Chapter 1

Crystal Lattices

1.1 The Solid State

X-ray crystallography deals with materials in the solid state, and it is there that we must begin. In principle, depending on the temperature, any material can exist as a solid, a liquid, or a gas — provided that it does not decompose before it attains a given state. Materials are comprised of either atoms, neutral molecules, or ions, which, in the gas phase, behave more or less independently, despite attractive forces between them. This is because these entities have sufficient kinetic energy that the attractive forces between them are comparatively very weak. While they may tug a bit on one another in passing, they readily escape from one another to continue their motion indefinitely.

As the material is cooled — as kinetic energy is removed from the atoms, molecules, or ions in the gas phase — a point is reached at which they no longer have sufficient kinetic energy to move freely, and the attractive forces begin to dominate, keeping the entities in close proximity to one another. While they have enough kinetic energy to move around, over, and under one another, they are now constrained to stay in the vicinity of one another, and the material is said to have condensed into the liquid state.

Finally, when the material has cooled sufficiently, the attractive forces become so dominant that the components hold onto one another tenaciously, tending to arrange themselves so that they can gain as much interaction as possible, resulting in a minimum energy configuration. In the process they lose their translational motion, and the material becomes rigid. We now consider the manner in which the components of these rigid materials arrange themselves in order to maximize the attractive forces between them.

Since the attractive forces increase as the distances between entities decrease it is probably not surprising that simple materials, consisting of spherical ions or atoms, will tend to pack the spheres together in an orderly arrangement in order to get as many spheres as possible touching one another. Fig. 1.1 shows the arrangement of the spherical atoms in metallic copper and the spherical ions in a typical salt, sodium chloride. Both provide an excellent illustration of the old idiom, "Nature abhors a vacuum." In both structures the organization is clearly evident – spheres

Understanding Single-Crystal X-Ray Crystallography. Dennis W. Bennett

ISBN: 978-3-527-32677-8 (HC), 978-3-527-32794-2 (SC)

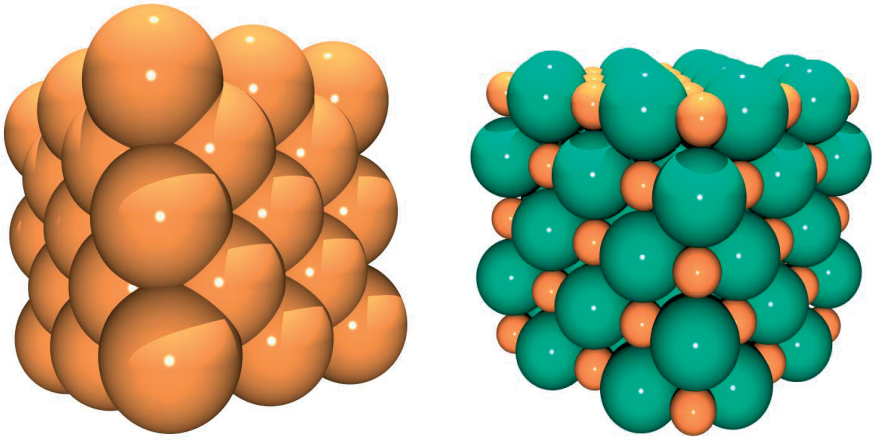

Figure 1.1 Arrangements of the atoms in copper and the ions in sodium chloride.

tend to arrange themselves into ordered three-dimensional arrays. The internal organization of such materials is reflected in their external appearance; both materials appear to the naked eye as cubes — or other polyhedra related to the cube — as shown in Fig. 1.2. In contrast to atoms and atomic ions, molecules and polyatomic ions are not spherical. Indeed, many molecules have virtually no symmetry of their own. Nevertheless, they generally form solid crystals with a symmetric appearance, suggesting that they also have ordered internal arrangements. In Fig. 1.3, a space-filling model of a sucrose molecule (table sugar) is compared to the macroscopic appearance of solid sucrose.

Figure 1.2 Naturally occurring crystals of metallic copper coated with tenorite and ionic sodium chloride (halite). Copper crystal specimen and photo courtesy of Rob Lavinsky; halite crystal photo courtesy of the Smithsonian Institution.

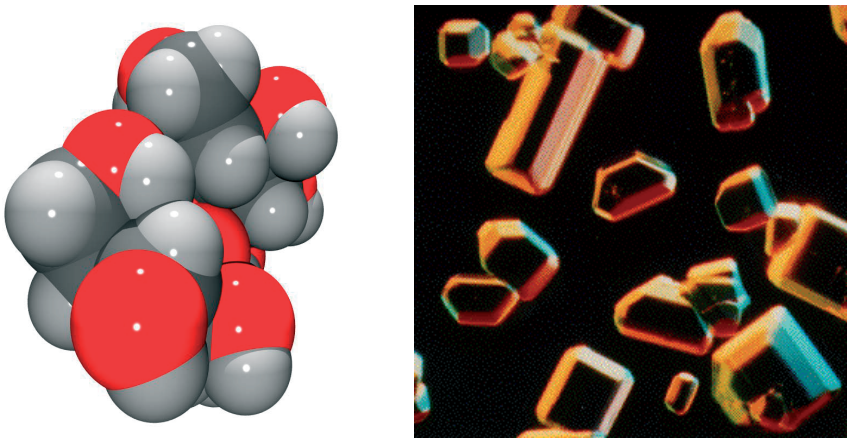

Figure 1.3 Space-filling model of a sucrose molecule and crystals of sucrose. Photo courtesy of Nicolas von Geijn and the Andrew van Hook Association.

The molecule has no observable symmetry, yet the solid material consists of symmetric crystals. A simpler molecule, 2-mercaptopyridine, illustrates how most molecules "fill the space," and pack next to one another in symmetric arrays. The space-filling representation of the molecule is shown on the left in Fig. 1.4. The arrangement of the molecules as they occur in the solid state indicates clearly that they arrange themselves in a periodically repeating sequence, which can be visualized as extending indefinitely throughout the three dimensions of the crystal. This periodicity is much easier to visualize if the space-filling representation of the molecule is replaced by a "ball and stick" model, indicating the locations of the atomic centers and their intramolecular connectivities (bonds), as illustrated in Fig. 1.5. Note that it is possible to construct a series of identical parallelepiped "cages" in

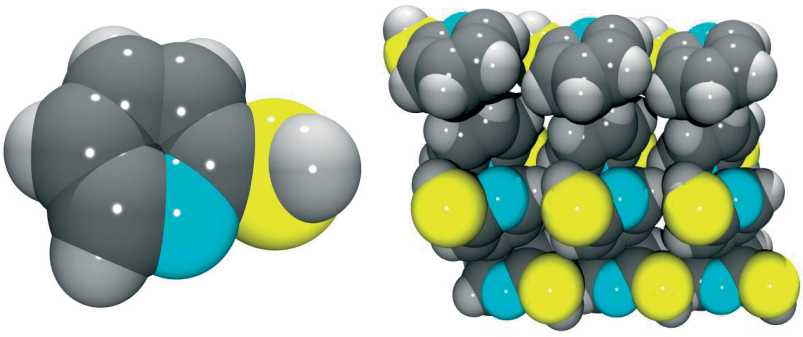

Figure 1.4 Space-filling model of 2-mercaptopyridine and the arrangement of the molecules in the solid state. The yellow atom is sulfur, the dark gray atoms are carbon, the light gray atoms are hydrogen, and the blue atom is nitrogen.

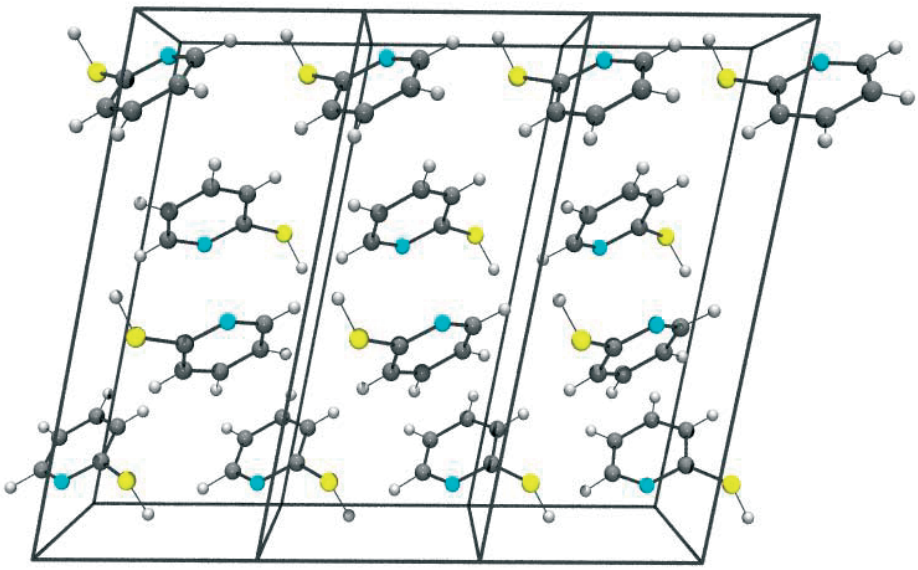

Figure 1.5 Crystallographic packing of 2-mercaptopyridine showing three unit cells translated along the a axis.

the structure, such that the contents and appearance of these cages are identical, again throughout the crystal.

We will come to know these parallelepipeds as the *unit cells* of the crystal and the framework that results from the entire array of these unit cells as the *crystal lattice*. Not all materials condense into these regular structures. Ideally, when a solid material is allowed to form slowly its component molecules or ions have sufficient opportunity to rearrange into a symmetric structure. However, if they lose kinetic energy too rapidly they will be trapped in less symmetric arrangements, and the rigid material will not exhibit the crystallinity of more ordered solids. These disordered materials are known as *glasses*. A common example is window glass, which is identical in chemical composition to quartz, except that window glass is formed from the rapid cooling of molten silicon dioxide, while quartz crystals are formed in nature from the same elements — slowly over millions of years. Fortunately for the crystallographer most substances, given the opportunity, tend to form crystalline solids.

1.2 The Crystal Lattice

1.2.1 Two-dimensional Lattices

It is often useful to begin a discussion in two dimensions, then extrapolate what we have learned to three dimensions. Consider a two-dimensional analog of the periodic arrangement of molecules described in Section 1.1 (Fig. 1.6), which we will call the *beta-structure*. The periodicity of this array, as in the previous three-

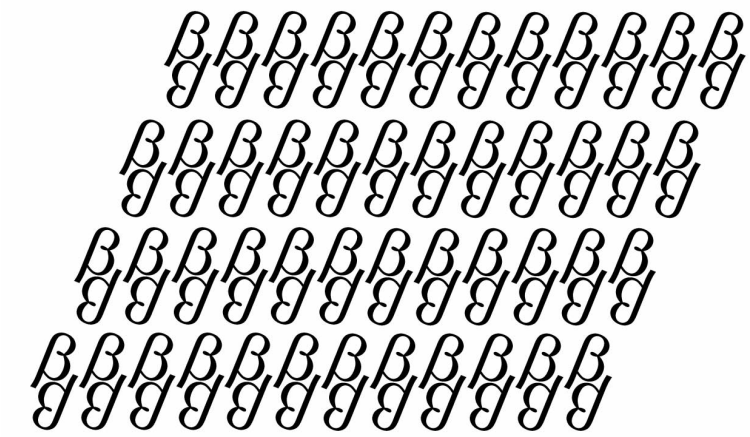

Figure 1.6 Two-dimensional periodic array.

dimensional case, can be represented by a two-dimensional lattice resulting from a series of parallelogram cells with identical contents (Fig. 1.7). The framework is created from two sets of equidistant parallel lines. The lattice is uniquely *defined* as the set of points at the intersections of these two sets of lines. The line segments

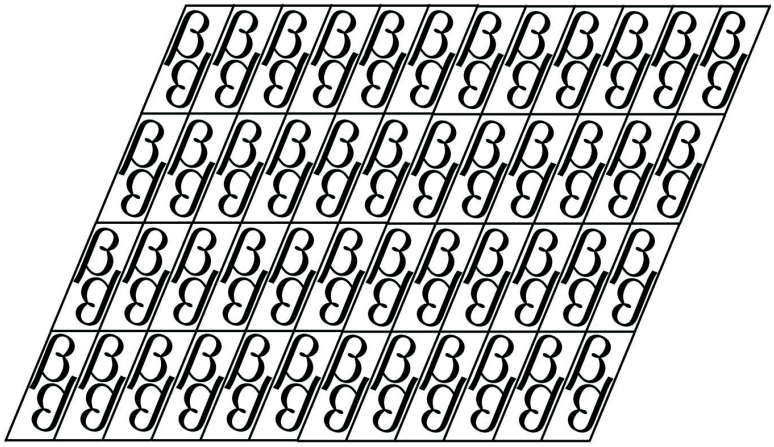

Figure 1.7 Lattice and contents of unit cells in a two-dimensional periodic array.

which make up the unit cell are called *axes*, since they will become the coordinate axes for locating points within the unit cell, much like the x and y axes in Cartesian coordinates. The unit cell is a parallelogram characterized by the lengths of its axes, a and b, and the angle between them, γ (Fig. 1.8). While the set of points defining

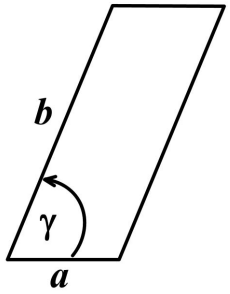

Figure 1.8 Unit cell for the beta-structure illustrating the crystallographic axes and the inter-axial angle.

the lattice is unique for a given structure, the sets of lines which define the unit cell, the repeating unit in the lattice, are the not the only sets of equidistant parallel lines which contain all of the points in the lattice, as illustrated in Fig. 1.9. Indeed, there are infinitely many sets of parallel lines in a given structure, each of which contains all of the points in the lattice. Since a given set of lines contains every point in the lattice and consists of equally spaced parallel lines, each set will divide a given unit cell axis into an integral number of equally spaced line segments.

In the example shown in Fig. 1.9, the red lines divide the a axis into 2 segments, and the b axis into 3 segments. The blue lines divide the a axis into 1 segment, and the b axis into 2 segments. Since each set of lines divides the axes differently, the two integers representing the number of divisions for both axes will be unique for a given set (subject to the caveat discussed below). It follows that each set of lines can be identified by these two integers, and that they can be used as *indices* to classify each set, provided that they are kept in "a, b" order. This scheme was devised by the Welsh crystallographer William Miller[19] (1801–1880), and the ordered integers are known as the *Miller indices* of the lines to which they correspond. They are generally placed between parentheses. In the example given in Fig. 1.9 the Miller indices of the red lines are (2 3), and(1 2) for the blue lines. By convention, the indices for a given set of lines are denoted (h k). This assignment is not unambiguous, since for every set of lines with a negative slope, there is another set of lines with indices of the same magnitude but with a positive slope, as indicated by the green lines in Fig 1.9. (which also divide a into one segment and b into two segments). We will find a convenient and formal way to deal with this ambiguity later on in the chapter by taking axial directions into account using vectors. For now we can deal with this informally by assigning positive directions along a and b to be to the right and up, and negative directions to be to the left and down. We select any lattice point as an origin and look for the next line away from the origin which crosses the axes. We note whether the crossing points are in positive or

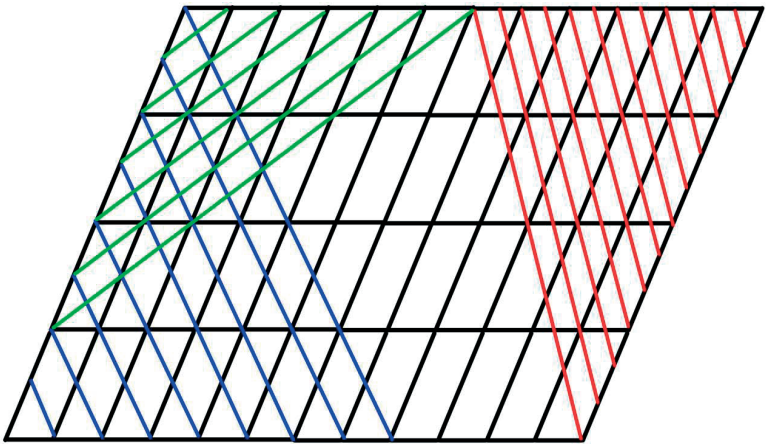

Figure 1.9 Three of the infinitely many sets of equidistant parallel lines that intersect every point in the beta-structure lattice.

negative directions along a and b, then assign the signs of those directions to the indices. Negative indices are indicated by placing a *bar* over the index. The lines in Fig. 1.9 serve as examples to clarify this. We find a blue line which intersects the axes one axial length to the *right* along a and, *for the same unit cell*, half of an axial length *up* along b. Indices $(h\ k) = (1\ 2)$ are therefore assigned to the blue lines. Note that there is also a blue line which crosses the a axis to the *left* and *down* along b, so that indices $(\bar{1}\ \bar{2})$ can also be assigned to the blue lines. While these represent the same lines, there are special cases in diffraction where *direction* is important, and we therefore keep them separate to account for this. Similarly, if we proceed in a positive direction along a to a green line ($h = +1$), we find that the line intersects the b axis halfway down — in a negative direction, and indices $(1\ \bar{2})$ are assigned to the green lines. Alternatively, if we proceed along a in a negative direction, we find a green line which intersects b in a positive direction, and we can equally assign indices $(\bar{1}\ 2)$ to the green lines — again retaining both sets of indices for the special cases that we will encounter later on. It is left to the reader to verify that the red lines have indices (2 3) – and $(\bar{2}\ \bar{3})$. The original set of lines collinear with the a axis divide the b axis into 1 segment, and do not divide the a axis at all; thus they are assigned indices (0 1) – and $(0\ \bar{1})$. Similarly, the lines collinear with the b axis are assigned indices(1 0) – and $(\bar{1}\ 0)$. Note that the larger the magnitude of an index, the more times the lines cross the axis corresponding to that index, and the shorter the resulting line segments become. *It follows that as the magnitudes of the Miller indices increase, the spacings between the correspondingly indexed parallel lines must decrease.* We will encounter this *reciprocal* relationship again in the next section (in three dimensions) and throughout the text — it lies at the heart of the diffraction experiment!

1.2.2 Three-dimensional Lattices

Two-dimensional lattices are created by the intersection of two sets of equidistant parallel *lines;* three-dimensional lattices are the result of the intersection of three

sets of intersecting parallel *planes* (Fig. 1.10). The unit cell thus becomes a parallelepiped, characterized by the length of its three axes, a, b, and c and the three corresponding inter-axial angles, α, β, and γ. By convention, α is the angle between the b and c axes, β is the angle between the a and c axes, and γ is the angle between the a and b axes, as illustrated for the 2-mercaptopyridine unit cell in Fig. 1.11(a).

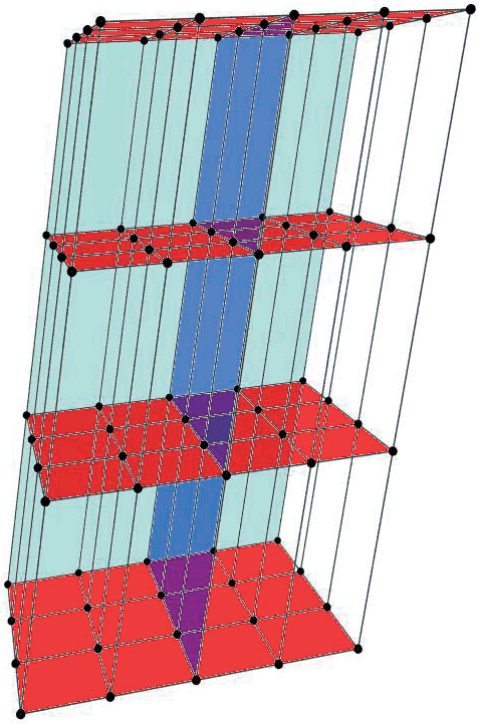

Figure 1.10 Three sets of intersecting equidistant parallel planes defining the lattice of the 2-mercaptopyridine structure. The lattice consists of the points indicated with small black spheres. Four a-b planes are shown in red, an a-c plane is shown in green, and a b-c plane is shown in blue.

