CHAPTER 2

Formal Geometry of Crystal Lattices

5. DESCRIPTION OF THE IDEAL CRYSTAL

Amorphous and quasi-crystalline forms of solid have been much investigated in recent years, but most solid metals are crystalline and some appreciation of crystallography is essential to a study of metallic transformations. The scientific concept of a crystal has evolved gradually from the original classification by external shape to modern views on the internal atomic arrangement. The recognition that the distinguishing feature of crystalline solids lies in their regular internal arrangement led to a description which Zachariasen (1945) has termed the macroscopic concept of a crystal. The macroscopic crystal is defined in terms of physical properties which have precise meaning (or, at least, are measurable) only for regions containing appreciable numbers of atoms. Such properties are invariant with respect to a translation within an infinite crystal, but (except for scalar properties) not with respect to a rotation. A crystal is thus a homogeneous, anisotropic solid; a noncrystalline, or amorphous, solid is both homogeneous and isotropic.

The development of X-ray methods enabled the structure of a crystal to be investigated on a finer scale. It was then found that crystals are not truly homogeneous, but the arrangement of atoms is periodic in three dimensions. This is the familiar modern picture, which we shall take as our starting point, but we emphasize here that it is only an abstraction from the much less ordered situation in a real crystal. In recent years, attention has been directed especially to the imperfections in real crystal structures. These imperfections represent comparatively small deviations from the mathematical concept of an ideal crystal, but they nevertheless control many of the most important physical properties. In this chapter we are concerned only with the ideal crystal; the nature of the approximations involved in this description, and the extent to which a real crystal may be considered to be an ideal crystal containing imperfections, will be considered in detail later.

The ideal crystal may be regarded as the repetition in three dimensions of some unit of structure, within which the position of each atom is specified exactly by a set of spatial coordinates. Let us choose an origin within the crystal. This will be one of an infinite set of points, each possessing an identical configuration of surrounding atoms. The positions of these points may be represented by the vectors

$$\mathbf{u} = u_i \mathbf{a}_i \qquad (i = 1, 2, 3), \tag{5.1}$$

where u_i have only integral values, and the summation convention is used. The translation between any two lattice points is a lattice vector, and the three non-coplanar lattice vectors

$\mathbf{a}_i$ outline a parallelepiped known as the unit cell. Any other set of lattice vectors, formed from linear combinations of the set $\mathbf{a}_i$, may also be used as basic vectors, so there is an infinite number of possible unit cells. The volume of the unit cell is given by the scalar triple product of the vectors which outline it; when this volume is a minimum the vectors are primitive vectors, and define a unit cell of the Bravais lattice. Each such primitive cell contains only one lattice point; cells of larger volume contain two or more lattice points. The primitive unit cell may still be chosen in an infinity of ways, since we have placed no restriction on its shape. The most useful unit cell is usually that in which the vectors $\mathbf{a}_i$ are as nearly as possible of equal scalar magnitude.

The quantities u_i of eqn. (5.1) give the position of the end point of $\mathbf{u}$ in an oblique Cartesian coordinate system, in which distances along the axes are measured in multiples of the lengths of the basic vectors $\mathbf{a}_i$. Such a coordinate system forms a natural framework for representation of the crystal lattice, and is generally preferable to the alternative method of using coordinate axes parallel to the vectors $\mathbf{a}_i$, but having unit measure lengths. In the latter system, the coordinates of a lattice point are $u_{(i)}\,|\,\mathbf{a}_{(i)}|$ (the brackets indicating suspension of the summation rule), and these are sometimes called the physical components of the vector $\mathbf{u}$. However, it is often convenient to use a coordinate system with orthogonal axes of equal base lengths, and this is called an orthonormal system. It is defined by the three unit vectors $\mathbf{i}_i$, which satisfy the relations

$$\mathbf{i}_i \cdot \mathbf{i}_j = \delta_{ij}, \tag{5.2}$$

where δ_{ij}, called the Kronecker delta, is equal to unity when $i = j$, and to zero when $i \neq j$. In terms of the orthonormal system, a vector $\mathbf{x}$ may be written as

$$\mathbf{x} = x_i \mathbf{i}_i. \tag{5.3}$$

In tensor analysis or matrix algebra,† the vector $\mathbf{u}$ is regarded as the array of numbers u_i, which form its components. We may, for example, write $\mathbf{u}$ as a single row matrix $(u_1\ u_2\ u_3)$ or as a single column matrix $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$, and these arrays would conventionally be given the same symbol as $\mathbf{u}$. This sometimes leads to confusion, and it is desirable to have a way of distinguishing between the vector $\mathbf{u}$, which is a physical entity, and its matrix representation, which depends on a particular coordinate system. We shall do this by using a notation in which **bold face type** is used for symbols representing physical quantities (vectors and tensors), and sans serif type is used for their matrix representations. In addition, we specify that column matrices may be written $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = [u_1\, u_2\, u_3]$ for convenience, whilst row matrices

† The formal theory of this chapter may be expressed in either tensor or matrix notation. We find it more convenient to use the latter, but many of the equations will be given in both forms to facilitate reference to other books and papers. A brief description of the more important features of tensor notation is given on pp. 35-8.

are written $(u_1\, u_2\, u_3)$. We thus have

$$\mathsf{u} = [u_1\, u_2\, u_3] \qquad \mathsf{x} = [x_1\, x_2\, x_3], \tag{5.4}$$

where u and x are representations of the vectors **u** and **x**. Wherever possible we shall keep to the practice that a letter such as **u** (and later **v**) will represent a vector which is most conveniently expressed in a lattice coordinate system, while a letter such as **x** (and later **y**) will represent a vector which is usually referred to an orthonormal system.

Any matrix of n rows and m columns may be formed into a new matrix of m rows and n columns (the *transposed* matrix) by interchanging the rows and columns. Clearly, the single row and single column matrix representations of **u** are the transposes of each other, so that eqn. (5.4) also implies the notation

$$\mathsf{u}' = (u_1\, u_2\, u_3) \qquad \mathsf{x}' = (x_1\, x_2\, x_3), \tag{5.5}$$

where u′ is the symbol for the transposed matrix formed from u.

Later in this chapter we need to refer the vector **u** to other sets of coordinate axes. We define a second set of axes by the base vectors $\mathbf{b}_i$, and the vector **u** then has a different matrix representation. We may distinguish between the two representations when required by describing the set of vectors $\mathbf{a}_i$ as the basis A, and the set of vectors $\mathbf{b}_i$ as the basis B. The column matrix representations are then written as ${}^{A}\mathsf{u} = [{}^{A}u_1\, {}^{A}u_2\, {}^{A}u_3]$ and ${}^{B}\mathsf{u} = [{}^{B}u_1\, {}^{B}u_2\, {}^{B}u_3]$. Symbols of this type are sufficient for most purposes, but occasionally extra clarity is achieved by use of an extended notation, such as that used by Bowles and Mackenzie (1954). In the extended notation, the representations of **u** as column matrices are

$${}^{A}\mathsf{u} = [\mathrm{A};\, \mathbf{u}] \qquad {}^{B}\mathsf{u} = [\mathrm{B};\, \mathbf{u}], \tag{5.6}$$

and the corresponding row matrices are

$${}^{A}\mathsf{u}' = (\mathbf{u};\, \mathrm{A}) \qquad {}^{B}\mathsf{u}' = (\mathbf{u};\, \mathrm{B}) \tag{5.7}$$

in which both the round brackets and the reversal of the order of the vector symbol **u** and the base symbol A or B signify the tranposition of the column matrices ${}^{A}\mathsf{u}$, ${}^{B}\mathsf{u}$. The same notation is applied to vectors referred to an orthonormal basis, for which we shall usually use the symbol I. Hence two different vectors **u** and **x** have representations

$${}^{I}\mathsf{u} = [\mathrm{I};\, \mathbf{u}] \qquad {}^{I}\mathsf{x} = [\mathrm{I};\, \mathbf{x}]$$

in such a basis.

The advantages of the extended notation will become apparent later, but it is often sufficient to use the sans serif symbols. When no confusion about the bases is possible, the identifying superscripts will be omitted, as in eqn. (5.4).

A clear distinction must be made between the lattice points of the unit cell, and the positions of the atoms within the cell. The simplest types of crystal structure are obtained by placing an atom at each point of the lattice, and this category includes two of the three common metallic structures. More generally, the primitive Bravais lattice only gives the interval over which the unit of pattern, or motif unit, is repeated. This unit may be a

single atom or a more complex atom group; in the latter case, the structure is said to have a basis. The repeating properties of the lattice then require that if an atom of kind A is situated in a given position with respect to the origin, a similar atom is similarly situated with respect to each of the lattice points.† If there are r atoms of this kind within the unit cell, their positions with respect to the origin of the cell may be specified by the vectors $\boldsymbol{\xi}_{A,1}$, $\boldsymbol{\xi}_{A,2}$, ..., $\boldsymbol{\xi}_{A,r}$, and hence the positions of all the atoms of type A are given by

$$\mathbf{u}_{A,n} = \mathbf{u}+\boldsymbol{\xi}_{A,n} \qquad (n = 1, 2, \ldots, r).$$

Often we do not need to specify the type of atom, so that the subscript A may be omitted and the equation may be written

$$\mathbf{u}_n = \mathbf{u}+\boldsymbol{\xi}_n = (u_i+\xi_{n,i})\mathbf{a}_i \qquad (n = 1, 2, \ldots, r). \tag{5.8}$$

Whilst the components u_i are all integers, the components $\xi_{n,i}$ are all less than unity. For a realistic choice of motif unit, the restriction $|\xi_{n,i}| \leqslant \frac{1}{2}$ will usually be valid. When the atomic arrangement is centrosymmetric, it is possible to choose an origin so that for each atom in a position $\boldsymbol{\xi}$, there is a similar atom in a position $-\boldsymbol{\xi}$.

A structure which contains r atoms in the unit cell may be discussed in two different ways. We may think of a single lattice framework, at each point of which is situated a motif unit of r atoms, or we may consider the whole structure to be composed of r interpenetrating simple Bravais lattices. In many inorganic and organic crystals, the motif units have some physical significance, since they are the molecules of the compound. This is not true for most metallic structures, and it is sometimes possible to choose motif units in various different ways, all having equal validity. The alternative description may then be useful, and we shall write of single, double, etc., lattice structures, meaning structures in which the primitive unit cell contains one, two, etc., atoms. All single and double lattice structures are centrosymmetric, and the structures of most metals fall into one of these two groups.

The ideal crystal is classified by considering the symmetry properties of the atomic arrangement. There are 230 space groups, or combinations of symmetry elements, but most of these are obtained from relations between the vectors $\boldsymbol{\xi}_{A,n}$, $\boldsymbol{\xi}_{B,n}$, etc. The symmetry properties of the lattice are much more restricted, and there are only fourteen Bravais lattices, obtained from relations between the vectors $\mathbf{a}_i$, Instead of the primitive unit cell, it is often convenient to use a larger unit cell which illustrates the symmetry of the lattice positions. For example, if the unit cell of the Bravais lattice has rhombohedral shape, and the angles between the axes are either 109°30′ or 60°, it is readily shown that the lattice positions have cubic symmetry. The conventional unit cells are cubic, and contain two and four points of the Bravais lattices respectively; the lattices are called body-centred and face-centred cubic. If we place an atom on each point of these lattices, we obtain the two common cubic metallic structures.

† This statement has to be modified in the case of a substitutional solid solution in which the possible atomic positions are occupied more or less randomly. The structure is then periodic only if the differences between the atoms of different species are ignored, so that for these atoms, $\boldsymbol{\xi}_A = \boldsymbol{\xi}_B =$, etc.

The set of integers u_i defines a translation from the origin to some lattice point. An alternative set of integers, $U_i = Cu_i$ gives a parallel translation to a lattice point which is C times more distant from the origin. If we assume that the set u_i contains no common factor, we may use the equation

$$\mathbf{u} = C_1 u_i \mathbf{a}_i = U_i \mathbf{a}_i \tag{5.9}$$

in which C_1 takes all possible integral values from $-\infty$ to $+\infty$ to represent all lattice points in a straight line passing through the origin. In a similar way, the equation

$$\mathbf{u} = \mathbf{u}_k + U_i \mathbf{a}_i \tag{5.10}$$

represents all lattice points in a parallel straight line passing through a lattice point with position vector $\mathbf{u}_k$. For any given set of rational values u_i, eqn. (5.10) represents all the lattice points. The lattice structure may thus be regarded as rows of points on parallel straight lines. In conventional crystallographic notation, the quantities u_i are called the direction indices of the line, and are enclosed in square brackets $[u_1 u_2 u_3]$. The direction indices of a line are thus given by its representation as a column matrix u. When the symmetry of the lattice requires that certain directions are equivalent, the whole set of such directions may be represented by the symbol $\langle u_1 u_2 u_3 \rangle$. When u_i have non-integral values, the direction they specify does not lie along a row of lattice points, and is called irrational.

Since the lattice points are arranged along straight lines, it is also possible to regard them as situated on planes. Consider first the plane defined by the basic vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. The normal to this plane is parallel to the vector $\mathbf{a}_1 \wedge \mathbf{a}_2$, and may be denoted by a vector $\mathbf{a}_3^* = (\mathbf{a}_1 \wedge \mathbf{a}_2)/v_a$, where v_a is a scalar constant. The area of the face of the unit cell formed by the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ is numerically equal to the length of the vector $\mathbf{a}_1 \wedge \mathbf{a}_2$, and since the volume of the unit cell is given by the scalar triple product $(\mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3) = \mathbf{a}_3 \cdot \mathbf{a}_1 \wedge \mathbf{a}_2$, we have the distance between adjacent lattice planes in the $\mathbf{a}_3^*$ direction is equal to $\mathbf{a}_3 \cdot \mathbf{a}_1 \wedge \mathbf{a}_2 / |\mathbf{a}_1 \wedge \mathbf{a}_2|$. For reasons which will soon be evident, it is convenient to take the constant v_a equal to the volume of the unit cell, so that $|\mathbf{a}_3^*|$ is equal to the reciprocal of the interplanar spacing. We have then $\mathbf{a}_3 \cdot \mathbf{a}_3^* = 1$. Proceeding in the same way for the other two faces of the unit cell, we obtain a set of vectors $\mathbf{a}_i^*$ perpendicular to the faces of the cell and satisfying the relations:

$$\mathbf{a}_i \cdot \mathbf{a}_j^* = \delta_{ij}. \tag{5.11}$$

The vectors $\mathbf{a}_i^*$ are said to be reciprocal to the vectors $\mathbf{a}_i$. We can now exactly reverse the above reasoning by writing $\mathbf{a}_1 = \mathbf{a}_2^* \wedge \mathbf{a}_3^* / v_a^*$, and since $\mathbf{a}_1 \cdot \mathbf{a}_1^* = 1$, it follows that $v_a^* = (\mathbf{a}_1^* \mathbf{a}_2^* \mathbf{a}_3^*)$, the volume of the parallelepiped formed by the reciprocal set $\mathbf{a}_i^*$. The relations between the two sets of vectors are thus symmetrical, and each is reciprocal to the other. The two volumes are also reciprocal and

$$v_a v_a^* = 1. \tag{5.12}$$

The set of reciprocal vectors, $\mathbf{a}_i^*$, may be used to define a new coordinate system, which we describe as basis A*, and we then have

$$\mathbf{u} = u_i^* \mathbf{a}_i^*,$$

so that $\mathbf{u}$ is represented by†

$$\mathbf{u}^* = [u_1^* \, u_2^* \, u_3^*] = [\mathrm{A}^*; \mathbf{u}].$$

Suppose we form the scalar product of $\mathbf{u}$ and one of the reciprocal vectors $\mathbf{a}_i^*$. This gives

$$\mathbf{u} \cdot \mathbf{a}_i^* = u_j \mathbf{a}_j \cdot \mathbf{a}_i^* = u_i \tag{5.13}$$

and, similarly,

$$\mathbf{u} \cdot \mathbf{a}_i = u_i^*. \tag{5.14}$$

These relations are shown in two dimensional form in Fig. 2.1. In the axis system A, the components u_i measured in units of $|\mathbf{a}_i|$ give the displacements parallel to the axes which

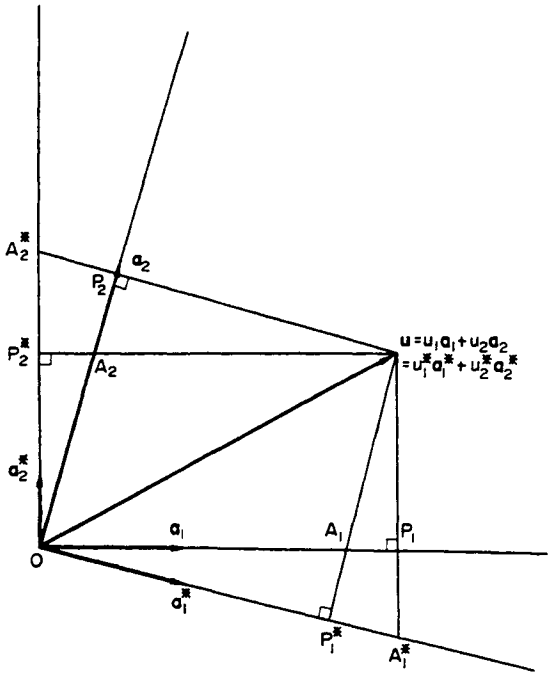

FIG. 2.1. Covariant and contravariant components of $\mathbf{u}$.

$OA_1 = u_1 a_1$ $\quad OA_1^* = u_1^* a_1^*$ $\quad OP_1 = u_1^*/a_1$ $\quad OP_1^* = u_1/a_1^*$

$OA_2 = u_2 a_2$ $\quad OA_2^* = u_2^* a_2^*$ $\quad OP_2 = u_2^*/a_2$ $\quad OP_2^* = u_2/a_2^*$

add together to give $\mathbf{u}$. The components u_i^* in the same basis are the projections of $\mathbf{u}$ along the axes, measured in units of $1/|\mathbf{a}_i|$. In tensor analysis (see p. 35), the quantities u_i are referred to as the contravariant components of $\mathbf{u}$, and u_i^* are the covariant components of $\mathbf{u}$. In the basis A^*, the interpretations of u_i and u_i^* are reversed, as shown in Fig. 2.1.

† In matrix algebra, the notation U^* commonly means the matrix which is the complex conjugate of the matrix U. In this book, all matrices are real, so no confusion arises.

We can now write useful expressions for the scalar and vector product of two vectors $\mathbf{u}$ and $\mathbf{v}$. For the scalar product

$$\mathbf{u}\cdot\mathbf{v} = (u_i\mathbf{a}_i)\cdot(v_j^*\mathbf{a}_j^*) = u_iv_i^* = u_i^*v_i. \tag{5.15}$$

In matrix form this equation is

$$\mathbf{u}\cdot\mathbf{v} = \mathsf{u}'\,\mathsf{v}^* = \mathsf{v}'\,\mathsf{u}^*$$

and, in particular, the length of the vector $\mathbf{u}$ is given by

$$|\mathbf{u}|^2 = \mathsf{u}'\,\mathsf{u}^*. \tag{5.16}$$

Equation (5.16) involves both sets of components u_i and u_i^*, but it is obvious that the length of the vector can be expressed in terms of the components u_i alone. In order to do this, we must find the relation between the bases A and A*, and this is discussed in the next section.

For the vector product of $\mathbf{u}$ and $\mathbf{v}$, we have

$$\mathbf{u}\wedge\mathbf{v} = (u_j\mathbf{a}_j)\wedge(v_k\mathbf{a}_k) = v_a\varepsilon_{ijk}u_jv_k\mathbf{a}_i^*, \tag{5.17}$$

where $\varepsilon_{ijk} = 0$ unless i, j, k are all different, and $\varepsilon_{ijk} = \pm 1$ according to whether i, j, k are even or odd permutations of 1, 2, 3. If we write $\mathbf{w} = (\mathbf{u}\wedge\mathbf{v})$, eqn. (5.17) gives the components of $\mathbf{w}$ in the basis A* as

$$w_i^* = v_a\varepsilon_{ijk}u_jv_k. \tag{5.18}$$

Clearly the components of $\mathbf{w}$ in the basis A are given by

$$w_i = v_a^*\varepsilon_{ijk}u_j^*v_k^*.$$

The orientation of any plane is completely specified by the relative lengths of the intercepts it makes on the axes $\mathbf{a}_i$. Consider a plane which intersects the axes at points distant h_1^{-1}, h_2^{-1}, h_3^{-1} from the origin. This plane contains the two vectors $(\mathbf{a}_1/h_1 - \mathbf{a}_2/h_2)$ and $(\mathbf{a}_2/h_2 - \mathbf{a}_3/h_3)$, and its normal is then parallel to their vector product, i.e. to a vector

$$C\mathbf{h} = \frac{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)}{h_1h_2h_3}(h_1\mathbf{a}_1^* + h_2\mathbf{a}_2^* + h_3\mathbf{a}_3^*).$$

Leaving out the scalar multiplying factor, we may use the normal vector

$$\mathbf{h} = h_i\mathbf{a}_i^* \tag{5.19}$$

to represent the plane. The numbers h_i are the components of the vector $\mathbf{h}$ in the basis A*, and for consistency of notation we should have written them h_i^*. The matrix representation of $\mathbf{h}$ is then

$$\mathsf{h}^* = [h_1^*\ h_2^*\ h_3^*] = [\mathrm{A}^*;\ \mathbf{h}].$$

Since we shall always refer vectors which represent plane normals to the reciprocal basis A*, we may omit the asterisks and use h_i for these components. Note that the magnitude of $\mathbf{h}$ has, at present, been chosen arbitrarily.

