c) where both reaction products are different in structure from the parent phase. If no suitable defects are available to allow precipitation of the equilibrium phase it is often found that intermediate metastable phases form with a structure and composition approximating that of the stable equilibrium phase. In fact the properties of most precipitation hardening alloys are based on such metastable phases. Structurally, the mechanism of a phase transformation depends, therefore, not only on preexisting microstructure and defects but also on available intermediate phases.

Acknowledgments

This work is supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract No. De-AC03-76SF00098.

Bibliography

Aaronson, H. I., Laughlin, D. E., Sekerka, R. F., and Wayman, C. M. (eds.) (1982). Proc. Conf. on "Solid-Solid Phase Transformations." AIME.

Boccara, N. (ed.) (1981). "Symmetries and Broken Symmetries in Condensed Matter Physics." IDSET, Paris.

Christian, J. W. (1975). "The Theory of Transformations in Metals and Alloys," Part I, 2nd ed. Pergamon Press, Oxford.

Hahn, T. (ed.) (1983). "International Tables for Crystallography," Vol. A., D. Reidel, Boston.

Kelly, A., and Groves, G. W. (1970). "Crystallography and Crystal Defects." Addison-Wesley, London.

Khachaturyan, A. G. (1983). "Theory of Structural Transformations in Solids." John Wiley, New York.

Nye, J. F. (1957). "Physical Properties of Crystals." Oxford University Press, New York.

Shubnikov, A. V. (1974). "Symmetry in Science and Art." Plenum, New York.

Wayman, C. M. (1964). "Introduction to the Crystallography of Martensitic Transformations." Macmillan, New York.

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720