The analogy with the two-dimensional lattice also holds for the Miller indices. Just as there are infinitely many sets of equidistant parallel lines intersecting every point in the two-dimensional lattice, there are an infinite number of sets of equidistant parallel planes which intersect every point in the three-dimensional lattice. Fig. 1.11(b) shows a set of planes that divides the a axis into two segments, the b axis into 2 segments and the c axis into 3 segments. As with each set of lines in two dimensions, this set of planes can be assigned Miller indices, (h k l) — in this example, (2 2 3). There are seven other sets of planes with indices ($\bar{2}$ $\bar{2}$ $\bar{3}$), (2 2 $\bar{3}$), ($\bar{2}$ $\bar{2}$ 3), (2 $\bar{2}$ 3), ($\bar{2}$ 2 $\bar{3}$), ($\bar{2}$ 2 3), and (2 2 $\bar{3}$), determined by assigning directions along a, b, and c, just as we did in Sec. 1.2.1. Each set of planes is

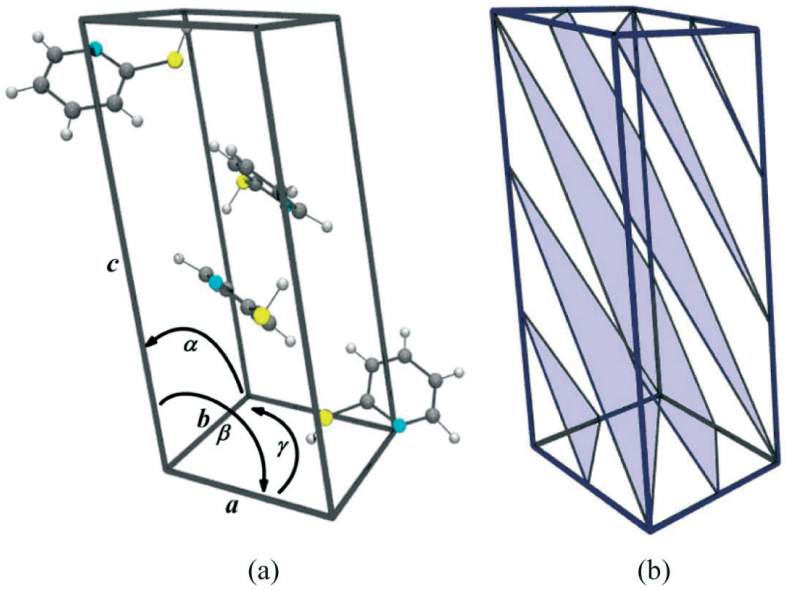

Figure 1.11 Unit cell for the 2-mercaptopyridine structure illustrating (a) conventional unit cell parameters and (b) intersection of the $(h\ k\ l) = (2\ 2\ 3)$ planes with the unit cell axes.

oriented differently in the lattice and is uniquely characterized by its indices.* As with the two-dimensional lattice, each set of planes with indices $(h\ k\ l)$ can also be indexed with $(\bar{h}\ \bar{k}\ \bar{l})$, so that there are actually four unique sets of planes; we retain the eight sets of indices to keep track of *direction*. The planes containing the b and c axes divide the a axis into one segment — and since they do not divide the b or c axes, they are assigned indices $(1\ 0\ 0)$ – and $(\bar{1}\ 0\ 0)$. Similarly, the a-c planes are assigned $(0\ 1\ 0)$— and $(0\ \bar{1}\ 0)$ indices, and the a-b planes are assigned $(0\ 0\ 1)$ — and $(0\ 0\ \bar{1})$ indices. As in the two-dimensional case, the larger a given index, the more the corresponding axis is divided, and the smaller the distance between the planes. *Again, there is a reciprocal relationship between the magnitude of the plane indices and the distances between the planes corresponding to those indices.*

*In the early days of crystallography, planes were indexed to describe crystal faces. For every set of planes with indices $(h\ k\ l)$, there is another set with indices $(2h\ 2k\ 2l)$, and another with indices $(3h\ 3k\ 3l)$, etc. All of these planes are parallel to one another, and macroscopically indistinguishable. For that reason Miller indices formally have no common factors, thus excluding the $(nh\ nk\ nl)$ planes for $n > 1$. Since we must include these planes, from this point on we will refer to the indices of all of the planes in the lattice as *general indices*, or more concisely, simply as *indices*.

1.3 Vectors in Crystallography

The discussion of indices in the previous section required us to consider both the *magnitudes* and *directions* of line segments in order to assign the indices unambiguously. Directed line segments are often referred to as *vectors*, although formally they are geometric *representations* of vectors. In order to develop the mathematics of crystallography and diffraction we will rely heavily on the use of vectors.

A vector is an ordered array of elements known as *components*. The number of components is referred to as the *dimension* of the vector. We will describe a vector by placing its components in square brackets: $[v_1\ v_2\ v_3 \ldots v_n]$, where v_i is the ith component of an n-dimensional vector. A boldface lower case letter will be used to indicate the vector (e.g., $\mathbf{v} = [v_1\ v_2\ v_3]$) *and* its geometric representation; the term *vector* will be used for both as well.

The sum of two vectors *of the same dimension* produces a third vector of that dimension,

$$\begin{aligned} \mathbf{v_i} + \mathbf{v_j} &= [v_{i,1}\ v_{i,2}\ v_{i,3} \ldots v_{i,n}] + [v_{j,1}\ v_{j,2}\ v_{j,3} \ldots v_{j,n}] \\ &= [(v_{i,1} + v_{j,1})\ (v_{i,2} + v_{j,2})\ (v_{i,3} + v_{j,3} \ldots (v_{i,n} + v_{j,n})]. \end{aligned} \tag{1.1}$$

A one-dimensional vector is known as a *scalar*. It is generally a single number which represents the *magnitude* of something, and is ordinarily not placed in brackets. The product of a vector of dimension n and a scalar results in a vector of the same dimension in which each component is multiplied by the scalar:

$$s\mathbf{v} = s[v_1\ v_2\ v_3 \ldots v_n] = [sv_1\ sv_2\ sv_3 \ldots sv_n]. \tag{1.2}$$

We have already encountered examples of two and three-dimensional vectors — the indices of sets of lines in two dimensions and sets of planes in three dimensions. The indices are descriptors, in the sense of identifying specific sets of planes. Indices are also vectors which occur commonly in the mathematics of diffraction. We will adopt the convention of placing indices in parentheses when referring specifically to sets of lattice planes, and in square brackets when they are to be employed as vectors.

Crystallography is chiefly concerned with directed line segments in two or three dimensions, and the discussion here will largely be confined to those vectors. A directed line segment can be described by specifying its length (*magnitude*) and direction, usually expressed as an angle or angles with respect to some standard direction, and indicated by an arrowhead at the leading end of the line segment. The leading end of the line segment is called its *head*, the trailing end is called its *tail*. A vector with its tail at point o and its head at point p will be indicated by placing an arrow over the symbols for the points, e.g., $\mathbf{v} = \overrightarrow{op}$. Unfortunately, there is no standard convention for indicating the magnitude of a vector. For example, the magnitude of the vector $\mathbf{v}$ can be found in various texts as $|\mathbf{v}|$, $\|\mathbf{v}\|$, or simply v. With the exception of cases where it might be confusing, a lower case script letter will be employed to indicate any scalar quantity, including the magnitude of a vector. When the vector is represented by the points which define it, its magnitude will be indicated by placing the symbol for the vector between vertical lines, e.g., $v = |\overrightarrow{op}|$.

1.3.1 Geometric Vector Addition and Multiplication

Geometric vectors are added and multiplied by an established set of rules, and we must consider those rules in order to link ordered pairs and triples of numbers with geometric vectors in two and three dimensions. Geometric vectors, $\mathbf{v_1}$ and $\mathbf{v_2}$, are added to one another by placing the tail of $\mathbf{v_2}$ at the head of $\mathbf{v_1}$, and creating a third *resultant* vector by connecting the tail of $\mathbf{v_1}$ to the head of $\mathbf{v_2}$ (Fig. 1.12). Note that this resultant vector is also obtained if we reverse the order in which we combine the vectors. For "regular" numbers (scalars) the order in which we add things does not alter the results, and we say that scalar addition is *commutative*. Vector addition behaves just like regular addition in this respect – vector addition is commutative.

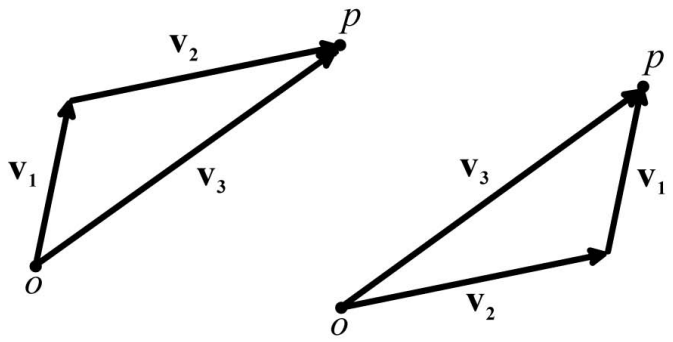

Figure 1.12 The addition of geometric vectors: $\mathbf{v_3} = \mathbf{v_1} + \mathbf{v_2} = \mathbf{v_2} + \mathbf{v_1}$. Note that merging the two representations creates a parallelogram with the resultant vector along the diagonal — providing a convenient method for vector addition.

Adding vectors in this manner makes intuitive sense if we consider our intent to make use of them in order to locate points inside the crystal lattice. Locating the point p relative to point o can be effected by traversing directly along $\mathbf{v_3}$, or by taking the path along $\mathbf{v_1}$, followed by $\mathbf{v_2}$. The combined effects of $\mathbf{v_1}$ and $\mathbf{v_2}$ are clearly additive in this respect, and equivalent to $\mathbf{v_3}$.

Vectors can be multiplied* in a number of ways. The simplest is multiplication by a scalar. Multiplication by a positive scalar produces a new vector in the same direction as the original with its magnitude multiplied by the scalar; multiplying by a negative scalar produces a new vector in the opposite direction. Fig. 1.13 illustrates a vector multiplied by the scalars 0.5 and -2.0, producing a new vector parallel to the original with half the length ($v_2 = 0.5v_1$) and an antiparallel vector with twice the length ($v_3 = -2.0v_1$).

Multiplying a vector by -1 produces the negative of that vector — a vector of equal magnitude but opposite direction. In order to subtract $\mathbf{v_2}$ from $\mathbf{v_1}$, we

**Multiplication* is used here in a more abstract context to indicate a *combination* of entities in accordance with a specific set of rules. In this sense *addition* is also considered a form of *multiplication*.

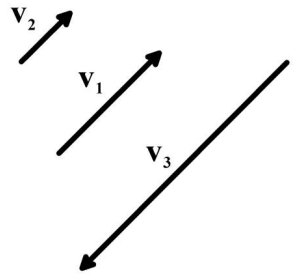

Figure 1.13 Geometric vector $\mathbf{v_1}$ multiplied by scalars: $\mathbf{v_2} = 0.5\mathbf{v_1}$; $\mathbf{v_3} = -2.0\mathbf{v_1}$

multiply $\mathbf{v_2}$ by -1 to negate it, then add it to $\mathbf{v_1}$. Note that vector subtraction is *not* commutative. As illustrated in Fig. 1.14, reversing the order of subtraction changes the sign of the resultant vector, but leaves the magnitude unchanged. As with addition, the subtraction of vectors parallels the non-commutative subtraction of ordinary numbers.

Vectors can also be multiplied together to produce a scalar. The *scalar product* of two vectors (also known as the *inner product* or *dot product*) is defined as the product of the magnitudes of the two vectors and the cosine of the angle between them:

$$\mathbf{v_1} \cdot \mathbf{v_2} = v_1 v_2 \cos\theta. \tag{1.3}$$

Since the magnitude of the original vectors and the angle between them does not change, $\mathbf{v_1} \cdot \mathbf{v_2} = \mathbf{v_2} \cdot \mathbf{v_1}$; the scalar product is commutative. The scalar product has three important properties which are very useful in vector analysis. First, the scalar product of a vector with itself gives us the square of the magnitude (length) of the vector:

$$\mathbf{v_1} \cdot \mathbf{v_1} = v_1 v_1 \cos(0) = v_1^2. \tag{1.4}$$

Second, the scalar product of two *orthogonal* (perpendicular) vectors is zero ($\theta = \pi/2 = 90°$) :

$$\mathbf{v_1} \cdot \mathbf{v_2} = v_1 v_2 \cos(\pi/2) = 0, \mathbf{v_1} \perp \mathbf{v_2}. \tag{1.5}$$

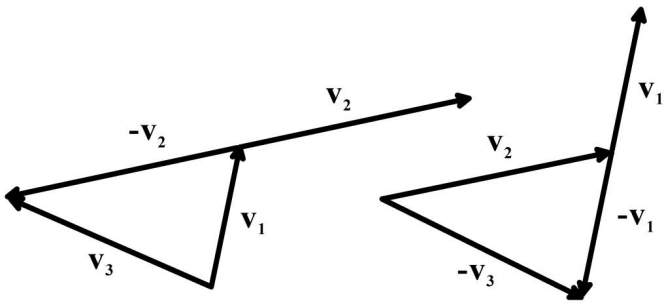

Figure 1.14 The subtraction of geometric vectors: $\mathbf{v_3} = \mathbf{v_1} - \mathbf{v_2}$; $\mathbf{v_3'} = -\mathbf{v_3} = \mathbf{v_2} - \mathbf{v_1}$.

The third property involves the *projection* of one vector onto another. This entails the construction of a perpendicular from one vector to the head of another, as illustrated in Fig. 1.15. $\mathbf{v_1}$ is said to have been projected onto $\mathbf{v_2}$. p_1 is the magnitude of the projection of $\mathbf{v_1}$ onto $\mathbf{v_2}$. p_1 can be determined from the scalar product $\mathbf{v_1} \cdot \mathbf{v_2}$ and the magnitude of $\mathbf{v_2}$:

$$\begin{aligned} \cos\theta &= \frac{v_1}{p_1} \\ p_1 &= v_1 \cos\theta \\ \mathbf{v_1} \cdot \mathbf{v_2} &= v_1 v_2 \cos\theta \\ \mathbf{v_1} \cdot \mathbf{v_2} &= v_2 p_1 \\ p_1 &= \frac{\mathbf{v_1} \cdot \mathbf{v_2}}{v_2}. \end{aligned} \tag{1.6}$$

We will find all three of these properties very useful as we analyze the vectors in both the crystal lattice and the X-ray diffraction pattern.

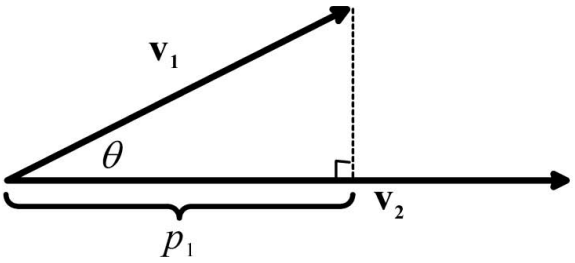

Figure 1.15 The projection of vector $\mathbf{v_1}$ onto the vector $\mathbf{v_2}$.

Finally, two vectors can be multiplied to produce a *vector product* (also known as an *outer product* or *cross product*) — a new vector created from the originals. The magnitude of the new vector is determined as the product of the magnitudes of the original two vectors and the sine of the angle between them:

$$\mathbf{v_3} = \mathbf{v_1} \times \mathbf{v_2}, \tag{1.7}$$

$$v_3 = v_1 v_2 \sin\theta. \tag{1.8}$$

This new vector is perpendicular to *both* of the original vectors, and is therefore perpendicular to the plane in which both vectors lie (Fig. 1.16). Since $\mathbf{v_3}$ can point in either of two directions, both perpendicular to the $\{\mathbf{v_1}, \mathbf{v_2}\}$ plane, we adopt the convention that it if we align the index finger of *the right hand* along $\mathbf{v_1}$, and the middle finger of the right hand along $\mathbf{v_2}$, then $\mathbf{v_3}$ will point in the direction of the right thumb. *This is referred to as the right hand rule.* Note that $\mathbf{v_2} \times \mathbf{v_1}$ produces a vector of the same magnitude pointing in the opposite direction to $\mathbf{v_1} \times \mathbf{v_2}$. Thus $\mathbf{v_2} \times \mathbf{v_1} = -(\mathbf{v_1} \times \mathbf{v_2})$; vector products are *not* commutative.

1.3.2 Basis Vectors and Coordinates

As the reader might suspect at this point, unit cell axes can be represented as vectors with tails sharing a common lattice point, which we will designate as the

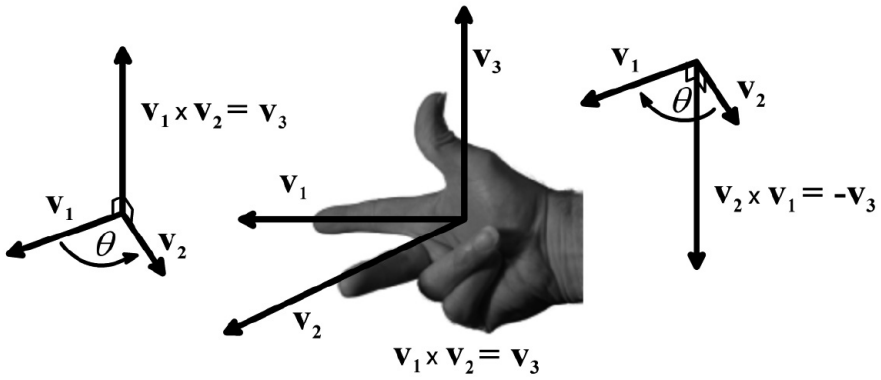

Figure 1.16 *Right hand rule* for the vector product: $\mathbf{v_3} = \mathbf{v_1} \times \mathbf{v_2}$, $-\mathbf{v_3} = \mathbf{v_2} \times \mathbf{v_1}$, $v_3 = v_1 v_2 \sin\theta$.

origin of the unit cell, illustrated in Fig. 1.17 for a two-dimensional unit cell. X-ray crystallography involves specification of locations inside the lattice, and the unit cell axes provide a convenient set of reference vectors for determining the location of a point anywhere in the lattice — as the end of a vector with its tail at the origin and its head coincident with the point. We can imagine taking a path to the point p along a vector $\mathbf{x}$ coincident with the $\mathbf{a}$ axis, then following a vector $\mathbf{y}$ parallel to the $\mathbf{b}$ axis to get to p — or alternatively taking the direct path along $\overrightarrow{\mathbf{op}}$ — which is clearly the vector sum of $\mathbf{x}$ and $\mathbf{y}$:

$$\overrightarrow{\mathbf{op}} = \mathbf{x} + \mathbf{y}$$

Because $\mathbf{x}$ is parallel to $\mathbf{a}$ it can be expressed as the vector $\mathbf{a}$ multiplied by a scalar, $\mathbf{x} = s_x\mathbf{a}$; similarly, $\mathbf{y} = s_y\mathbf{b}$:

$$\overrightarrow{\mathbf{op}} = s_x\mathbf{a} + s_y\mathbf{b}$$

Note that *any* vector in two-dimensions can be described as a vector sum of reference vectors multiplied by appropriate scalars. Since the sum contains only linear terms (no exponents), it is called a *linear combination*; the reference vectors are known as *basis* vectors, and the set of two reference vectors (in two-dimensions) is called a *basis set*, denoted $\{\mathbf{v_1}, \mathbf{v_2}\}$ — in our case $\{\mathbf{a}, \mathbf{b}\}$. In the general case, the selection of the initial basis set is arbitrary, provided that the two vectors are not parallel. We can express two parallel vectors as a linear combination of one another (by multiplying one of the vectors by a scalar and the other by zero). Such vectors are said to be *linearly dependent*. Basis vectors which are not parallel to one another cannot be expressed as linear combinations of one another, and are termed *linearly independent* vectors. We can now state these observations more formally — *any vector in two-dimensional space can be expressed as a linear combination of two linearly independent basis vectors.*

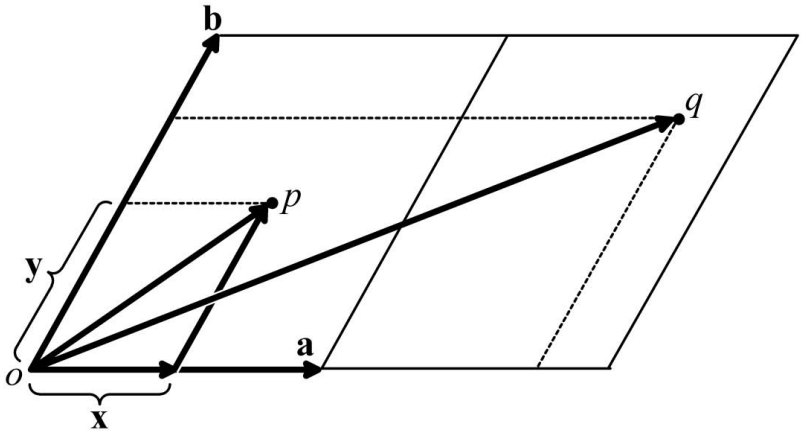

Figure 1.17 Two-dimensional unit cells illustrating the axes **a** and **b** as basis vectors for two-dimensional space and the vectors locating points p and q as linear combinations of the axial basis vectors.

A logical way to determine s_x and s_y is to express them as fractions of the lengths of the basis vectors:

$$\begin{aligned} s_x &= \frac{x}{a} = x_f \\ s_y &= \frac{y}{b} = y_f \\ \overrightarrow{\mathbf{op}} &= x_f\mathbf{a} + y_f\mathbf{b} \end{aligned}$$

Once we have selected our basis vectors, $\{\mathbf{a}, \mathbf{b}\}$, we can locate any point in the lattice by determining the ordered pair of numbers, $[x_f\ y_f]$, defining a vector from the origin of the unit cell to that point. The determination of a basis set and the expression of vectors as linear combinations of the basis vectors provides the link between vectors and their geometric representations in two dimensions:

$$\overrightarrow{\mathbf{op}} = [x_f\ y_f]. \tag{1.9}$$

The components of the vector, x_f and y_f, are called the *coordinates* of the vector, and when expressed as fractions of the magnitudes of the basis vectors are known as *fractional coordinates*. For example, in Fig. 1.17,

$$\begin{aligned} \overrightarrow{\mathbf{op}} &= \tfrac{1}{2}\,\mathbf{a} + \tfrac{1}{2}\,\mathbf{b} = [\tfrac{1}{2}\ \tfrac{1}{2}] \\ \overrightarrow{\mathbf{oq}} &= 1\tfrac{3}{4}\,\mathbf{a} + \tfrac{3}{4}\,\mathbf{b} = [1\tfrac{3}{4}\ \tfrac{3}{4}] \\ \mathbf{x} &= \tfrac{1}{2}\,\mathbf{a} + 0\,\mathbf{b} = [\tfrac{1}{2}\ 0] \\ \mathbf{y} &= 0\,\mathbf{a} + \tfrac{1}{2}\,\mathbf{b} = [0\ \tfrac{1}{2}] \\ \mathbf{a} &= 1\,\mathbf{a} + 0\,\mathbf{b} = [1\ 0] \qquad (1.10) \\ \mathbf{b} &= 0\,\mathbf{a} + 1\,\mathbf{b} = [0\ 1]. \qquad (1.11) \end{aligned}$$

The product of a vector and a scalar can now be expressed in terms of the components of the vector. Fig. 1.18 illustrates the formation of $\mathbf{v_2} = [v_{2x}\ v_{2y}]$ as the product of $\mathbf{v_1} = [v_{1x}\ v_{1y}]$ and a scalar, s. The triangles with sides (v_1, v_{1x}, v_{1y}) and (v_2, v_{2x}, v_{2y}) are similar triangles, which means that the ratios of the lengths of similar sides are a constant. Since $\mathbf{v_2} = s\mathbf{v_1}$,

$$
\begin{array}{rcl}
v_2 & = & sv_1 \\
\dfrac{v_{1x}}{v_1} & = & \dfrac{v_{2x}}{v_2} = \dfrac{v_{2x}}{sv_1} \qquad \dfrac{v_{1y}}{v_1} = \dfrac{v_{2y}}{v_2} = \dfrac{v_{2y}}{sv_1} \\
v_{2x} & = & sv_{1x} \qquad\qquad v_{2y} = sv_{1y} \\
\mathbf{v_2} & = & s\mathbf{v_1} = [v_{2x}\ v_{2y}] = s[v_{1x}\ v_{1y}] = [sv_{1x}\ sv_{1y}].
\end{array}
$$

Multiplying a vector by a scalar produces a new vector formed by multiplying each component of the original vector by the scalar.