The triad of numbers h_i represent the orientations of all parallel planes which intersect the axes at distances C_2/h_i. The scalar equation of these planes is

$$h_1u_1+h_2u_2+h_3u_3-C_2 = 0, \tag{5.20}$$

where each u_i is now regarded as a continuous variable. We are interested in those planes which contain sets of equivalent lattice points. A plane containing three lattice points includes two lattice directions, and hence an infinite set of points. The coordinates of any three points are $[U_1\ U_2\ U_3]$, $[V_1\ V_2\ V_3]$, and $[W_1\ W_2\ W_3]$. From eqn. (5.20) and the three equations obtained by substituting these values into it, we obtain the determinal equation

$$\begin{vmatrix} u_1 & u_2 & u_3 & -1 \\ U_1 & U_2 & U_3 & -1 \\ V_1 & V_2 & V_3 & -1 \\ W_1 & W_2 & W_3 & -1 \end{vmatrix} = 0, \tag{5.21}$$

and since the quantities U_i, V_i, W_i are all integers, it follows from (5.20) and (5.21) that h_1, h_2, h_3 and C_2 are also integral. Planes in which h_i are not integral do not contain lattice points, and are termed irrational. Suppose first that h_i contains no common factor. For each integral value of C_2 between $-\infty$ and $+\infty$, there is an infinite number of points $u_i = U_i$ satisfying eqn. (5.20). As C_2 varies, (5.20) thus represents an infinite set of equally spaced parallel planes, containing all the lattice points. One of these planes ($C_2 = 0$) passes through the origin, and the next plane of the set has $C_2 = 1$. As already shown, the vector $\mathbf{h}$ is perpendicular to the set of planes, and the interplanar spacing is thus given by the projection of the vector $\mathbf{a}_1/h_1$ on the direction of $\mathbf{h}$, i.e. by the relation

$$\mathrm{d}_h = \frac{\mathbf{h}\cdot\mathbf{a}_1/h_1}{|\mathbf{h}|} = \frac{1}{|\mathbf{h}|}. \tag{5.22}$$

The vectors $\mathbf{h}$ are thus not only perpendicular to the planes having indices h_i, but are of length equal to the reciprocal of the interplanar spacing. The orientation of a plane is usually specified in crystallography by placing the indices h_i in round brackets, $(h_1\,h_2\,h_3)$, and this is equivalent to writing the vector as a row matrix $\mathrm{h}' = (\mathbf{h};\ \mathrm{A}^*)$. When there are several equivalent planes, they are indicated by the symbol $\{h_1\,h_2\,h_3\}$.

We now consider the set of equations obtained by replacing the h_i in eqn. (5.19) by quantities $H_i = C_3h_i$ (C_3 integral), it being assumed as before that the set h_i contains no common factor. As C_3 varies from $-\infty$ to $+\infty$, we again obtain an infinite set of planes parallel to the first set but spaced C_3 times more closely. Equation (5.20) can now be satisfied with quantities $u_i = U_i$ only if C_2 is an integral multiple of C_3. When this happens, the plane is one of the original set h_i. We thus see that the vector:

$$\mathbf{h} = C_3h_i\mathbf{a}_i^* = H_i\mathbf{a}_i^* \tag{5.23}$$

represents a set of parallel planes of spacing, $1/|\mathbf{h}|$, but that only every C_3th plane of the set passes through equivalent lattice points.

Equation (5.23) shows that the vectors **h** define a lattice, and this is called the reciprocal lattice. Each point of the reciprocal lattice corresponds to an infinite set of parallel crystal planes; similarly, each set of reciprocal lattice planes is associated with a point of the real lattice. The linear vector space which is defined by the vectors $\mathbf{a}_i^*$ is called reciprocal space. The concepts of reciprocal space, and of the reciprocal lattice, have proved very useful in crystal geometry and in the theory of X-ray diffraction.

6. LINEAR TRANSFORMATIONS OF THE COORDINATE SYSTEM

The position of any lattice point is given by its coordinates u_i with respect to some chosen origin and set of base vectors A. As previously emphasized, the choice of basis is arbitrary, and it is often desirable to use a new basis B defined by the vectors $\mathbf{b}_i$. We then have to develop the appropriate transformation formulae connecting ${}^{A}u_i$ and ${}^{B}u_i$ for any direction, and ${}^{A}h_i$ and ${}^{B}h_i$ for any plane.

The new basic vector $\mathbf{b}_1$ will differ from $\mathbf{a}_1$ both in magnitude and direction. Since any four vectors are linearly dependent, however, we can write $\mathbf{a}_1$ as a linear function of the new vectors $\mathbf{b}_i$:

$$\mathbf{a}_1 = J_{11}\mathbf{b}_1 + J_{21}\mathbf{b}_2 + J_{31}\mathbf{b}_3 .$$

This relation signifies that the direction indices of the vector $\mathbf{a}_1$ referred to the basis B are $[J_{11}\,J_{21}\,J_{31}]$.† The transformation from one basis to the other may thus be represented completely by three equations of the above form, or in the usual summation convention

$$\mathbf{a}_i = J_{ji}\mathbf{b}_j . \tag{6.1}$$

Note that although written in subscript form, this is a vector equation and not a relation between vector components. When we deal with base vectors, the subscript identifies a particular vector of the set A or B, rather than a particular component. The vector $\mathbf{a}_i$ may be represented in its own basis system A by the column matrix $\mathbf{a}_i$; obviously the kth component of $\mathbf{a}_i$ in this representation is $(a_i)_k = \delta_{ik}$.

If we now write the set of vectors $\mathbf{a}_i$ which constitute the basis A as a column matrix of *vectors*, $\mathsf{A} = [\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3]$, and the set $\mathbf{b}_i$ as a similar matrix B, the three equations may be combined in the matrix equation

$$\mathsf{A} = \mathsf{J}'\,\mathsf{B}. \tag{6.2}$$

Here J is the 3×3 array with elements J_{ij}, and the transposed matrix J′ is the corresponding array with elements J_{ji}. A is obtained by multiplying J′ and B in accordance with the laws of matrix algebra.

The components of each individual vector of the set $\mathbf{a}_i$ are given by the representations

$$[\mathsf{A};\,\mathbf{a}_i] = [\delta_{1i}\delta_{2i}\delta_{3i}] \quad \text{and} \quad [\mathsf{B};\,\mathbf{a}_i] = [J_{1i}J_{2i}J_{3i}].$$

† We write the direction indices of $\mathbf{a}_1$ in this form, rather than as $[J_{11}\,J_{12}\,J_{13}]$ in order that the matrix array $\mathsf{J} = J_{ij}$ introduced below shall give directly the transformation of vector *components*.

By comparing coefficients, it is readily seen that these quantities are related by the matrix equation

$$[\mathrm{B};\ \mathbf{a}_i] = \mathrm{J}[\mathrm{A};\ \mathbf{a}_i]. \tag{6.3}$$

We shall now show that this is a particular example of the general equation relating the components ${}^{A}u_i$. ${}^{B}u_i$ of any vector $\mathbf{u}$. Writing $\mathbf{u} = {}^{A}u_i\,\mathbf{a}_i = {}^{B}u_i\,\mathbf{b}_i$, and substituting for $\mathbf{a}_i$ from (6.1),

$${}^{B}u_i\,\mathbf{b}_i = {}^{A}u_j\,J_{ji}\mathbf{b}_j = {}^{A}u_j\,J_{ij}\mathbf{b}_i$$

since both i and j are dummy indices (to be summed) on the right-hand side. The expression is now an identity, and the coefficients of corresponding vectors $\mathbf{b}_i$ on both sides may be equated to give

$${}^{B}u_i = J_{ij}\,{}^{A}u_j \tag{6.4}$$

or in matrix form

$${}^{B}\mathrm{u} = \mathrm{J}\,{}^{A}\mathrm{u}. \tag{6.5}$$

This equation reduces to (6.3) when we put $\mathbf{u} = \mathbf{a}_i$. The matrix J is a representation of the transformation from the basis A to the basis B; its columns are the direction indices of $\mathbf{a}_i$ referred to B. Although this notation is often adequate, it is sometimes necessary to show explicitly that J represents the operation changing the reference basis, and we then use the extended notation

$$\mathrm{J} = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})$$

so that (6.5) may be written in full as

$$[\mathrm{B};\ \mathbf{u}] = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{u}]. \tag{6.6}$$

The bold face symbol **J** is to be regarded as an operator which transforms the representation ${}^{A}\mathrm{U}$ into ${}^{B}\mathrm{U}$; equivalently, we may say that **J** is a function of A and B which gives the transforming matrix J.† The fact that eqn. (6.6) represents a change of axes is emphasized by the use of different base symbols on either side of **J**; note also that the base symbols A occur in juxtaposition on the right-hand side.

Provided that the vectors $\mathbf{a}_i$, $\mathbf{b}_i$ are non-coplanar sets (which is a necessary condition for them to define unit cells), the determinant $|\mathrm{J}|$ of the matrix array J does not vanish. The three simultaneous eqn. (6.5) may then be solved for ${}^{A}\mathrm{u}$ in terms of ${}^{B}\mathrm{u}$, giving the reverse transformation

$${}^{A}\mathrm{u} = \mathrm{J}^{-1}\,{}^{B}\mathrm{u}, \tag{6.7}$$

where the *reciprocal* matrix J^{-1} has elements $J^{-1}{}_{ij} = J^{ji}/|\mathrm{J}|$ and J^{ji} means the cofactor of the element J_{ij} of the matrix J. In the fuller matrix notation, the reverse transformation is

$$[\mathrm{A};\ \mathbf{u}] = (\mathrm{A}\,\mathbf{J}\,\mathrm{B})\,[\mathrm{B};\ \mathbf{u}]$$

so that our notation implies

$$(\mathrm{A}\,\mathbf{J}\,\mathrm{B}) = \mathrm{J}^{-1} = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})^{-1}. \tag{6.8}$$

† An alternative notation sometimes used (e.g. Bullough and Bilby, 1956) is to write (B J A) as (B/A). This has the slight disadavantage of not allowing the use of the single symbol J when convenient.

Similarly, the transformation between sets of vectors has an inverse

$$\mathrm{B} = \mathrm{J}'^{-1}\mathrm{A}.$$

If we wish to write eqn. (6.5) in terms of row vectors, we must transpose both sides of the equation. From the rule for transposing matrix products, this gives

$${}^{B}\mathrm{u}' = {}^{A}\mathrm{u}'\,\mathrm{J}',$$

but if we write this in extended notation, it is convenient to have a symbol for the transpose of $(\mathrm{B}\,\mathbf{J}\,\mathrm{A})$ which will preserve the juxtaposition of like base symbols. We thus introduce the notation

$$(\mathrm{B}\,\mathbf{J}\,\mathrm{A})' = (\mathrm{A}\,\mathbf{J}'\,\mathrm{B}) = \mathrm{J}'$$

and the transpose of (6.5) is then

$$(\mathbf{u};\mathrm{B}) = (\mathbf{u};\mathrm{A})\,(\mathrm{A}\,\mathbf{J}'\,\mathrm{B}). \tag{6.9}$$

The new base vectors $\mathbf{b}_i$ are associated with a new reciprocal set of vectors $\mathbf{b}_i^*$ where

$$\mathbf{b}_i \cdot \mathbf{b}_j^* = \delta_{ij}.$$

The square of the length of $\mathbf{u}$ may be written from (5.16) as

$$|\mathbf{u}|^2 = (\mathbf{u};\mathrm{B}^*)\,[\mathrm{B};\mathbf{u}] = (\mathbf{u};\mathrm{B}^*)\,(\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,[\mathrm{A};\mathbf{u}] = (\mathbf{u};\mathrm{A}^*)\,[\mathrm{A};\mathbf{u}]$$

so that

$$(\mathbf{u};\mathrm{A}^*) = (\mathbf{u};\mathrm{B}^*)\,(\mathrm{B}\,\mathbf{J}\,\mathrm{A}). \tag{6.10}$$

However,

$$(\mathbf{u};\mathrm{A}^*) = (\mathbf{u};\mathrm{B}^*)\,(\mathrm{B}^*\,\mathbf{J}'\,\mathrm{A}^*),$$

so that our notation implies

$$(\mathrm{B}^*\,\mathbf{J}'\,\mathrm{A}^*) = (\mathrm{B}\,\mathbf{J}\,\mathrm{A}) \quad \text{and} \quad (\mathrm{A}\,\mathbf{J}'\,\mathrm{B}) = (\mathrm{A}^*\,\mathbf{J}\,\mathrm{B}^*). \tag{6.11}$$

The law for the transformation of vector components referred to the reciprocal bases is thus identical with that for transforming base vectors (eqns. (6.1) and (6.2)); if J is the representation of the change $\mathrm{A} \rightarrow \mathrm{B}$, J′ is the corresponding representation of the change $\mathrm{B}^* \rightarrow \mathrm{A}^*$.

From eqn. (6.10), the reverse transformation may be written

$$[\mathrm{B}^*;\mathbf{u}] = (\mathrm{A}\,\mathbf{J}'\,\mathrm{B})^{-1}\,[\mathrm{A}^*;\mathbf{u}] = (\mathrm{B}\,\mathbf{J}'\,\mathrm{A})\,[\mathrm{A}^*;u]. \tag{6.12}$$

In particular, the quantities ${}^{A}h_i$ must transform like this, since they are the components of $\mathbf{h}$ referred to the basis A^* (see p. 29). The transformation law is, however, usually more conveniently expressed in the form (6.10), since it is generally preferable to write the vector $\mathbf{h}$ representing a plane normal as a row matrix h′. This gives

$$(\mathbf{h};\mathrm{B}^*) = (\mathbf{h};\mathrm{A}^*)\,(\mathrm{A}\,\mathbf{J}\,\mathrm{B}) \tag{6.13}$$

or in component form

$$^{B}h_j = J^{-1}{}_{ij}\, {}^{A}h_i .$$

We also have

$$\mathbf{h} = {}^{B}h_i \mathbf{b}_i^* = {}^{A}h_j \mathbf{a}_j^* = J_{ij}\, {}^{B}h_i \mathbf{a}_j^* ,$$

so that

$$\mathbf{b}_i^* = J_{ij}\mathbf{a}_j^*$$

or

$$\mathsf{B}^* = \mathsf{J}\,\mathsf{A}^*, \tag{6.14}$$

which gives the relation between reciprocal vectors, corresponding to (6.2) for lattice vectors.

Finally, it is useful to examine the transformation from a base A to its own reciprocal base A*. Let the components of a vector **u** referred to the two bases be related by

$$\mathsf{u}^* = \mathsf{G}\,\mathsf{u}, \tag{6.15}$$

where $\mathsf{G} = (\mathsf{A}^*\,\mathbf{G}\,\mathsf{A})$ is the matrix representation of the transformation from A to A*. From eqn. (5.15) the scalar product of two vectors may be written

$$\mathbf{u}\cdot\mathbf{v} = u_j^* v_j = G_{ji} u_i v_j ,$$

but this product may also be expanded as

$$\mathbf{u}\cdot\mathbf{v} = u_i \mathbf{a}_i \cdot v_j \mathbf{a}_j = u_i v_j \mathbf{a}_i \cdot \mathbf{a}_j$$

so that the elements of the matrix G are

$$G_{ij} = G_{ji} = \mathbf{a}_i \cdot \mathbf{a}_j . \tag{6.16}$$

The function symbol **G** is called the metric,† and is represented by the symmetrical square matrix G. Its importance arises from its fundamental connection with the distance between two points, since the square of the length of a vector is

$$|\mathbf{u}|^2 = \mathsf{u}'\,\mathsf{u}^* = \mathsf{u}'\,\mathsf{G}\,\mathsf{u} = G_{ij} u_i u_j . \tag{6.17}$$

The reverse transformation is

$$\mathsf{u} = \mathsf{G}^{-1}\,\mathsf{u}^*, \tag{6.18}$$

where G^{-1} has components $G^{-1}{}_{ij} = G^{-1}{}_{ji} = \mathbf{a}_i^* \cdot \mathbf{a}_j^*$. The length of **u** may then also be expressed as

$$|\mathbf{u}|^2 = \mathsf{u}^{*\prime}\,\mathsf{G}^{-1}\,\mathsf{u}^* = G^{-1}{}_{ij}\, u_i^* u_j^* .$$

Similarly, eqn. (5.22) gives for the interplanar spacing d of the planes **h**

$$(d)^{-2} = |\mathbf{h}|^2 = G^{-1}{}_{ij}\, h_i h_j . \tag{6.19}$$

† Since the relation between A and A* is defined by A alone, **G** may also be regarded as a second order tensor (see next section) and the matrix $\mathsf{G} = G_{ij}$ then gives the components of this tensor in the system A. **G** is usually called the metric tensor. The function symbol **J**, connecting arbitrary axis systems, is not a tensor.

It follows from a result to be proved later (Section 7, p.47-8) that the determinant of the matrix G is equal to the ratio of the volume of the cell defined by $\mathbf{a}_i$ to that defined by $\mathbf{a}_i^*$. Combining this result with eqn. (5.12) shows that

$$|\mathsf{G}|^{1/2} = v_a \quad \text{and similarly} \quad |\mathsf{G}^{-1}|^{1/2} = v_a^*, \tag{6.20}$$

where, as before, v_a, v_a^* are the volumes of the respective unit cells.

Since G is symmetric, $\mathsf{G}' = \mathsf{G}$, or in full

$$(\mathrm{A}\,\mathsf{G}'\,\mathrm{A}^*) = (\mathrm{A}^*\,\mathsf{G}\,\mathrm{A}).$$

We now write the scalar product of $\mathbf{u}$ and $\mathbf{v}$ in the basis A and transform to the basis B using eqns. (6.17) and (6.6)

$$\mathbf{u}\cdot\mathbf{v} = (\mathbf{u};\mathrm{A})(\mathrm{A}^*\,\mathsf{G}\,\mathrm{A})[\mathrm{A};\mathbf{v}] = (\mathbf{u};\mathrm{B})(\mathrm{B}\,\mathsf{J}'\,\mathrm{A})(\mathrm{A}^*\,\mathsf{G}\,\mathrm{A})(\mathrm{A}\,\mathsf{J}\,\mathrm{B})[\mathrm{B};\mathbf{v}]$$

so that

$$(\mathrm{B}^*\,\mathsf{G}\,\mathrm{B}) = (\mathrm{B}\,\mathsf{J}'\,\mathrm{A})(\mathrm{A}^*\,\mathsf{G}\,\mathrm{A})(\mathrm{A}\,\mathsf{J}\,\mathrm{B}), \tag{6.21}$$

which gives the relation between the representations of G in two different bases. We could also have derived (6.21) by making use of the identity $(\mathrm{B}^*\,\mathsf{G}\,\mathrm{B}) = (\mathrm{B}^*\,\mathsf{J}\,\mathrm{A}^*)(\mathrm{A}^*\,\mathsf{G}\,\mathrm{A})(\mathrm{A}\,\mathsf{J}\,\mathrm{B})$ and substituting for $(\mathrm{B}^*\,\mathsf{J}\,\mathrm{A}^*)$ from (6.11).

Consider now the orthonormal basis I. The basic vectors $\mathbf{i}_i$ then satisfy eqn. (5.2) and $\mathsf{G}^{-1} = \mathsf{G} = \mathsf{I}$, the unit matrix having elements $I_{ij} = \delta_{ij}$. The metric of an orthonormal system is thus unity. This corresponds to the fact that the bases I and I* are identical in such a system.

Examination of the above equations shows that the laws of transformation are of two kinds. Quantities which transform like the vector components u_i or the reciprocal base vectors $\mathbf{a}_i^*$ are called *contravariant* in tensor analysis; those which transform like the base vectors $\mathbf{a}_i$ or the plane indices (reciprocal vector components) h_i are called *covariant*. In addition, there are scalar quantities which are *invariant* with respect to an axis transformation. Thus u_i form the components of a contravariant tensor of the first order (a vector), or more simply, may be described as the contravariant components of the real vector $\mathbf{u}$; u_i^* are the covariant components of $\mathbf{u}$. Now let ${}^{A}u_i$, ${}^{B}u_i$ represent continuous variables along the axes $\mathbf{a}_i$, $\mathbf{b}_i$, so that they define coordinate systems. The linear relation between the coordinates, given by (6.4), may be written

$${}^{B}u_i = (\partial {}^{B}u_i/\partial {}^{A}u_j)\,{}^{A}u_j, \tag{6.22}$$

where $J_{ij} = \partial {}^{B}u_i/\partial {}^{A}u_j$, etc. This is the law for the transformation of the contravariant components of $\mathbf{u}$. The corresponding law for the covariant components u_i^* is

$${}^{B}u_i^* = (\partial {}^{A}u_j/\partial {}^{B}u_i)\,{}^{A}u_j^*. \tag{6.23}$$

In tensor notation, contravariant quantities are distinguished by writing the identifying suffix as a superscript, and there is then no need for a separate notation for the bases A and A*. Thus eqn. (5.1) would be written

$$\mathbf{u} = u^i\mathbf{a}_i, \tag{6.24}$$

where the summation convention applies as before. Note that the superscript i is not a power index.

Covariant quantities are written with the suffix as subscript, so that the covariant components u_i^* are now written simply u_i. From eqn. (5.13) it follows that the covariant components of $\mathbf{u}$ in base A are equal to the orthogonal projections of $\mathbf{u}$ on the axes $\mathbf{a}_i$ (see Fig. 2.1). From (5.12) we also have

$$\mathbf{u} = u_i \mathbf{a}^i, \tag{6.25}$$

where $\mathbf{a}^i$ are the reciprocal vectors, formerly written $\mathbf{a}_i^*$. The covariant components of $\mathbf{u}$ in the base A are thus the contravariant components in the base A*, and vice versa. In general, we may alter any equation of the form (6.24) by lowering the dummy index in one place and raising it in another.

In terms of the new notation, the transformation laws (6.22) and (6.23) become

$$^B u^i = (\partial\, ^B u^i / \partial\, ^A u^j)\, ^A u^j \tag{6.26}$$

and

$$^B u_i = (\partial\, ^A u^j / \partial\, ^B u^i)\, ^A u_j = (\partial\, ^B u_i / \partial\, ^A u_j)\, ^A u_j. \tag{6.27}$$

In the above discussion, we have been careful to write about the covariant and contravariant components of one vector $\mathbf{u}$, since we wish to emphasize the idea of the vector as the physical entity. In tensor algebra, it is customary to treat the components as separate covariant and contravariant vectors, and this leads to economy of description. The vector $\mathbf{u}$ is sometimes called the real vector.