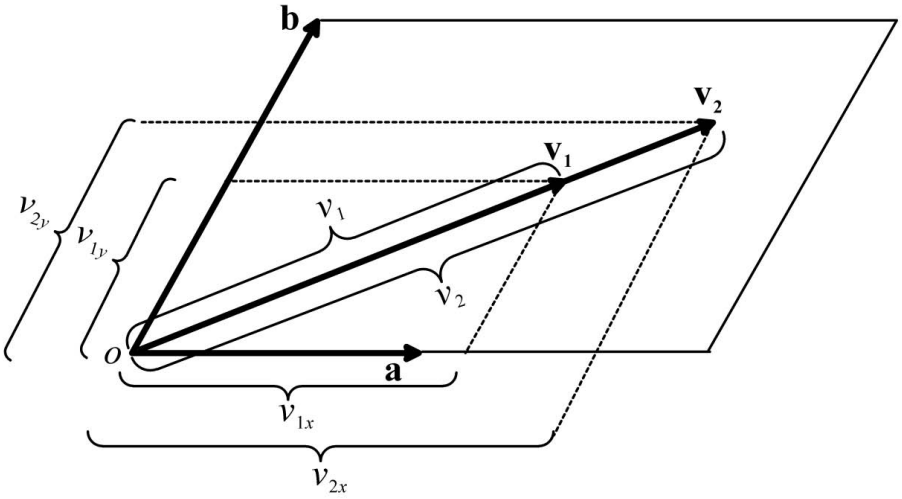

Figure 1.18 The components of the product of a vector and a scalar: $\mathbf{v_2} = s\mathbf{v_1}$.

Vector addition can also be expressed in terms of vector components. Fig. 1.19 illustrates the formation of $\mathbf{v_3} = [v_{3x}\ v_{3y}]$ as the sum of $\mathbf{v_1} = [v_{1x}\ v_{1y}]$ and $\mathbf{v_2} = [v_{2x}\ v_{2y}]$. In order to move the tail of $\mathbf{v_2}$ to the head of $\mathbf{v_1}$, $\mathbf{v_2}$ must be translated by adding v_{1x} to its $\mathbf{a}$ component and v_{1y} to its $\mathbf{b}$ component:

$$
\begin{array}{rcl}
\mathbf{v_1} & = & v_{1x}\mathbf{a} + v_{1y}\mathbf{b} = [v_{1x}\ v_{1y}] \\
\mathbf{v_2} & = & v_{2x}\mathbf{a} + v_{2y}\mathbf{b} = [v_{2x}\ v_{2y}] \\
\mathbf{v_3} & = & v_{3x}\mathbf{a} + v_{3y}\mathbf{b} = [v_{3x}\ v_{3y}] \\
v_{3x} & = & v_{1x} + v_{2x} \\
v_{3y} & = & v_{1y} + v_{2y} \\
\mathbf{v_3} & = & (v_{1x} + v_{2x})\mathbf{a} + (v_{1y} + v_{2y})\mathbf{b} = [(v_{1x} + v_{2x})\ (v_{1y} + v_{2y})]
\end{array}
\tag{1.12}
$$

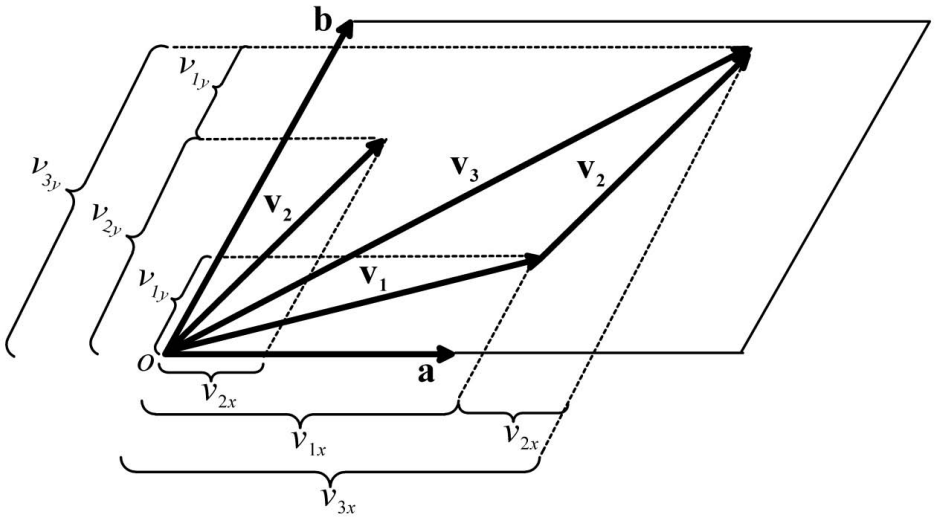

Figure 1.19 The components of the sum of two vectors: $\mathbf{v_3} = \mathbf{v_1} + \mathbf{v_2}$.

Finally, the subtraction of vectors can be accomplished using the vector components by combining scalar multiplication and vector addition:

$$\begin{aligned} -\mathbf{v_1} &= -1(v_{1x}\mathbf{a} + v_{1y}\mathbf{b}) = [-v_{1x} \ -v_{1y}] \\ \mathbf{v_2} - \mathbf{v_1} &= \mathbf{v_2} + (-\mathbf{v_1}) = (v_{2x} - v_{1x})\mathbf{a} + (v_{2y} - v_{1y})\mathbf{b} \\ &= [(v_{2x} - v_{1x}) \ (v_{2y} - v_{1y})] \end{aligned} \tag{1.13}$$

Adding or subtracting the components of two vectors produces a new vector which is the vector sum or difference of the original vectors.

Extrapolation of these concepts to a three-dimensional lattice is straightforward. *Any vector in three-dimensional space can be expressed as a linear combination of three linearly independent basis vectors.* The **c** axis becomes a third linearly independent basis vector, creating a three-dimensional basis set, $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, as illustrated in Fig. 1.20. A point p in the lattice is now found by traversing along vector x, coincident with the **a** axis, then along a vector y parallel to the **b** axis, and finally along z parallel to the **c** axis, such that

$$\overrightarrow{\mathbf{op}} = \mathbf{x} + \mathbf{y} + \mathbf{z}$$

$$\overrightarrow{\mathbf{op}} = s_x\mathbf{a} + s_y\mathbf{b} + s_z\mathbf{c}$$

$$s_x = \frac{x}{a} = x_f \tag{1.14}$$

$$s_y = \frac{y}{b} = y_f \tag{1.15}$$

$$s_z = \frac{z}{c} = z_f \tag{1.16}$$

$$\overrightarrow{\mathbf{op}} = x_f\mathbf{a} + y_f\mathbf{b} + z_f\mathbf{c} = [x_f \ y_f \ z_f]. \tag{1.17}$$

In Fig. 1.21(a) the vector $\overrightarrow{\mathbf{op}} = [\frac{1}{2} \ \frac{1}{2} \ \frac{1}{2}]$ defines the point at the center of the unit cell.

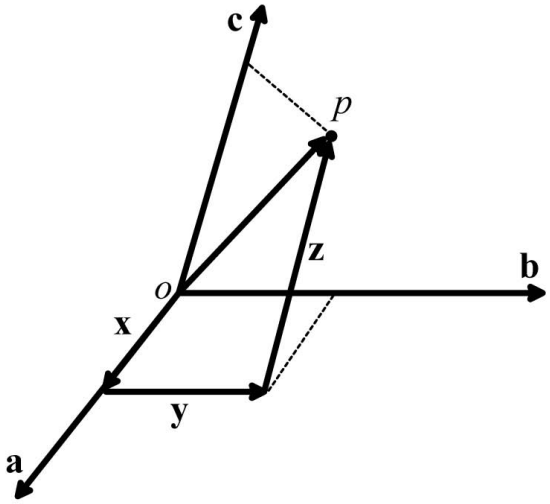

Figure 1.20 Three-dimensional unit cell illustrating the axes **a**, **b**, and **c** as basis vectors for three-dimensional space and the vector locating point p as a linear combination of the axial basis vectors.

The indices of a given set of lattice planes are conveniently determined from the fractional coordinates of the three vectors from the origin of the unit cell to the initial points of intersection of the planes along each coordinate axis. Recall that the indices were defined by the number of equal segments into which each axis was divided (Fig. 1.11). This led to an ambiguity in the assignment of the indices, since there were eight sets of planes which divided the axes into the same increments. These planes were differentiated by taking into account directions in the lattice, and the use of vectors provides a natural way to accomplish this. For a given set of planes, the **a**, **b**, and **c** axes are divided into h, k and l segments, respectively. The fractional coordinates to the first point of intersection along each axis are therefore $[(1/h\ 0\ 0]$, $[0\ 1/k\ 0]$, and $[0\ 0\ 1/l]$, as illustrated in Fig. 1.21(b). The signs of the indices are now taken as the signs of the vectors, determined from their directions in the right-handed unit cell coordinate system.

As we observed in two dimensions, multiplying a vector by a scalar in three dimensions produces a new vector formed by multiplying each component of the original vector by the scalar; adding or subtracting the components of two vectors produces the sum or difference of of the two vectors. For $\mathbf{v_1} = [v_{1x}\ v_{1y}\ v_{1z}]$ and $\mathbf{v_2} = [v_{2x}\ v_{2y}\ v_{2z}]$,

$$\mathbf{v_2} = s\mathbf{v_1} = [v_{2x}\ v_{2y}\ v_{2z}] = s[v_{1x}\ v_{1y}\ v_{1z}] = [sv_{1x}\ sv_{1y}\ v_{1z}] \tag{1.18}$$

$$\begin{aligned}\mathbf{v_3} &= \mathbf{v_1} + \mathbf{v_2} = (v_{1x} + v_{2x})\mathbf{a} + (v_{1y} + v_{2y})\mathbf{b} + (v_{1z} + v_{2z})\mathbf{c} \\ &= [(v_{1x} + v_{2x})\ (v_{1y} + v_{2y})\ (v_{1z} + v_{2z})]\end{aligned} \tag{1.19}$$

$$\begin{aligned}\mathbf{v'_3} &= \mathbf{v_2} - \mathbf{v_1} = (v_{2x} - v_{1x})\mathbf{a} + (v_{2y} - v_{1y})\mathbf{b} + (v_{2z} - v_{1z})\mathbf{c} \\ &= [(v_{2x} - v_{1x})\ (v_{2y} - v_{1y})\ (v_{2z} - v_{1z})].\end{aligned} \tag{1.20}$$

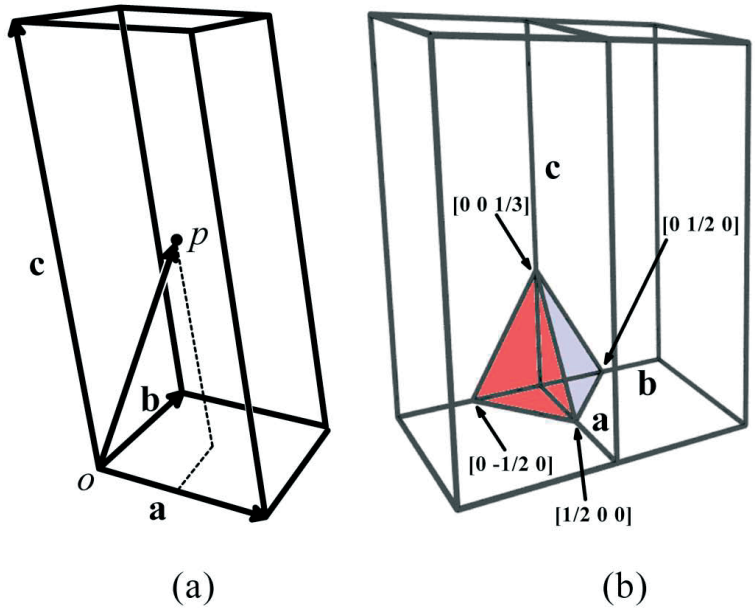

Figure 1.21 2-mercaptopyridine unit cell showing (a) p in the center of the cell at $[x_f\ y_f\ z_f] = [\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}]$ and (b) the fractional coordinates of the axial intersection vectors for the (2 2 3) planes (violet) and the (2 $\bar{2}$ 3) planes (red).

The multiplication of a vector by a scalar and the addition of vectors, both effected by operating on the vector components, is general for any basis set. This is *not* the case for the scalar and vector products. The formation of these two products from the components is transparent only in specific coordinate systems in which the basis vectors are vectors of unit length (*normalized*) which are perpendicular to one another(*orthogonal*). Such basis sets are known as *orthonormal* bases.

1.3.3 Orthonormal Bases

The unit cell axes are a natural choice for basis vectors when describing locations in the lattice — every coordinate consists of an integer (to define the specific unit cell) plus a fraction (to specify the position within that unit cell). Furthermore, since every unit cell is identical, the fractional coordinates alone uniquely define the internal structure of the crystal; the integers serve only to define the specific cell in which a targeted location exists. For example, the fractional coordinates [10.5 232.5 19.5] tell us that the point of interest lies in the center of a unit cell 10 a unit cell lengths along the **a** axial direction, 232 b units in the **b** direction, and 19 c units in the **c** direction. The periodic nature of the crystal lattice is reflected in this basis set, and we will see that its use greatly simplifies the mathematics of diffraction.

However, we will also need to perform numerical calculations, specifically concerned with molecular parameters such as bond lengths and angles. In addition,

we will find the need to transform the natural coordinates of the crystal lattice into a laboratory reference frame in the form of a Cartesian coordinate system in order to collect diffraction data. This will require us to describe the vectors in the lattice with a basis set which will allow us to use the coordinates to compute scalar and vector products. As alluded to above, this necessitates the use of a unique basis set consisting of orthonormal vectors – vectors of unit length which are perpendicular to one another.

An orthonormal basis set is shown in Fig. 1.22. We will also refer to a basis set as a *coordinate system*. An orthonormal coordinate system is a special case of a Cartesian coordinate system in which the basis vectors are of unit length ($i = j = k = 1$), referred to as *unit vectors*. The **i**, **j**, and **k** vectors are mutually perpendicular and are chosen so that **k** points in the direction of $\mathbf{i} \times \mathbf{j}$. A point p is located in this coordinate system by traversing a distance x_c along **i**, then y_c along a vector parallel to **j**, and z_c along a vector parallel to **k**, exactly as was done in the general coordinate system; its location is determined by the vector $\overrightarrow{\mathbf{op}} = [x_c\ y_c\ z_c]$. We will use the subscript "c" to refer to a *Cartesian coordinate* throughout the text. A vector in this coordinate system is a linear combination of orthonormal basis vectors:

$$\overrightarrow{\mathbf{op}} = x_c\mathbf{i} + y_c\mathbf{j} + z_c\mathbf{k} = [x_c\ y_c\ z_c] \tag{1.21}$$

The multiplicative properties of vectors in an orthonormal coordinate system are unique, allowing for the determination of scalar and vector products from vector components, thus providing a powerful tool for determining distances and angles

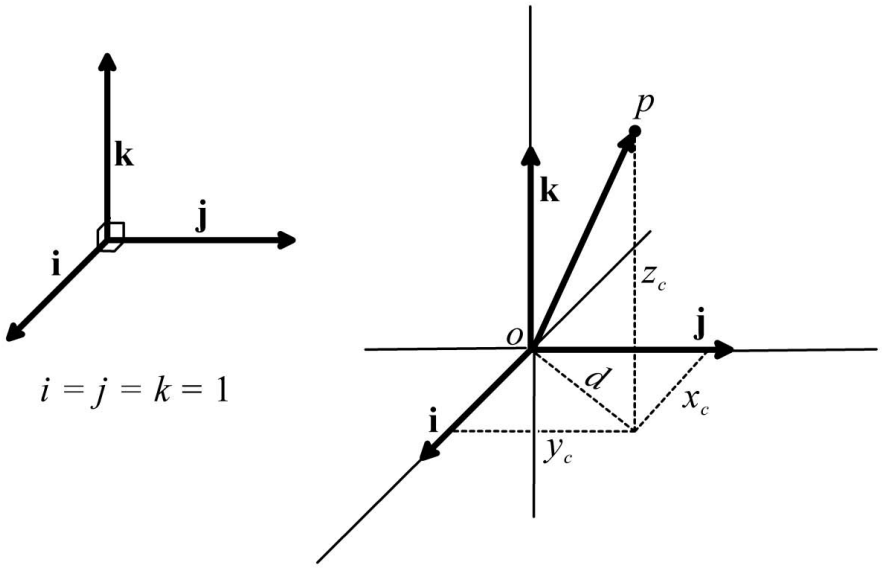

Figure 1.22 Orthonormal basis set — the point p is located at the end of the vector defined by $[x_c\ y_c\ z_c]$.

within the crystal lattice. This uniqueness arises from the multiplicative properties of the basis vectors themselves. The scalar products of **i**, **j**, and **k** are summarized below:

$$\begin{aligned}
\mathbf{i}\cdot\mathbf{i} &= ii\cos(0) = 1 \\
\mathbf{j}\cdot\mathbf{j} &= jj\cos(0) = 1 \qquad (1.22)\\
\mathbf{k}\cdot\mathbf{k} &= kk\cos(0) = 1 \\
\mathbf{i}\cdot\mathbf{j} &= ij\cos(\tfrac{\pi}{2}) = 0 \\
\mathbf{i}\cdot\mathbf{k} &= ik\cos(\tfrac{\pi}{2}) = 0 \qquad (1.23)\\
\mathbf{j}\cdot\mathbf{k} &= jk\cos(\tfrac{\pi}{2}) = 0.
\end{aligned}$$

The vector product of **i**, **j**, or **k** with itself produces a vector of zero length (a point) known as the *null* vector:

$$\begin{aligned}
|\mathbf{i}\times\mathbf{i}| &= ii\sin(0) = 0 \implies \mathbf{i}\times\mathbf{i} = \mathbf{0} \\
|\mathbf{j}\times\mathbf{j}| &= jj\sin(0) = 0 \implies \mathbf{j}\times\mathbf{j} = \mathbf{0} \qquad (1.24)\\
|\mathbf{k}\times\mathbf{k}| &= kk\sin(0) = 0 \implies \mathbf{k}\times\mathbf{k} = \mathbf{0}
\end{aligned}$$

The remaining vector products depend on the order in which the basis vectors are multiplied. We have defined the coordinate system such that the vector product $\mathbf{i}\times\mathbf{j}$ points in the direction of **k**. A coordinate system defined in this manner is termed a *right-handed coordinate system*. Since $|\mathbf{i}\times\mathbf{j}| = ij\sin(\frac{\pi}{2}) = 1$, the vector product of **i** and **j** is *exactly* **k**. Recalling that reversing the order of multiplication changes the sign of the vector product, the remaining vector products are determined in the same manner:

$$\begin{aligned}
\mathbf{i}\times\mathbf{j} &= \mathbf{k} \qquad & \mathbf{j}\times\mathbf{i} &= -\mathbf{k} \\
\mathbf{j}\times\mathbf{k} &= \mathbf{i} \qquad & \mathbf{k}\times\mathbf{j} &= -\mathbf{i} \qquad (1.25)\\
\mathbf{k}\times\mathbf{i} &= \mathbf{j} \qquad & \mathbf{i}\times\mathbf{k} &= -\mathbf{j}
\end{aligned}$$

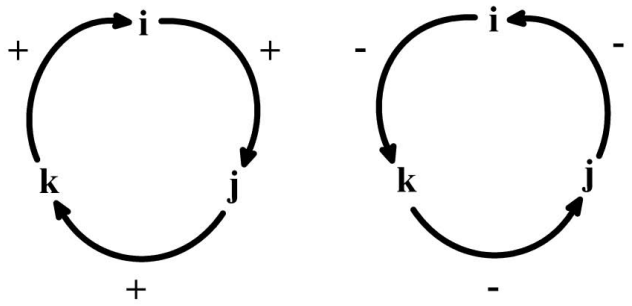

Figure 1.23 Mnemonic device for recalling the vector products of orthonormal basis vectors. Clockwise rotations are positive; counterclockwise rotations are negative. For example, $\mathbf{k}\times\mathbf{i} = \mathbf{j}$ and $\mathbf{j}\times\mathbf{i} = -\mathbf{k}$.

Although these vector products can always be generated by referring to the right-handed coordinate system illustrated in Fig. 1.22, a convenient mnemonic device for remembering them is shown in Fig. 1.23. Using this scheme, clockwise rotations produce positive vector products, counter-clockwise rotations produce negative vector products.

1.3.4 The Scalar Product in an Orthonormal Coordinate System

The right angles of the orthogonal coordinate system make it possible to use the Pythagorean theorem to determine the length of a vector with its tail at the origin from the components of the vector. Referring to Fig. 1.22, the diagonal d is the hypotenuse of a right triangle, as is the vector $\overrightarrow{\mathbf{op}}$:

$$\begin{aligned} d^2 &= x_c^2 + y_c^2 \\ |\overrightarrow{\mathbf{op}}|^2 &= d^2 + z_c^2 = x_c^2 + y_c^2 + z_c^2 \\ |\overrightarrow{\mathbf{op}}| &= (x_c^2 + y_c^2 + z_c^2)^{\frac{1}{2}}. \end{aligned} \tag{1.26}$$

Now, consider the arbitrary vector, $\mathbf{v_1}$ in Fig. 1.24. In order to determine its length we construct two vectors emanating from the origin, $\mathbf{v_2}$ with its head at the tail of $\mathbf{v_1}$, and $\mathbf{v_3}$ with its head at the head of $\mathbf{v_1}$. Clearly $\mathbf{v_3} = \mathbf{v_2} + \mathbf{v_1}$ and $\mathbf{v_1} = \mathbf{v_3} - \mathbf{v_2}$. The process of negating $\mathbf{v_2}$ and adding it to $\mathbf{v_3}$ translates $\mathbf{v_1}$ it to the origin. Indeed, this is just the reverse of the process of vector addition, where $\mathbf{v_1}$ is translated from the origin to the head of the vector $\mathbf{v_2}$ to which it is added. The components of the translated vector $\mathbf{v_1} = [v_{1x}\ v_{1y}\ v_{1z}]$ can therefore provide us with the length of the vector. Let $\mathbf{v_2} = [v_{2x}\ v_{2y}\ v_{2z}]$ and $\mathbf{v_3} = [v_{3x}\ v_{3y}\ v_{3z}]$. Then, $\mathbf{v_1} = [v_{1x}\ v_{1y}\ v_{1z}] = [(v_{3x} - v_{2x})\ (v_{3y} - v_{2y})\ (v_{3z} - v_{2z})]$. The subtraction of the coordinates of $\mathbf{v_2}$ from those of $\mathbf{v_3}$ translates $\mathbf{v_1}$ back to the origin, allowing for the determination of the squares of the lengths of all three vectors:

$$\begin{aligned} v_3^2 &= v_{3x}^2 + v_{3y}^2 + v_{3z}^2 \\ v_2^2 &= v_{2x}^2 + v_{2y}^2 + v_{2z}^2 \\ v_1^2 &= (v_{3x} - v_{2x})^2 + (v_{3y} - v_{2y})^2 + (v_{3z} - v_{2z})^2. \end{aligned} \tag{1.27}$$

The third equation allows us to compute distances between any two points in the lattice, provided that we have the coordinates for the two points in an orthonormal basis. We could, for example, imagine an atom with coordinates $[v_{3x}\ v_{3y}\ v_{3z}]$ and another at $[v_{2x}\ v_{2y}\ v_{2z}]$. The distance between the two atoms can be determined by subtracting the coordinates of the two points, squaring the differences, summing the squares, and taking the square root of the sum to give

$$v_1 = ((v_{3x} - v_{2x})^2 + (v_{3y} - v_{2y})^2 + (v_{3z} - v_{2z})^2)^{\frac{1}{2}}. \tag{1.28}$$

Finally, expanding Eqn. 1.27 provides an expression for the scalar product in terms of the coordinates *in an orthonormal basis*:

$$\begin{aligned} v_1^2 &= v_{3x}^2 - 2v_{3x}v_{2x} + v_{2x}^2 + v_{3y}^2 - 2v_{3y}v_{2y} + v_{2y}^2 + v_{3z}^2 - 2v_{3z}v_{2z} + v_{2z}^2 \\ v_1^2 &= (v_{2x}^2 + v_{2y}^2 + v_{2z}^2) + (v_{3x}^2 + v_{3y}^2 + v_{3z}^2) - 2(v_{2x}v_{3x} + v_{2y}v_{3y} + v_{2z}v_{3z}) \\ v_1^2 &= v_2^2 + v_3^2 - 2(v_{2x}v_{3x} + v_{2y}v_{3y} + v_{2z}v_{3z}). \end{aligned} \tag{1.29}$$

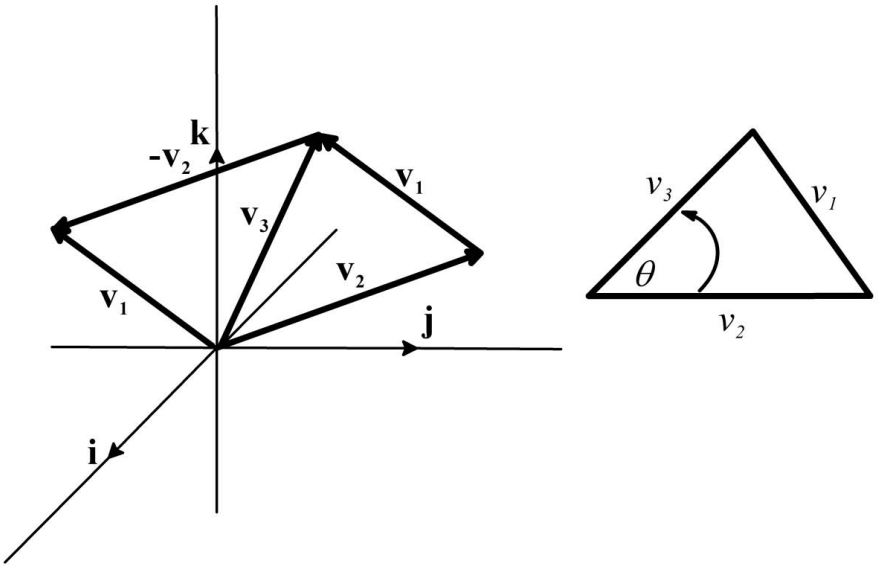

Figure 1.24 Vectors used to determine the length of a general vector in an orthonormal coordinate system. The Law of Cosines determines the length of $\mathbf{v_1}$, $v_1^2 = v_2^2 + v_3^2 - 2v_2v_3\cos\theta$.

Referring to the triangle to the right in Fig. 1.24, θ is the angle between $\mathbf{v_2}$ and $\mathbf{v_3}$. The Law of Cosines produces the square of the length of the third side of a triangle from the lengths of the other two sides and the angle between them. In this case we are interested in the square of the length of $\mathbf{v_1}$: $v_1^2 = v_2^2 + v_3^2 - 2v_2v_3\cos\theta$. Equating these two expressions for v_1^2 gives

$$\begin{aligned} v_2^2 + v_3^2 - 2v_2v_3\cos\theta &= v_2^2 + v_3^2 - 2(v_{2x}v_{3x} + v_{2y}v_{3y} + v_{2z}v_{3z}) \\ v_2v_3\cos\theta &= v_{2x}v_{3x} + v_{2y}v_{3y} + v_{2z}v_{3z}. \end{aligned} \tag{1.30}$$

We have arrived at a very useful result — $v_2v_3\cos\theta$ is the scalar product of $\mathbf{v_2}$ and $\mathbf{v_3}$:

$$\mathbf{v_2}\cdot\mathbf{v_3} = v_{2x}v_{3x} + v_{2y}v_{3y} + v_{2z}v_{3z}. \tag{1.31}$$

In an orthonormal coordinate system the scalar product of two vectors is given by the sum of the products of the components of each vector. The length of a vector with its tail at the origin is now readily determined by taking the scalar product of a vector with itself:

$$\begin{aligned} \mathbf{v_1} &= v_{1x}\mathbf{i} + v_{1y}\mathbf{j} + v_{1z}\mathbf{k} = [v_{1x}\ v_{1y}\ v_{1z}] \\ \mathbf{v_1}\cdot\mathbf{v_1} &= v_{1x}v_{1x} + v_{1y}v_{1y} + v_{1z}v_{1z} \\ v_1v_1\cos(0) &= v_1^2 = v_{1x}^2 + v_{1y}^2 + v_{1z}^2 \\ v_1 &= (v_{1x}^2 + v_{1y}^2 + v_{1z}^2)^{\frac{1}{2}} \\ v_1 &= (\mathbf{v_1}\cdot\mathbf{v_1})^{\frac{1}{2}}. \end{aligned} \tag{1.32}$$

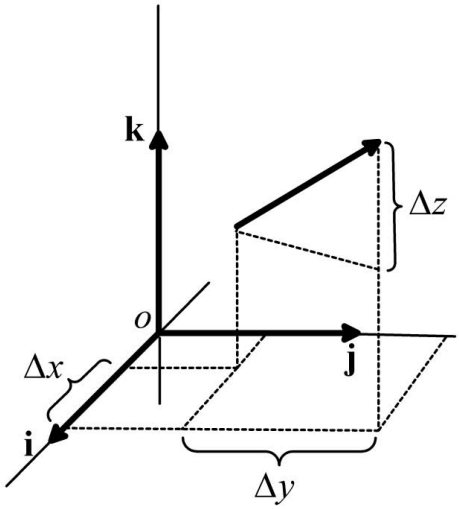

Figure 1.25 Vector in orthonormal coordinates illustrating the changes in the components obtained from subtracting the coordinates of the head of the vector from those of the tail. Translating the vector to the origin places the tail at [0 0 0] and the head at $[\Delta x\ \Delta y\ \Delta z]$.

The scalar product also yields the angle between two vectors emanating from the origin in an orthonormal system:

$$\begin{aligned}
\mathbf{v_1} &= v_{1x}\mathbf{i} + v_{1y}\mathbf{j} + v_{1z}\mathbf{k} = [v_{1x}\ v_{1y}\ v_{1z}] \\
\mathbf{v_2} &= v_{2x}\mathbf{i} + v_{2y}\mathbf{j} + v_{2z}\mathbf{k} = [v_{2x}\ v_{2y}\ v_{2z}] \\
v_1 &= (\mathbf{v_1}\cdot\mathbf{v_1})^{\frac{1}{2}} = (v_{1x}^2 + v_{1y}^2 + v_{1z}^2)^{\frac{1}{2}} \\
v_2 &= (\mathbf{v_2}\cdot\mathbf{v_2})^{\frac{1}{2}} = (v_{2x}^2 + v_{2y}^2 + v_{2z}^2)^{\frac{1}{2}} \\
\mathbf{v_1}\cdot\mathbf{v_2} &= v_1 v_2 \cos\theta = v_{1x}v_{2x} + v_{1y}v_{2y} + v_{1z}v_{2z} \\
\cos\theta &= \frac{\mathbf{v_1}\cdot\mathbf{v_2}}{v_1 v_2} = \frac{v_{1x}v_{2x} + v_{1y}v_{2y} + v_{1z}v_{2z}}{(v_{1x}^2 + v_{1y}^2 + v_{1z}^2)^{\frac{1}{2}}(v_{2x}^2 + v_{2y}^2 + v_{2z}^2)^{\frac{1}{2}}}.
\end{aligned} \tag{1.33}$$

The origin constraint is not as restrictive as it first appears, since we can translate a vector to the origin by subtracting the coordinates of its head and tail, as we did earlier to determine the length of a general vector. Fig. 1.25 provides an alternative way to envision this. Consider the head of the vector to have coordinates $[x_h y_h z_h]$, and the tail to have coordinates $[x_t\ y_t\ z_t]$ (These would be the components of $\mathbf{v_2}$ and $\mathbf{v_3}$ in the previous examples). Subtracting the components of the tail from those of the head gives the *change* in the coordinates along each of the axes: $[(x_h - x_t)\ (y_h - y_t)\ (z_h - z_t)] = [\Delta x\ \Delta y\ \Delta z]$. If we translate the vector by maintaining its magnitude and direction and moving its tail to the origin, the components of the vector will then be $[\Delta x\ \Delta y\ \Delta z]$, and the magnitude of the vector will be $((\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2)^{\frac{1}{2}}$. For two vectors, $\mathbf{v_1}$ and $\mathbf{v_2}$, the scalar product will be obtained by translating both vectors to the origin: $\mathbf{v_1}\cdot\mathbf{v_2} = \Delta x_1\Delta x_2 + \Delta y_1\Delta y_2 + \Delta z_1\Delta z_2$.

The parallel between the scalar product of vectors and the multiplication of ordinary numbers has already been established for the commutative property of multiplication, i.e., the product remains unchanged when the order of multiplication is changed. The associate property of multiplication also applies to ordinary numbers. In this case multiplications are always binary operations between pairs of numbers, and the property tells us that we can create any pairs that we wish to obtain their product: $abc = (ab)c = a(bc)$, etc. The analogous scalar product makes no sense, since the scalar product of any pair of vectors will result in a scalar, making a second scalar product impossible. However there is a close analogy in the multiplication of a scalar and a scalar product:

$$\begin{aligned} \mathbf{v_1} \cdot (s\mathbf{v_2}) &= v_1(sv_2)\cos\theta \\ &= s(v_1 v_2)\cos\theta \;=\; s(\mathbf{v_1} \cdot \mathbf{v_2}) \end{aligned} \tag{1.34}$$

The scalar product is associative with respect to multiplication by a scalar. The multiplication of ordinary numbers is also distributive: $a(b + c) = ab + ac$. The scalar product analogy is $\mathbf{v_1} \cdot (\mathbf{v_2} + \mathbf{v_3})$. In orthonormal coordinates,

$$\begin{aligned} \mathbf{v_1} \cdot (\mathbf{v_2} + \mathbf{v_3}) &= [v_{1x}\ v_{1y}\ v_{1z}] \cdot [(v_{2x} + v_{3x})\ (v_{2y} + v_{3y})\ (v_{2z} + v_{3z})] \\ &= v_{1x}(v_{2x} + v_{3x}) + v_{1y}(v_{2y} + v_{3y}) + v_{1z}(v_{2z} + v_{3z}) \\ &= v_{1x}v_{2x} + v_{1x}v_{3x} + v_{1y}v_{2y} + v_{1y}v_{3y} + v_{1z}v_{2z} + v_{1z}v_{3z} \\ &= (v_{1x}v_{2x} + v_{1y}v_{2y} + v_{1z}v_{2z}) + (v_{1x}v_{3x} + v_{1y}v_{3y} + v_{1z}v_{3z}) \\ &= (\mathbf{v_1} \cdot \mathbf{v_2}) + (\mathbf{v_1} \cdot \mathbf{v_3}). \end{aligned} \tag{1.35}$$

The scalar product is distributive with respect to vector addition. We will find the use of scalar products invaluable in the calculation of interatomic distances and angles inside the crystal lattice.

A Note on Terminology: Scalar/Dot/Inner Product. The terms *scalar product*, *dot product*, and *inner product* are commonly used interchangeably. As demonstrated above, the scalar resulting from $\mathbf{v_1} \cdot \mathbf{v_2} = v_1 v_2 \cos\theta$ is equal to the scalar resulting from $\mathbf{v_1} \cdot \mathbf{v_2} = v_{1x}v_{2x} + v_{1y}v_{2y} + v_{1z}v_{2z}$, only if the components are defined in an orthonormal basis.

In most cases the use of various terms to describe both of these scalars is unambiguous, either because the coordinate system is orthonormal, or because there is no confusion in the context of their use. However, in crystallography, the transformation of vectors between orthonormal and non-orthonormal coordinate systems occurs regularly, and it is important to have a descriptive term for a scalar that does not vary when these transformations occur. We initially referred to this invariant term as *the scalar product*, and we will continue to do so (expanding the concept a bit when discussing reciprocal lattices).

On the other hand, the value of the sum of the products of the components of two vectors does vary with the coordinate system, but we will find this scalar useful in representing the sums of products of the components of vectors in the basis of the crystal lattice with the components of vectors defined in the basis of the lattice of the diffraction pattern. In addition, the scalar provides a convenient generalization of the matrix product of a vector and its transpose, encountered later on in this chapter. In order to differentiate this scalar from the other, we will adopt the convention that the sum of the products of the components is the *inner product*

of the vectors. Whenever there is no need to distinguish between the two entities, the term *dot product* will apply.

1.3.5 The Vector Product in an Orthonormal Coordinate System

The vector product can also be expressed in terms of the coordinates in an orthonormal basis. Fig. 1.26 illustrates the vector product $\mathbf{v_1} = \mathbf{v_2} \times \mathbf{v_3}$ for two vectors with tails at the origin (translated as discussed previously, if necessary).

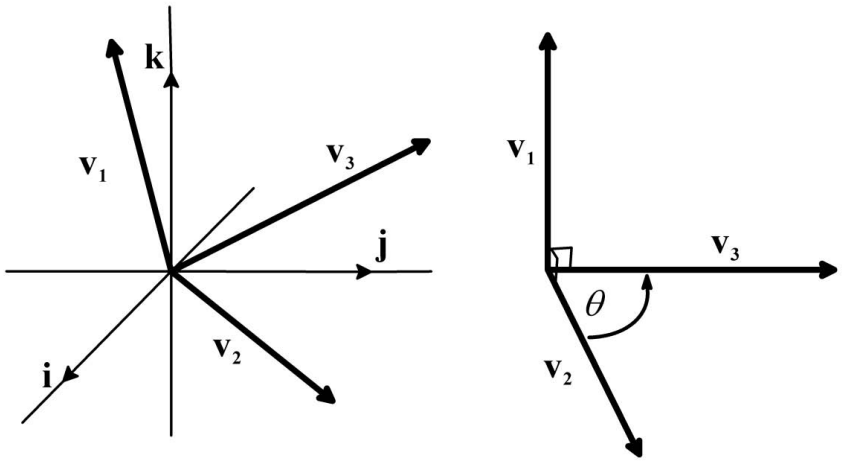

Figure 1.26 Vector product in orthonormal coordinates: $\mathbf{v_1} = \mathbf{v_2} \times \mathbf{v_3}$. $v_1 = v_2 v_3 \sin\theta$ with $\mathbf{v_1} \perp \mathbf{v_2}$ and $\mathbf{v_1} \perp \mathbf{v_3}$.

From the definition of the vector product,

$$\begin{aligned} v_1 &= v_2 v_3 \sin\theta \\ v_1^2 &= v_2^2 v_3^2 \sin^2\theta = v_2^2 v_3^2 (1 - \cos^2\theta) \\ &= v_2^2 v_3^2 - v_2^2 v_3^2 \cos^2\theta = v_2^2 v_3^2 - (v_2 v_3 \cos\theta)^2 = v_2^2 v_3^2 - (\mathbf{v_2} \cdot \mathbf{v_3})^2 \\ &= (v_{2x}^2 + v_{2y}^2 + v_{2z}^2)(v_{3x}^2 + v_{3y}^2 + v_{3z}^2) - (v_{2x} v_{3x} + v_{2y} v_{3y} + v_{2z} v_{3z})^2 \end{aligned}$$

Expansion and collection of the terms on the right gives

$$\begin{aligned} v_1^2 &= (v_{2y}^2 v_{3z}^2 - 2 v_{2y} v_{3y} v_{2z} v_{3z} + v_{2z}^2 v_{3y}^2) \\ &+ (v_{2z}^2 v_{3x}^2 - 2 v_{2z} v_{3z} v_{2x} v_{3x} + v_{2x}^2 v_{3z}^2) \\ &+ (v_{2x}^2 v_{3y}^2 - 2 v_{2x} v_{3x} v_{2y} v_{3y} + v_{2y}^2 v_{3x}^2) \\ v_1^2 &= (v_{2y} v_{3z} - v_{2z} v_{3y})^2 + (v_{2z} v_{3x} - v_{2x} v_{3z})^2 + (v_{2x} v_{3y} - v_{2y} v_{3x})^2. \end{aligned} \quad (1.36)$$

Since the square of the magnitude of $\mathbf{v_1}$ is the sum of the squares of its components, the square roots of the three terms in Eqn. 1.36 are clearly the components of the vector product of $\mathbf{v_2}$ and $\mathbf{v_3}$. Unfortunately, the equation does not tell us which

coordinate axis each component corresponds to. To determine which is which we apply the added constraint that $\mathbf{v_1} \perp \mathbf{v_2}$ and $\mathbf{v_1} \perp \mathbf{v_3} \Longrightarrow \mathbf{v_1} \cdot \mathbf{v_2} = 0$ and $\mathbf{v_1} \cdot \mathbf{v_3} = 0$. In order for the terms in the scalar products to cancel to zero, every number triple in the scalar product sum must contain an x, y, and z contribution. Of the six possible ways of permuting the components, only when $(v_{2y}v_{3z} - v_{2z}v_{3y})$ is multiplied by v_{2x} or v_{3x}, $(v_{2z}v_{3x} - v_{2x}v_{3z})$ is multiplied by v_{2y} or v_{3y}, and $(v_{2x}v_{3y} - v_{2y}v_{3x})$ is multiplied by v_{2z} or v_{3z} will this be the case. We have thus arrived at an expression for the vector product of two vectors in terms of their orthonormal components:

$$\begin{aligned} \mathbf{v_2} \times \mathbf{v_3} &= (v_{2y}v_{3z} - v_{2z}v_{3y})\mathbf{i} + (v_{2z}v_{3x} - v_{2x}v_{3z})\mathbf{j} + (v_{2x}v_{3y} - v_{2y}v_{3x})\mathbf{k} \\ &= [(v_{2y}v_{3z} - v_{2z}v_{3y})\ (v_{2z}v_{3x} - v_{2x}v_{3z})\ (v_{2x}v_{3y} - v_{2y}v_{3x})]. \end{aligned} \tag{1.37}$$

In an orthonormal coordinate system the vector product of two vectors is a vector with components comprised of all the possible pairs of cross terms of the components of the original vectors; the vector product is often called the *cross* product.

It has already been established that the vector product is not commutative. It is instructive to revisit this in terms of the orthonormal components of the vector.

$$\begin{aligned} \mathbf{v_3} \times \mathbf{v_2} &= (v_{3y}v_{2z} - v_{3z}v_{2y})\mathbf{i} + (v_{3z}v_{2x} - v_{3x}v_{2z})\mathbf{j} + (v_{3x}v_{2y} - v_{3y}v_{2x})\mathbf{k} \\ &= -(v_{2y}v_{3z} - v_{2z}v_{3y})\mathbf{i} - (v_{2z}v_{3x} - v_{2x}v_{3z})\mathbf{j} - (v_{2x}v_{3y} - v_{2y}v_{3x})\mathbf{k} \\ &= -(\mathbf{v_2} \times \mathbf{v_3}). \end{aligned} \tag{1.38}$$

As demonstrated for the general coordinate system previously, the sign of the vector has changed, but the magnitude, $v_2v_3 \sin\theta$, remains the same. In addition to being commutative, the *scalar* product is also associative with respect to multiplication by a scalar, and distributive with respect to vector addition. Despite its non-commutative nature, does the vector product share any analogous properties with ordinary numbers? Consider the vector product $\mathbf{v_2} \times s\mathbf{v_3}$:

$$\begin{aligned} \mathbf{v_2} \times s\mathbf{v_3} &= (v_{2y}sv_{3z} - v_{2z}sv_{3y})\mathbf{i} + (v_{2z}sv_{3x} - v_{2x}sv_{3z})\mathbf{j} + (v_{2x}sv_{3y} - v_{2y}sv_{3x})\mathbf{k} \\ &= s(v_{2y}v_{3z} - v_{2z}v_{3y})\mathbf{i} + s(v_{2z}v_{3x} - v_{2x}v_{3z})\mathbf{j} + s(v_{2x}v_{3y} - v_{2y}v_{3x})\mathbf{k} \\ &= s((v_{2y}v_{3z} - v_{2z}v_{3y})\mathbf{i} + (v_{2z}v_{3x} - v_{2x}v_{3z})\mathbf{j} + (v_{2x}v_{3y} - v_{2y}v_{3x}))\mathbf{k} \\ &= s(\mathbf{v_2} \times \mathbf{v_3}). \end{aligned} \tag{1.39}$$

The vector product is associative with respect to scalar multiplication.