Anticipating the results of the next section a little, we may also form covariant, contravariant and mixed tensors of the second order, with components T_{ij}, T^{ij}, and T_j^i respectively. These tensors are arrays of nine quantities, each depending on two directions of the coordinate axes, which represent a linear relation between two vectors (e.g. the stress tensor relates the linear force on a surface element and the vector normal to the element). When the vectors are referred to another coordinate system, the quantities in the tensor transform according to laws which are contravariant for both suffices, covariant for both suffices, or mixed, thus:

$$\left.\begin{aligned}
&\text{(Contravariant)} && ^B T^{ij} = (\partial\, ^B u^i / \partial\, ^A u^k)(\partial\, ^B u^j / \partial\, ^A u^l)\, ^A T^{kl}, \\
&\text{(Covariant)} && ^B T_{ij} = (\partial\, ^A u^k / \partial\, ^B u^i)(\partial\, ^A u^l / \partial\, ^B u^j)\, ^A T_{kl} \\
&&& \quad = (\partial\, ^B u_i / \partial\, ^A u_k)(\partial\, ^B u_j / \partial\, ^A u_l)\, ^A T_{kl}, \\
&\text{(Mixed)} && ^B T_j^i = (\partial\, ^B u^i / \partial\, ^A u^k)(\partial\, ^A u^l / \partial\, ^B u^j)\, ^A T_l^k \\
&&& \quad = (\partial\, ^B u^i / \partial\, ^A u^k)(\partial\, ^B u_j / \partial\, ^A u_l)\, ^A T_l^k.
\end{aligned}\right\} \tag{6.28}$$

In the remainder of this book we shall not use tensor notation, preferring the matrix representation when it is necessary to distinguish covariance from contravariance. When components are required, they will therefore be written with all suffices as subscripts, except for the standard notation for the cofactor of a matrix element, used previously. Tensor notation is especially powerful in problems where curvilinear coordinates are required, so that the linear relations (6.2) and (6.4) have to be replaced by the more general

$$^B u^i = f^i(^A u^1, ^A u^2, ^A u^3)$$

so that

$$\mathrm{d}^B u^i = (\partial {}^B u^i / \partial {}^A u^j)\, \mathrm{d}^A u^j . \tag{6.29}$$

The general properties of contravariance or covariance are then still defined by eqns. (6.26)–(6.28), but is should be noted that the coordinates u^i themselves no longer transform according to the contravariant law. The transformation of coordinates is only contravariant when the relation (6.29) is linear of type (6.4); this is called an affine transformation (see next section) and the coordinate systems are all Cartesian. In all other cases, it is clearly to some extent arbitrary whether we write the coordinates as u^i or u_i.

It is perhaps of interest to show that the components of the matrix **G** may be regarded as the representation of a tensor. If the coordinates of two neighbouring points in the reference system A are ${}^A u^i$ and ${}^A u^i + \mathrm{d}^A u^i$, we know that the separation of the points, $\mathrm{d}w$, is given by

$$(\mathrm{d}w)^2 = {}^A G_{ij}\, \mathrm{d}^A u^i\, \mathrm{d}^A u^j .$$

In the system B, the corresponding equation is given from (6.29) as

$$(\mathrm{d}w)^2 = {}^B G_{ij}\, \mathrm{d}^B u^i\, \mathrm{d}^B u^j = {}^B G_{ij} (\partial {}^B u^i / \partial {}^A u^k)\, (\partial {}^B u^j / \partial {}^A u^l)\, \mathrm{d}^A u^k\, \mathrm{d}^A u^l ,$$

and since the separation is an invariant,

$${}^A G_{ij} = (\partial {}^B u^i / \partial {}^A u^k)\, (\partial {}^B u^j / \partial {}^A u^l)\, {}^B G_{ij} . \tag{6.30}$$

Comparison with (6.28) shows that the components of **G** transform according to the law for covariant tensors of the second order, so that we are justified in regarding **G** as a tensor. In effect, **G** relates two vectors such that the contravariant components of the second are the covariant components of the first. In the same way, we can show that the components of the matrix $\mathbf{G}^{-1}$ transform according to the law for contravariant tensors (6.28), and would be written G^{ij} in tensor notation.

Note that if both sets of axes are orthogonal and Cartesian, the covariant and contravariant components coincide, and the two transformation laws are identical. This is why contravariance and covariance are not distinguished in elementary vector analysis. We shall frequently use oblique Cartesian coordinates, but curvilinear coordinates are not needed for most problems in crystal geometry. As already emphasized, we shall now employ matrix rather than tensor notation.

For reference, we give in Table II a summary comparison of the main features of the different notations, and of the standard equations of crystal geometry expressed in these notations. Most of the results in the table have been obtained in the text above, but there is one additional point to note.

Equations (6.17) and (6.19) for the length of a vector and the spacing of a set of planes respectively are valid for any choice of the base vectors $\mathbf{a}_i$, including vectors which do not define a primitive unit cell. However, we may need to know the distance between adjacent lattice points (identity distance) in the direction **u**, or between adjacent lattice planes normal to **h**. It we use components u_i, h_i which are integers with no common factor, these dis-

TABLE II. SUMMARY OF LATTICE GEOMETRY

	Matrix notation	Tensor notation
Base vectors	$\mathbf{A} = \mathbf{a}_i$	$\mathbf{a}_i$
Reciprocal base vectors	$\mathbf{A}^* = \mathbf{a}_i^*$	$\mathbf{a}^i$
Contravariant components of $\mathbf{u}$	$\mathbf{u} = [\mathbf{A}; \mathbf{u}]$	u^i
Covariant components of $\mathbf{u}$	$\mathbf{u}^* = [\mathbf{A}^*; \mathbf{u}]$	u_i
Covariant components of plane normal, $\mathbf{h}$	$\mathbf{h}'$ (strictly $\mathbf{h}^{*\prime}$) $= (\mathbf{h}; \mathbf{A}^*)$	h_i
Metric $\mathbf{a}_i \cdot \mathbf{a}_j$	$\mathbf{G} = (\mathbf{A}^* \mathbf{G} \mathbf{A})$	G_{ij}
Reciprocal metric, $\mathbf{a}^i \cdot \mathbf{a}^j$	$\mathbf{G}^{-1} = (\mathbf{A} \mathbf{G}^{-1} \mathbf{A}^*)$	G^{ij}
Volume of cell, v_a	$\lvert \mathbf{G} \rvert^{1/2}$	
Volume of reciprocal cell, v_a^*	$\lvert \mathbf{G}^{-1} \rvert^{1/2}$	
Scalar product, $\mathbf{u} \cdot \mathbf{v}$	$\mathbf{u}'\mathbf{v}^* = \mathbf{u}' \mathbf{G} \mathbf{v}$ $= \mathbf{u}^{*\prime} \mathbf{G}^{-1} \mathbf{v}^*$	$G_{ij} u^i v^j = G^{ij} u_i v_j$
Contravariant components of $\mathbf{w} = \mathbf{u} \wedge \mathbf{v}$		$w^i = \lvert \mathbf{G}^{-1} \rvert^{1/2} \varepsilon^{ijk} u_j v_k$ †
Covariant components of $\mathbf{w} = \mathbf{u} \wedge \mathbf{v}$		$w_i = \lvert \mathbf{G} \rvert^{1/2} \varepsilon_{ijk} u^j v^k$ †
Repeat distance along $\mathbf{u}$ (u^i relatively prime)	$I'(\mathbf{u}' \mathbf{G} \mathbf{u})^{1/2}$	$I'(G_{ij} u^i u^j)^{1/2}$
Interplanar spacing, d of planes with h_i relatively prime	$I(\mathbf{h}' \mathbf{G}^{-1} \mathbf{h})^{-1/2}$	$I(G^{ij} h_i h_j)^{-1/2}$
Cosine of angle between $\mathbf{u}$ and $\mathbf{v}$	$\dfrac{\mathbf{u}' \mathbf{G} \mathbf{v}}{(\mathbf{u}' \mathbf{G} \mathbf{u})^{1/2} (\mathbf{v}' \mathbf{G} \mathbf{v})^{1/2}}$	$\dfrac{G_{ij} u^i v^j}{(G_{ij} u^i u^j)^{1/2} (G_{ij} v^i v^j)^{1/2}}$
Cosine of angle between planes $\mathbf{h}$ and $\mathbf{k}$	$\dfrac{\mathbf{h}' \mathbf{G}^{-1} \mathbf{k}}{(\mathbf{h}' \mathbf{G}^{-1} \mathbf{h})^{1/2} (\mathbf{k}' \mathbf{G}^{-1} \mathbf{k})^{1/2}}$	$\dfrac{G^{ij} h_i k_j}{(G^{ij} h_i h_j)^{1/2} (G^{ij} k_i k_j)^{1/2}}$
Cosine of angle between direction $\mathbf{u}$ and plane $\mathbf{h}$	$\dfrac{\mathbf{h}' \mathbf{u}}{(\mathbf{h}' \mathbf{G}^{-1} \mathbf{h})^{1/2} (\mathbf{u}' \mathbf{G} \mathbf{u})^{1/2}}$	$\dfrac{u^i h_j}{(G_{ij} u^i u^j)^{1/2} (G^{ij} h_i h_j)^{1/2}}$
Zone axis of $\mathbf{h}$ and $\mathbf{k}$	$\mathbf{u}$ parallel $\mathbf{h} \wedge \mathbf{k}$	$u^i \propto$ to $\varepsilon^{ijk} h_j h_k$
Plane containing $\mathbf{u}$ and $\mathbf{v}$	$\mathbf{h}$ parallel $\mathbf{u} \wedge \mathbf{v}$	$h_i \propto$ to $\varepsilon_{ijk} u^j v^k$

† The factors $\lvert \mathbf{G} \rvert^{1/2}$ and $\lvert \mathbf{G}^{-1} \rvert^{1/2}$ are often included in the definition of ε_{ijk} and ε^{ijk} respectively.

tances are given by the same expressions, (6.17) and (6.19), provided the basis A is primitive. As we have already stated, however, it is often convenient to use centred (non-primitive) cells, so that the vectors $u_i \mathbf{a}_i$ with u_i taking all possible integral values do not represent *all* the lattice points. For example, in the body-centred cubic structure with the conventional cubic basis, the lattice points are given by the vectors $\frac{1}{2} u_i \mathbf{a}_i$ with the restriction that the quantities u_i must be all odd or all even integers. This complication is allowed for in Table II by introducing the "cell factors" I, I' into the equations for identity distance and planar spacing. The cell factors for the important structures are given separately in Table III.

TABLE III. CELL FACTORS FOR CENTRED LATTICES

Type of cell	Plane indices	I	Direction indices	I'
Primitive	All h_i	1	All u_i	1
Body-centred	Σh_i even	1	u_i all odd	$\frac{1}{2}$
	Σh_i odd	$\frac{1}{2}$	u_i mixed odd and even	1
Base-centred	h_1+h_2 even	1	u_1+u_2 even	$\frac{1}{2}$
	h_1+h_2 odd	$\frac{1}{2}$	u_1+u_2 odd	1
Face-centred	h_i all odd	1	Σu_i even	$\frac{1}{2}$
	h_i mixed odd and even	$\frac{1}{2}$	Σu_i odd	1

We conclude this section by referring to the well-known Miller–Bravais four-axis system for indexing directions and planes in hexagonal structures. The four-axis convention is described in most elementary textbooks on crystallography; it has the advantage that symmetry-related planes and directions are obtained by a simple permutation of indices, but it is clearly inconvenient when a hexagonal lattice has to be related to a lattice of different symmetry. However, the relations between the direct and reciprocal lattices when a four-axis system is used, together with associated problems of crystal geometry, are not obvious, and we shall briefly describe some aspects of these relations in order to complete our description of lattice geometry.

The hexagonal lattice may be regarded as a series of planar hexagonal nets of edge a stacked vertically above each other with a separation c. The most useful bases of conventional type are defined by the vector sets $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{c}$ and $\mathbf{a}_1$, $\mathbf{a}_1+2\mathbf{a}_2$, $\mathbf{c}$, which describe respectively a primitive cell and a C-centred orthorhombic cell. Various rhombohedral bases may also be used, but for hexagonal lattices their disadvantages outweigh their advantages. The non-conventional four-axis basis consists of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{c}$ and a vector direction $\mathbf{u}$ is represented by

$$\mathbf{u} = u_i\mathbf{a}_i \qquad (i = 1, \ldots, 4), \tag{6.31}$$

where it is a requirement that

$$u_1+u_2+u_3 = 0. \tag{6.32}$$

Rational lattice directions are represented by relatively prime integral values of u_i, and the coordinates of all the lattice points along such a direction through the origin are then given by $I'nu_i$, where n is any integer and I' is equal to $\frac{1}{3}$ when u_1-u_2 is divisible by 3 and otherwise is unity. (I' is analogous to the cell factors of Table II.)

If lattice planes are now defined in terms of their reciprocal intercepts h_i on the axes of the direct basis, we find correspondingly that

$$h_1+h_2+h_3 = 0. \tag{6.33}$$

However, some complications arise when we represent plane normals as directions of the reciprocal lattice. With a four-axis system it is not possible to define a reciprocal basis by

means of eqns. (5.11), since no vector can be simultaneously perpendicular to three non-coplanar vectors. Although a suitable four-axis basis for the reciprocal lattice can readily be found, the fact that this basis is not strictly reciprocal to the direct basis has caused much confusion in the literature.

Clearly the reciprocal lattice itself is completely defined by the direct lattice; it may be introduced through eqns. (5.11) applied to a conventional direct basis, or by using the geometrical interpretation discussed at the end of Section 5, which is independent of any choice of basis. The reciprocal lattice is also hexagonal and consists of planar nets of edge $(2/3)^{-2}a^{-1}$ stacked at a separation of c^{-1}. The hexagonal nets of the direct and reciprocal lattices have parallel normals, but are rotated through 30° relative to each other. The appropriate four-axis reciprocal basis is defined by the vectors

$$\mathbf{a}_i^+ = (\tfrac{2}{3}a^{-2})\mathbf{a}_i \qquad (i = 1, \ldots, 3), \qquad \mathbf{a}_4^+ = c^{-2}\mathbf{a}_4, \tag{6.34}$$

where, following Nicholas and Segall (1970), we have used the notation $\mathbf{a}_i^+$ rather than $\mathbf{a}_i^*$ because eqns. (5.11) are not valid. The normal to the plane h_i is now a vector $\mathbf{h}$ of the reciprocal lattice, where

$$\mathbf{h} = h_i\mathbf{a}_i^+ \qquad (i = 1, \ldots, 4). \tag{6.35}$$

This equation is strictly analogous to (5.19). It also follows that if the direction u_i is contained in the plane h_i, then

$$h_iu_i = 0 \qquad (i = 1, \ldots, 4), \tag{6.36}$$

which is the obvious analogue of the three-axis condition.

It should be noted that the four-axis basis of the direct lattice defines a hexagonal prism, of volume $3^{\frac{3}{2}}\, a^2c/2$, which contains three lattice points; this is why the cell factor I' has to be introduced when repeat distances are calculated. However, the vectors $\mathbf{a}_i^+$ define a hexagonal prism of the reciprocal lattice which contains only one reciprocal lattice point, and the reciprocal spacings of lattice planes are therefore given by $|\mathbf{h}|$, where h_i are relatively prime integers, without the need to introduce a cell factor I.

Since eqns. (5.11) are invalid, the existence of the reciprocal basis (6.34) for which eqns. (6.35) and (6.36) are satisfied, appears almost fortuitous. That this is not so was shown by Frank (1965) and Nicholas and Segall (1970). Frank's approach is to consider a four-dimensional lattice with an orthogonal basis $\boldsymbol{\alpha}_i$ with $|\boldsymbol{\alpha}_i| = \alpha$ for $i = 1, \ldots, 3$ and $|\boldsymbol{\alpha}_4| = \gamma$. Equations (5.11), with a range of $1, \ldots, 4$, for i, then enable an orthogonal reciprocal basis $\boldsymbol{\alpha}_i^*$ to be defined. The real three-dimensional lattice is obtained either by projecting the four-dimensional lattice along some direction or by sectioning it in a particular "hyperplane". Frank showed that a projection along [1110] gives the hexagonal lattice with parameters $a = (\frac{2}{3})^{1/2}\,\alpha$ and $c = \gamma$. Similarly, a section of the four-dimensional reciprocal lattice in (1110) gives the reciprocal of the real hexagonal lattice. Moreover the four-axis bases of the direct and reciprocal lattices are simply obtained by projecting the orthogonal bases $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}_i^*$. The significance of the above procedure is readily appreciated by lowering the dimensionality by one; it then becomes equivalent to projecting a simple cubic lattice along [111] or sectioning its reciprocal lattice in (111) to give the direct and reciprocal nets, and

the three-axis bases in these nets are obtained from the projections of the cubic direct and reciprocal axes.

Nicholas and Segall (1970) have given a more complete description of the general case of a redundant base vector, i.e. the use in an n-dimensional space of a basis $\mathbf{a}_i$ with $i = 1, \ldots n+1$. They show that (6.36) is universally valid (whatever i) provided that h_i are defined by the reciprocal intercepts of a hyperplane on $\mathbf{a}_i$, and also that it is always possible to define a reciprocal basis $\mathbf{a}_i^+$ for which (6.35) is valid. Moreover, this choice is not unique; in the hexagonal lattice, for example, any other basis $\mathbf{a}_i^* = \mathbf{a}_i^+ + \mathbf{d}^+$ $(i = 1, \ldots 3)$, $\mathbf{a}_4^* = \mathbf{a}_4^+$, where $\mathbf{d}^+$ is any vector, will not disturb the validity of (6.35). The real justification for the choice of axes which is customarily made (and hence for the conditions (6.32) and (6.33)) is in order to ensure that cyclically permuted indices represent crystallographically equivalent directions and planes.

The advantage of linking the four-axis systems to the orthogonal vectors of a four-dimensional lattice is that crystallographic formulae can be derived rather simply for the four-dimensional lattice and then transformed to the four-axis representation of a three-

TABLE IIIA. SUMMARY OF LATTICE GEOMETRY WITH MILLER–BRAVAIS INDICES

Base vectors	$\mathbf{a}_i \quad (i = 1,\ldots,4)$
Reciprocal lattice basis	$\mathbf{a}_i^+ \quad (i = 1,\ldots,4)$
Direction vector, **u**	$u_i\mathbf{a}_i$
Plane normal, **h**	$h_i\mathbf{a}_i^+$
Metric of four-space, **G**	$G_{ij}^{-1} = \text{diag}\{\frac{3}{2}a^2, \frac{3}{2}a^2, \frac{3}{2}a^2, c^2\}$
Reciprocal metric, $\mathbf{G}^{-1}$	$G_{ij}{}^{1} = \text{diag}\{\frac{2}{3}a^{-2}, \frac{2}{3}a^{-2}, \frac{2}{3}a^{-2}, c^{-2}\}$
Volume of hexagonal prism (3 lattice points) defined by $\mathbf{a}_i$	$3^{3/2}\,a^2c/2$
Volume of primitive hexagonal prism defined by $\mathbf{a}_i^+$	$(\frac{2}{3})^{\frac{1}{2}}a^{-2}c^{-1}$
Scalar product, $\mathbf{u}\cdot\mathbf{v}$	$G_{ij}u_iv_j$
Repeat distance along **u** (u_i relatively prime)	$I'(G_{ij}u_iu_j)^{1/2}$
"Cell factor"	$I' = \frac{1}{3}$ if $u_1 - u_2 = 3n$ $I' = 1$ if $u_1 - u_2 \neq 3n$
Interplanar spacing (h_i relatively prime)	$(G_{ij}^{-1}\,h_ih_j)^{-1/2}$
Cosine of angles between **u** and **v**, **h** and **k**, **u** and **h**	See Table II – the equations are identical with $i,j = 1, \ldots, 4$
Zone axis, **u**, of **h** and **k**	$u_1 \propto h_4(k_2-k_3)-k_4(h_2-h_3)$ $u_2 \propto h_4(k_3-k_1)-k_4(h_3-h_1)$ $u_3 \propto h_4(k_1-k_2)-k_4(h_1-h_2)$ $u_4 \propto -3(h_1k_2-k_1h_2)$
Normal, **h**, to plane containing **u** and **v**	$h_1 \propto u_4(v_2-v_3)-v_4(u_2-u_3)$ $h_2 \propto u_4(v_3-v_1)-v_4(u_3-u_1)$ $h_3 \propto u_4(v_1-v_2)-v_4(u_1-u_2)$ $h_4 \propto -3(u_1v_2-v_1u_2)$

dimensional lattice. Particular equations (not usually derived in this way) are given in various papers (e.g. Otte and Crocker, 1965, 1966; Nicholas, 1966; Neumann, 1966; Okamoto and Thomas, 1968); most of these papers contain mistakes as pointed out by Nicholas (1970). The most important formulae are summarized in Table IIIA, which may be looked upon as a supplement to Table II. By introducing the metric of the four-dimensional space, formulae containing scalar products may be written in vector form, exactly as in Table II. Vector products, however, cannot be expressed quite so simply because the two bases are not properly reciprocal, and we have therefore written out the components of two typical vector products in full. The full tensor notation with contravariance and covariance distinguished by superscript and subscript indices has not been used in Table IIIA, but may readily be derived as in Table II.

Finally, it should perhaps be emphasized that the whole of this decriptions is concerned with the geometry of the hexagonal *lattice*. The hexagonal close-packed (h.c.p.) structure is derived from that lattice by placing two atoms around each lattice point, e.g. in sites 0, 0, 0, 0 and $\frac{1}{3}$, 0, $^{-}\frac{1}{3}$, $\frac{1}{2}$ and their equivalents (see footnote on p. 122).

7. AFFINE TRANSFORMATIONS: HOMOGENEOUS DEFORMATION

Equations (6.5) or (6.7) are commonly interpreted in two ways. In the first of these, used in the last section, the quantities ${}^{A}u_i$, ${}^{B}u_i$ are the components of the same vector **u** referred to two different sets of base vectors. An alternative interpretation is to suppose that we have a fixed reference system $\mathbf{a}_i$, and the equations then represent a physical transformation which changes a vector ${}^{A}\mathbf{u}$ into another vector ${}^{B}\mathbf{u}$, where ${}^{B}u_i$ are the components of ${}^{B}\mathbf{u}$ in the fixed reference system. This second interpretation may be used, for example, to specify the relations between two different crystal lattices, which are in fixed orientations with respect to each other. The possibility of interpreting (6.5) in these two ways is a result of the linear nature of both axis transformations and homogeneous deformations. Nevertheless, some confusion may arise if the equation is freely interpreted in either sense, as is sometimes done, and one advantage of the extended matrix notation is that it completely avoids ambiguity of this kind.