The *scalar product* provides for a convenient test for perpendicular vectors. If $\mathbf{v_1} \perp \mathbf{v_2}$ then $\mathbf{v_1} \cdot \mathbf{v_2} = v_1v_2\cos(\pi/2) = 0$. The associative property of the *vector product* allows for a convenient test for parallel vectors. Suppose $\mathbf{v_1} \| \mathbf{v_2} \Longrightarrow \mathbf{v_2} = s\mathbf{v_1}$. Thus $\mathbf{v_1} \times \mathbf{v_2} = \mathbf{v_1} \times s\mathbf{v_1} = s(\mathbf{v_1} \times \mathbf{v_1}) = \mathbf{0}$, the null vector, with components [0 0 0]. *If the scalar product is zero then the vectors are perpendicular. If the vector*

product is $\mathbf{0}$ *then the vectors are parallel.* Finally, consider the vector product of a vector with the sum of two vectors:

$$\begin{aligned}
\mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) &= (v_{1y}(v_{2z} + v_{3z})) - v_{1z}(v_{2y} + v_{3y}))\mathbf{i} \\
&+ (v_{1z}(v_{2x} + v_{3x})) - v_{1x}(v_{2z} + v_{3z}))\mathbf{j} \\
&+ (v_{1x}(v_{2y} + v_{3y})) - v_{1y}(v_{2y} + v_{3x}))\mathbf{k} \\
\mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) &= (v_{1y}v_{2z} + v_{1y}v_{3z} - v_{1z}v_{2y} - v_{1z}v_{3y})\mathbf{i} \\
&+ (v_{1z}v_{2x} + v_{1z}v_{3x} - v_{1x}v_{2z} - v_{1x}v_{3z})\mathbf{j} \\
&+ (v_{1x}v_{2y} + v_{1x}v_{3y} - v_{1y}v_{2y} - v_{1y}v_{3x})\mathbf{k} \\
\mathbf{v_1} \times (\mathbf{v_2} + \mathbf{v_3}) &= \\
(v_{1y}v_{2z} - v_{1z}v_{2y})\mathbf{i} &+ (v_{1z}v_{2x} - v_{1x}v_{2z})\mathbf{j} + (v_{1x}v_{2y} - v_{1y}v_{2y})\mathbf{k} \\
+ (v_{1y}v_{3z} - v_{1z}v_{3y})\mathbf{i} &+ (v_{1z}v_{3x} - v_{1x}v_{3z})\mathbf{j} + (v_{1x}v_{3y} - v_{1y}v_{3x})\mathbf{k} \\
&= (\mathbf{v_1} \times \mathbf{v_2}) + (\mathbf{v_1} \times \mathbf{v_3}).
\end{aligned} \tag{1.40}$$

The vector product is distributive with respect to vector addition.

1.4 Matrices in Crystallography

As emphasized in the previous section, a basis set consisting of the unit cell axes is the natural coordinate system for the treatment of lattice symmetry, the topic of the next chapter, as well as for the mathematics of diffraction, treated in Chapter 3. On the other hand, computing distances and angles requires an orthonormal basis set. It follows that we will need a means to transform vectors expressed in one coordinate system to the other. It is also necessary to transform vectors in order to relocate coordinates when considering lattice symmetry. In general, these vector transformations are accomplished using matrices.

1.4.1 Matrix Definitions

A matrix is a rectangular array of numeric or algebraic *elements* arranged in m rows and n columns:

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{21} & d_{22} & d_{23} & \cdots & d_{2n} \\ d_{31} & d_{32} & d_{33} & \cdots & d_{3n} \\ d_{41} & d_{42} & d_{43} & \cdots & d_{4n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{m1} & d_{m2} & d_{m3} & \cdots & d_{mn} \end{bmatrix} \tag{1.41}$$

We describe the array as an $m \times n$ matrix, where the matrix element d_{ij} is the element at the intersection of the ith row and the jth column; m and n are called the *row dimension* and *column dimension*, respectively, of the matrix. In general, a boldface capital letter will be used to indicate a matrix, except in the special cases when $m = 1$ or $n = 1$.

When $m = 1$ or $n = 1$ the resulting matrices *are* vectors. When $m = 1$ the vector takes the form of a single row which we refer to as a *row vector*; when $n = 1$ the matrix takes the form of a single column referred to as a *column vector*:

$$\text{row vector } \mathbf{v} = \begin{bmatrix} v_1 & v_2 & v_3 & \cdots & v_n \end{bmatrix} \qquad \text{column vector } \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ \vdots \\ v_m \end{bmatrix}$$

It is often useful to express an $m \times n$ matrix as a matrix of n column vectors, each with m components:

$$\mathbf{D} = \begin{bmatrix} \mathbf{d_1} & \mathbf{d_2} & \mathbf{d_3} & \cdots & \mathbf{d_n} \end{bmatrix}, \qquad \mathbf{d_i} = \begin{bmatrix} d_{1i} \\ d_{2i} \\ d_{3i} \\ d_{4i} \\ \vdots \\ d_{mi} \end{bmatrix} \tag{1.42}$$

The *column rank* of the matrix is defined as the number of these vectors which are linearly independent (i.e., cannot be written as linear combinations of other column vectors in the matrix). The *row rank* of the matrix is defined as the number of row vectors in the matrix which are linearly independent. We will generally find that the rank of the matrices in crystallography will be $n = m = 3$. When $n = m$ the matrix is refereed to as a *square matrix*.

The *transpose* of an $m \times n$ matrix is generated by switching its rows and columns, creating an $n \times m$ matrix:

$$\mathbf{D}^T = \begin{bmatrix} d_{11} & d_{21} & d_{31} & \cdots & d_{m1} \\ d_{12} & d_{22} & d_{32} & \cdots & d_{m2} \\ d_{13} & d_{23} & d_{33} & \cdots & d_{m3} \\ d_{14} & d_{24} & d_{34} & \cdots & d_{m4} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{1n} & d_{2n} & d_{3n} & \cdots & d_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{d_1}^T \\ \mathbf{d_2}^T \\ \mathbf{d_3}^T \\ \mathbf{d_4}^T \\ \cdots \\ \mathbf{d_n}^T \end{bmatrix}, \mathbf{d_i}^T = \begin{bmatrix} d_{1i} & d_{2i} & d_{3i} & d_{4i} & \cdots & d_{mi} \end{bmatrix} \tag{1.43}$$

Note that the column vectors in Eqn. 1.42 and the row vectors in Eqn. 1.43 are transposes of one another. Transposing a matrix switches d_{ij} and d_{ji}. A square matrix which remains unchanged when it is transposed (i.e. $d_{ij} = d_{ji}$ and $\mathbf{D} = \mathbf{D}^T$) appears identical above and below the diagonal elements of the matrix and is termed a *symmetric* matrix.

1.4.2 Matrix Operations

The product of a scalar and a matrix multiplies every element in the matrix by the scalar:

$$s\mathbf{D} = s\begin{bmatrix} d_{11} & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{21} & d_{22} & d_{23} & \cdots & d_{2n} \\ d_{31} & d_{32} & d_{33} & \cdots & d_{3n} \\ d_{41} & d_{42} & d_{43} & \cdots & d_{4n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{m1} & d_{m2} & d_{m3} & \cdots & d_{mn} \end{bmatrix} = \begin{bmatrix} sd_{11} & sd_{12} & sd_{13} & \cdots & sd_{1n} \\ sd_{21} & sd_{22} & sd_{23} & \cdots & sd_{2n} \\ sd_{31} & sd_{32} & sd_{33} & \cdots & sd_{3n} \\ sd_{41} & sd_{42} & sd_{43} & \cdots & sd_{4n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ sd_{m1} & sd_{m2} & sd_{m3} & \cdots & sd_{mn} \end{bmatrix} \tag{1.44}$$

The multiplication of a matrix and a scalar is a commutative process, since $sd_{ij} = d_{ij}s$.

The sum of two matrices is obtained by adding corresponding matrix elements:

$$\begin{aligned} \mathbf{D} + \mathbf{E} &= \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ d_{31} & d_{32} & \cdots & d_{3n} \\ d_{41} & d_{42} & \cdots & d_{4n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{m1} & d_{m2} & \cdots & d_{mn} \end{bmatrix} + \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ e_{31} & e_{32} & \cdots & e_{3n} \\ e_{41} & e_{42} & \cdots & e_{4n} \\ \vdots & \vdots & \vdots & \vdots \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{bmatrix} \\ &= \begin{bmatrix} d_{11}+e_{11} & d_{12}+e_{12} & \cdots & d_{1n}+e_{1n} \\ d_{21}+e_{21} & d_{22}+e_{22} & \cdots & d_{2n}+e_{2n} \\ d_{31}+e_{31} & d_{32}+e_{32} & \cdots & d_{3n}+e_{3n} \\ d_{41}+e_{41} & d_{42}+e_{42} & \cdots & d_{4n}+e_{4n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{m1}+e_{m1} & d_{m2}+e_{m2} & \cdots & d_{mn}+e_{mn} \end{bmatrix} \end{aligned} \tag{1.45}$$

Because corresponding matrix elements are added the matrices must have identical row and column dimensions. Matrix addition is also commutative, since $d_{ij} + e_{ij} = e_{ij} + d_{ij}$.

Unlike matrix-scalar multiplication and matrix addition, multiplying matrices with one another has no simple numerical analogy, and is a bit more complicated and abstract. However, vectors are transformed by matrix multiplication and we will find the ability to create *matrix products* to be critical to the development of the mathematics of crystallography. As we shall soon observe, matrix products are not commutative, and we must keep track of the matrix "on the left" and the one "on the right". For the product **DE**, we say that **D** *premultiplies* **E** and that **E** *postmultiplies* **D**. In order to form the matrix product $\mathbf{F} = \mathbf{DE}$ the number of columns in **D** must equal the number of rows in **E**. For **D** $m \times q$ and **E** $q \times n$, the resulting product matrix **F** will be an $m \times n$ matrix with its elements defined by $f_{ij} = \sum_{k=1}^{q} d_{ik}e_{kj}$. The ij^{th} element in the product matrix is formed from the ith row of the matrix on the left and the jth column on the right by multiplying the first element in row i of **D** by the first element in column j of **E**, then multiplying the second elements, then the third ... and so forth until there are q such products. These products are then added together to form the ij^{th} element in **F**. While this may seem a bit confusing, this is precisely the process that was undertaken in the formation of the scalar product, which, for a general coordinate system becomes

the *inner* product (see the previous section for a discussion of this terminology). The ij^{th} element of the product matrix is just the inner product of the ith row vector of the matrix on the left and the jth column vector of the matrix on the right, both of which contain q components. Thus,

$$\mathbf{DE} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1q} \\ d_{21} & d_{22} & \cdots & d_{2q} \\ d_{31} & d_{32} & \cdots & d_{3q} \\ d_{41} & d_{42} & \cdots & d_{4q} \\ \vdots & \vdots & \vdots & \vdots \\ d_{m1} & d_{m2} & \cdots & d_{mq} \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ e_{31} & e_{32} & \cdots & e_{3n} \\ e_{41} & e_{42} & \cdots & e_{4n} \\ \vdots & \vdots & \vdots & \vdots \\ e_{q1} & e_{q2} & \cdots & e_{qn} \end{bmatrix} \tag{1.46}$$

$$= \begin{bmatrix} (\sum_{k=1}^{q} d_{1k}e_{k1}) & (\sum_{k=1}^{q} d_{1k}e_{k2}) & \cdots & (\sum_{k=1}^{q} d_{1k}e_{kn}) \\ (\sum_{k=1}^{q} d_{2k}e_{k1}) & (\sum_{k=1}^{q} d_{2k}e_{k2}) & \cdots & (\sum_{k=1}^{q} d_{2k}e_{kn}) \\ (\sum_{k=1}^{q} d_{3k}e_{k1}) & (\sum_{k=1}^{q} d_{3k}e_{k2}) & \cdots & (\sum_{k=1}^{q} d_{3k}e_{kn}) \\ (\sum_{k=1}^{q} d_{4k}e_{k1}) & (\sum_{k=1}^{q} d_{4k}e_{k2}) & \cdots & (\sum_{k=1}^{q} d_{4k}e_{kn}) \\ \vdots & \vdots & \vdots & \vdots \\ (\sum_{k=1}^{q} d_{mk}e_{k1}) & (\sum_{k=1}^{q} d_{mk}e_{k2}) & \cdots & (\sum_{k=1}^{q} d_{mk}e_{kn}) \end{bmatrix}.$$

A simpler 3×3 example should serve to clarify this:

$$\mathbf{DE} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} =$$

$$\begin{bmatrix} (d_{11}e_{11} + d_{12}e_{21} + d_{13}e_{31}) & (d_{11}e_{12} + d_{12}e_{22} + d_{13}e_{32}) & (d_{11}e_{13} + d_{12}e_{23} + d_{13}e_{33}) \\ (d_{21}e_{11} + d_{22}e_{21} + d_{23}e_{31}) & (d_{21}e_{12} + d_{22}e_{22} + d_{23}e_{32}) & (d_{21}e_{13} + d_{22}e_{23} + d_{23}e_{33}) \\ (d_{31}e_{11} + d_{32}e_{21} + d_{33}e_{31}) & (d_{31}e_{12} + d_{32}e_{22} + d_{33}e_{32}) & (d_{31}e_{13} + d_{32}e_{23} + d_{33}e_{33}) \end{bmatrix}.$$

For example,

$$\begin{aligned} \mathbf{DE} &= \begin{bmatrix} -1 & -2 & 3 \\ 2 & -1 & 2 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 3 & 2 & 0 \\ 2 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} (-1 \cdot 1 - 2 \cdot 3 + 3 \cdot 2) & (-1 \cdot 1 - 2 \cdot 2 + 3 \cdot 1) & (-1 \cdot 2 - 2 \cdot 0 + 3 \cdot 2) \\ (2 \cdot 1 - 1 \cdot 3 + 2 \cdot 2) & (2 \cdot 1 - 1 \cdot 2 + 2 \cdot 1) & (2 \cdot 2 - 1 \cdot 0 + 2 \cdot 2) \\ (0 \cdot 1 + 1 \cdot 3 - 2 \cdot 2) & (0 \cdot 1 + 1 \cdot 2 - 2 \cdot 1) & (0 \cdot 2 + 1 \cdot 0 - 2 \cdot 2) \end{bmatrix} \\ &= \begin{bmatrix} -1 & -2 & 4 \\ 3 & 2 & 8 \\ -1 & 0 & -4 \end{bmatrix}. \end{aligned}$$

Finally, consider two column vectors with the same dimension, $\mathbf{v_a}$ and $\mathbf{v_b}$. Transposing $\mathbf{v_a}$ and premultiplying $\mathbf{v_b}$ $(n \times 1)$ by $\mathbf{v_a}^T$ $(1 \times n)$ results in a 1×1 matrix — a scalar. Thus the inner product, $\mathbf{v_a} \cdot \mathbf{v_b}$, can be represented as a matrix product, $\mathbf{v_a}^T\mathbf{v_b}$:

$$\mathbf{v_a}^T\mathbf{v_b} = \begin{bmatrix} v_{a1} & v_{a2} & \cdots & v_{an} \end{bmatrix} \begin{bmatrix} v_{b1} \\ v_{b2} \\ \vdots \\ v_{bn} \end{bmatrix} = (v_{a1}v_{b1} + v_{a2}v_{b2} + \cdots + v_{an}v_{bn}). \tag{1.47}$$

This provides a very convenient description of matrix multiplication. Expressing $\mathbf{D}$ as a column vector consisting of m row vectors of dimension q, and $\mathbf{E}$ as a row vector consisting of n column vectors of dimension q,

$$\mathbf{DE} = \begin{bmatrix} \mathbf{d_1}^T \\ \mathbf{d_2}^T \\ \mathbf{d_3}^T \\ \vdots \\ \mathbf{d_m}^T \end{bmatrix} \begin{bmatrix} \mathbf{e_1} & \mathbf{e_2} & \mathbf{e_3} & \cdots & \mathbf{e_n} \end{bmatrix} = \begin{bmatrix} \mathbf{d_1}^T\mathbf{e_1} & \mathbf{d_1}^T\mathbf{e_2} & \mathbf{d_1}^T\mathbf{e_3} & \cdots & \mathbf{d_1}^T\mathbf{e_n} \\ \mathbf{d_2}^T\mathbf{e_1} & \mathbf{d_2}^T\mathbf{e_2} & \mathbf{d_2}^T\mathbf{e_3} & \cdots & \mathbf{d_2}^T\mathbf{e_n} \\ \mathbf{d_3}^T\mathbf{e_1} & \mathbf{d_3}^T\mathbf{e_2} & \mathbf{d_3}^T\mathbf{e_3} & \cdots & \mathbf{d_3}^T\mathbf{e_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{d_m}^T\mathbf{e_1} & \mathbf{d_m}^T\mathbf{e_2} & \mathbf{d_m}^T\mathbf{e_3} & \cdots & \mathbf{d_m}^T\mathbf{e_n} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{d_1}\cdot\mathbf{e_1} & \mathbf{d_1}\cdot\mathbf{e_2} & \mathbf{d_1}\cdot\mathbf{e_3} & \cdots & \mathbf{d_1}\cdot\mathbf{e_n} \\ \mathbf{d_2}\cdot\mathbf{e_1} & \mathbf{d_2}\cdot\mathbf{e_2} & \mathbf{d_2}\cdot\mathbf{e_3} & \cdots & \mathbf{d_2}\cdot\mathbf{e_n} \\ \mathbf{d_3}\cdot\mathbf{e_1} & \mathbf{d_3}\cdot\mathbf{e_2} & \mathbf{d_3}\cdot\mathbf{e_3} & \cdots & \mathbf{d_3}\cdot\mathbf{e_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{d_m}\cdot\mathbf{e_1} & \mathbf{d_m}\cdot\mathbf{e_2} & \mathbf{d_m}\cdot\mathbf{e_3} & \cdots & \mathbf{d_m}\cdot\mathbf{e_n} \end{bmatrix}. \tag{1.48}$$

1.4.3 Matrix Transformations

Premultipling an $n \times 1$ column vector by an $m \times n$ matrix produces a new $m \times 1$ column vector:

$$\mathbf{Dv} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{m1} & d_{m2} & \cdots & d_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} (d_{11}v_1 + d_{12}v_2 + \cdots + d_{1n}v_n) \\ (d_{21}v_1 + d_{22}v_2 + \cdots + d_{2n}v_n) \\ \vdots \\ (d_{m1}v_1 + d_{m2}v_2 + \cdots + d_{mn}v_n) \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{d_1}^T \\ \mathbf{d_2}^T \\ \vdots \\ \mathbf{d_m}^T \end{bmatrix} \mathbf{v} = \begin{bmatrix} \mathbf{d_1}^T\mathbf{v} \\ \mathbf{d_2}^T\mathbf{v} \\ \vdots \\ \mathbf{d_m}^T\mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{d_1}\cdot\mathbf{v} \\ \mathbf{d_2}\cdot\mathbf{v} \\ \vdots \\ \mathbf{d_m}\cdot\mathbf{v} \end{bmatrix}. \tag{1.49}$$

Whenever the matrix is a square matrix, the dimension of the vector does not change and we say that we have *transformed* the original vector into a new one. In crystallography the square *transformation matrix* is nearly always a 3×3 matrix. By selecting appropriate elements for the transformation matrix we can perform specific *operations* on the vector to transform it in a specific manner. Transformation matrices are often called *matrix operators*. For example, consider the vector at point $(x_c\ y_c\ z_c)$ in an orthonormal basis in Fig. 1.27(a). Suppose that we wish to determine the coordinates of the point after it has been *reflected* across the xy (**ij**) plane. To accomplish this geometrically we project the head of the vector onto the plane, then extend it by the length of the perpendicular to the other side of the plane. The matrix equation that will accomplish the same task has the following form:

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \\ -z_c \end{bmatrix}.$$

The x component in the new vector remains unchanged: $t_{11}x_c + t_{12}y_c + t_{13}z_c = x_c$. Thus $t_{11} = 1$, $t_{12} = 0$, and $t_{13} = 0$. The y coordinate also does not change. It

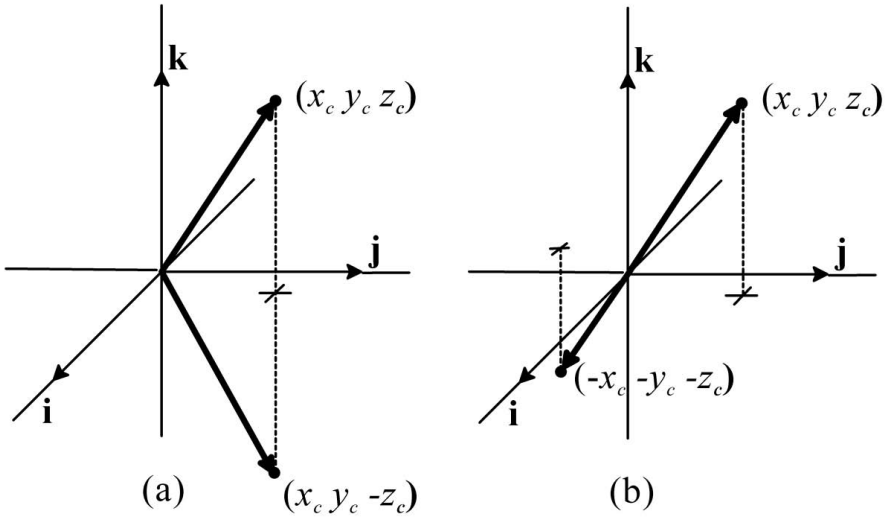

Figure 1.27 (a) Reflection of a vector at point $(x_c\ y_c\ z_c)$ across the xy plane. (b) Inversion of the vector $[x_c\ y_c\ z_c]$ through the origin.

follows that $t_{21} = 0$, $t_{22} = 1$, and $t_{23} = 0$. The z component changes in sign with no contributions from the original x and y components. Hence $t_{31} = 0$, $t_{32} = 0$, and $t_{33} = -1$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \\ -z_c \end{bmatrix}. \tag{1.50}$$

The transformation matrices for reflections across the xz and yz planes are, respectively,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.51}$$

Operations such as *reflections* are necessary to describe lattice symmetry. Fig. 1.27(b) illustrates another important *symmetry operation* — the *inversion* of a vector through the origin. This transformation changes the sign of *all* three components, and therefore results in the following matrix equation:

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} -x_c \\ -y_c \\ -z_c \end{bmatrix}. \tag{1.52}$$

In addition to reflections and inversions, a vector can be rotated by fixing its tail at the origin, maintaining its length, and moving its head to another position by rotating the vector around an axis. The *rotation* operation often takes place around one of the coordinate axes. Fig. 1.28(a) illustrates the rotation of a vector at point

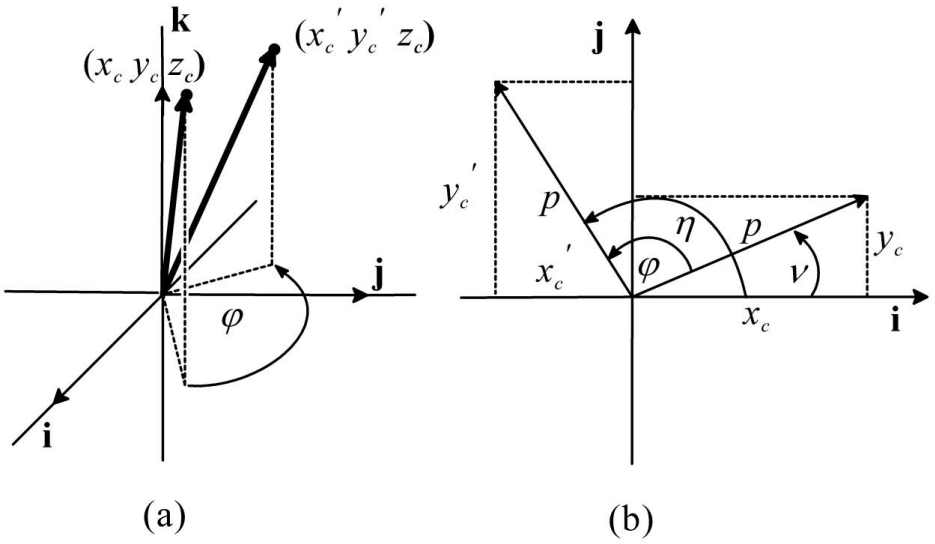

Figure 1.28 (a) Rotation of a vector at point $(x_c\ y_c\ z_c)$ through angle φ around the z axis. (b) Projection of the vector rotation onto the xy plane "looking down" the z axis. p is the length of the projected vector.

$(x_c\ y_c\ z_c)$ rotated around the z axis through an angle φ. The z coordinate remains constant in the rotation. Fig. 1.28(b) is the projection of the vector before and after rotation onto the xy plane, viewing down the z axis in the negative direction. Because we are operating in a right-handed coordinate system, pointing the thumb in the direction of the rotation axis, **k** in this case, assigns a positive rotation in the direction of the curl of the right hand fingers, counterclockwise when "looking down" the axis. The vector **p** is the rotated vector projected onto the plane, with magnitude p. The projected vector is rotated in the xy plane through the angle φ. If ν is the original angle of the projected vector with respect to the x axis, and η is the final angle, then

$$\begin{aligned} x_c &= p\cos\nu \\ y_c &= p\sin\nu \\ x_c' &= p\cos\eta = p\cos(\nu+\varphi) \\ y_c' &= p\sin\eta = p\sin(\nu+\varphi). \\ z_c' &= z_c \end{aligned}$$

From the trigonometric identities for angle sums,

$$\begin{aligned} x_c' &= p\cos\nu\cos\varphi - p\sin\nu\sin\varphi = x_c\cos\varphi - y_c\sin\varphi \\ y_c' &= p\cos\nu\sin\varphi + p\sin\nu\cos\varphi = x_c\sin\varphi + y_c\cos\varphi. \end{aligned}$$

The matrix equation to rotate a vector at $(x_c\ y_c\ z_c)$ through angle φ around the z axis is then

$$\begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} x_c' \\ y_c' \\ z_c \end{bmatrix}. \tag{1.53}$$

Rotation around the x axis produces essentially the same equations, this time transforming the y and z coordinates in exactly the same manner. Rotation about y, however, is positive when the x coordinates are reversed (see Fig. 1.29). In order to keep the signs of the trigonometric functions consistent in each of the quadrants the rotation must occur in the opposite direction; φ must be replaced by $-\varphi$ in the derivation. The resulting transformation matrices for x-rotation and y-rotation are therefore, respectively,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix} \text{ and } \begin{bmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix}. \tag{1.54}$$

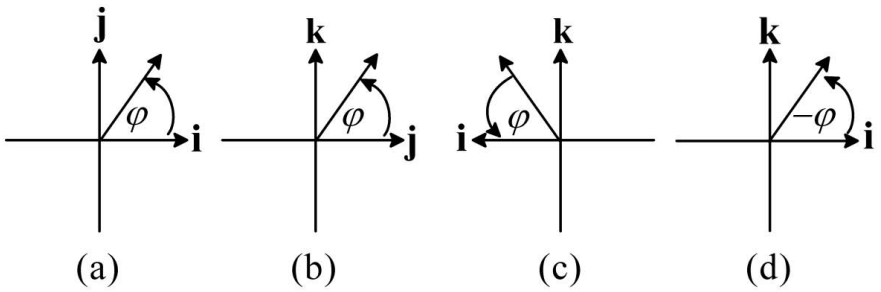

Figure 1.29 Views down the axes showing the signs of the rotational angle: (a)z axis, (b)x axis, and (c) y axis. (d) Inverted axes for the y-rotation showing the effect on the sign of the rotational angle.

It is important to note that each of the matrix operations described in this section maintained the magnitude of the vector transformed by them. If the columns of these matrices are treated as vectors, it can readily be shown that these vectors are orthogonal to one another. They also have unit magnitudes, and are therefore orthonormal vectors. A matrix with orthonormal columns is known as an *orthonormal matrix*. For example, consider the z-rotation matrix:

$$\mathbf{v_1} = \begin{bmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{bmatrix} \qquad \mathbf{v_2} = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{bmatrix} \qquad \mathbf{v_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{1.55}$$

$$
\begin{aligned}
v_1 &= (\cos^2\varphi + \sin^2\varphi + 0^2)^{\frac{1}{2}} = 1 \\
v_2 &= (\sin^2\varphi + \cos^2\varphi + 0^2)^{\frac{1}{2}} = 1 \\
v_3 &= (0^2 + 0^2 + 1^2)^{\frac{1}{2}} = 1 \\
\mathbf{v_1}\cdot\mathbf{v_2} &= (-\cos\varphi\sin\varphi + \cos\varphi\sin\varphi + 0) = 0 \\
\mathbf{v_1}\cdot\mathbf{v_3} &= (0 + 0 + 0) = 0 \\
\mathbf{v_2}\cdot\mathbf{v_3} &= (0 + 0 + 0) = 0
\end{aligned}
$$

Now, consider the transformation of a vector $\mathbf{v}$ by a general orthonormal transformation matrix, $\mathbf{T} = [\mathbf{t_1}\ \mathbf{t_2}\ \mathbf{t_3}]$:

$$
\mathbf{Tv} = \mathbf{v}' = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} x_c' \\ y_c' \\ z_c' \end{bmatrix} = \begin{bmatrix} t_{11}x_c + t_{12}y_c + t_{13}z_c \\ t_{21}x_c + t_{22}y_c + t_{23}z_c \\ t_{31}x_c + t_{32}y_c + t_{33}z_c \end{bmatrix}
$$

$$
\begin{aligned}
v^2 &= x_c^2 + y_c^2 + z_c^2 \\
v'^2 &= (t_{11}x_c + t_{12}y_c + t_{13}z_c)^2 + (t_{21}x_c + t_{22}y_c + t_{23}z_c)^2 + (t_{31}x_c + t_{32}y_c + t_{33}z_c)^2 \\
&= (t_{11}^2 + t_{21}^2 + t_{31}^2)x_c^2 + (t_{12}^2 + t_{22}^2 + t_{32}^2)y_c^2 + (t_{13}^2 + t_{23}^2 + t_{33}^2)z_c^2 \\
&\quad + 2(t_{11}t_{12} + t_{21}t_{22} + t_{31}t_{32})x_cy_c \\
&\quad + 2(t_{11}t_{13} + t_{21}t_{23} + t_{31}t_{33})x_cz_c \\
&\quad + 2(t_{12}t_{13} + t_{22}t_{23} + t_{32}t_{33})y_cz_c \\
&= (\mathbf{t_1}\cdot\mathbf{t_1})x_c^2 + (\mathbf{t_2}\cdot\mathbf{t_2})y_c^2 + (\mathbf{t_3}\cdot\mathbf{t_3})z_c^2 \\
&\quad + (\mathbf{t_1}\cdot\mathbf{t_2})x_cy_c + (\mathbf{t_1}\cdot\mathbf{t_3})x_cz_c + (\mathbf{t_2}\cdot\mathbf{t_3})y_cz_c \\
&= (1)x_c^2 + (1)y_c^2 + (1)z_c^2 + 0 + 0 + 0 = x_c^2 + y_c^2 + z_c^2 = v^2. \qquad (1.56)
\end{aligned}
$$

When an orthonormal matrix operates on a vector it changes the direction of the vector but leaves its magnitude unchanged.

1.4.4 The Determinant of a Matrix

The *determinant* is a scalar that results from the signed products of the permutations of all the matrix elements in a matrix. Adding these products together produces a number with properties which tell us something about the matrix (hence the name). The determinant of a matrix is represented by placing the symbol of the matrix between vertical bars. For an $n \times n$ square matrix, $\mathbf{D}$,

$$
|\mathbf{D}| = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix} = \sum (-1)^k d_{1i_1} d_{2i_2} d_{3i_3} \cdots d_{ni_n}. \qquad (1.57)
$$

The sum is over all products of n-fold permutations of the indices. $i_1, i_2, \cdots i_n$, is a permutation of the indices 1,2,$\cdots n$; k is the number of inversions ("switches") of

indices necessary to achieve the permutation, e.g., $123 \xrightarrow{1} 132 \xrightarrow{2} 312 \xrightarrow{3} 321,\ k = 3$. While this seems abstract, the example for $n = 3$ illustrates that it is only tedious:

$$\begin{aligned} |\mathbf{D}| &= \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix} \\ &= (-1)^0 d_{11}d_{22}d_{33} + (-1)^1 d_{11}d_{23}d_{32} + (-1)^1 d_{12}d_{21}d_{33} \\ &\quad +(-1)^2 d_{12}d_{23}d_{31} + (-1)^2 d_{13}d_{21}d_{32} + (-1)^3 d_{13}d_{22}d_{31}. \end{aligned} \tag{1.58}$$

Four important properties of the determinant arise directly from the definition. Rather than cover the relatively abstract general proofs, the discussion here will be limited to the $n = 3$ case, although in most instances the extrapolation to larger matrices is a logical one.

Determinant Property 1. *Switching two columns or rows in a matrix changes the sign of its determinant.*

$$\begin{aligned} |\mathbf{D}'| &= \begin{vmatrix} d_{11} & d_{13} & d_{12} \\ d_{21} & d_{23} & d_{22} \\ d_{31} & d_{33} & d_{32} \end{vmatrix} \\ &= (-1)^0 d_{11}d_{23}d_{32} + (-1)^1 d_{11}d_{22}d_{33} + (-1)^1 d_{13}d_{21}d_{32} \\ &\quad +(-1)^2 d_{13}d_{22}d_{31} + (-1)^2 d_{12}d_{21}d_{33} + (-1)^3 d_{12}d_{23}d_{31} \\ &= -|\mathbf{D}|. \end{aligned} \tag{1.59}$$

The permutations are the same, but the number of inversions for each has changed. In every case $(-1)^k$ has changed sign; the magnitude of the determinant has not changed, but the sign of the determinant has. *The sign of the determinant changes if a row or column in the matrix is switched with another.*

Determinant Property 2. *The determinant of a matrix that contains column or row vectors that are linearly dependent (not linearly independent) is exactly zero.* Consider the case where two columns (or rows) of the matrix are linearly dependent, that is, one column is a scalar multiple of the other (If the columns are treated as geometric vectors, the vectors would be parallel to one another):

$$\begin{aligned} |\mathbf{D}''| &= \begin{vmatrix} d_{11} & d_{12} & sd_{12} \\ d_{21} & d_{22} & sd_{22} \\ d_{31} & d_{32} & sd_{32} \end{vmatrix} \\ &= (-1)^0 d_{11}d_{22}sd_{32} + (-1)^1 d_{11}sd_{22}d_{32} + (-1)^1 d_{12}d_{21}sd_{32} \\ &\quad +(-1)^2 d_{12}sd_{22}d_{31} + (-1)^2 sd_{12}d_{21}d_{32} + (-1)^3 sd_{12}d_{22}d_{31} \\ &= s\{(-1)^0 d_{11}d_{22}d_{32} + (-1)^1 d_{11}d_{22}d_{32} + (-1)^1 d_{12}d_{21}d_{32} \\ &\quad +(-1)^2 d_{12}d_{22}d_{31} + (-1)^2 d_{12}d_{21}d_{32} + (-1)^3 d_{12}d_{22}d_{31}\} \\ &= s \begin{vmatrix} d_{11} & d_{12} & d_{12} \\ d_{21} & d_{22} & d_{22} \\ d_{31} & d_{32} & d_{32} \end{vmatrix} \end{aligned} \tag{1.60}$$

If we switch the second and thirds columns of the matrix the determinant clearly remains unchanged, yet, according to Eqn. 1.59 it must change sign. This is only possible if the determinant is zero:

$$|\mathbf{D}''| = -|\mathbf{D}''| = 0 \tag{1.61}$$

A matrix with a zero determinant is called a *singular* matrix. Eqn. 1.60 also illustrates another property of the determinant: *Multiplying a column or row of a determinant by a scalar multiplies the determinant by the scalar.*

Determinant Property 3. *The determinant of a matrix and its transpose are equal.* While this might appear intuitive, it is not obvious if one considers the definition of the determinant as a sum of permutations. The proof of this in the general case is based on arguments of the equivalency of various permutations, and would take us far afield. Fortunately, since we will only deal with 3×3 determinants, this can be demonstrated explicitly. Let $\mathbf{E} = \mathbf{D}^T$ and $e_{ij} = d_{ji}$. Then

$$\begin{aligned}
|\mathbf{D}| &= \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{n2} & d_{33} \end{vmatrix} = \sum (-1)^k d_{1i_1} d_{2i_2} d_{3i_3} \\
|\mathbf{E}| &= \begin{vmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} = \sum (-1)^k e_{1i_1} e_{2i_2} e_{3i_3} \\
&= \sum (-1)^k d_{i_1 1} d_{i_2 2} d_{i_3 3}
\end{aligned}$$

Since each term in the sum has one index for each row and column there will be a permutation in $|\mathbf{E}|$ matching every permutation in $|\mathbf{D}|$. Thus the only possible difference is in the signs of the matching permutations. The following table demonstrates that the signs do not change:

k	$\|\mathbf{D}\|$	$\|\mathbf{E}\|$	$\|\mathbf{D}^T\|$
0	$d_{11}d_{22}d_{33}$	$e_{11}e_{22}e_{33}$	$d_{11}d_{22}d_{33}$
1	$d_{11}d_{23}d_{32}$	$e_{11}e_{23}e_{32}$	$d_{11}d_{32}d_{23}$
1	$d_{12}d_{21}d_{33}$	$e_{12}e_{21}e_{33}$	$d_{21}d_{12}d_{33}$
2	$d_{12}d_{23}d_{31}$	$e_{12}e_{23}e_{31}$	$d_{21}d_{32}d_{13}$
2	$d_{13}d_{21}d_{32}$	$e_{13}e_{21}e_{32}$	$d_{31}d_{12}d_{23}$
2	$d_{13}d_{22}d_{31}$	$e_{13}e_{22}e_{31}$	$d_{31}d_{22}d_{13}$

Since every permutation in $|\mathbf{D}^T|$ has the same sign as the corresponding permutation in $|\mathbf{D}|$,

$$|\mathbf{D}^T| = |\mathbf{D}| \tag{1.62}$$

Determinant Property 4. *The determinant of the product of two matrices is the product of the determinants of the matrices.* This is another determinant property

that appears deceptively intuitive on first glance. The proof demonstrates that it is not! Let $\mathbf{P} = \mathbf{DE}$.

$$\begin{aligned}
|\mathbf{P}| &= |\mathbf{DE}| = \sum (-1)^k p_{1i_1} p_{2i_2} p_{3i_3}, \text{ where} \\
p_{1i_1} &= \sum_{j_1}^{3} d_{1j_1} e_{j_1 1},\ p_{2i_2} = \sum_{j_2}^{3} d_{2j_2} e_{j_2 2}, \text{ and } p_{3i_3} = \sum_{j_3}^{3} d_{3j_3} e_{j_3 3}. \\
|\mathbf{P}| &= \sum (-1)^k \sum_{j_1}^{3} d_{1j_1} e_{j_1 1} \sum_{j_1}^{3} d_{2j_2} e_{j_2 2} \sum_{j_3}^{3} d_{3j_3} e_{j_3 3} \\
&= \sum_{j_1}^{3} \sum_{j_2}^{3} \sum_{j_3}^{3} d_{1j_1} d_{2j_2} d_{3j_3} \underbrace{\sum (-1)^k e_{j_1 1} e_{j_2 2} e_{j_3 3}}_{\boldsymbol{E}_{j_1 j_2 j_3}}.
\end{aligned}$$

For every set of indices $j_1 j_2 j_3$, $\boldsymbol{E}_{j_1 j_2 j_3}$ is a signed sum of permutations — *a determinant*:

$$\boldsymbol{E}_{j_1 j_2 j_3} = \sum (-1)^k e_{j_1 1} e_{j_2 2} e_{j_3 3} = \begin{vmatrix} e_{j_1 1} & e_{j_1 2} & e_{j_1 3} \\ e_{j_2 1} & e_{j_2 2} & e_{j_3 3} \\ e_{j_3 1} & e_{j_3 2} & e_{j_3 3} \end{vmatrix}.$$

There will be a term in the complete sum for every set of indices $j_1 j_2 j_3$. However if $j_1 = j_2$ or $j_1 = j_3$ or $j_2 = j_3$ then two rows of the determinant for that term in the sum will be identical – and $\boldsymbol{E}_{j_1 j_2 j_3} = 0$. It follows that the only terms that survive in the complete sum are terms in which $j_1 \neq j_2 \neq j_3$, that is $(j_1 j_2 j_3)$ = (1,2,3), (2,3,1), etc. — all of the permutations of the indices $(j_1 j_2 j_3)$. Thus $\sum_{j_1}^{3} \sum_{j_2}^{3} \sum_{j_3}^{3} d_{1j_1} d_{2j_2} d_{3j_3}$ is a sum which includes all the permutations of the indices — *it would be a determinant if the permutations were multiplied by the appropriate signs*:

$$|\mathbf{P}| = \sum d_{1j_1} d_{2j_2} d_{3j_3} \boldsymbol{E}_{j_1 j_2 j_3} \tag{1.63}$$

If $(j_1 j_2 j_3)$ = (1,2,3) then

$$\boldsymbol{E}_{123} = \begin{vmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} = |\mathbf{E}|.$$

Another permutation of $(j_1 j_2 j_3)$ generates the same matrix with one or more rows switched. For example, if $(j_1 j_2 j_3)$ = (1,3,2),

$$\boldsymbol{E}_{132} = \begin{vmatrix} e_{11} & e_{12} & e_{13} \\ e_{31} & e_{32} & e_{33} \\ e_{21} & e_{22} & e_{23} \end{vmatrix} = -|\mathbf{E}|.$$

For any permutation we must switch the rows to get back to (1,2,3) order. The sign changes for each switch. In general, where k' is the number of switches to restore the matrix to (1,2,3) order, $\boldsymbol{E}_{j_1 j_2 j_3} = (-1)^{k'} |\mathbf{E}|$. The number of row switches necessary to get $\boldsymbol{E}_{j_1 j_2 j_3}$ into (123) order is exactly the same as the number of switches necessary to get $d_{1j_1} d_{2j_2} d_{3j_3}$ into (123) order, and that is the number of

switches necessary to create the signed permutation of $\mathbf{D}$ from (123) order. $(-1)^{k'}$ is the sign of the permutation:

$$\begin{aligned} |\mathbf{P}| &= \sum d_{1j_1} d_{2j_2} d_{3j_3} (-1)^{k'} |\mathbf{E}| \\ &= \sum (-1)^{k'} d_{1j_1} d_{2j_2} d_{3j_3} |\mathbf{E}| \\ &= |\mathbf{D}||\mathbf{E}| \\ |\mathbf{DE}| &= |\mathbf{D}||\mathbf{E}| \end{aligned} \tag{1.64}$$

The formal method for evaluating a determinant, which guarantees that all of the permutations with appropriate signs are obtained, is called *cofactor expansion*. It is undertaken in two steps. First, a row of the matrix is selected (or a column) — then a set of sub-matrices is formed by selecting each element in the row, "crossing out" the row and column that the element intersects, and evaluating the determinant of the remaining sub-matrix (called the *minor* of the element). In the second step each determinant is multiplied by $(-1)^{(i+j)}$, where i and j are the indices of the element which generates the minor. This new signed determinant,$\mathcal{D}_{ij}$, is called the *cofactor* of the matrix element, d_{ij}. Again, this apparently complex notion is clarified by an example. We select the first row of the determinant in Eqn. 1.58, determine the minors of each element in the row, and determine the cofactors. The cofactor of d_{11} is

$$\begin{vmatrix} \cancel{d_{11}} & \cancel{d_{12}} & \cancel{d_{13}} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix} \Longrightarrow \mathcal{D}_{11} = -1^{(1+1)} \begin{vmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{vmatrix} = \begin{vmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{vmatrix}. \tag{1.65}$$