Consider the general linear transformation

$$\mathbf{v} = \mathbf{S}\mathbf{u}, \tag{7.1}$$

where **S** is a physical entity (a tensor) which converts the vector **u** into a new vector **v**. In matrix notation, this equation is written

$$\mathrm{v} = \mathrm{S}\,\mathrm{u} \quad \text{or} \quad [\mathrm{A};\mathbf{v}] = (\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,[\mathrm{A};\mathbf{u}]. \tag{7.2}$$

The extended notation emphasizes that all physical quantities, **u**, **v**, and **S** are referred to the basis A, and a clear distinction is obtained between square matrix quantities of the type (A S A), which are the representations of a tensor in some particular coordinate system, and those of type (B S A), which are the representations of a function operator connecting two coordinate systems.

During a transformation of type (7.1), points which were originally collinear remain collinear, and lines which were originally coplanar remain coplanar. Such a transformation is called affine; it represents a homogeneous deformation of space, or of the crystal lattice.

Equation (7.2) is the matrix representation of a homogeneous deformation referred to the system A. In the basis B, there will be a corresponding representation

$$[\mathrm{B};\ \mathbf{v}] = (\mathrm{B}\,\mathbf{S}\,\mathrm{B})\,[\mathrm{B};\ \mathbf{u}]. \tag{7.3}$$

Suppose the relation between A and B is given by

$$[\mathrm{B};\ \mathbf{u}] = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{u}]$$

so that (7.2) may be written

$$(\mathrm{A}\,\mathbf{J}\,\mathrm{B})\,[\mathrm{B};\ \mathbf{v}] = (\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{J}\,\mathrm{B})\,[\mathrm{B};\ \mathbf{u}]$$

and multiplying both sides by (B J A)

$$[\mathrm{B};\ \mathbf{v}] = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{J}\,\mathrm{B})\,[\mathrm{B};\ \mathbf{u}].$$

Comparing this equation with (7.3) we see that

$$(\mathrm{B}\,\mathbf{S}\,\mathrm{B}) = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{J}\,\mathrm{B}), \tag{7.4}$$

and this is called the similarity transform of (A S A) into (B S B), both of these matrices being representations of the tensor S. In shortened form, the equation is

$$^{B}\mathsf{S} = \mathsf{J}\,^{A}\mathsf{S}\,\mathsf{J}^{-1}.$$

The usefulness of the juxtaposition of the bases in the extended form of the equation should be noted.†

Now consider **u** to be any vector in the plane having normal **h** so that $\mathbf{h}\cdot\mathbf{u} = \mathsf{h}'\,\mathsf{u} = 0$. After deformation, the vector **u** is changed into a vector **v**, and the plane has a new normal **k** where $\mathbf{k}\cdot\mathbf{v} = 0$. Writing the matrix representations of these two equations in the basis A, we have

$$\mathsf{h}'\,\mathsf{u} = \mathsf{k}'\,\mathsf{S}\,\mathsf{u} = 0 \quad \text{or} \quad \mathsf{k}' = \mathsf{h}'\,\mathsf{S}^{-1}. \tag{7.5}$$

This gives the effect of the tensor S on the components of vectors normal to planes, and may be written in full as

$$(\mathbf{k};\ \mathrm{A}^{*}) = (\mathbf{h};\ \mathrm{A}^{*})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})^{-1}.$$

Let us now consider the properties of the most general form of homogeneous deformation. The following statements will be taken as self-evident, though formal proofs occur incidentally later in this section. If we imagine a sphere inscribed in the material before deformation, this would be distorted into a triaxial ellipsoid, called the strain ellipsoid. The principal axes of the strain ellipsoid would be mutually perpendicular before deformation, and would thus have suffered no relative change in orientation, although in general

† It will be noted that the components of the matrix S are those of a mixed tensor, and the tensor form of (7.2) is $v^{i} = S^{i}_{j}u^{j}$.

each would have been rotated from its original position, and changed in length. It follows also that we could have inscribed an ellipsoid in the material before deformation, such that after deformation it became a sphere; this is called the reciprocal strain ellipsoid. Lines in the directions of the axes of the reciprocal strain ellipsoid before deformation lie in the directions of the axes of the strain ellipsoid after deformation; the axes of the reciprocal strain ellipsoid are called the directions of principal strain.

In the most general deformation, all vectors change their length, but there is at least one vector (and generally three) which is unrotated. We may prove this as follows. Suppose for some vector $\mathbf{u}$ the transformation leaves $\mathbf{v}$ parallel to $\mathbf{u}$. Then the only effect of $\mathbf{S}$ is to multiply the components of $\mathbf{u}$ by a constant scalar factor, say λ_i. Then

$$\mathsf{S}\,\mathsf{u} = \lambda_i\,\mathsf{u} \quad \text{or} \quad (\mathsf{S}-\lambda_i\,\mathsf{I})\,\mathsf{u} = 0, \tag{7.6}$$

where I is the unit matrix. This equation has non-trivial solutions ($\mathsf{u}_i \neq 0$ for all i) only when

$$|\mathsf{S}-\lambda_i\,\mathsf{I}| = 0. \tag{7.7}$$

Equation (7.7) is a cubic in λ_i, and is called the characteristic equation of the matrix S. If S is real, there are three roots, of which one must necessarily be real; if the matrix S is symmetric, all three must be real. There is thus always one possible solution of (7.6), giving a vector which is unchanged in direction (if λ_i is negative, it is reversed in sign, but this does not correspond to a physically achievable deformation). If all three roots are real, there are three such directions. The values of the roots, which are called the eigenvalues of the matrix, are given by the following equations, obtained by expanding (7.7):

$$\left.\begin{aligned} \lambda_1+\lambda_2+\lambda_3 &= S_{11}+S_{22}+S_{33} = S_{ii}, \\ \lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1 &= S_{11}S_{22}+S_{22}S_{33}+S_{33}S_{11}-S_{12}S_{21}-S_{23}S_{32}-S_{31}S_{13} \\ \lambda_1\lambda_2\lambda_3 &= |\mathsf{S}|. \end{aligned}\right\} \tag{7.8}$$

It is readily proved that two matrices related by a similarity transformation have the same eigenvalues. Since the eigenvalues of a diagonal matrix are simply its non-zero components, it follows that provided the matrix has three distinct, real eigenvalues, it can always be reduced to diagonal form by a similarity transformation. Physically, this corresponds to an axis transformation to a new set of coordinates lying along the unrotated directions. It is obvious that the matrix representation of S in this system will be a diagonal matrix.

Having found the eigenvalues, we can determine an axis transformation which will diagonalize S as follows. Choose the first root, λ_1, and write (7.6) in the form:

$$(\mathsf{S}-\lambda_1\,\mathsf{I})\,\bar{\mathsf{a}}_1 = 0 \quad \text{or} \quad \mathsf{E}\,\bar{\mathsf{a}}_1 = 0, \tag{7.9}$$

where $\bar{\mathsf{a}}_1$ is the matrix representation of one of the unrotated vectors, which we now call $\bar{\mathbf{a}}_1$. Since (7.9) represents three linear homogeneous equations, we cannot determine the components $(\bar{a}_1)_i$ of $\bar{\mathbf{a}}_1$ uniquely, but we may find their ratio. If eqns. (7.9) are written in full, we see that a possible solution is

$$(\bar{a}_1)_i = CE^{ki}, \tag{7.10}$$

where C is an undetermined constant, and k may be 1, 2, or 3. Some of the cofactors E^{ki} may be zero, but there are always sufficient non-zero ones to give a solution for the vector components $(\bar{\mathbf{a}}_1)_i$. We repeat this process with the other two roots, and obtain the components of two other unrotated vectors, $\bar{\mathbf{a}}_2$ and $\bar{\mathbf{a}}_3$. The vectors $\bar{\mathbf{a}}_i$ are called the eigenvectors of S, and may be used as a new basis $\bar{\mathrm{A}}$. If the transformation from the new basis to the old is represented by $\mathsf{J} = (\mathrm{A}\,\mathbf{J}\,\bar{\mathrm{A}})$, the matrix J has columns J_{ji} consisting of the components $(\bar{a}_i)_j$ of the eigenvectors $\bar{\mathbf{a}}_i$ (see p. 32). The deformation is represented in the new basis by the matrix $\bar{\mathsf{S}} = (\bar{\mathrm{A}}\,\mathbf{S}\,\bar{\mathrm{A}})$, and from (7.4)

$$\bar{\mathsf{S}} = (\bar{\mathrm{A}}\,\mathbf{J}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{J}\,\bar{\mathrm{A}}) = \mathsf{J}^{-1}\,\mathsf{S}\,\mathsf{J}.$$

From (7.9)

$$(S\,J)_{ik} = S_{ij}J_{jk} = (S_{ij}-\lambda_k\delta_{ij})J_{jk}+\lambda_k\delta_{ij}J_{jk} = \lambda_k J_{ik},$$

so that

$$\mathsf{J}^{-1}\mathsf{S}J = \lambda_k\delta_{jk}. \tag{7.11}$$

The matrix $\bar{\mathsf{S}}$ is thus a diagonal representation of the deformation, as concluded above. The quantities λ_i give the ratios of the lengths of the vectors $\bar{\mathbf{a}}_i$ after the deformation to their lengths before deformation.

We have seen that the components of J are not determined absolutely, since each eigenvector contains an arbitrary constant. The diagonal matrix $\bar{\mathsf{S}}$ represents the strain S in all coordinate systems with axes parallel to $\bar{\mathbf{a}}_i$, and the magnitudes of the measure lengths may be chosen arbitrarily. There is also an arbitrary choice of the order of the columns of J, and correspondingly of the elements of $\bar{\mathsf{S}}$, since we may label any root of (7.7) as λ_1. This arises because we are free to label our coordinate axes in the basis $\bar{\mathrm{A}}$ in any way we please, giving six different axis transformations from A to $\bar{\mathrm{A}}$. We shall return to this question later, when discussing the idea of correspondence between directions and planes in different lattices.

When there are three real roots of (7.7), the unrotated directions may be used to specify three unrotated planes. Alternatively, these may be obtained by considering the condition that the vectors **h** and **k** which represent a plane normal before and after deformation are parallel. If h and k are the representations of **h** and **k** as column matrices in the bases A^* (as before) we have

$$\mathsf{k}' = \mathsf{h}'\,\mathsf{S}^{-1} = (1/\lambda_i)\,\mathsf{h}', \tag{7.12}$$

which corresponds to (7.6) and has non-trivial solutions only when (7.7) is satisfied. The vectors obtained by substituting the roots λ_i into (7.12) give the directions of the unrotated plane normals, and are reciprocal to the eigenvectors $\bar{\mathbf{a}}_i$. The quantities λ_i give the ratios of the initial to the final spacings of the unrotated planes.

Note that if any $\lambda_i = 1$, the corresponding eigenvector is an invariant line, i.e. a direction which is both unrotated and undistorted, and there is correspondingly a plane with an invariant normal. If there are two invariant lines, they define an invariant plane, and the basis $\bar{\mathrm{A}}$ is no longer unique. The normal to an invariant plane is necessarily unrotated, but need not itself be invariant.

A special case of homogeneous deformation arises when there are three orthogonal directions which are unrotated by the deformation. The axes of the strain ellipsoid and the reciprocal strain ellipsoid then coincide, and the deformation is said to be a pure strain. The formal definition of a pure strain is that it is a deformation such that the three orthogonal directions which remain orthogonal retain their directions and senses.

Let us now consider a deformation $\mathbf{v} = \mathbf{P}\mathbf{u}$ which is represented in an orthonormal basis I by the equation

$$[\mathrm{I};\ \mathbf{v}] = [\mathrm{I}\,\mathbf{P}\,\mathrm{I}]\,[\mathrm{I};\ \mathbf{u}]. \tag{7.13}$$

Consider a second orthonormal basis K, related to I by

$$[\mathrm{K};\ \mathbf{u}] = (\mathrm{K}\,\mathbf{L}\,\mathrm{I})\,[\mathrm{I};\ \mathbf{u}]. \tag{7.14}$$

The matrix $(\mathrm{K}\,\mathbf{L}\,\mathrm{I}) \equiv \mathsf{L}$ represents an axis transformation which is merely equivalent to rotating the basis I into a new position.† The components L_{ij} are thus the cosines of the angles between $\mathbf{k}_i$ and $\mathbf{i}_j$. By writing $\mathbf{i}_j = L_{ij}\mathbf{k}_i$, we obtain

$$\mathbf{i}_j \cdot \mathbf{i}_k = L_{ij}L_{lk}\mathbf{k}_i \cdot \mathbf{k}_l = L_{ij}L_{lk}\delta_{il} = L_{ij}L_{ik}$$

and similarly $\mathbf{k}_j \cdot \mathbf{k}_k = L_{ji}L_{ki}$. Since both these scalar products are equal to δ_{jk}, we obtain the well known relations between the direction cosines of the axes

$$L_{ij}L_{ik} = L_{ji}L_{ki} = \delta_{jk},$$

or in matrix form

$$\mathsf{L}\,\mathsf{L}' = \mathsf{I}, \quad \mathsf{L}' = \mathsf{L}^{-1}, \quad |\mathsf{L}| = \pm 1. \tag{7.15}$$

Such a matrix is called orthogonal. When $|\mathsf{L}| = +1$, the bases I and K both correspond to right-handed (or left-handed) sets of base vectors, and the transformation of axes represented by L is a proper rotation. When $|\mathsf{L}| = -1$, a right-handed set of base vectors is converted into a left-handed set, and vice versa; this is equivalent to a rotation plus a reflection in some plane, and is called an improper rotation. We consider only proper rotations.

If the strain (7.13) is now referred to the basis K, its representation is given by (7.4)

$$(\mathrm{K}\,\mathbf{P}\,\mathrm{K}) = (\mathrm{K}\,\mathbf{L}\,\mathrm{I})\,(\mathrm{I}\,\mathbf{P}\,\mathrm{I})\,(\mathrm{I}\,\mathbf{L}\,\mathrm{K}) = (\mathrm{K}\,\mathbf{L}\,\mathrm{I})\,(\mathrm{I}\,\mathbf{P}\,\mathrm{I})\,(\mathrm{I}\,\mathbf{L}'\,\mathrm{K}).$$

Hence taking the transpose of both sides

$$(\mathrm{K}\,\mathbf{P}'\,\mathrm{K}) = (\mathrm{K}\,\mathbf{L}\,\mathrm{I})\,(\mathrm{I}\,\mathbf{P}'\,\mathrm{I})\,(\mathrm{I}\,\mathbf{L}'\,\mathrm{K}),$$

and it follows that if $(\mathrm{I}\,\mathbf{P}\,\mathrm{I})$ is a symmetric matrix, so also is $(\mathrm{K}\,\mathbf{P}\,\mathrm{K})$. Symmetric matrices thus remain symmetric as a result of an orthogonal transformation.

If the deformation is a pure strain, we may choose the vectors $\mathbf{k}_i$ to lie along the principal axes of this strain (i.e. $\mathrm{K} = \bar{\mathrm{I}}$). The matrix $(\mathrm{K}\,\mathbf{P}\,\mathrm{K}) = \bar{\mathsf{P}}$ is then diagonal, and $(\mathrm{I}\,\mathbf{P}\,\mathrm{I})$ must therefore be symmetrical. A pure strain is thus characterized by a symmetric representation of the tensor $\mathbf{P}$ in any orthonormal basis.

† We use the symbol $\mathbf{L}$ for axis transformations which are pure rotations, and $\mathbf{J}$ for more general axis transformations. A rotation contains three independent quantities, two to specify the axis of rotation and one to fix the magnitude of the rotation about this axis. It may also be described by a Rodrigues vector $\mathbf{r} \tan(\theta/2)$ used in a Rodrigues-Frank map (Frank, 1988) or by a quarternion $[r_1 \sin\theta/2,\ r_2 \sin\theta/2,\ r_3 \sin\theta/2,\ \cos\theta/2]$ (Handscomb 1958).

A pure deformation is equivalent to simple extensions or contractions along the three principal axes of strain. If the material consists of a rectangular parallelepiped with edges along the principal axes, it will thus remain a rectangular parallelepiped after deformation. The ratio of the new volume to the original volume is then $\lambda_1\lambda_2\lambda_3 = |\mathsf{P}|$, where P is the matrix representation of the strain in any orthonormal basis, and λ_i are the eigenvalues of P.

In simple cases, the pure strain matrix P may be reduced to diagonal form by inspection. Thus if the basis I may be transformed into the basis $\bar{\mathrm{I}}$ by a single rotation about one of the vectors $\mathbf{i}_i$ the components L_{ij} $(i \neq j)$ of the rotation matrix are zero. In the general case, the problem is equivalent to finding the principal axes of a quadric surface. We use the procedure of p. 44, and since P is symmetric, the roots of the equation

$$(\mathsf{P} - \lambda_i \mathsf{I})\,\mathsf{x} = 0, \tag{7.16}$$

are necessarily all real. This equation is identical with an equation known as the discriminating cubic of the geometrical surface with scalar equation

$$\mathsf{x}'\,\mathsf{P}\,\mathsf{x} = P_{ij}x_i x_j = \text{const.} \tag{7.17}$$

As before, the columns of the matrix J which diagonalizes P are multiplied by undetermined constants. However, we wish the basis $\bar{\mathrm{I}}$ to be orthonormal, so we must normalize the eigenvectors by choosing the constants so that eqns. (7.15) are satisfied. In this way, J becomes an orthogonal matrix L which is unique, except for the order of its columns. Three of the six ways in which the columns can be arranged correspond to improper rotations; of the remaining three, the most obvious choice is to label the principal axis which makes the smallest angle with $\mathbf{i}_1$ as $\bar{\mathbf{i}}_1$, etc.

A pure strain is one of the two component deformations into which any homogeneous deformation may be analysed. The other type is a pure rotation, characterized by the condition that all vectors remain the same length. Obviously, a pure rotation is given by a tensor relation $\mathbf{v} = \mathbf{R}\mathbf{u}$, in which the components of the tensor $\mathbf{R}$ form an orthogonal matrix in an orthonormal basis. This follows since if v and u are the vector representations in the orthonormal basis

$$|\mathbf{v}|^2 = \mathsf{v}'\,\mathsf{v} = \mathsf{u}'\,\mathsf{R}'\,\mathsf{R}\,\mathsf{u} = \mathsf{u}'\,\mathsf{u} = |\mathbf{u}|^2,$$

provided R is orthogonal.

Any homogeneous deformation may be regarded as the result of a pure strain combined with a pure rotation. Thus we may write $\mathbf{v} = \mathbf{S}\mathbf{u}$ as

$$\mathsf{v} = \mathsf{S}\,\mathsf{u} = \mathsf{P}_1\,\mathsf{R}_1\,\mathsf{u} = \mathsf{R}\,\mathsf{P}\,\mathsf{u},$$

where P_1, P represent pure strains, and R_1, R pure rotations.

Note that $\mathsf{P}_1 \neq \mathsf{P}$ and $\mathsf{R}_2 \neq \mathsf{R}$; there are two ways of resolving the deformation, depending on whether the rotation or the pure strain is considered to occur (mathematically) first. For the present, we find it convenient to use the second resolution, in which:

$$\mathsf{y} = \mathsf{R}\,\mathsf{P}\,\mathsf{x}, \tag{7.18}$$

means that the vector **x** is first given a pure strain **P** and then a pure rotation **R** to form a new vector **y**.

We have already seen that the ratio of the transformed volume to the original volume during the pure strain is given by $|P|$. Since the rotation cannot change the volume, this quantity is also equal to the volume ratio for the whole deformation **S**. Moreover, since $|R| = 1$, $|S| = |P|$, so that for any affine transformation, the volume ratio is given by $|S|$. Finally, we note from eqn. (7.4) that for any axis transformation, $|^{B}S| = |^{A}S|$. The volume ratio is thus given by the determinant of *any* matrix representation of **S**.

The geometrical relations involved in the general deformation may be appreciated by reference to Fig. 2.2, which is, however, two-dimensional. As a result of the deformation **S**,

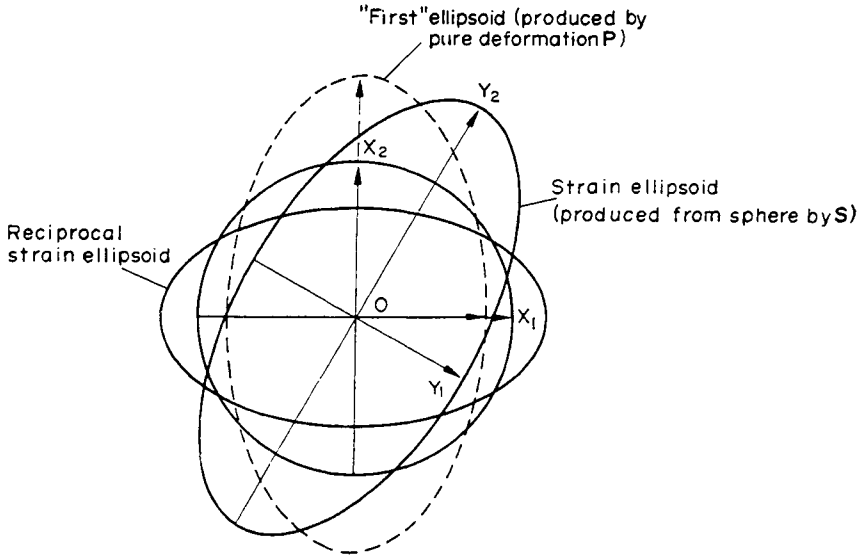

FIG. 2.2. Geometrical relations in homogeneous deformation.
The axis about which the first ellipsoid is rotated into the strain ellipsoid may have any orientation, but is assumed to be OX_3 in order to give a two-dimensional figure.

a sphere is distorted into the strain ellipsoid, and the vectors OX_1, OX_2 (and OX_3 not shown) become OY_1, etc. The deformation is regarded as taking place in two parts; during the pure strain, the vectors OX_1, etc. undergo simple extension or contraction to produce an ellipsoid which is shown dotted. There is no standard name for this figure, but we shall refer to it as the *first* ellipsoid. Following this, there is a rotation (not necessarily in the plane of OX_1 and OX_2) to the positions OY_1, etc.