Similarly,

$$\mathcal{D}_{12} = -1^{(1+2)} \begin{vmatrix} d_{21} & d_{23} \\ d_{31} & d_{33} \end{vmatrix} = - \begin{vmatrix} d_{21} & d_{23} \\ d_{31} & d_{33} \end{vmatrix} \tag{1.66}$$

and

$$\mathcal{D}_{13} = -1^{(1+3)} \begin{vmatrix} d_{21} & d_{22} \\ d_{31} & d_{32} \end{vmatrix} = \begin{vmatrix} d_{21} & d_{22} \\ d_{31} & d_{32} \end{vmatrix}. \tag{1.67}$$

The determinant is evaluated as the sum of the products of the matrix elements and their cofactors:

$$\begin{aligned} |\mathbf{D}| &= d_{11}\mathcal{D}_{11} + d_{12}\mathcal{D}_{12} + d_{13}\mathcal{D}_{13} \\ &= d_{11} \begin{vmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{vmatrix} - d_{12} \begin{vmatrix} d_{21} & d_{23} \\ d_{31} & d_{33} \end{vmatrix} + d_{13} \begin{vmatrix} d_{21} & d_{22} \\ d_{31} & d_{32} \end{vmatrix}. \end{aligned} \tag{1.68}$$

The cofactor determinants must now be evaluated in the same manner. For a large matrix this process is cumbersome, but for a 3×3 matrix we are left with only 2×2 determinants. The cofactors for these determinants contain only 1×1 determinants, which are obviously scalars. For a general 2×2 determinant:

$$\begin{aligned} |\mathbf{E}| &= \begin{vmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{vmatrix} = e_{11}(-1)^{(1+1)}|e_{22}| + e_{12}(-1)^{(1+2)}|e_{21}| \\ &= e_{11}e_{22} - e_{12}e_{21}, \\ &\quad \text{the difference of the products of the diagonal elements.} \end{aligned} \tag{1.69}$$

Evaluating each of the 2×2 determinants results in the evaluation of the determinant of the original matrix, which is identical to Eqn. 1.58:

$$\begin{aligned} |\mathbf{D}| &= d_{11}(d_{22}d_{33} - d_{23}d_{32}) - d_{12}(d_{21}d_{33} - d_{23}d_{31}) \\ &\quad + d_{13}(d_{21}d_{32} - d_{22}d_{31}) \\ &= d_{11}d_{22}d_{33} - d_{11}d_{23}d_{32} - d_{12}d_{21}d_{33} + d_{12}d_{23}d_{31} \\ &\quad + d_{13}d_{21}d_{32} - d_{13}d_{22}d_{31}. \end{aligned} \tag{1.70}$$

The determinant in the form of Eqn. 1.68 contains all of the possible permutations of the elements of the second and third rows of the matrix in the form of 2×2 determinants. Recall that the vector product contained all of these permutations as well. Writing the determinant as three row vectors gives us a convenient representation of the vector product in an orthonormal coordinate system:

$$\mathbf{v_2} \times \mathbf{v_3} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_{2x} & v_{2y} & v_{2z} \\ v_{3x} & v_{3y} & v_{3z} \end{vmatrix} = \mathbf{i} \begin{vmatrix} v_{2y} & v_{2z} \\ v_{3y} & v_{3z} \end{vmatrix} - \mathbf{j} \begin{vmatrix} v_{2x} & v_{2z} \\ v_{3x} & v_{3z} \end{vmatrix} + \mathbf{k} \begin{vmatrix} v_{2x} & v_{2y} \\ v_{3x} & v_{3y} \end{vmatrix}. \tag{1.71}$$

Expansion of Eqn. 1.71 results in the previously determined expression for the vector product (Eqn. 1.37). Note that $\mathbf{v_3} \times \mathbf{v_2}$ reverses the row vectors in the determinant, thus changing its sign! The representation of the vector product as a determinant provides a useful way to handle vector products – and will lead us to a simple method for computing the unit cell volume from the orthonormal coordinates of the unit cell axes.

1.4.5 The Inverse of a Matrix

In the previous section the cofactor expansion of a 3×3 matrix was shown to produce the determinant of the matrix. The first row of the matrix was selected to generate the expansion. It is left to the reader to verify that selection of the second or third row produces the same result. Formal proofs for general square matrices often turn out to be cumbersome. Proofs involving determinants and other properties of general matrices can be found in any number of linear algebra books (for example see Campbell, 1971[20]). For the remainder of the chapter we will take the simpler path of exhibiting properties for 3×3 matrices wherever the extrapolation to larger matrices appears logical.

In ordinary algebra, there is an inverse operation to multiplication which reverses the effect of the multiplication – the reciprocal. Multiplying a number by x, then by x^{-1} leaves the number unchanged: $x^{-1} \cdot (x \cdot s) = (x^{-1} \cdot x) \cdot s = (1)s = s$. The reciprocal *inverts* the original multiplication x, and is therefore also called the *inverse* of x. In number theory the number "1" is a multiplier which leaves the number that it multiplies unchanged, and is referred to as an *identity* element. Since we have discovered that matrices transform vectors, is there a matrix that will be analogous to the reciprocal in number theory which will undo a transformation? We begin to answer this question be determining the matrix that is analogous to the identity element in number theory, that is, a matrix that will multiply a vector and leave the vector unchanged. The only matrix that fulfills this function has

diagonal elements of one, and off-diagonal elements of zero. The matrix is called *the identity matrix*, and is referred to by the symbol **I**:

$$\mathbf{Iv} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{1.72}$$

Consider the transformation $\mathbf{Dv} = \mathbf{v'}$. To reverse this transformation, we seek a matrix $\mathbf{D'}$ such that $\mathbf{D'v'} = \mathbf{v}$, leaving $\mathbf{v}$ unchanged:

$$\begin{aligned} \mathbf{D'v'} &= \mathbf{Iv} = \mathbf{v} \\ \mathbf{D'}(\mathbf{Dv}) &= (\mathbf{D'D})\mathbf{v} = \mathbf{Iv} \\ \Longrightarrow \mathbf{D'D} &= \mathbf{I}. \end{aligned} \tag{1.73}$$

Thus, presuming that matrix multiplication is associative (which we will demonstrate in the next section), the analog of the inverse in number theory is clearly the matrix which reverses the transform, $\mathbf{D'}$, which we will denote as the *inverse matrix* (or alternatively, just the *inverse*) of $\mathbf{D}$, $\mathbf{D^{-1}}$. Unfortunately, determining the inverse of a matrix is much less straightforward than determining the reciprocal from ordinary numbers. We begin the task by returning to the *cofactors* discussed in the last section. Every matrix element has a cofactor (Eqns. 1.65–1.67). It follows that we can generate the cofactor for each matrix element and create a new matrix which we call the *cofactor matrix* of the original matrix. For the matrix $\mathbf{D}$, we denote the cofactor matrix of $\mathbf{D}$ as $\mathbf{D_c}$:

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \qquad \mathbf{D_c} = \begin{bmatrix} \mathcal{D}_{11} & \mathcal{D}_{12} & \mathcal{D}_{13} \\ \mathcal{D}_{21} & \mathcal{D}_{22} & \mathcal{D}_{23} \\ \mathcal{D}_{31} & \mathcal{D}_{32} & \mathcal{D}_{33} \end{bmatrix} \tag{1.74}$$

We now evaluate the matrix product,

$$\mathbf{P} = \mathbf{D}\mathbf{D}_\mathbf{c}^T = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} \mathcal{D}_{11} & \mathcal{D}_{21} & \mathcal{D}_{31} \\ \mathcal{D}_{12} & \mathcal{D}_{22} & \mathcal{D}_{32} \\ \mathcal{D}_{13} & \mathcal{D}_{23} & \mathcal{D}_{33} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}.$$

$$\begin{aligned} p_{11} = d_{11}\mathcal{D}_{11} + d_{12}\mathcal{D}_{12} + d_{13}\mathcal{D}_{13} &= |\mathbf{D}| \\ p_{22} = d_{21}\mathcal{D}_{21} + d_{22}\mathcal{D}_{22} + d_{23}\mathcal{D}_{23} &= |\mathbf{D}| \\ p_{33} = d_{31}\mathcal{D}_{31} + d_{32}\mathcal{D}_{32} + d_{33}\mathcal{D}_{33} &= |\mathbf{D}|. \end{aligned} \tag{1.75}$$

p_{11} is the cofactor expansion of the determinant using the first row of the matrix, p_{22} is the cofactor expansion using the second row, and p_{33} is the cofactor expansion using the third row. The off-diagonal elements are sums of products of the elements of one row and the cofactors of another row. This creates a determinant with two identical rows, which must equal zero. Using p_{21} as an example,

$$\begin{aligned} p_{21} &= d_{21}\mathcal{D}_{11} + d_{22}\mathcal{D}_{12} + d_{23}\mathcal{D}_{13} \\ &= d_{21}\begin{vmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{vmatrix} - d_{22}\begin{vmatrix} d_{21} & d_{23} \\ d_{31} & d_{33} \end{vmatrix} + d_{23}\begin{vmatrix} d_{21} & d_{22} \\ d_{31} & d_{32} \end{vmatrix} \\ &= \begin{vmatrix} d_{21} & d_{22} & d_{23} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix} = 0. \end{aligned}$$

The remaining off-diagonal elements suffer the same fate. It follows that the product matrix consists of identical scalars, $|\mathbf{D}|$, along the diagonal, and zeros everywhere else:

$$\mathbf{P} = \mathbf{D}\mathbf{D}_\mathbf{c}^T = \begin{bmatrix} |\mathbf{D}| & 0 & 0 \\ 0 & |\mathbf{D}| & 0 \\ 0 & 0 & |\mathbf{D}| \end{bmatrix} = |\mathbf{D}| \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = |\mathbf{D}|\mathbf{I}. \tag{1.76}$$

Since $|\mathbf{D}|$ is a scalar, *provided that it is not zero*,

$$\mathbf{D}\frac{\mathbf{D}_\mathbf{c}^T}{|\mathbf{D}|} = \mathbf{I} \quad \Longrightarrow \quad \mathbf{D^{-1}} = \frac{\mathbf{D}_\mathbf{c}^T}{|\mathbf{D}|}. \tag{1.77}$$

In order to determine the inverse of a matrix we generate a cofactor matrix and transpose it (the transpose of the cofactor matrix is called the *adjoint* matrix), then divide each element of the resulting matrix by the determinant of the original matrix. It follows that *the inverse of a matrix can exist only if the determinant of the matrix is non-zero.* Thus, *if a matrix is singular, it does not have an inverse.* The inverse of a 3×3 matrix will be especially important throughout the remainder of the book. Given the determinant, $|\mathbf{D}|$, from Eqn. 1.70,

$$\mathbf{D^{-1}} = \frac{1}{|\mathbf{D}|} \begin{bmatrix} (d_{22}d_{33} - d_{32}d_{23}) & (d_{32}d_{13} - d_{12}d_{33}) & (d_{12}d_{23} - d_{22}d_{13}) \\ (d_{31}d_{23} - d_{21}d_{33}) & (d_{11}d_{33} - d_{31}d_{13}) & (d_{21}d_{13} - d_{11}d_{23}) \\ (d_{21}d_{32} - d_{31}d_{22}) & (d_{31}d_{12} - d_{11}d_{32}) & (d_{11}d_{22} - d_{21}d_{12}) \end{bmatrix}. \tag{1.78}$$

The inverse of the transformation matrices described earlier will be especially important in the discussion of symmetry at the end of this chapter. Recall that these matrices were orthonormal matrices. Consider the product of an orthonormal matrix $\mathbf{T} = [\mathbf{t_1}\ \mathbf{t_2}\ \mathbf{t_3}]$ and its transpose:

$$\begin{aligned} \mathbf{T}^T\mathbf{T} &= \begin{bmatrix} t_{11} & t_{21} & t_{31} \\ t_{12} & t_{22} & t_{32} \\ t_{13} & t_{23} & t_{33} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{t_1}\cdot\mathbf{t_1}) & (\mathbf{t_1}\cdot\mathbf{t_2}) & (\mathbf{t_1}\cdot\mathbf{t_3}) \\ (\mathbf{t_2}\cdot\mathbf{t_1}) & (\mathbf{t_2}\cdot\mathbf{t_2}) & (\mathbf{t_2}\cdot\mathbf{t_3}) \\ (\mathbf{t_3}\cdot\mathbf{t_1}) & (\mathbf{t_3}\cdot\mathbf{t_2}) & (\mathbf{t_3}\cdot\mathbf{t_3}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}. \end{aligned} \tag{1.79}$$

The transpose of an orthonormal matrix is its inverse. Note also that the identity matrix is its own inverse:

$$\mathbf{II} = \mathbf{I} \implies \mathbf{I^{-1}} = \mathbf{I}. \tag{1.80}$$

If a matrix is symmetric ($\mathbf{T} = \mathbf{T}^T$), its inverse will also be symmetric:

$$\begin{aligned} &\mathbf{T}\mathbf{T^{-1}} = \mathbf{I} \\ &(\mathbf{T^{-1}})^T\mathbf{T}^T = \mathbf{I} \\ &(\mathbf{T^{-1}})^T\mathbf{T} = \mathbf{I} \implies \\ &(\mathbf{T^{-1}})^T = \mathbf{T^{-1}}. \end{aligned} \tag{1.81}$$

1.4.6 The Rules of Matrix Algebra

Matrix transformations are *linear* operations since they change the vectors upon which they operate linearly, i.e., they do not modify vectors with exponents or other functions. The matrix operations defined previously provide the basis for formulating the algebraic rules for combining and manipulating matrices. The rules of this *linear algebra* that are important to crystallography are summarized and rationalized in this section.

Rule 1. *Matrix additions are associative.* For $\mathbf{G} = (\mathbf{D} + \mathbf{E}) + \mathbf{F}$, the $ijth$ matrix element of $\mathbf{G}$ is $g_{ij} = (d_{ij} + e_{ij}) + f_{ij}$. For $\mathbf{H} = \mathbf{D} + (\mathbf{E} + \mathbf{F})$, the $ijth$ matrix element of $\mathbf{H}$ is $h_{ij} = d_{ij} + (e_{ij} + f_{ij})$. Clearly, $g_{ij} = h_{ij}$ and

$$(\mathbf{D} + \mathbf{E}) + \mathbf{F} = \mathbf{D} + (\mathbf{E} + \mathbf{F}) \tag{1.82}$$

Rule 2. *Matrix additions are commutative.* For $\mathbf{F} = \mathbf{D} + \mathbf{E}$, the $ijth$ matrix element of $\mathbf{F}$ is $f_{ij} = d_{ij} + e_{ij}$. For $\mathbf{G} = \mathbf{E} + \mathbf{D}$, the $ijth$ matrix element of $\mathbf{G}$ is $g_{ij} = e_{ij} + d_{ij}$. Again, it is obvious that $f_{ij} = g_{ij}$ and

$$\mathbf{D} + \mathbf{E} = \mathbf{E} + \mathbf{D} \tag{1.83}$$

Rule 3. *Matrix additions are distributive with respect to scalar multiplication.* Let $\mathbf{F} = s(\mathbf{D} + \mathbf{E})$. Then $f_{ij} = s(d_{ij} + e_{ij}) = sd_{ij} + se_{ij}$. It follows that

$$s(\mathbf{D} + \mathbf{E}) = s\mathbf{D} + s\mathbf{E}. \tag{1.84}$$

Also, if $\mathbf{G} = (s_1 + s_2)\mathbf{D}$, then $g_{ij} = (s_1 + s_2)d_{ij} = s_1 d_{ij} + s_2 d_{ij}$ and

$$(s_1 + s_2)\mathbf{D} = s_1\mathbf{D} + s_2\mathbf{D}. \tag{1.85}$$

Rule 4. *Matrix multiplications are associative.* For $\mathbf{G} = (\mathbf{DE})\mathbf{F}$ and $\mathbf{H} = \mathbf{D}(\mathbf{EF})$, Let $(\mathbf{DE})\mathbf{F} = \mathbf{PF}$ and $\mathbf{D}(\mathbf{EF}) = \mathbf{DQ}$. For simplicity, assume that $\mathbf{D}$ and $\mathbf{E}$ are 3×3 matrices and $\mathbf{F}$ is either a 3×3 matrix or a 3×1 column vector. Then

$$\begin{aligned}
g_{ij} &= \sum_{k=1}^{3} p_{ik} f_{kj} \\
p_{ik} &= \sum_{m=1}^{3} d_{im} e_{mk} \\
g_{ij} &= \sum_{k=1}^{3} \left(\sum_{m=1}^{3} d_{im} e_{mk} \right) f_{kj} \\
g_{ij} &= \sum_{m=1}^{3} d_{im} \left(\sum_{m=1}^{3} e_{mk} f_{kj} \right) \\
g_{ij} &= \sum_{m=1}^{3} d_{im} q_{mj} = h_{ij} \implies \mathbf{G} = \mathbf{H}. \text{ Thus,} \\
(\mathbf{DE})\mathbf{F} &= \mathbf{D}(\mathbf{EF}).
\end{aligned} \tag{1.86}$$

Rule 5. *Matrix multiplications are, in general,* ***not*** *commutative.* Consider, for example, rotation of the vector $\mathbf{v}$ around the z axis by 90° with $\mathbf{R_Z}$, followed by a rotation around the x axis of 90° with $\mathbf{R_X}$, $\mathbf{R_X}(\mathbf{R_Z v}) = (\mathbf{R_X R_Z})\mathbf{v}$ (Rule 4). The matrix product creates a new transformation matrix, $\mathbf{R_1} = \mathbf{R_Z R_X}$:

$$\begin{aligned}
\mathbf{R_1} &= \begin{bmatrix} \cos(\pi/2) & -\sin(\pi/2) & 0 \\ \sin(\pi/2) & \cos(\pi/2) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\pi/2) & -\sin(\pi/2) \\ 0 & \sin(\pi/2) & \cos(\pi/2) \end{bmatrix} \\
&= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\
\mathbf{R_1 v} &= \mathbf{R_1} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} z_c \\ x_c \\ y_c \end{bmatrix}
\end{aligned}$$

If we reverse the order of the operations, rotating around the x axis followed by rotation around the z axis we have $\mathbf{R_Z}(\mathbf{R_X v}) = (\mathbf{R_Z R_X})\mathbf{v}$. The new transformation matrix is now $\mathbf{R_2} = \mathbf{R_X R_Z}$:

$$\begin{aligned}
\mathbf{R_2} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\pi/2) & -\sin(\pi/2) \\ 0 & \sin(\pi/2) & \cos(\pi/2) \end{bmatrix} \begin{bmatrix} \cos(\pi/2) & -\sin(\pi/2) & 0 \\ \sin(\pi/2) & \cos(\pi/2) & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\
\mathbf{R_2 v} &= \mathbf{R_2} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} -y_c \\ z_c \\ x_c \end{bmatrix}
\end{aligned}$$

Thus $\mathbf{R_Z R_X} \neq \mathbf{R_X R_Z}$. In general, for matrices $\mathbf{DE}$ and $\mathbf{ED}$,

$$\mathbf{DE} \neq \mathbf{ED}. \tag{1.87}$$

The order in which we multiply matrices matters! It follows that for the equality $\mathbf{D} = \mathbf{E}$, *pre-multiplying or post-multiplying both sides of the matrix equation by another matrix retains the equality*: $\mathbf{FD} = \mathbf{FE}$ and $\mathbf{DF} = \mathbf{EF}$, *but only in special cases does* $\mathbf{DF} = \mathbf{FE} = \mathbf{FD} = \mathbf{EF}$. When the order of multiplication of two matrices does not alter the product we say that the matrices *commute*.

Rule 6. *A matrix commutes with the identity matrix and its own inverse.* The identity matrix commutes with all matrices:

$$\mathbf{DI} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \mathbf{D}$$

$$\mathbf{ID} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \mathbf{D}$$

$$\mathbf{DI} = \mathbf{ID} = \mathbf{D}. \tag{1.88}$$

A matrix also commutes with its inverse:

$$\begin{aligned}
\mathbf{D^{-1}D} &= \mathbf{I} \\
\mathbf{D^{-1}D^{-1}D} &= \mathbf{D^{-1}I} = \mathbf{D^{-1}} \\
\mathbf{D^{-1}DD^{-1}} &= \mathbf{ID^{-1}} = \mathbf{D^{-1}} \\
\mathbf{D^{-1}D^{-1}D} &= \mathbf{D^{-1}DD^{-1}} \\
\mathbf{D^{-1}D} &= \mathbf{DD^{-1}} = \mathbf{I}.
\end{aligned} \tag{1.89}$$

Rule 7. *Matrix multiplications are distributive.* Let $\mathbf{S} = \mathbf{E} + \mathbf{F}$, $\mathbf{T} = \mathbf{D}(\mathbf{E} + \mathbf{F}) = \mathbf{DS}$, $\mathbf{V} = \mathbf{DE}$, and $\mathbf{W} = \mathbf{DF}$. Then

$$\begin{aligned}
t_{ij} &= \sum_{k=1}^{3} d_{ik}s_{kj} = \sum_{k=1}^{3} d_{ik}(e_{kj} + f_{kj}) \\
&= \sum_{k=1}^{3} d_{ik}e_{kj} + d_{ik}f_{kj} = \sum_{k=1}^{3} d_{ik}e_{kj} + \sum_{k=1}^{3} d_{ik}f_{kj} \\
&= v_{ij} + w_{ij} \Longrightarrow \mathbf{T} = \mathbf{V} + \mathbf{W}
\end{aligned}$$

$$\mathbf{D}(\mathbf{E} + \mathbf{F}) = \mathbf{DE} + \mathbf{DF}. \tag{1.90}$$

In this case the matrix sum is pre-multiplied by a matrix, and this is formally called the *left distributive property of matrix multiplication.* By a similar argument, it is easy to demonstrate the *right distributive property of matrix multiplication*, in which the matrix sum is post-multiplied by a matrix: $(\mathbf{E} + \mathbf{F})\mathbf{D} = \mathbf{ED} + \mathbf{FD}$.