The reciprocal strain ellipsoid will obviously be changed into a sphere by the pure strain **P**, since **R** produces no change in shape. The axes of this ellipsoid will thus be coincident with the principal axes of strain, as previously assumed. Given the matrix representation of the general deformation, in any set of orthonormal axes, we may resolve into P and R as follows. Any vector **x** is converted into $\mathbf{y} = \mathsf{S}\,\mathbf{x}$. Now suppose **x** represents a radius vector

of the reciprocal strain ellipsoid. Then after deformation, the components y satisfy the equation

$$y_1^2+y_2^2+y_3^2 = \text{const.}$$

or in matrix form

$$\mathsf{y}'\,\mathsf{y} = \text{const.}$$

Substituting for y in terms of the original components x, we obtain the scalar equation of the reciprocal strain ellipsoid

$$\mathsf{x}'\,\mathsf{S}'\,\mathsf{S}\,\mathsf{x} = S_{ki}S_{kj}x_ix_j = \text{const.} \tag{7.19}$$

By expanding S into its components R and P

$$\mathsf{S}'\,\mathsf{S} = \mathsf{P}'\,\mathsf{R}'\,\mathsf{R}\,\mathsf{P} = \mathsf{P}'\,\mathsf{P} = \mathsf{P}^2 \tag{7.20}$$

and this set of equations is sufficient to determine P. The symmetric matrix S′ S has eigenvalues λ_i^2 which are the squares of the eigenvalues of P. The reciprocal strain ellipsoid has the property that the ratio of the length of any deformed vector to its original length is proportional to the inverse radius vector of the ellipsoid drawn in the original direction.

From the eigenvalues of S′ S, we can construct the matrix $\bar{\mathsf{P}}$ which is the diagonal representation of **P** in an orthonormal system along the principal axes, and we can also find the orthogonal transformation $\mathsf{L} = (\bar{\mathsf{I}}\,\mathbf{L}\,\mathsf{I})$ which transforms the basis I into the principal basis $\bar{\mathsf{I}}$. The representations of the components of S in the original basis are then given by

and

$$\left.\begin{aligned}\mathsf{P} &= \mathsf{L}^{-1}\bar{\mathsf{P}}\,\mathsf{L}\\ \mathsf{R} &= \mathsf{S}\,\mathsf{P}^{-1}.\end{aligned}\right\} \tag{7.21}$$

When referred to principal axes, the whole deformation takes the form

$$\bar{\mathsf{y}} = \bar{\mathsf{S}}\,\bar{\mathsf{x}} = \mathsf{L}\,\mathsf{R}\,\mathsf{L}^{-1}\,\bar{\mathsf{P}}\,\bar{\mathsf{x}}. \tag{7.22}$$

We can also show that the above procedure gives the principal axes without explicit reference to the reciprocal strain ellipsoid. Suppose we have two vectors, $\mathbf{x}_1$ and $\mathbf{x}_2$, which are converted into two orthogonal vectors, $\mathbf{y}_1$, and $\mathbf{y}_2$. Then in the basis I

$$\mathbf{y}_1\cdot\mathbf{y}_2 = \mathsf{y}_1'\,\mathsf{y}_2 = \mathsf{x}_1'\,\mathsf{S}'\,\mathsf{S}\,\mathsf{x}_2 = 0.$$

Now if the two vectors were perpendicular before the deformation, $\mathsf{x}_1'\,\mathsf{x}_2 = 0$, so that

$$\mathsf{x}_1'\,\mathsf{S}'\,\mathsf{S}\,\mathsf{x}_2 = \lambda_i^2\,\mathsf{x}_1'\,\mathsf{x}_2.$$

where λ_i is a scalar. The vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are then both solutions of the equation

$$(\mathsf{S}'\,\mathsf{S}-\lambda_i^2\,\mathsf{I})\,\mathsf{x} = 0 \tag{7.23}$$

and for non-trivial solutions

$$|\mathsf{S}'\,\mathsf{S}-\lambda_i^2\,\mathsf{I}| = 0. \tag{7.24}$$

This gives three orthogonal vectors which define the principal axes, and the procedure is equivalent to diagonalizing the matrix S′ S, as described above. Moreover, if the three roots

are λ_1^2, λ_2^2, λ_3^2, we have

$$|\mathbf{y}_1|^2 = \mathbf{y}_1' \mathbf{y}_1 = \lambda_1^2 \mathbf{x}_1' \mathbf{x}_1,$$

so that the principal deformations are λ_i.

The above equations have to be modified when the deformation is expressed in a general basis A. The scalar product of $\mathbf{y}_1$ and $\mathbf{y}_2$ is now written

$$\mathbf{y}_1 \cdot \mathbf{y}_2 = (\mathbf{y}_1;\ \mathrm{A})\,(\mathrm{A}^*\,\mathbf{G}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{y}_2] = (\mathbf{x}_1;\ \mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})'\,(\mathrm{A}^*\,\mathbf{G}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{x}_2]$$

so that the general equation corresponding to (7.23) is

$$\{(\mathrm{A}\,\mathbf{S}\,\mathrm{A})'\,(\mathrm{A}^*\,\mathbf{G}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A}) - \lambda_i^2(\mathrm{A}^*\,\mathbf{G}\,\mathrm{A})\}\,[\mathrm{A};\ \mathbf{x}] = 0, \tag{7.25}$$

and there is an obvious corresponding equation for finding the characteristic roots.

The strain ellipsoid has a scalar equation which may be found by making $|\mathbf{x}|^2$ constant. Thus in the basis I

$$\mathbf{y}'\,(\mathbf{S}^{-1})'\,\mathbf{S}^{-1}\,\mathbf{y} = \text{const.} \tag{7.26}$$

This surface has the geometrical interpretation that the ratio of the length of a deformed vector to its original length is proportional to the radius vector of the ellipsoid drawn in the final direction of the vector.

Finally, we note that the surface

$$\mathbf{x}'\,(\mathbf{S} - \mathbf{I})\,\mathbf{x} = \text{const}, \tag{7.27}$$

is called the elongation quadric, or in linear elasticity theory (Section 10), the strain quadric. The surface may be either an ellipsoid or a hyperboloid; it has axes in the same directions as those given by eqn. (7.17), and in general these do not coincide with the principal axes of strain. The elongation quadric has the geometrical interpretation that the *extension* of any line, resolved in the original direction of that line, is inversely proportional to the square of the radius vector to the surface, drawn in the original direction of the line.

8. TWIN CRYSTALS

Solid metals are usually composed of a compact mass of separate crystals or grains, joined along arbitrary internal surfaces, and randomly orientated with respect to each other. The orientation relation between any two grains having the same crystal structure may be expressed by a tensor relation representing a pure rotation, and the transformation may always be achieved by a proper rotation, with $|\mathbf{R}| = +1$. In crystals of fairly high symmetry, the relation may also be expressed as an improper rotation, if so desired. In certain crystals which possess no centre of symmetry and few planes of symmetry, an improper rotation may produce an atomic arrangement which is not obtainable by a proper rotation. Such arrangements are called optical isomorphs, and show optical activity, i.e. the ability to rotate the plane of polarized light. They do not occur in metals.

The relation between two randomly orientated crystals thus requires three degrees of freedom for its specification, since there are three independent quantities in a rotation mat-

rix. If these crystals meet along a grain boundary surface, two further parameters are needed to specify the orientation of this surface at any point. The general grain boundary thus has five degrees of freedom. The concept of grain boundaries really belongs to the subject of crystal imperfections, and is considered in Chapter VIII. We may usefully consider here, however, the transformation between two orientations which are related in a well specified manner to the symmetry of the structure, so that the crystals are said to be twins of each other. Two crystals in twinned orientation may still be joined along any surface, but there is always some plane which will give a boundary of very low energy, and this composition plane then has no degrees of freedom.

Two crystals are twins of each other when they may be brought into coincidence either by a rotation of 180° about some axis (the twinning axis) or by reflection across some plane (the twinning plane).† The possible orientation relations may be further classified by the relation of the axis of symmetry to the composition plane of low energy along which the twin crystals are usually joined. The rotation axis may be normal to this composition plane (normal twins) or parallel to the composition plane (parallel twins). If the atomic structure is centrosymmetric, it follows that a normal rotation twin is equivalent to a reflection in the composition plane, which is then the twinning plane. For non-centrosymmetric structures, the operations of normal rotation and reflection will produce different twins having the same composition plane. Similarly, a parallel rotation twin is equivalent to a twin produced by a reflection in a plane normal to the rotation axis, and hence to the composition plane, for centrosymmetric structures, but these two operations produce different twins in non-centrosymmetric structures. There are thus four possible types of orientation relation between two twins, reducing to two equivalent pairs for centrosymmetric structures (and, of course, lattices).

Since most metallic structures are centrosymmetric, articles on twinning sometimes refer only to two types of twin orientation. These are then designated as "reflection" twins (equivalent to normal rotation twins) and "rotation" twins (equivalent to reflection in a plane normal to the composition plane). We have presented these results in axiomatic form, but they follow naturally from the condition that the two lattices fit together exactly along the composition plane. We shall emphasize this aspect of twinning in a more complete treatment in a later chapter.

Obviously, it is not possible to have reflection twins in which the twinning plane is a plane of symmetry in the crystal structure, since the twinning operation then merely reproduces the original orientation. In the same way, the twinning axis in a rotation twin can never be an even axis of symmetry. When two crystals are joined in twin orientation, the twinning plane and/or the twinning axis become pseudo-symmetry elements of the composite structure.

Most metallic crystals form twins which may be regarded both as reflection and normal rotation twins, though there are other types in metals of low symmetry. We could specify the twin relations by the rotation matrix R required to bring the crystals into coincidence

† A more general description of a mechanical twin (Crocker, 1962) is any region of the parent which has undergone a homogeneous shear to give a re-orientated region with the same crystal structure. The above orientation relations are then not necessarily valid; this is discussed in Chapter 20.

with each other, but it is often more useful to employ another type of deformation tensor. In many metals, and some non-metallic crystals, twins may be formed by a physical deformation of the original structure, and this is known as glide twinning. The process is macroscopically equivalent to a homogeneous shear of the original structure, and we therefore use this kind of deformation to describe the twinning law.

Figure 2.3 shows a section through a three-dimensional lattice which has undergone glide twinning. The open circles represent the lattice points in their original positions; the

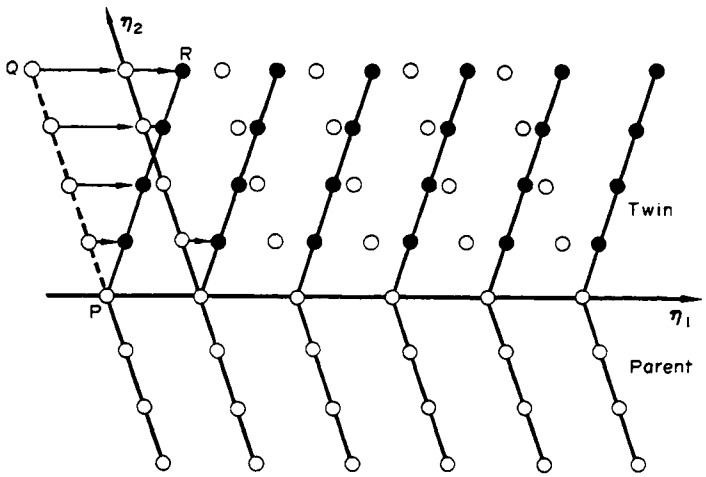

FIG. 2.3. Simple illustration of glide twinning.
PQ, *PR* are corresponding directions in parent and twin.

filled circles are the final positions to which they move. It will be seen that, in the twinned region, each lattice point moves in the same direction through a distance proportional to its distance from the composition plane, which is parallel to the direction of movement. In the simple example shown in the figure, the twinned structure is a reflection of the original structure in the composition plane, and the twinned lattice is obtained by a homogeneous simple shear of the original lattice. The composition, or twinning, plane is conventionally denoted K_1 and the direction of shear η_1. The plane containing η_1 and perpendicular to K_1 (i.e. the plane of the diagram) is called the plane of shear.† Note that each lattice point moves only a fraction of the lattice repeat distance.

In the last paragraph, we emphasized that the two *lattices* in glide twinning are related by a homogeneous shear. The two *structures* are not so related unless the primitive unit cell contains only one atom, i.e. the structure may be obtained by placing an atom on each point of the lattice. More generally, some of the atoms must move in different directions from the lattice points; inhomogeneous movements of this kind are sometimes called atomic "shuffles". The *macroscopic* effect of the deformation is unaffected by the shuffles.

† The K_1 plane is sometimes referred to as the shearing plane. This usage is better avoided, because of the possibility of confusion between "shearing plane" and "plane of shear".

We can thus describe the twin orientation by a matrix representing a simple shear, providing we confine our attention to the lattice and ignore the vectors $\xi_{A,n}$ of eqn. (5.8).

Figure 2.4 shows the section of an original sphere which becomes an ellipsoid after the deformation. The section is in the plane of shear; vectors perpendicular to this plane are unaffected by the deformation, so the problem is essentially two-dimensional. If we use an orthonormal basis, with axes x_i parallel to η_{11}, perpendicular to the plane of shear and perpendicular to K_1 respectively, we may specify the twin relation as

$$\mathbf{y} = \mathbf{S}\,\mathbf{x},$$

where

$$\mathbf{S} \equiv \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{8.1}$$

and s, the amount of shear, is the distance moved by a lattice point at unit distance from the plane K_1. We note that $|\mathbf{S}| = 1$, so there is no volume change in the transformation, as is physically obvious.

From Fig. 2.4., we see that the original sphere and the strain ellipsoid meet in two circ-

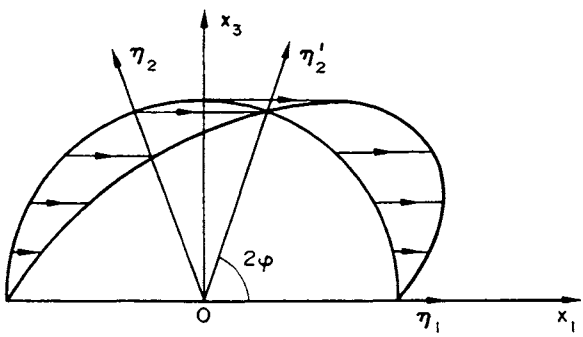

FIG. 2.4. Geometry of glide twinning.
s is the displacement at unit distance along OX_3.

les, which thus represent planes which are undistorted by the transformation. One of these is the composition plane K_1 and this is also unrotated; the second plane is rotated through an angle $\pi - 4\varphi = 2 \arctan(s/2)$ where 2φ is the angle between the undistorted planes in their final positions. The second undistorted plane is denoted K_2, and its intersection with the plane of shear is in the direction η_2. The relation between s and φ is

$$s = 2 \cot 2\varphi \tag{8.2}$$

and the equation of the K_2 plane in its initial and final positions is

$$x_1/x_3 = \mp s/2. \tag{8.3}$$

The matrix S is not symmetrical, and the deformation is thus not a pure shear. Using the method of pp. 44-7 or more simply from the geometry of Fig. 2.4, we find that the prin-

cipal axes of strain are obtained by a right-handed rotation of x_i through an angle φ about x_2. Thus maximum extension takes place in a direction at $\pi/2-\varphi = \arctan[\frac{1}{2}\{s+(s^2+4)^{1/2}\}]$ to η_1, and the change in length is in the ratio $[1+\frac{1}{2}s\{s+(s^2+4)^{1/2}\}]^{1/2} : 1$. Maximum contraction takes place in a direction at an angle $\varphi = \arctan[\frac{1}{2}\{s-(s^2+4)^{1/2}\}]$ to η_1, and the change in length is in the ratio $[1+\frac{1}{2}s\{s-(s^2+4)^{1/2}\}]^{1/2} : 1$.

The simple shear is equivalent to a pure shear, specified by the above extension and contraction, followed by a right-handed rotation of $\pi/2-2\varphi$ about x_2. During the pure shear, the K_1 and K_2 planes rotate through equal angles $\pi/2-2\varphi$ in opposite directions. It is thus clear that if we superimpose a left-handed rotation of this amount on the pure shear, the plane hitherto denoted K_2 will be unrotated in the total deformation. The role of the K_1 and K_2 planes, and also of the η_1 and η_2 directions is thus interchanged; twins so related are called reciprocal or conjugate twins. Two crystals in twin orientation may be transformed into reciprocal twin orientation by a relative rotation of $\pi-4\varphi = 2 \arctan(s/2)$ about the normal to the plane of shear.

The above specification of glide twinning does not yet include the most important condition, namely that the twinning deformation should produce an equivalent lattice to the original lattice. If the twin is a mirror image in the composition plane, it is geometrically obvious that lattice points which are reflections of each other after twinning must have been separated by a vector parallel to η_2 before twinning. For this type of twinning, therefore, η_2 must be a rational direction parallel to a row of lattice points. The lattice structure will obviously be preserved if any three non-coplanar lattice vectors in the parent crystal are transformed into vectors in the twin which retain their lengths and mutual inclinations. All vectors which remain unchanged in length lie in either K_1 or K_2, so we must select one lattice vector from one of these planes, and two from the other. This means that either K_1 or K_2 must be a rational plane. The angle between any two vectors in either K_1 or K_2 is unchanged by the twinning shear, but in general a vector in K_1 and a vector in K_2 change their relative inclination. However, the angle between η_1 and any vector in K_2 is unchanged, as is the angle between η_2 and any vector in K_1. Our three vectors may thus be η_2 and any two lattice vectors in K_1, or η_1 and any two lattice vectors in K_2. In the first case, we have η_2 and K_1 rational, and this is called a twin of the first kind; a twin of the second kind has η_1 and K_2 rational.†

The two possibilities are illustrated in Figs. 2.5 and 2.6. Twins of the first kind are a simple reflection of the parent crystal in the K_1 plane, and the twin *lattice* may equivalently be obtained by a rotation of 180° about the normal to K_1. As mentioned above, these two descriptions of the *structure* are also equivalent if this is centrosymmetric. Twins of the second kind are related by a rotation of 180° about η_1, and for centrosymmetric structures, this is equivalent to a mirror reflection in the plane perpendicular to η_1, as shown in the figure.

† If the more general definition of mechanical twinning mentioned on p. 52 is adopted, twins with all four elements irrational are possible. In practice, such twinning modes probably result only from 'double twinning', that is, from the combination of two twinning operations of the type discussed in the text. More rigorous proofs of the statements in the text are given in Chapter 20.

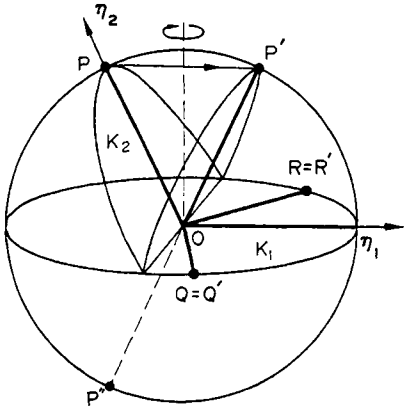

FIG. 2.5. Type I twinning.

P, *Q*, *R* represent lattice vectors before twinning. *P'*, *Q'*, *R'* represent the corresponding vectors after twinning (after Cahn, 1953).

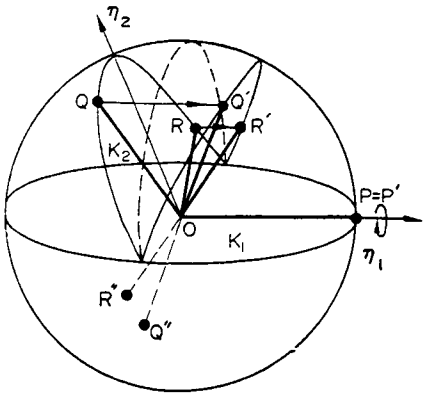

FIG. 2.6. Type II twinning.

P, *Q*, *R* represent lattice vectors before twinning. *P'*, *Q'*, *R'* represent the corresponding vectors after twinning (after Cahn, 1953).

If the basic vectors used to define the twinning elements are all primitive vectors of the lattice, the homogeneous deformation will produce the twinned lattice, as assumed above. However, it is possible to produce twinned structures even when all these vectors are not primitive. In this case, some of the atomic movements must be inhomogeneous, even in a structure without a basis, since only a fraction of the lattice points move to their twinned positions as the result of the shear. In effect, the macroscopic shear now converts a superlattice of the structure into its twinned orientation.

For any crystal structure, the twinning elements and the amount of shear are completely specified by either η_2 and K_1 or by η_1 and K_2. All four elements may be rational, and we then have compound twins. Most twins are compound in metals of cubic, trigonal and tetragonal symmetry, but both type I and type II twins have been found in orthorhombic α-uranium. In metals of the highest symmetry, K_1 and K_2 and η_1 and η_2 are frequently crystallographically equivalent, so that the twin and its reciprocal represent equivalent twinning modes.

If the deformation is represented in a basis A of the parent crystal by S, we may now use the transformation formulae previously derived to find the new indices of any plane or lattice direction relative to A. If these new indices are $\mathsf{k} = [\mathrm{A}^*; \mathbf{k}]$ and $\mathsf{v} = [\mathrm{A}; \mathbf{v}]$ where k and v are the new plane normal and lattice vector, we have

$$\mathsf{k}' = \mathsf{h}'\,\mathsf{S}^{-1} \quad \text{and} \quad \mathsf{v} = \mathsf{S}\,\mathsf{u}.$$

In general, both directions and planes in the twin have irrational indices when referred to a basis in the parent. We then find it useful to introduce the idea of *correspondence* between the two lattices. Refer the vector v to a new basis which forms a unit cell in the twin. Since twin and parent have the same structure, it will be natural to choose a unit cell of the same size and shape as that outlined by the vectors $\mathbf{a}_i$; the new basis B will then differ from A

only by a rotation (a proper rotation if both sets of base vectors are right-handed). We now have

$$\text{and} \qquad \begin{aligned} {}^{B}\mathbf{v} &= [\mathrm{B};\, \mathbf{v}] = (\mathrm{B}\,\mathbf{L}\,\mathrm{A})\,[\mathrm{A};\, \mathbf{v}] \\ {}^{B}\mathbf{v} &= \mathbf{L}\,{}^{A}\mathbf{v} = \mathbf{L}\,\mathbf{S}\,{}^{A}\mathbf{u} = \mathbf{C}\,{}^{A}\mathbf{u}. \end{aligned} \tag{8.4}$$

The matrix $\mathbf{C} = (\mathrm{B}\,\mathbf{C}\,\mathrm{A})$ is called the correspondence between the two lattices; it combines the effects of the deformation and the change of basis. Clearly, since $\mathbf{v}$ is rational in B and $\mathbf{u}$ is rational in A, ${}^{B}\mathbf{v}$ and ${}^{A}\mathbf{u}$ both have rational components, and the components of C are all rational. Moreover, since the bases A and B both refer to a unit cell of the same volume, $|\mathbf{C}| = \pm 1$. From the result on p. 45, it follows that the columns of $(\mathrm{B}\,\mathbf{C}\,\mathrm{A})$ are the components referred to base B of the vectors which are formed from the base vectors $\mathbf{a}_i$ of A by the transformation S. If we use the transformed base vectors of A as the base vectors of B (symbolically, B = CA), the correspondence matrix is $(\mathrm{CA}\,\mathbf{C}\,\mathrm{A}) = \mathbf{I}$. We can thus always establish a unitary correspondence between direction indices in the parent and in a suitable basis in the twin. In type I twinning, for example, we could choose the basis A to have $\mathbf{a}_1$ and $\mathbf{a}_2$ in K_1 and $\mathbf{a}_3 = \mathbf{h} - \frac{1}{2}\mathbf{s}$ where $\mathbf{h}$ is the vector normal to K_1 and $\mathbf{s}$ is a vector in K_1 chosen so that $\mathbf{a}_3$ is parallel to η_2. After twinning, the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are unchanged and $\mathbf{a}_3$ becomes $\mathbf{h} + \frac{1}{2}\mathbf{s}$. If these three vectors are used as the basis B, then $\mathbf{C} = \mathbf{I}$. However, it is sometimes more convenient to choose bases so that the correspondence matrix is not the unit matrix. When this is done, the transformation of direction indices is given by eqn. (8.4), and the correspondence between plane indies is specified by

$$\left.\begin{aligned} (\mathbf{k};\, \mathrm{B}^*) &= (\mathbf{k};\, \mathrm{A}^*)\,(\mathrm{A}^*\,\mathbf{L}\,\mathrm{B}^*) \\ &= (\mathbf{h};\, \mathrm{A}^*)\,(\mathrm{A}\,\mathbf{S}^{-1}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{L}\,\mathrm{B}) \\ &= (\mathbf{h};\, \mathrm{E}^*)\,(\mathrm{A}\,\mathbf{C}\,\mathrm{B}) \end{aligned}\right\} \tag{8.5}$$

using the result of (6.11). This equation may be written more briefly as

$${}^{B}\mathbf{k}' = {}^{A}\mathbf{h}'\,\mathbf{C}^{-1}.$$

9. RELATIONS BETWEEN DIFFERENT LATTICES

In deformation twinning, the magnitude of the shear is fixed, once the twinning elements have been specified. An arbitrary shear would produce a new lattice with a different symmetry, but having the same volume. More generally, any space lattice may be converted into any other lattice by a homogeneous deformation S. The deformation determines not only the symmetry and parameters of the new lattice, but also its orientation relative to the original lattice. The orientation of the interphase boundary (i.e. the surface of separation) must be specified separately.

When we have two crystal structures in contact with each other in a fixed relative orientation, it is often convenient to choose a deformation tensor S to describe the relation between the lattices. We have already emphasized that whilst any two unit cells of the lattices may be connected by a suitable homogeneous deformation, the positions of the atoms cannot necessarily all be described in this way. If the primitive unit cells of the two structures

contain different numbers of atoms, the conversion of one structure into another involves a net loss or gain of lattice points. In such cases, it seems sensible to choose a tensor **S** which relates unit cells containing the same number of atoms, and this is essential in transformations where there is a correspondence (see below).

In the case of twinning, we emphasized that shuffles may be produced either because the structure contains more than one atom per unit cell, or because the simple shear **S** does not relate all the lattice points. The same conclusion applies to more general lattice deformations; the smallest unit cells related by **S** need not be primitive cells of either structure, and shuffling is then required to complete the transformation even if both structures have only one atom per lattice point. A distinction is sometimes made between these latter type "lattice shuffles", which are determined by choice of **S**, and the more general "structure shuffles", which arise from the atomic position vectors $\boldsymbol{\xi}$ of (5.8).

For complete generality, the relation between any two lattices should be written in the form

$$\mathbf{v} = \mathbf{t} + \mathbf{S}\mathbf{u}, \tag{9.1}$$

where the tensor **S** specifies the relative sizes and orientations of the unit cells, and the constant vector **t** represents a translation of the lattice points of one crystal relative to those of the other. In general, such translations are not detectable by ordinary crystallographic methods, and are of interest only when the actual atomic positions in two lattices separated by an interface are being considered (Section 36). A special case is when $\mathbf{S} = \mathbf{I}$; the relation then describes a surface defect, known as a stacking fault, in a single lattice (Section 16). For the remainder of this section, we assume $\mathbf{t} = 0$ and consider only the orientation and structural relations specified by **S**.

In any structure, there is an infinite number of operations which will bring the lattice into self-coincidence. Correspondingly, there is an infinite number of deformations **S** which will convert a specified unit cell into another specified unit cell with given orientation relation. It is clear that as long as we wish merely to give a formal statement of the relative positions of the two sets of lattice points, any deformation **S** which gives the desired relation is valid. In some phase transformations, however, the atoms in a region of a product crystal have moved from their original positions in the parent crystal in such a way that the lattice of the parent has been effectively deformed into the lattice of the product. There is then a particular tensor **S** which not only specifies the relations of the lattices, but also the way in which one lattice may change into another. This tensor has usually to be selected by some external physical assumption, e.g. that each point of the original lattice moves to the nearest point of the final lattice. The simplest example of the physical significance of the choice of **S** occurs in twinning. A particular twinning law could be represented by a rotation about a suitable axis by any odd multiple of π, or alternatively by a simple shear deformation. For mechanical twinning, the latter statement of the law is more meaningful, since one lattice is physically sheared into the other.

A deformation which is physically significant implies a one to one correspondence between vectors in the two lattices. Each vector in one lattice may be associated unambiguously

with a "corresponding" vector of the other lattice into which it is converted by the transformation. We summarize these relationships by means of a correspondence matrix, as already used for twin crystals.

Suppose we use bases A and B in the two lattices which we call α and β respectively; the unit cells defined by A and B need not contain the same numbers of atoms. The relation between vectors in the two lattices is given by

$$[\mathrm{A};\ \mathbf{v}] = (\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{u}],$$

and the relation between the two bases is

$$[\mathrm{B};\ \mathbf{u}] = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{u}].$$

Combining these two equations,

$$[\mathrm{B};\ \mathbf{v}] = (\mathrm{B}\,\mathbf{J}\,\mathrm{A})\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{u}] = (\mathrm{B}\,\mathbf{C}\,\mathrm{A})\,[\mathrm{A};\ \mathbf{u}],$$

or briefly

$$^{B}\mathsf{v} = \mathsf{C}\,{}^{A}\mathsf{u}. \tag{9.2}$$

As on p. 56, the columns of the correspondence matrix $\mathbf{C} = (\mathrm{B}\,\mathbf{C}\,\mathrm{A})$ are the components referred to basis B of the vectors which are formed from the base vectors $\mathbf{a}_i$ by the deformation S. If these transformed vectors are used to define the new unit cell (B = CA), the correspondence matrix is I. However, this will not usually happen if B and A are derived from conventional unit cells in the two structures.

On pp. 47-8 we showed that the determinant of $(\mathrm{A}\,\mathbf{S}\,\mathrm{A})$ gives the ratio of the volume of the β structure to that of the α structure. If the bases A and B contain the same number of atoms, the determinant of $(\mathrm{B}\,\mathbf{J}\,\mathrm{A})$ will give the ratio of the volume per atom in the α structure to that in the β structure, and hence the determinant of $(\mathrm{B}\,\mathbf{C}\,\mathrm{A})$ will be unity.† Correspondingly, if there are different numbers of atoms in the two bases (as may often happen if $\mathbf{a}_i$ and $\mathbf{b}_i$ define primitive or conventional unit cells), $|\mathbf{C}|$ will equal the ratio of the number of atoms in the unit cell defined by A to the number in the unit cell defined by B. The elements of $(\mathrm{B}\,\mathbf{C}\,\mathrm{A})$ will all be rational; that is, they are small integers or fractions.

The reverse transformation is clearly specified by

$$[\mathrm{A};\ \mathbf{u}] = (\mathrm{A}\,\mathbf{C}\,\mathrm{B})\,(\mathrm{B};\ \mathbf{v}),$$

where $(\mathrm{A}\,\mathbf{C}\,\mathrm{B})$ is the reciprocal matrix $(\mathrm{B}\,\mathbf{C}\,\mathrm{A})^{-1}$.

The correspondence of directions also implies a one to one correspondence of lattice planes, since from eqns. (7.5) and (6.10)

$$(\mathbf{k};\ \mathrm{A}^*) = (\mathbf{h};\ \mathrm{A}^*)\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})^{-1},$$

and

$$(\mathbf{k};\ \mathrm{B}^*) = (\mathbf{h};\ \mathrm{A}^*)\,(\mathrm{A}\,\mathbf{S}\,\mathrm{A})^{-1}\,(\mathrm{A}\,\mathbf{J}\,\mathrm{B}) = (\mathbf{h};\ \mathrm{A}^*)\,(\mathrm{A}\,\mathbf{C}\,\mathrm{B}), \tag{9.3}$$

or briefly

$$\mathsf{k}' = \mathsf{h}'\,\mathsf{C}^{-1}.$$

† The determinant $|\mathbf{C}|$ is ± 1 depending on whether or not both bases are defined by equal-handed sets of vectors.

In simple shear, the lattice points move parallel to the K_1 plane through distances which are proportional to their separation from this plane. This is an example of plane strain in which all displacements are coplanar. A more general type of homogeneous plane strain is shown in Fig. 2.7. Once again all lattice points move in the same direction through distances which are proportional to their separation from a fixed plane, but the direction of movement is no longer parallel to this plane. There are again two undistorted planes; one is also unrotated, and corresponds to the K_1 plane in mechanical twinning. The whole

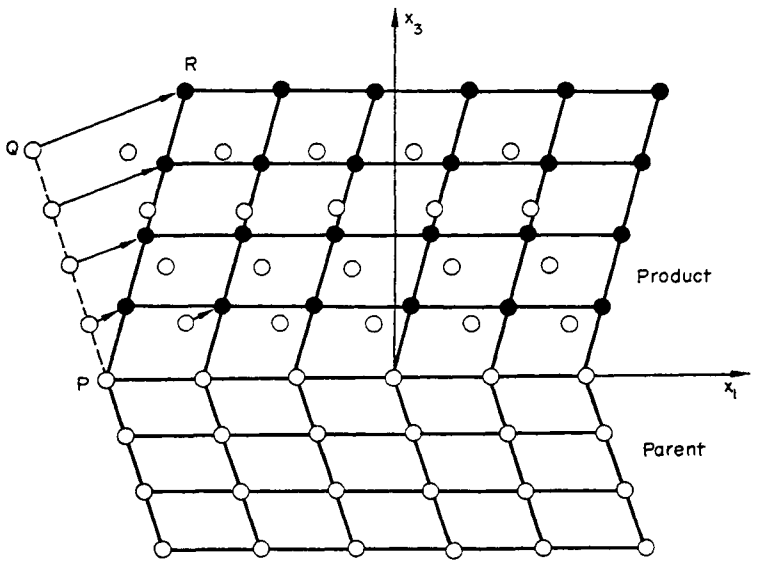

FIG. 2.7. Simple illustration of an invariant plane strain connecting two lattices. *PQ*, *PR* are corresponding directions in parent and product.

deformation may be considered as the relative displacement of a stack of such planes in a fixed direction.

The type of deformation shown in Fig. 2.7 is called an invariant plane strain. We can readily prove that the general condition for S to represent such a strain is that one principal strain be zero, and the other two have opposite signs. Thus in an orthonormal basis in which the deformation is $\mathbf{y} = \mathbf{S}\mathbf{x} = \mathbf{R}\mathbf{P}\mathbf{x}$, we refer the representations to the principal axes of strain (basis $\bar{\mathbf{i}}$) and obtain (see eqn. (7.22))

$$\bar{\mathbf{y}} = \mathbf{L}\mathbf{R}\mathbf{L}\bar{\mathbf{P}}\bar{\mathbf{x}},$$

where $\bar{\mathbf{P}} = (\lambda_i \delta_{ij})$ is a diagonal matrix. A vector is unchanged in length if

$$\bar{\mathbf{y}}'\bar{\mathbf{y}} = \bar{\mathbf{x}}'\bar{\mathbf{P}}'\bar{\mathbf{P}}\bar{\mathbf{x}} = \bar{\mathbf{x}}'\bar{\mathbf{x}},$$

since L and R are both orthogonal. In scalar form, this equation is

$$(\lambda_i^2 - 1)\bar{x}_i^2 = 0, \tag{9.4}$$

and gives the locus of the surface containing lines of unchanging length. If there are two such lines which are coplanar, they define an undistorted plane. By considering the intersection of (9.4) with the plane $\bar{x}_k = 0$, we obtain the equations

$$\frac{\lambda_i^2 - 1}{\lambda_j^2 - 1} = -\frac{\bar{x}_j^2}{\bar{x}_i^2}. \tag{9.5}$$

For a real solution, each $\lambda_i^2 - 1$ must differ in sign from each of the other two, because $\bar{x}_i^2$ are necessarily positive. This is only possible if one of the $\lambda_i^2 - 1$ is zero, and the other two have opposite signs. Note that in contrast to this, the only condition for a plane strain is that one λ_i should equal 1. The invariant plane strain is thus not the most general kind of plane strain.

Equation (9.5) now represents a straight line which is the intersection of the plane surface

$$(\lambda_1^2 - 1)\bar{x}_1^2 + (\lambda_3^2 - 1)\bar{x}_3^2 = 0, \quad \bar{x}_2 = \text{const}, \tag{9.6}$$

with the plane $\bar{x}_2 = 0$. This plane is a plane of zero distortion. During the pure deformation P, the undistorted plane will rotate through some angle, so an invariant plane strain is obtained by combining P with a rotation R which returns the undistorted plane to its original position. The total deformation S is thus determined, once P has been specified.

The tensor S has a particularly simple representation if we choose axes x_i in Fig. 2.7 to correspond to those previously used for simple shear in deformation twinning. The whole deformation is then equivalent to a simple shear on the x_1x_2 plane, combined with a uniaxial expansion or contradiction perpendicular to this plane. In this system the deformation is thus given by

$$\mathsf{S} = \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & 0 \\ 0 & 0 & 1+\Delta \end{pmatrix}, \tag{9.7}$$

where s and Δ specify the shear and the expansion, and $1+\Delta$ is the volume ratio of the transformation.

More generally, we now find the form of the tensor S in any basis A. A vector **u** in the undistorted plane must satisfy the relation

$$\mathsf{S}\,\mathsf{u} = \mathsf{u}. \tag{9.8}$$

The normal to the undistorted plane is ν where $\nu' \mathsf{u} = 0$, and after deformation the normal is given by

$$\nu' \mathsf{S}^{-1} = \{1/(1+\Delta)\}\nu', \tag{9.9}$$

where Δ is a scalar.

Quite generally, we may write $\mathsf{S} = \mathsf{I} + \mathsf{S}^\circ$, and since $\mathsf{S}^\circ \mathsf{u} = 0$ it follows that S° must have the form $\mathsf{e}\nu'$, where **e** is any constant vector. The above relations are thus satisfied if

$$\mathsf{S} = \mathsf{I} + \mathsf{e}\nu' \tag{9.10}$$

and this is the general representation of the invariant plane strain on the plane with normal ν. In suffix form

$$S_{ij} = \delta_{ij} + e_i \nu_j . \tag{9.11}$$

We also see from (9.9) that the scalar product $\nu' \mathsf{e} = \Delta$. By expansions of $|\mathsf{S}|$, the volume ratio of the new phase to the old is found to be $1+\Delta$; Δ is called the cubical dilatation, or simply the dilatation.

The reciprocal matrix to S is found similarly to be

$$\mathsf{S}^{-1} = \mathsf{I} - \frac{\mathsf{e}\,\nu'}{1+\Delta} . \tag{9.12}$$

Any point with coordinates given by the position vector **u** moves to a position **Su**, and its displacement is thus

$$\mathsf{S}\,\mathsf{u} - \mathsf{u} = \mathsf{e}\,\nu'\,\mathsf{u}, \tag{9.13}$$

which is always in the direction **e**. The amount of displacement is proportional to the perpendicular distance from the plane $\nu' \mathsf{u} = 0$.

Similarly, any normal **h** becomes a vector **k** where $\mathsf{k}' = \mathsf{h}'\,\mathsf{S}^{-1}$, and the displacement of the end point of this vector is thus

$$\mathsf{h}'\left(\mathsf{I} - \frac{\mathsf{e}\,\nu'}{1+\Delta}\right) - \mathsf{h}' = -\frac{\mathsf{h}'\,\mathsf{e}\,\nu'}{1+\Delta} . \tag{9.14}$$

This is always in the direction ν, and hence each plane normal rotates in a plane containing ν.

Crystals of a new phase frequently form inside an existing solid phase in the form of flat plates. In nucleation and growth reactions, this shape is usually adopted to minimize the elastic strain energy due to the volume change in the transformation, and the boundary surface need not have special significance in any representation of the orientation relations. In martensitic transformations, however, we have emphasized that some choice of S has physical significance in defining the atom movements. By analogy with the situation in mechanical twinning, we might expect that when the correct choice is made, S has the form (9.10), and the plane ν specifies the interfacial boundary or habit plane of the martensite crystals. However, although there are an infinite number of ways in which two lattices may be related by a tensor S, it is seldom possible to find a representation of type (9.10). This is because the lattice parameters of the two structures are determined mainly by short range interactions, and the condition (9.5) will be satisfied for two different structures only coincidentally. In fact, similar planes of almost identical atomic arrangement do occur in certain transformations between the closely related f.c.c. and h.c.p. structures and the interface plane is then rational in both lattices. Apart from these isolated examples, however, the interpretation of the martensite habit plane cannot be as simple as the interpretation of the K_1 plane in mechanical twinning.

As previously described, a general homogeneous deformation is characterized by at least one and possibly three unrotated planes. Jaswon and Wheeler (1948) pointed out that if the

martensite lattice is produced from the parent lattice by a homogeneous deformation, the associated disturbance of the untransformed parent phase in the region round a martensite crystal will be very large, unless the boundary between the phases is unrotated. They therefore suggested that the habit plane is an unrotated plane, and should be found by diagonalizing S in accordance with the procedure on p. 44. However, these unrotated planes are distorted, and the rotation of rows of atoms within them would equally give rise to an extremely large strain energy.

In considering martensitic transformation, we are thus faced with the difficulty that finite homogeneous deformation of a macroscopic region of the parent crystal is only feasible if the boundary between the deformed and undeformed regions at any stage at least approximates to an invariant plane of the deformation. This requirement can be reconciled with the impossibility of relating the lattices by a homogeneous deformation of this kind only by assuming that a martensite plate is not produced physically by a *homogeneous* deformation of the parent lattice. It is now generally accepted that the atom movements during transformation are such that the deformation of the parent lattice into the martensite lattice is homogeneous only over a localized region. An adjacent region of the plate is formed by a different physical deformation, which also, of course, generates the martensite structure. Each of these deformations may be factorized into two components, of which one is common to both regions, and the other is opposite in the two regions, in the sense that the combined effect of these components is to produce no net change in shape or volume. The whole transformation thus consists of a homogeneous deformation, together with deformations which are locally homogeneous, but which have zero macroscopic effect.

The geometry and crystallography of martensite transformations is considered in detail in Part II, Chapter 22; the subject is introduced here only in order to show the importance of analysing lattice deformations S into component deformations. From the above description, we see that the component of the lattice deformation common to all regions of a plate must approximate closely to an invariant plane strain. The component with zero macroscopic effect must be such that there is no dilatation (i.e. the determinant of its matrix representation must be unity), since any volume change has to be identical in all regions, and would thus accumulate to produce a macroscopic effect. It is generally assumed that this component is also a plane strain, that is, in this case, a simple shear. The description of this component depends on the order in which the separate deformations are considered to be applied, and this emphasizes the purely mathematical character of the factorization. The terms "first" and "second" strain, which are often used for convenience, have no physical significance, and this is true of all factorizations of this type.

Two invariant plane strains applied successively will not give a resultant deformation with an invariant plane unless the plane ν or the direction $\mathbf{e}$ (eqn. (9.10)) is common to both components. This may be seen by writing the components D, T in subscript form

$$D_{ik} = \delta_{ik} + (e_1)_i (\nu_1)_k , \qquad T_{kj} = \delta_{kj} + (e_2)_k (\nu_2)_j .$$

The product $S_{ij} = D_{ik}T_{kj}$ can only be written in the form $\delta_{ij} + (e_3)_i (\nu_3)_j$ if either $\mathbf{e}_1 = \mathbf{e}_2 = \mathbf{e}_3$ or $\nu_1 = \nu_2 = \nu_3$. There must, however, always be an invariant line in such a

transformation, since one line is common to the two invariant planes. Any deformation tensor which is to be factorized into two invariant plane strains must thus be an invariant line strain. For any deformation, there is at least one unrotated line, but this need not be unchanged in length, and an invariant line strain is not the most general form of homogeneous deformation.

The condition for two lattices to be related by an invariant line strain is simply that one principal strain either be zero, or have opposite sign from the other two. This follows immediately from eqn. (9.5) since we then have for any point on an unextended line

$$(\overline{x_1})^2 = -\frac{(\lambda_2^2-1)(\overline{x_2})^2-(\lambda_3^2-1)(\overline{x_3})^2}{(\lambda_1^2-1)}, \tag{9.15}$$

and if the two principal strains λ_2-1, λ_3-1 have the same sign, λ_1-1 must have the opposite sign. Provided this condition is satisfied, eqn. (9.15) represents a cone of directions of unchanging length. By addition of a suitable rotation, any line in the cone may be returned to its original position, so that it is an invariant line of the whole deformation. The rotation only affects the orientations of the two lattices, so that the condition above is sufficient to ensure that two lattices can be related by an invariant line strain.

In discussing invariant plane and line strains, we have imposed no restrictions on the nature of these planes or lines, since (in contrast to the situation in twinning) we make no general assumptions about the relations of the lattice symmetries or constants. It follows that the invariant planes may be irrational.

10. INFINITESIMAL DEFORMATIONS

The deformations considered previously in this chapter have been homogeneous and finite. In this section we briefly consider the relation of the results to the theory of linear elasticity, which deals with infinitesimal inhomogeneous displacements of the atoms. Let us first consider an affine transformation. Throughout this section, we shall use an orthogonal basis I in which the representation of a general position vector is x. Since we shall not be concerned with the crystal lattice, we use the symbol **x** rather than **u** for the vector itself, and the deformation

$$\mathbf{y} = \mathbf{S}\mathbf{x}$$

is represented by the matrix equation

$$\mathsf{y} = \mathsf{S}\,\mathsf{x},$$

in the basis I. As a result of the deformation, a point with coordinates x_i moves to a position with coordinates $y_i = S_{ij}x_j$. We define the displacement vector **w** of this point as the vector joining its initial and final positions. The displacement vector thus has components $w_i = y_i - x_i$, and is given by the equation

$$\left.\begin{aligned} &\mathbf{w} = \mathbf{S}\mathbf{x}-\mathbf{x}, \\ \text{i.e.}\quad &\mathbf{w} = (\mathsf{S}-\mathsf{I})\,\mathsf{x} \\ \text{or}\quad &w_i = (S_{ij}-\delta_{ij})x_j. \end{aligned}\right\} \tag{10.1}$$

Now consider a further small deformation with matrix representation $\mathbf{z} = \mathbf{T}\mathbf{y}$. The total deformation is thus

$$\mathbf{z} = \mathbf{T}\mathbf{S}\mathbf{x}$$

or in suffix form,

$$z_i = T_{ik}S_{kj}x_j = [(T_{ik}-\delta_{ik})(S_{kj}-\delta_{kj})+(S_{ij}+T_{ij})-\delta_{ij}]x_j. \tag{10.2}$$

Now suppose the deformations **S** and **T** are both infinitesimal, in the sense that the elements of the displacement vectors $(\mathbf{S}-\mathbf{I})\mathbf{x}$, $(\mathbf{T}-\mathbf{I})\mathbf{y}$ are so small that the product of any term in the matrix $(\mathbf{S}-\mathbf{I})$ with any other term, or with any term in the matrix $(\mathbf{T}-\mathbf{I})$, may be neglected. This means that the first term in the square brackets of eqn. (10.2) is zero. The total displacement is thus given by

$$w_i = z_i - x_i = (S_{ij}+T_{ij}-2\delta_{ij})x_j$$

or, in matrix form,

$$\mathbf{w} = (\mathbf{S}+\mathbf{T}-2\mathbf{I})\mathbf{x}. \tag{10.3}$$

The components of the displacement vector for two successive strains may thus be obtained by adding the components of the vectors for the separate strains. This result is obviously only valid for small displacements, and is known as the principle of superposition. When it is applicable, the resultant displacement is independent of the order of the strains. Conversely, we may factorize any infinitesimal deformation into two separate and independent components.

In principle, the infinitesimal affine deformation applied to any lattice will produce a new lattice. However, when the displacements are small enough for the principle of superposition to be applied, each lattice point in the deformed structure can be associated clearly and unambiguously with its original position. It is then more useful to regard the deformed structure as a slightly imperfect (strained) version of the original lattice, rather than as a new lattice. This point of view is, moreover, essential when later in this section we consider inhomogeneous displacements. In their deformed positions, the lattice points then no longer constitute a space lattice in our former mathematical use of the term. They may still be regarded as forming a slightly imperfect lattice of the original type.

Consider the equation

$$\mathbf{w} = (\mathbf{S}-\mathbf{I})\mathbf{x} = \mathbf{Q}\mathbf{x}, \tag{10.4}$$

where the components of **Q** are all infinitesimal. In general, **Q** will not be symmetric, since **S** need not represent a pure deformation. However, we may always write

$$\left.\begin{aligned} \mathbf{Q} &= \mathbf{e}+\boldsymbol{\omega}, \\ e_{ij} &= \tfrac{1}{2}(Q_{ij}+Q_{ji}), \\ \omega_{ij} &= \tfrac{1}{2}(Q_{ij}-Q_{ji}). \end{aligned}\right\} \tag{10.5}$$

where

The components e_{ij} form a symmetric matrix **e** which is a representation of the strain tensor, whilst the components ω_{ij} form an antisymmetric matrix $\boldsymbol{\omega}$. We shall now show that for infinitesimal deformations, an antisymmetric matrix is the representation of a rigid-body rotation.

The condition that the length of a vector should remain unchanged as the result of a deformation R was shown on p. 47 to be

$$\mathbf{x}'\,\mathsf{R}'\,\mathsf{R}\,\mathbf{x} = \mathbf{x}'\,\mathbf{x},$$

i.e.

$$R_{ki}R_{kj}x_ix_j = \delta_{ij}x_ix_j.$$

Expanding $R_{ki}R_{kj}$, this gives

$$\delta_{ij} = (R_{ki}-\delta_{ki})(R_{kj}-\delta_{kj})+R_{ji}+R_{ij}-\delta_{ij}. \tag{10.6}$$

Our previous assumption shows that the first term on the right may be neglected, so that

$$R_{ij}+R_{ji} = 2\delta_{ij}. \tag{10.7}$$

If the matrix R satisfies this condition, the displacement vector is given by

$$\mathbf{w} = (\mathsf{R}-\mathsf{I})\,\mathbf{x},$$

where the matrix R−I is antisymmetric. An antisymmetric displacement matrix thus represents an equal rotation of all vectors about the origin of coordinates, i.e. a rigid-body rotation.

From the Principle of Superposition, the general displacement vector may now be analysed into two components. The change $\mathbf{w}_1 = \boldsymbol{\omega}\,\mathbf{x}$ represents a rotation, whilst the change $\mathbf{w}_2 = \mathbf{e}\,\mathbf{x}$ represents a pure deformation. Since only three independent quantities are needed to specify the antisymmetric tensor $\boldsymbol{\omega}$, the elements ω_{ij} $(i \neq j)$ can also be regarded as the components of an axial vector

$$\boldsymbol{\omega} = [\mathrm{I};\boldsymbol{\omega}] = [\omega_{32}\omega_{13}\omega_{21}]. \tag{10.8}$$

The linear transformation $\mathbf{w}_1 = \boldsymbol{\omega}\mathbf{x}$, where $\boldsymbol{\omega}$ is the representation of the antisymmetric tensor, may also be written

$$\mathbf{w}_1 = \boldsymbol{\omega}\wedge\mathbf{x}, \tag{10.9}$$

where $\boldsymbol{\omega}$ is the vector of (10.8), as may readily be seen by comparing the coefficients of $\mathbf{w}_1$. It is unimportant whether we use the tensor or vector methods of representing the small rotation; note that the components of the vector give the component rotations about the three axes. For the infinitesimal rotations considered in linear elasticity theory, the component rotations may be considered as vectors, and added to give the net rotation.

The geometrical interpretation of the components of the strain tensor is readily obtained. Neglecting products of these components, we find that the diagonal elements represent the extensions (i.e. changes in length per unit length) of lines originally parallel to the coordinate axes. The components e_{ij} $(i \neq j)$ give half the cosines of the angles between vectors originally parallel to x_i and x_j; since e_{ij} is small, $2e_{ij}$ thus gives the relative rotation (the change in the mutual orientation) of such vectors. The quantities $2e_{ij}$ are commonly called the shear strains.

During the deformation, a vector of original length $|\mathbf{x}|$ changes into a vector of length $|\mathbf{y}| = |\mathbf{x}|+\delta|\mathbf{x}|$. Since $\delta|\mathbf{x}|$ is small, $|\mathbf{y}|^2-|\mathbf{x}|^2 = 2|\mathbf{x}|\,\delta|\mathbf{x}|$, and

$$2|\mathbf{x}|\,\delta|\mathbf{x}| = \mathbf{x}'\,\mathsf{S}'\,\mathsf{S}\,\mathbf{x}-\mathbf{x}'\,\mathbf{x} = (S_{ij}+S_{ji}-2\delta_{ij})x_ix_j,$$

where cross products of the terms of S have again been neglected. Since $S_{ij}+S_{ji} = 2(e_{ij}+\delta_{ij})$, we have

$$|\mathbf{x}|\,\delta|\mathbf{x}| = e_{ij}x_ix_j.$$

Consider the surface

$$e_{ij}x_ix_j = \text{const.} \tag{10.10}$$

By comparison with the previous equation, we see that

$$\frac{\delta|\mathbf{x}|}{|\mathbf{x}|} = \frac{\text{const}}{|\mathbf{x}|^2},$$

where the coordinates of the end point of $\mathbf{x}$ satisfy eqn. (10.10). The surface (10.10) thus has the property that the extension of any vector is proportional to the inverse square of the radius vector to the surface in the corresponding direction. The direction of the displacement is normal to the tangent plane to the surface at the point $\mathbf{x}$. We have already met this equation in the theory of finite deformations, and we noted there that $\delta|\mathbf{x}|$ should strictly be replaced by the resolved elongation of $\mathbf{x}$. When the displacements are infinitesimal, the elongation quadric is known as the strain quadric, and its axes are the principal axes of the strain.

So far, we have assumed that the deformations, though infinitesimal, are homogeneous. The theory may be extended to inhomogeneous infinitesimal deformations in the following way. A point with coordinates x_i will move during deformation to a new position y_i, and the displacement may again be specified by

$$\mathbf{w} = \mathbf{y}-\mathbf{x}.$$

Consider a neighbouring point with coordinates $x_i+\zeta_i$ (Fig. 2.8), so that before deformation the two points were related by a vector $\zeta = [1; \zeta]$. This point will move to a position

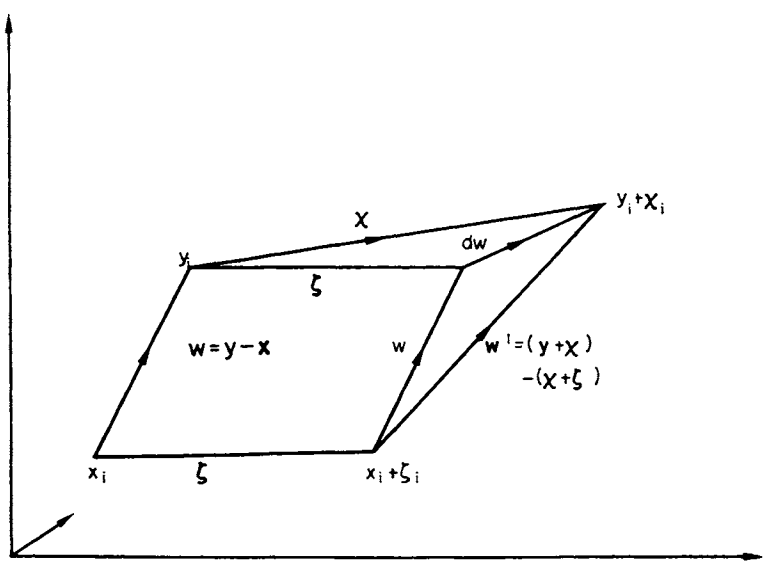

FIG. 2.8. Displacements in infinitesimal deformation.

with coordinates $y_i+\chi_i$, so that its displacement vector is $\mathbf{w}' = (\mathbf{y}+\boldsymbol{\chi})-(\mathbf{x}+\boldsymbol{\zeta})$. We may write the components of this displacement as a Taylor series

$$(y_i+\chi_i)-(x_i+\zeta_i) = (y_i-x_i)+\frac{\partial(y_i-x_i)}{\partial x_j}\zeta_j+0(\zeta_j^2), \quad \text{etc.}$$

The differentials in this expression are to be evaluated at the point $\mathbf{x}$. If the components of $\boldsymbol{\zeta}$ are small enough for the squares to be neglected, this gives

$$\chi_i-\zeta_i = \frac{\partial w_i}{\partial x_j}\zeta_j,$$

and the change in the separation of the two points is

$$\mathrm{d}\mathbf{w} = \mathbf{w}'-\mathbf{w} = \boldsymbol{\chi}-\boldsymbol{\zeta} = \mathbf{Q}\boldsymbol{\zeta}, \tag{10.11}$$

where the matrix $\mathbf{Q}$ has components $\partial w_i/\partial x_j$. Comparison of eqn. (10.11) with (10.4) shows that they are of the same form; the vector $\mathbf{dw}$ represents the displacement of points near $\mathbf{x}$ relative to an origin which moves with $\mathbf{x}$. The transformation is affine in the small region round $\mathbf{x}$, and $\mathbf{Q}$ may again be separated into a strain tensor $\mathbf{e}$ and a rotation tensor $\boldsymbol{\omega}$, where now

$$\left.\begin{aligned} e_{ij} &= \tfrac{1}{2}(\partial w_i/\partial x_j+\partial w_j/\partial x_i), \\ \omega_{ij} &= \tfrac{1}{2}(\partial w_i/\partial x_j-\partial w_j/\partial x_i). \end{aligned}\right\} \tag{10.12}$$

In the region around $\mathbf{x}$, $\mathbf{e}$ and $\boldsymbol{\omega}$ may be interpreted in the same way as we previously interpreted them for the whole crystal. An inhomogeneous deformation may thus still be considered affine in a small region, and is characterized by the pure strain and rotation of each region. In contrast to the previous results, the components of $\mathbf{e}$ and $\boldsymbol{\omega}$ are not constants, but are functions of the coordinates x_i.

If we use the vector representation of the rotation $\boldsymbol{\omega}$, we have the components of this vector as

$$\omega_i = \tfrac{1}{2}(\partial w_k/\partial x_j-\partial w_j/\partial x_k) \qquad (i \neq j \neq k)$$

or

$$2\boldsymbol{\omega} = \operatorname{curl} \mathbf{w}. \tag{10.13}$$

The rotation is then represented by a vector which is a function of the coordinates x_i, but is independent of the choice of coordinate system.

Utilizing the result on p. 48, it follows from (10.11) that the ratio of the new volume of a small region to its old volume is given by

$$\frac{v+\Delta v}{v} = |1+\mathbf{Q}| = 1+\frac{\partial w_i}{\partial x_i},$$

since all the remaining terms vanish for an infinitesimal deformation. The quantity

$$\operatorname{div} \mathbf{w} = \frac{\partial w_i}{\partial x_i} = e_{ii} = \frac{\Delta v}{v} = \Delta, \tag{10.14}$$

is called the elastic dilatation, and may be given the same symbol Δ, already used for finite homogeneous deformations. For infinitesimal deformations, Δ is clearly a scalar property of the vector **w**, and is independent of the coordinate system. If **e** is *any* representation of the strain tensor **e**, then

$$\Delta = \text{trace}\,(\mathbf{e}) = e_{ii}. \tag{10.15}$$

The components of **e** are not able to vary in an arbitrary manner, since they determine the displacements **w** of the points of the material. Provided the volume considered is singly connected (i.e. any closed curve can be shrunk continuously down to a point without crossing the boundaries), **w** must be everywhere single-valued and continuous. This restriction results in six equations which must be satisfied by the second differential coefficients of the tensor components e_{ij}; they are known as the equations of compatibility. We shall not derive them here, but they may all be expressed in the shortened form

$$\frac{\partial^2 e_{ij}}{\partial x_k \partial x_l} + \frac{\partial^2 e_{kl}}{\partial x_i \partial x_j} - \frac{\partial^2 e_{ik}}{\partial x_j \partial x_l} - \frac{\partial^2 e_{jl}}{\partial x_i \partial x_k} = 0. \tag{10.16}$$

11. STRESS–STRAIN RELATIONS: THEORY OF ELASTICITY

The equations of elastic equilibrium do not, of course, form part of the geometry of crystal lattices, but it is convenient to consider them here, since the theory of linear elasticity is so dependent on the notion of infinitesimal strain introduced in the last section. We first clarify the concept of stress. Any volume element v of a body is subject in general to forces of two kinds. There are first forces which act on all the particles of the volume element, and these are called body forces. They are caused by the presence of some external field of force, for example the gravitational field. In general, this force will not be uniform, but its effect on the whole of the volume element may be summed to give a resultant body force $\int_v \mathbf{g}\,dv$, where **g** is the body force vector, and a resultant moment about any origin of $\int_v (\mathbf{g} \wedge \mathbf{x})\,dv$, where **x** is the position vector with respect to the origin.

The second type of force acting on the volume v arises from internal forces between particles of the material. Thus, consider a small planar element δO of the (interior) surface O separating v from the rest of the body. In general, the two parts of the body on either side of δO will be in a strained condition, and will exert equal and opposite forces on each other. These forces are used to specify the state of stress at a point within the element δO. Let the outward normal to δO be **n**, a unit vector. In this section, we use **n** rather than **h** as the symbol for a vector normal to a plane, since we are interested in both rational and irrational planes. For the most part, we are not concerned with the crystallographic nature of the structure, so we take **n** as a unit normal. The forces exerted by the material on one side of the element δO across δO will then reduce to a single force acting at a point within the element, and a moment about some axis. As $\delta O \to 0$, the direction of the resultant force approaches a fixed value, though the magnitude of both force and couple tend, of course, to zero. However, the force divided by the area remains finite, and approaches a

limit as $\delta O \to 0$, whilst the moment divided by δO tends to zero. The limiting vector, having the dimensions force per unit area, is called the stress vector $\mathbf{p}$, or simply the stress acting across δO. In a fluid, the only surface force is a uniform hydrostatic pressure, and $\mathbf{p}$ is necessarily along the direction $-\mathbf{n}$. In a solid, however, $\mathbf{p}$ may make any angle with the direction $\mathbf{n}$. The quantity $\mathbf{p}\,\delta O$ when δO is small, is called the traction across δO. Forces of this second kind are called surface forces.

The resultant of the surface forces on the volume v is given by the surface integral $\int_{\delta} \mathbf{p}\,\mathrm{d}O$. This follows since the internal surface forces across all elements δv within v cancel out. In the same way, there will be a resultant moment of $\int_{\delta} (\mathbf{p} \wedge \mathbf{x})\,\mathrm{d}O$ on the origin, as a result of the surface forces.

The complete specification of the surface forces at a point requires a knowledge of the tractions across all planar elements at the point. We shall now show that this specification may be obtained by a set of quantities X_{ij} which form a symmetric matrix, and which are a representation of a second order tensor called the stress tensor. Consider a rectangular parallelepiped with faces perpendicular to the coordinate axes. The surface forces exerted by the material outside the parallelepiped on the material inside it are specified by a stress vector $\mathbf{p}$ at each point of the surface. At any point on the surface normal to the axis x_i which has a positive normal (i.e. outward normal along $+x_i$), we write the stress vector as

$$\mathbf{p}_i = X_{ij}\mathbf{i}_j, \tag{11.1}$$

where $\mathbf{i}_i$ define the orthonormal system of coordinates x_i. The quantity $X_{(i)(i)}$ thus gives the stress acting normal to the face at the point considered, and the other two components are shear stresses.

Note that if $\mathbf{p}$ is directed outwards from the face, $X_{(i)(i)}$ is positive, and hence the sign convention is such that tensile stresses exerted by the surrounding material on the volume element are positive. Equally, if we had taken a point on the surface normal to $\mathbf{i}_i$ and having negative surface normal, we should have written

$$\mathbf{p}_i = -X_{ij}\mathbf{i}_j,$$

so that tensile forces would again be specified by positive $X_{(i)(i)}$.

Now consider a small tetrahedron at the point P (Fig. 2.9), formed from three planes normal to the coordinate axes, and a fourth plane, ABC, which has unit normal $\mathbf{n}$ and area δO. The areas of the other three faces of the tetrahedron are $\delta O_i = n_i\,\delta O$, where n_i are the components of $\mathbf{n}$. The outward normals to all these faces are negative, if the components of $\mathbf{n}$ are all positive. In equilibrium, there will be no resultant force on the whole tetrahedron, so that

$$\int_v \mathbf{g}\,\mathrm{d}v + \int_O \mathbf{p}\,\mathrm{d}O = 0. \tag{11.2}$$

Suppose now that the volume of the tetrahedron is allowed to shrink continuously to zero. The forces proportional to the volume then vanish more rapidly than those proportional to the surface, and in the limit we need consider only the latter. The traction across

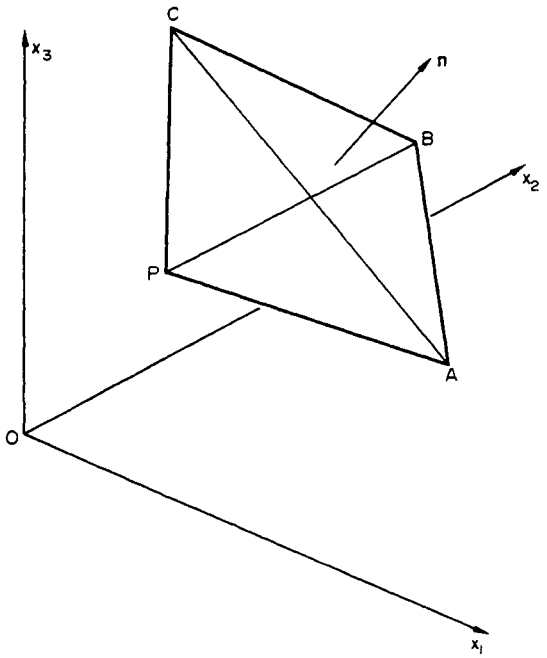

FIG. 2.9. To illustrate the meaning of the stress components at *P*.

each plane surface of the tetrahedron approaches the limit of the product of its area and a constant stress. Thus across the surfaces normal to the axes, the sum of the forces acting is

$$\mathbf{p}_i\,\delta O_i = -n_i X_{ij}\mathbf{i}_j\,\delta O.$$

If we suppose that the stress vector across *ABC* has the value **p** in the limit, the equilibrium condition is

$$\mathbf{p}\,\delta O + \mathbf{p}_i\,\delta O_i = 0$$

or

$$p_j = X_{ij}n_i. \tag{11.3}$$

In matrix form, we may write this equation as

$$(\mathbf{p};\,\mathrm{I}) = (\mathbf{n};\,\mathrm{I})\,(\mathrm{I}\,\mathbf{X}\,\mathrm{I})$$

or

$$\mathsf{p}' = \mathsf{n}'\,\mathsf{X}. \tag{11.4}$$

The stress vector across any plane at a point *P* is thus specified by the positive normal to the plane, and the array of quantities which represent the stress tensor at this point.

We now proceed to calculate the restrictive conditions which the components of **X** must satisfy if the body is to be in internal equilibrium. The condition of equilibrium within any volume v in the material is expressed by eqn. (11.2). If we take the component of force in the direction $\mathbf{i}_i$,

$$\int_v g_i\,\mathrm{d}v + \int_O X_{ji}n_j\,\mathrm{d}O = 0.$$

The second part of this expression may be transformed into a volume integral by means of Gauss's theorem to give

$$\int_v (g_i + \partial X_{ji}/\partial x_j)\, \mathrm{d}v = 0. \tag{11.5}$$

Since the region v is arbitrary, the expression in the integral must vanish identically, and we therefore have a set of equations

$$\partial X_{ji}/\partial x_j = -g_i. \tag{11.6}$$

which are known as the equations of equilibrium.

In many important physical problems the particles of the body are not held in static equilibrium but are in states of motion. Returning to Fig. 2.9 we see that if there is a resultant force on the tetrahedron, its centre of mass will have an acceleration in the direction of this resultant. Taking the component of acceleration in the direction $\mathbf{i}_i$, eqn. (11.5) is replaced by

$$\int_v (g_i + \partial X_{ji}/\partial x_j)\, \mathrm{d}v = \varrho v\, \partial^2 x_i/\partial t^2,$$

where ϱ is the density of the material at the point P. This gives a set of equations

$$\frac{\partial X_{ji}}{\partial x_j} + g_i = \varrho \frac{\partial^2 x_i}{\partial t^2}, \tag{11.7}$$

which are known as the equations of motion.

When the body is in equilibrium, the resultant moment of the forces acting on v must vanish at any point. Choose an arbitrary origin, and let $\mathbf{x}$ be the position vector of some point within the volume v. The moment about the origin is zero if

$$\int_v (\mathbf{g} \wedge \mathbf{x})\, \mathrm{d}v + \int_O (\mathbf{p} \wedge \mathbf{x})\, \mathrm{d}O = 0, \tag{11.8}$$

and taking the component in the direction $\mathbf{i}_1$,

$$\int_v (x_3 g_2 - x_2 g_3)\, \mathrm{d}v + \int_O (x_3 X_{i2} n_i - x_2 X_{i3} n_i)\, \mathrm{d}O = 0. \tag{11.9}$$

By using Gauss's theorem and eqn. (11.6), the second part of this expression may be written as the volume integral:

$$\int_v [-(g_2 x_3 - g_3 x_2) + X_{32} - X_{23}]\, \mathrm{d}v,$$

and substituting this into (11.9), finally,

$$X_{23} = X_{32}. \tag{11.10}$$

By similarly writing down the other resolved components of (11.8), we can show that the quantities X_{ij} form a symmetric matrix with $X_{ij} = X_{ji}$. Equation (11.4) can thus be written in the equivalent form

$$\mathsf{p} = \mathsf{X}\,\mathsf{n}. \tag{11.11}$$

Using a suitable orthogonal transformation, $\mathbf{X}$ may be referred to a new set of axes in which it has diagonal form. The values of the diagonal elements, i.e. the eigenvalues of $\mathbf{X}$, are called the principal stresses.

In the same way as we defined the strain quadric, we may define a surface by the relation

$$X_{ij}x_ix_j = \text{const}, \tag{11.12}$$

and this is called the stress quadric. It has the property that if a vector is drawn from the origin to a point **x** of the stress quadric in the direction of the normal **n** to any plane, the stress vector acting across that plane is in the direction of the normal to the tangent surface at x, and the normal component of the stress vector is proportional to the inverse square of the radius vector **x**.

For any body in equilibrium under the action of external surface forces and body forces, the six components of the stress tensor must satisfy the three eqns. (11.6) at all points in the interior. In addition, eqn. (11.11) must be satisfied over the external surface, **p** now being the externally applied stresses. This gives three boundary conditions. These equations are insufficient to determine the state of stress; the information which is still required is the connection between the state of stress and the state of strain. This relation is obtained in the generalized form of Hooke's law, which states that the stress components are linear functions of the strain components, and vice versa. For our purpose, this is best regarded as an empirical law, based on experiment; it is approximately valid for small deformations.

Since the stress components X and the strain components e are both representations of second rank tensors, the general linear relation is of the form

$$X_{ij} = c_{ijkl}e_{kl}, \qquad e_{ij} = s_{ijkl}X_{kl}, \tag{11.13}$$

where the quantities c_{ijkl}, s_{ijkl} are representations of fourth rank tensors. Fortunately, however, the 81 components of the representations c and s are readily reduced to more manageable numbers. In the first place, both X and e are symmetric matrices with only 6 independent components. This reduces the number of independent quantities in c and s to 36. It is usual to adopt a simplified notation, due to Voigt, and write the independent components of X and e as

$$\mathrm{X} = \begin{pmatrix} X_1 & X_4 & X_5 \\ X_4 & X_2 & X_6 \\ X_5 & X_6 & X_3 \end{pmatrix} \quad \text{and} \quad \mathrm{e} = \begin{pmatrix} e_1 & \frac{1}{2}e_4 & \frac{1}{2}e_5 \\ \frac{1}{2}e_4 & e_2 & \frac{1}{2}e_6 \\ \frac{1}{2}e_5 & \frac{1}{2}e_6 & e_3 \end{pmatrix}. \tag{11.14}$$

The factors of one half are introduced into the off-diagonal elements of the strain tensor to conform to long-established standard notation. Before the formal theory of tensors was developed, the shear strain components e_4, e_5 and e_6 were defined so as to give directly the changes in mutual inclination of vectors along the coordinate axes. These quantities are commonly called the shear strains, or simply the shears; to avoid confusion, it is better to designate them the engineering shear strains, and to distinguish them from the tensor shear strains $e_{ij}\,(i \neq j)$. Note that the array of engineering strains

$$\begin{pmatrix} e_1 & e_4 & e_5 \\ e_4 & e_2 & e_6 \\ e_5 & e_6 & e_3 \end{pmatrix}$$

does not constitute a tensor, and it is thus rather unfortunate that the older terminology is so well established. Although it is now common practice to use the tensor representation, the older component strains have been retained in defining the relations between stress and strain, which may now be written

$$\left.\begin{aligned} X_i &= c_{ij}e_j \\ e_i &= s_{ij}X_j \end{aligned} \qquad (i,j = 1, 2, \ldots, 6),\right\} \tag{11.15}$$

where, in contrast to previous equations, the range of i and j is from 1 to 6. The quantities c_{ij} and s_{ij} have been variously called elastic constants, moduli, or coefficients, with little agreement amongst different authors about which names are appropriate to the two sets. All these names are avoided in the modern (and descriptive) American terminology, in which the quantities c_{ij} are called stiffness constants, and the quantities s_{ij} are called compliances. In general, it is easy to examine a body under a uniaxial stress, but it is not possible to produce a uniaxial strain. The compliances may thus be determined directly by experiment; the stiffness constants only indirectly. With the above notation, the tensors **c** and **s** may be represented by 6×6 square matrices c and s.

Now suppose the deformation at any point is changed by an infinitesimal amount, so that the strains vary from e_i to $e_i+\mathrm{d}e_i$. During this change, work in done by the external forces, and the potential energy in the deformed region increases. The work per unit volume may be obtained by summing the products of each stress component and the corresponding change in strain, as may be seen by considering first a unit cube of material with only one stress component, X_i, acting. If the state of strain changes by an infinitesimal amount, the only work done by X_i is $X_i\,\mathrm{d}e_i$, where e_i is the corresponding strain component. The total work done on the unit cube in increasing the stress from 0 to X_i whilst the corresponding strain component increases from 0 to e_i is thus $\frac{1}{2}X_{(i)}e_{(i)}$. The principle of superposition now allows us to treat the work done by a system of forces acting on a unit cube as the sum of the amounts attributable to each stress component acting separately, so that for an incremental strain, the increase in strain energy per unit volume may be written

$$\mathrm{d}W_s/v = X_i\,\mathrm{d}e_i \qquad (i = 1, \ldots, 6). \tag{11.16}$$

The strain energy W_s must be a single valued function of the strains, so that

$$\mathrm{d}W_s = \frac{\partial W_s}{\partial e_i}\,\mathrm{d}e_i$$

and the condition for $\mathrm{d}W_s$ to be a perfect differential is

$$\frac{\partial^2 W_s}{\partial e_i\,\partial e_j} = \frac{\partial^2 W_s}{\partial e_j\,\partial e_i},$$

i.e. from (11.16)

$$\partial X_i/\partial e_j = \partial X_j/\partial e_i.$$

From eqn. (11.15), we see that this implies

$$c_{ij} = c_{ji} \tag{11.17}$$

and the stiffness constants thus form a symmetric matrix. Similarly, the compliances s_{ij} form a symmetric matrix s. The form of the strain energy function may be written

$$W_s/v = \tfrac{1}{2}c_{ij}e_ie_j + \text{const.} \tag{11.18}$$

The constant is zero if the zero energy is the undeformed state. If W_s is known, the components of the stress tensor may be obtained from the relation

$$X_i = \frac{1}{v}\frac{\partial W_s}{\partial e_i} = \frac{1}{2}(c_{ij}+c_{ji})e_j = c_{ij}e_j. \tag{11.19}$$

Leaving out the constant, the quantity W_s/v gives the strain energy per unit volume stored in the material. We note in passing that we have not specified the conditions for the change dW_s. The deformation may be carried out either adiabatically or isothermally; the function W_s will be different in the two cases, and the corresponding stiffness and compliance constants are also different.

The symmetry of the components of c and s reduces the number of independent quantities in each to 21, and this number of parameters is needed to specify the elastic properties of crystals of the lowest symmetry. Further reductions in the number of independent terms are due either to the symmetry properties of the crystal structure, or to an assumed law of force. A possible assumption is that the forces between atoms are all central, i.e. they act along the lines joining two atoms, and are a function only of the separation of these atoms. If all forces are of this kind, and in addition the crystal structure is such that each atom is at a centre of symmetry, the following equations may be derived:

$$\left.\begin{aligned} &c_{44} = c_{23}, \quad c_{55} = c_{31}, \quad c_{66} = c_{12}, \\ &c_{56} = c_{14}, \quad c_{64} = c_{25}, \quad c_{45} = c_{36}. \end{aligned}\right\} \tag{11.20}$$

For cubic crystals, as shown below, the equations all reduce to $c_{44} = c_{12}$. The equations are known as the Cauchy relations; they are not valid for metals.

The existence of symmetry elements in a crystal structure leads to a reduction in the number of independent elastic stiffness constants. Thus an n-fold axis of symmetry means that the crystal is brought into self-coincidence by a rotation of $2\pi/n$ about the axis, and such a rotation must therefore leave the elastic properties unchanged. If, therefore, we refer X and e to the new axes obtained by such a rotation, we obtain relations between the c_{ij} (or the s_{ij}). We shall not work through the examples, but merely quote the results.

If the x_3 axis is a twofold axis of symmetry (the x_1x_2 plane is a plane of symmetry), then

$$c_{14} = c_{15} = c_{24} = c_{25} = c_{34} = c_{35} = c_{64} = c_{65} = 0. \tag{11.21}$$

This symmetry element thus reduces the number of independent stiffness constants to 13. If the x_3 axis is a threefold axis of symmetry

$$\left.\begin{aligned} &c_{16} = c_{26} = c_{34} = c_{35} = c_{36} = c_{45} = 0, \\ &c_{11} = c_{22}, \quad c_{13} = c_{23}, \quad c_{14} = -c_{24} = c_{56}, \\ &c_{15} = -c_{25} = -c_{46}, \quad c_{44} = c_{55}, \quad c_{66} = \tfrac{1}{2}(c_{11}-c_{12}), \end{aligned}\right\} \tag{11.22}$$

and there are seven independent constants. For a fourfold axis along x_3, the following relations are additional to those for a twofold axis:

$$c_{36} = c_{45} = 0, \quad c_{11} = c_{22}, \quad c_{13} = c_{23}, \quad c_{16} = -c_{26}, \quad c_{44} = c_{55}, \tag{11.23}$$

which again gives seven independent constants. A sixfold axis about x_3 is obtained by combining the two- and threefold axis relations, and leaves five stiffness constants.

With the aid of these equations, the form of the matrix c for any symmetry can be determined. Thus a cubic crystal has fourfold axes x_1, x_2 and x_3. If the coordinate axes coincide with the edges of the cubic unit cell, the matrix has the form

$$\mathsf{c} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{44} \end{pmatrix} \tag{11.24}$$

and there are thus only three independent components. The behaviour of cubic crystals is most readily considered in relation to the combinations c_{44}, $(c_{11}-c_{12})/2$ and $(c_{11}+2c_{12})/3$. These measure respectively the resistance to deformation when a shearing stress is applied across a {100} plane in a ⟨010⟩ direction, across a {110} plane in a ⟨110⟩ direction, and when a uniform hydrostatic pressure is applied to the crystal. The cubic elastic properties are thus represented by two shear moduli and a bulk modulus. The two shear moduli become equal in an elastically isotropic material, and the ratio $2c_{44}/(c_{11}-c_{12})$ may thus be used as an elastic anisotropy factor for cubic crystals. Values of this factor are shown in Table IV, and provide an estimate of the validity of calculations using isotropic elastic theory for the metals concerned.

In an elastically isotropic body, the stiffness constants and compliances must be invariant with respect to any rotation of the axes. This leads to the further relation $c_{44} = (c_{11}-c_{12})/2$, and there are only two independent quantities to be specified. These are usually given the symbols

$$\mu = c_{44}, \quad \lambda = c_{12}.$$

μ is called the shear modulus; λ has no special name. If we now write out the components of stress for the isotropic medium, we find they can be expressed in the form

$$X_i = c_{ij}e_j = \lambda(e_1+e_2+e_3)+2\mu e_i = \lambda\Delta+2\mu e_j \qquad (i = 1, 2, 3)$$

and

$$X_i = c_{(i)(i)}e_i = \mu e_i \qquad (i = 4, 5, 6). \tag{11.25}$$

Using eqn. (11.14) to transform back to the components of the stress and strain tensors,

$$X_{ij} = \lambda\delta_{ij}\Delta+2\mu e_{ij}, \tag{11.26}$$

and, in particular,

$$X_{ii} = X_{11}+X_{22}+X_{33} = (3\lambda+2\mu)\Delta.$$

TABLE IV. VALUES OF ANISOTROPY FACTOR, $2c_{44}/(c_{11}-c_{12})$ FOR CUBIC METALS

(Elastic stiffnessses taken from Huntington, 1958)

Metal	Structure	Anisotropy factor
Lithium	b.c.c.	9·39
Sodium	b.c.c.	8·14
Potassium	b.c.c.	6·34
Copper	f.c.c.	3·20
Silver	f.c.c.	3·01
Gold	f.c.c.	2·90
Aluminium	f.c.c.	1·29
Silicon	Diamond cubic	1·56
Germanium	Diamond cubic	1·66
Lead	f.c.c.	3·89
Nickel	f.c.c.	2·51
Thorium	f.c.c.	3·62
Molybdenum	b.c.c.	0·77
Tungsten	b.c.c.	1·00
β-Brass	b.c.c.	8·40

The components of the strain tensor, e_{ij}, may similarly be obtained in terms of the stress components, using (11.13) which gives

$$e_{ij} = -\frac{\lambda}{2\mu(3\lambda+2\mu)}\delta_{ij}X_{kk}+\frac{1}{2\mu}X_{ij}. \tag{11.27}$$

If e is a diagonal matrix, X is also diagonal, so that the principal axes of stress and strain must coincide for an isotropic body.

Real crystals are not elastically isotropic, although some cubic crystals have small anisotropy factors. However, very many important problems are concerned not with the properties of isolated crystals, but with polycrystalline aggregates. If the distribution of orientations is effectively random, and the grain size small enough, these assemblies behave as isotropic bodies. Experimental results on elastic properties are then usually expressed in the form of various moduli of elasticity. If the body is deformed by a uniaxial stress, then all the X_{ij} except X_{11} are zero. The corresponding strains are

$$e_{11} = \frac{\lambda+\mu}{\mu(3\lambda+2\mu)}X_{11}, \quad e_{22} = e_{33} = -\frac{\lambda}{2\mu(3\lambda+2\mu)}X_{11}, \quad e_{ij} = 0 \qquad (i \neq j).$$

The ratio $Y = X_{11}/e_{11} = \mu(3\lambda+2\mu)/(\lambda+\mu)$ is called Young's modulus of elasticity. The ratio of the lateral contraction to the longitudinal expansion in such an experiment, $\nu = -e_{22}/e_{11} = \lambda/2(\lambda+\mu)$ is called Poisson's ratio.

A body subject to pure shear has $X_{23} = X_{32}$, and all the other components of stress are zero. This gives

$$e_{23} = e_{32} = X_{23}/2\mu.$$

The quantity μ thus represents the ratio of the shearing stress to the change in angle between the x_2 and x_3 axes; this latter quantity is the (engineering) shear strain, and μ is the shear

modulus. It is also readily seen that when a uniform hydrostatic pressure is applied, so that $X_{11} = X_{22} = X_{33}$, the ratio of the compressive stress to the cubical compression, $-\Delta$, is given by $K = \lambda + 2\mu/3$, where K is called the bulk modulus.

The theory of elasticity has to be applied to two principal classes of problem. These are the calculation of the distribution of stress and the displacements in the interior of an elastic body when the body forces are known, and either (a) the displacements over the surface of the body or (b) the forces acting on the surface, are specified. The equations to be solved are the stress equations of equilibrium (in more general problems, the equations of motion), together with the auxiliary conditions represented by the stress–strain relations, and where required (in the second type of problem), the compatibility conditions. For isotropic media, we may substitute eqns. (11.26) into the equations of equilibrium, to give

$$(\lambda+\mu)\frac{\partial \Delta}{\partial x_i} + \mu \nabla^2 w_i = -g_i \tag{11.28}$$

or, using (10.14),

$$(\lambda+\mu)\ \text{grad div}\ \mathbf{w} + \mu \nabla^2 \mathbf{w} = -\mathbf{g}. \tag{11.29}$$

Equations (11.28) and (11.29) are very useful in many elementary problems of the first type. Sometimes, however, it is convenient to use the identity $\nabla^2 = \text{grad div} - \text{curl curl}$ to transform (11.29) into

$$\left.\begin{aligned} (\lambda+2\mu)\ \text{grad div}\ \mathbf{w} - \mu\ \text{curl curl}\ \mathbf{w} &= -\mathbf{g} \\ \text{or} \qquad (\lambda+2\mu)\ \text{grad}\ \Delta - \mu\ \text{curl}\ \boldsymbol{\omega} &= -\mathbf{g}. \end{aligned}\right\} \tag{11.30}$$

The first problem is completely solved if solutions of one of the equivalent forms (11.28)–(11.30) are obtained, subject to the boundary conditions in which $\mathbf{w}$ is specified over the limiting surface of the solid considered.

To solve the second type of problem, the differential equations have to be expressed in terms of stresses rather than displacements. Since not every solution of the stress equations of equilibrium represents a possible state of strain, it is necessary to incorporate the compatibility equations. After some manipulation, eqns. (10.16) and (11.27) give the following set of six differential equations:

$$\nabla^2 X_{ii} + \frac{1}{1+\nu}\frac{\partial^2}{\partial x_i \partial x_j}(X_{ii}) = -\frac{\nu}{1+\nu}\delta_{ij}\ \text{div}\ \mathbf{g} - \left(\frac{\partial g_i}{\partial x_j} + \frac{\partial g_j}{\partial x_i}\right). \tag{11.31}$$

These are known as the Beltrami–Michell compatibility equations. The second main class of elastic problem then involves the solution of these equations, subject to the boundary conditions in which the forces acting over the surface of the body are specified. In this book, we shall not be concerned with problems of this type, nor with the mixed class when the boundary conditions are specified partially as forces and partially as displacements.

REFERENCES

BOWLES, J. S. and MACKENZIE, J. K. (1954) *Acta metall.* **2**, 129, 138.
BULLOUGH, R. and BILBY, B. A. (1956) *Proc. Phys. Soc. B*, **69**, 1276.
CAHN, R. W. (1953) *Acta metall.* **1**, 49.
CROCKER, A. G. (1962) *Phil. Mag.* **7**, 1901.
FRANK, F. C. (1965) *Acta crystallogr.* **18**, 862.
FRANK, F.C. (1988) *Met. Trans.* **19A**, 403.
HANDSCOMB, D.C. (1958) *Can.J.Math.* **10**, 85.
HUNTINGTON, H. B. (1958) *Solid State Phys.* **7**, 214.
JASWON, M. A. and WHEELER, J. A. (1948) *Acta crystallogr.* **1**, 216.
NEUMANN, P. (1966) *Phys. stat. sol.* **17**, K71.
NICHOLAS, J. F. (1966) *Acta crystallogr.* **21**, 880; (1970); *Phys. stat. sol.* (a) **1**, 563.
NICHOLAS, J. F. and SEGALL, R. L. (1970) *Acta crystallogr. A***26**, 522.
NYE, J. F. (1957) *Physical Properties of Crystals*, Clarendon Press, Oxford.
OKAMATO, P. R. and THOMAS, G. (1968) *Phys. stat. sol.* **25**, 81.
OTTE, H. M. and CROCKER, A. G. (1965) *Phys. stat. sol.* **9**, 44; (1966) *Ibid.* **16**, K25.
SOKOLNIKOFF, I. S. (1946) *Mathematical Theory of Elasticity*, McGraw-Hill, New York; (1951) *Tensor Analysis*, Wiley, New York.
ZACHARIASEN, W. H. (1945) *Theory of X-Ray Diffraction in Crystals*, Wiley, New York.