Rule 8. *The inverse of the product of an ordered array of matrices is the product of the inverses of the individual matrices* ***in reverse order***.

$$\begin{aligned}
\mathbf{D^{-1}D} &= \mathbf{D^{-1}ID} = \mathbf{D^{-1}}(\mathbf{E^{-1}E})\mathbf{D} \\
&= (\mathbf{D^{-1}E^{-1}})(\mathbf{ED}) = \mathbf{I} \\
&\Longrightarrow (\mathbf{ED})^{-1} = \mathbf{D^{-1}E^{-1}} \\
(\mathbf{D^{-1}E^{-1}})(\mathbf{ED}) &= (\mathbf{D^{-1}E^{-1}})\mathbf{I}(\mathbf{ED}) \\
&= (\mathbf{D^{-1}E^{-1}})(\mathbf{F^{-1}F})(\mathbf{ED}) \\
&= (\mathbf{D^{-1}E^{-1}F^{-1}})(\mathbf{FED}) = \mathbf{I} \\
&\Longrightarrow (\mathbf{FED})^{-1} = \mathbf{D^{-1}E^{-1}F^{-1}}.
\end{aligned}$$

By induction,

$$(\mathbf{QR}\ldots\mathbf{FED})^{-1} = \mathbf{D^{-1}E^{-1}F^{-1}}\ldots\mathbf{Q^{-1}R^{-1}}. \tag{1.91}$$

Rule 9. *The transpose of the product of an ordered array of matrices is the product of the transposes of the individual matrices* ***in reverse order***. Let $\mathbf{U} = \mathbf{DE}$,

$\mathbf{V} = (\mathbf{DE})^T$, $\mathbf{S} = \mathbf{E}^T, \mathbf{T} = \mathbf{D}^T$, and $\mathbf{W} = \mathbf{E}^T\mathbf{D}^T = \mathbf{ST}$. Again, for simplicity, assume that all matrices are 3×3.

$$
\begin{aligned}
u_{ij} &= \sum_{k=1}^{3} d_{ik}e_{kj} \\
v_{ij} = u_{ij}^T &= u_{ji} = \sum_{k=1}^{3} d_{jk}e_{ki} \\
w_{ij} &= \sum_{i=1}^{3} s_{ik}t_{kj},\ s_{ik} = e_{ki} \text{ and } t_{kj} = d_{jk} \\
w_{ij} &= \sum_{i=1}^{3} e_{ki}d_{jk} = \sum_{i=1}^{3} d_{jk}e_{ki} = v_{ij} \Longrightarrow \mathbf{V} = \mathbf{W} \\
\mathbf{E}^T\mathbf{D}^T &= (\mathbf{DE})^T \\
\mathbf{F}^T\mathbf{E}^T\mathbf{D}^T &= \mathbf{F}^T(\mathbf{DE})^T = (\mathbf{DEF})^T
\end{aligned}
$$

By induction,

$$(\mathbf{QR}\ldots\mathbf{FED})^T = \mathbf{D^T E^T F^T}\ldots\mathbf{Q^T R^T}. \tag{1.92}$$

Rule 10. *The transpose of the sum of two matrices is the sum of the transposes of the individual matrices.* Let $\mathbf{D} = \mathbf{A} + \mathbf{B}$.

$$
\begin{aligned}
\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} \end{bmatrix}
\end{aligned}
\tag{1.93}
$$

$$
\begin{aligned}
\mathbf{D}^T = \begin{bmatrix} d_{11} & d_{21} & d_{31} \\ d_{12} & d_{22} & d_{32} \\ d_{13} & d_{23} & d_{33} \end{bmatrix} &= \begin{bmatrix} a_{11}+b_{11} & a_{21}+b_{21} & a_{31}+b_{31} \\ a_{12}+b_{12} & a_{22}+b_{22} & a_{32}+b_{32} \\ a_{13}+b_{13} & a_{23}+b_{23} & a_{33}+b_{33} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \\
&= \mathbf{A}^T + \mathbf{B}^T.
\end{aligned}
\tag{1.94}
$$

Rule 11. *The transpose of the inverse of a matrix is the inverse of the transpose of the matrix.*

$$\begin{aligned}
\mathbf{G}^{-1}\mathbf{G} &= \mathbf{I} \\
\mathbf{G}^T(\mathbf{G}^{-1})^T &= \mathbf{I}^T = \mathbf{I} \\
\mathbf{G}^T(\mathbf{G}^T)^{-1} &= \mathbf{I} \\
\mathbf{G}^T(\mathbf{G}^T)^{-1} &= \mathbf{G}^T(\mathbf{G}^{-1})^T \\
(\mathbf{G}^T)^{-1} &= (\mathbf{G}^{-1})^T. \qquad (1.95)
\end{aligned}$$

1.4.7 The Eigenvectors and Eigenvalues of a Matrix

In general, when a matrix operates on a vector, the vector changes in both direction and magnitude:

$$\mathbf{D}\mathbf{v} = \mathbf{v}'. \qquad (1.96)$$

Among the infinitely many vectors that a matrix can transform, there is a subset of those vectors, *characteristic* of the matrix, that *may* change in magnitude, but either remain parallel or become antiparallel to the original vectors when operated on by the matrix. These characteristic vectors, $\{\mathbf{e}\}$, are known as *eigenvectors* (*eigen* in German roughly translates as *characteristic* or *innate*). When a matrix operates on its eigenvectors, the vector is stretched, shrunk, or unmodified by a scalar, but otherwise it either retains its direction or reverses it:

$$\mathbf{D}\mathbf{e}_i = \lambda_i \mathbf{e}_i, \quad ß = 1, 2, \ldots, n. \qquad (1.97)$$

The scalar multiplier of each eigenvector, λ, is also characteristic of the matrix, and is known as the *eigenvalue* of the eigenvector.

Note that if an eigenvector is multiplied or divided by a constant the resulting vector is still an eigenvector of the matrix with the same eigenvalue:

$$\mathbf{D}(q\mathbf{e}_i) = \lambda_i(q\mathbf{e}_i). \qquad (1.98)$$

Thus the "eigenvector" for a given eigenvalue is actually an infinite set of vectors with all possible magnitudes, all pointing in the same direction. The most useful eigenvector is the one of unit length, obtained by dividing any one of the set of eigenvectors by its length:

$$\begin{aligned}
\mathbf{D}\frac{\mathbf{e}_i}{e_i} &= \lambda_i \frac{\mathbf{e}_i}{e_i} \\
\mathbf{D}\mathbf{e}_{\mathbf{u}i} &= \lambda_i \mathbf{e}_{\mathbf{u}i}, \quad e_{ui} = 1. \qquad (1.99)
\end{aligned}$$

For an $n \times n$ matrix there are n unit eigenvectors and n corresponding eigenvalues. For convenience we will focus on the three-dimensional case. The three resulting equations (Eqn. 1.99) can be represented by a single matrix equation:

$$\begin{aligned}
\mathbf{D}\mathbf{E}_{\mathbf{u}} &= \mathbf{D}\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \\
\mathbf{D}\mathbf{E}_{\mathbf{u}} &= \mathbf{E}_{\mathbf{u}}[\lambda]. \qquad (1.100)
\end{aligned}$$

The column vectors of $\mathbf{E_u}$ are the unit eigenvectors of $\mathbf{D}$, and the diagonal elements of $[\lambda]$ are the eigenvalues. The matrix consisting of the eigenvectors of $\mathbf{D}$, $\mathbf{E}$, is referred to as the *modal matrix* of $\mathbf{D}$. Eqn. 1.100 is equally valid for general eigenvectors, i.e., $\mathbf{DE} = \mathbf{E}\,[\lambda]$. Note that the equation must be written in this manner; $\mathbf{E}$ and $[\lambda]$ *do not commute.*

The treatment of atom displacements in Chapter 5 will involve the eigenvectors and eigenvalues of symmetric matrices. In preparation for this, we derive an important property for *symmetric* matrices: *If the matrix,* $\mathbf{D}$*, is symmetric, then its modal matrix is orthogonal.* To prove this we show that any pair of the eigenvectors that compose a modal matrix of $\mathbf{D}$, $\mathbf{e}_i$ and $\mathbf{e}_j$, with distinct eigenvalues $\lambda_i \neq \lambda_j$, have a scalar product of zero, given that $\mathbf{D} = \mathbf{D}^T$:

$$\begin{aligned}
&\mathbf{D}\,\mathbf{e}_i = \lambda_i\,\mathbf{e}_i \\
&\mathbf{e}_j^T\mathbf{D}\,\mathbf{e}_i = \lambda_i\,\mathbf{e}_j^T\,\mathbf{e}_i \qquad (1.101)\\
&\mathbf{D}\,\mathbf{e}_j = \lambda_j\,\mathbf{e}_j \\
&(\mathbf{D}\,\mathbf{e}_j)^T = \mathbf{e}_j^T\mathbf{D}^T = \mathbf{e}_j^T\mathbf{D} = \lambda_j\,\mathbf{e}_j^T \\
&\mathbf{e}_j^T\mathbf{D}\,\mathbf{e}_i = \lambda_j\,\mathbf{e}_j^T\,\mathbf{e}_i \qquad (1.102)
\end{aligned}$$

Subtracting Eqn. 1.101 from Eqn. 1.102 results in $(\lambda_j - \lambda_i)(\mathbf{e}_j^T\,\mathbf{e}_i) = 0$. Since $\lambda_j - \lambda_i \neq 0$, $\mathbf{e}_j^T\,\mathbf{e}_i = 0$; the eigenvectors of the modal matrix of $\mathbf{D}$ are orthogonal.

The eigenvectors and eigenvalues of matrix $\mathbf{D}$ (for $n = 3$) are determined by solving Eqn. 1.97:

$$\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \lambda \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

$$\begin{aligned}
&d_{11}\,e_1 + d_{12}\,e_2 + d_{13}\,e_3 = \lambda\,e_1 \\
&d_{21}\,e_1 + d_{22}\,e_2 + d_{23}\,e_3 = \lambda\,e_2 \\
&d_{31}\,e_1 + d_{32}\,e_2 + d_{33}\,e_3 = \lambda\,e_3
\end{aligned}$$

The result is three homogenous (all equal to zero) linear equations:

$$(d_{11} - \lambda)\,e_1 + d_{12}\,e_2 + d_{13}\,e_3 = 0 \qquad (1.103)$$
$$d_{21}\,e_1 + (d_{22} - \lambda)\,e_2 + d_{23}\,e_3 = 0 \qquad (1.104)$$
$$d_{31}\,e_1 + d_{32}\,e_2 + (d_{33} - \lambda)\,e_3 = 0 \qquad (1.105)$$

$$\begin{bmatrix} d_{11} - \lambda & d_{12} & d_{13} \\ d_{21} & d_{22} - \lambda & d_{23} \\ d_{31} & d_{32} & d_{33} - \lambda \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{D}_\lambda\mathbf{e} = \mathbf{0}. \qquad (1.106)$$

If $\mathbf{D}_\lambda$ has an inverse then $\mathbf{e} = \mathbf{D}_\lambda^{-1}\mathbf{0} = \mathbf{0}$, a trivial solution (correct, but useless). Thus for there to be a non-trivial solution, $\mathbf{D}_\lambda$ cannot have an inverse. Recall that in Sec. 1.4.5 it was shown that a matrix has an inverse if and only if its determinant

is not singular (equal to 0). Thus the criterion for a nontrivial solution is that $|\mathbf{D}_\lambda| = 0$:

$$\begin{aligned} |\mathbf{D}_\lambda| \;=\; & (d_{11}-\lambda)(d_{22}-\lambda)(d_{33}-\lambda) - (d_{11}-\lambda)d_{23}d_{32} - d_{12}d_{21}(d_{33}-\lambda) \\ & -d_{13}(d_{22}-\lambda)d_{31} + d_{12}d_{23}d_{31} + d_{13}d_{21}d_{32} = 0. \end{aligned}$$

Expanding and collecting the terms produces a third order polynomial (cubic) equation:

$$\begin{aligned} & \lambda^3 - (d_{11}+d_{22}+d_{33})\lambda^2 \\ & \quad +(d_{22}d_{33} - d_{32}d_{23} + d_{11}d_{33} - d_{31}d_{13} + d_{11}d_{22} - d_{21}d_{12})\lambda \\ & \quad -(d_{11}d_{22}d_{33} - d_{11}d_{23}d_{32} - d_{12}d_{21}d_{33} \\ & \quad +d_{12}d_{23}d_{31} + d_{13}d_{21}d_{32} - d_{13}d_{22}d_{31}) = 0, \end{aligned} \tag{1.107}$$

conveniently written in terms of the trace, the cofactors of the diagonal elements, and the determinant of $\mathbf{D}$:

$$\lambda^3 - (d_{11}+d_{22}+d_{33})\lambda^2 + (\mathcal{D}_{11}+\mathcal{D}_{22}+\mathcal{D}_{33})\lambda - |\mathbf{D}| = 0. \tag{1.108}$$

The cubic equation, of the form $f(\lambda) = a\,\lambda^3 + b\,\lambda^2 + c\,\lambda + d = 0$, has three roots, λ_1, λ_2, and λ_3, which are the eigenvalues of $\mathbf{D}$.* Inserting each eigenvalue into Eqns. 1.103–1.105 allows us to solve for the corresponding eigenvector. The solutions are not unique — there are an infinite number of eigenvectors for each eigenvalue, differing by their magnitudes (Eqn. 1.99). Because of this each eigenvector solution will have one arbitrary component (there will be one vector in the infinite set that has this component). For example, setting $e_3 = 1$ and inserting λ_1 into Eqns. 1.103–1.105 finds e_2 in terms of e_1 from Eqn. 1.103. Substituting the expression for e_2 into Eqn. 1.104 then provides a value for e_1, and subsequently, e_2. The eigenvector has magnitude $e = \sqrt{(e_1^2 + e_2^2 + 1^2)}$, and $\mathbf{e}_u = [e_1/e \;\; e_2/e \;\; 1/e]^T$.* The remaining two unit eigenvectors are determined by substituting λ_2 and λ_3, respectively.

If a matrix is orthogonal, it does not change the magnitude of any vector that it transforms, including it eigenvectors. It follows that the real† eigenvalues of an orthogonal matrix must be ± 1:

$$\mathbf{D}\mathbf{e}_i = \pm\mathbf{e}_i, \quad i = 1, 2, \ldots, n. \tag{1.109}$$

*There are analytical formulas for the three roots of a cubic equation,[21] just as there are for the two roots of a quadratic equation. The formulae are quite complex, involving a number of operations. The solutions for cubic and higher order equations are generally obtained numerically, by searching for values of the independent variable that set the function very close to zero (within some preselected tolerance limit).

*The representation of a vector in a text line requires it to be written as a row vector. Most of the vector operations throughout the book will be undertaken with column vectors, formally requiring the row vector in the text line to be written as its transpose. To avoid having to repeat this notation continually throughout the book, unless specifically indicated, *a row vector in a text line will be assumed to be a column vector.*

†An orthogonal rotation matrix can also have one real eigenvalue of +1 and two complex eigenvalues, $e^{i\theta} = \cos\theta + i\sin\theta$ and $e^{-i\theta} = \cos\theta - i\sin\theta$; the eigenvectors in these cases have imaginary components and do not change direction in "complex space."

1.5 Coordinate Systems in Crystallography

As discussed earlier, the ability to treat vectors in either the unit cell basis $\{\mathbf{a}\ \mathbf{b}\ \mathbf{c}\}$ or an orthonormal basis $\{\mathbf{i}\ \mathbf{j}\ \mathbf{k}\}$ requires a method for transforming the vector between the two bases. The matrix algebra described in the last section provides us with just that! We begin with the components of a vector defined in one coordinate system and set out to determine a transformation matrix that will generate the coordinates of the vector in a new coordinate system — a *change of basis*.

1.5.1 Change of Basis

Fig. 1.30 illustrates a vector $\mathbf{v}$ described in two different coordinate systems, one defined by the basis set $\{R\} = \{\mathbf{r_x}\ \mathbf{r_y}\ \mathbf{r_z}\}$ and the other defined by the basis set $\{S\} = \{\mathbf{s_x}\ \mathbf{s_y}\ \mathbf{s_z}\}$. In the $\{R\}$ basis $\mathbf{v} = [x\ y\ z]$ — in fractional coordinates $\mathbf{v} = [x_r\ y_r\ z_r]$ where $x_r = x/r_x$, $y_r = y/r_y$, and $z_r = z/r_z$. In the $\{S\}$ basis $\mathbf{v} = [x'\ y'\ z']$ — in fractional coordinates $\mathbf{v} = [x_s\ y_s\ z_s]$ where $x_s = x'/s_x$, $y_s = y'/s_y$, and $z_s = z'/s_z$.

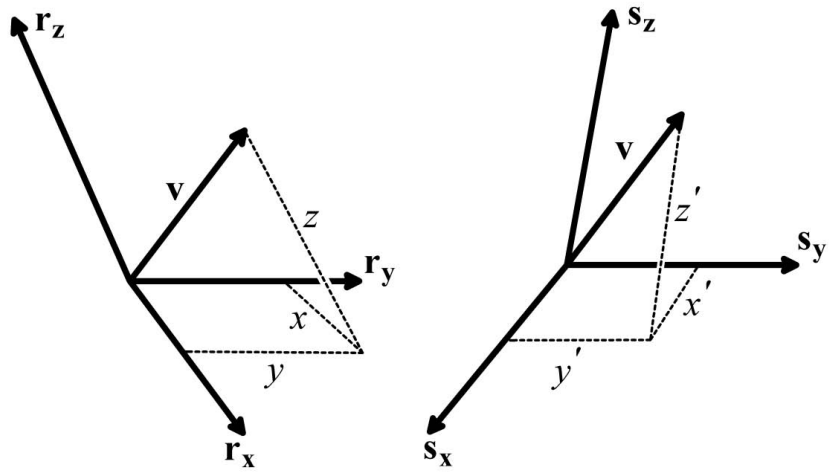

Figure 1.30 Vector $\mathbf{v}$ in two different coordinate systems, $\{R\} = \{\mathbf{r_x}\ \mathbf{r_y}\ \mathbf{r_z}\}$ and $\{S\} = \{\mathbf{s_x}\ \mathbf{s_y}\ \mathbf{s_z}\}$. $[x\ y\ z]$ are the displacement distances along each of the axes in $\{R\}$, $[x'\ y'\ z']$ are the displacements along each of the axes in $\{S\}$.

Suppose that we have described the vector in $\{R\}$, and wish to determine its coordinates in $\{S\}$. Thus we know $[x_r\ y_r\ z_r]$ and can express the vector in terms of the coordinates:

$$\mathbf{v} = x_r\mathbf{r_x} + y_r\mathbf{r_y} + z_r\mathbf{r_z}. \tag{1.110}$$

We wish to determine $[x_s\ y_s\ z_s]$ such that

$$\mathbf{v} = x_s\mathbf{s_x} + y_s\mathbf{s_y} + z_s\mathbf{s_z}. \tag{1.111}$$

In $\{R\}$ the components of the basis vectors are

$$\begin{aligned}
\mathbf{r_x} &= 1\mathbf{r_x} + 0\mathbf{r_y} + 0\mathbf{r_z} = [1\ 0\ 0] \\
\mathbf{r_y} &= 0\mathbf{r_x} + 1\mathbf{r_y} + 0\mathbf{r_z} = [0\ 1\ 0] \\
\mathbf{r_z} &= 0\mathbf{r_x} + 0\mathbf{r_y} + 1\mathbf{r_z} = [0\ 0\ 1].
\end{aligned}$$

Since the basis vectors in $\{R\}$ are vectors in 3-space, each of them can also be described as a linear combination of the basis vectors in $\{S\}$:

$$\begin{aligned}
\mathbf{r_x} &= r_{11}\mathbf{s_x} + r_{12}\mathbf{s_y} + r_{13}\mathbf{s_z} \\
\mathbf{r_y} &= r_{21}\mathbf{s_x} + r_{22}\mathbf{s_y} + r_{23}\mathbf{s_z} \\
\mathbf{r_z} &= r_{31}\mathbf{s_x} + r_{32}\mathbf{s_y} + r_{33}\mathbf{s_z}.
\end{aligned}$$

Thus the vector in $\{R\}$ can be expressed as

$$\begin{aligned}
\mathbf{v} = {} & x_r\ (r_{11}\mathbf{s_x} + r_{12}\mathbf{s_y} + r_{13}\mathbf{s_z}) + y_r\ (r_{21}\mathbf{s_x} + r_{22}\mathbf{s_y} + r_{23}\mathbf{s_z}) \\
& + z_r\ (r_{31}\mathbf{s_x} + r_{32}\mathbf{s_y} + r_{33}\mathbf{s_z}).
\end{aligned}$$

Expanding this expression and collecting terms gives

$$\begin{aligned}
\mathbf{v} = {} & (x_r r_{11} + y_r r_{21} + z_r r_{31})\ s_x + (x_r r_{12} + y_r r_{22} + z_r r_{32})\ s_y \\
& + (x_r r_{13} + y_r r_{23} + z_r r_{33})\ s_z.
\end{aligned}$$

This is $\mathbf{v}$ *expressed as a linear combination of the basis vectors in* $\{S\}$ We have determined the components of the vector in $\{S\}$ in terms of its components and the basis vectors in $\{R\}$, provided that we know the components of the basis vectors in $\{R\}$ in the $\{S\}$ basis:

$$\begin{aligned}
x_s &= (x_r r_{11} + y_r r_{21} + z_r r_{31}) \\
y_s &= (x_r r_{12} + y_r r_{22} + z_r r_{32}) \\
z_s &= (x_r r_{13} + y_r r_{23} + z_r r_{33}).
\end{aligned} \tag{1.112}$$

Eqns. 1.112 are easily recognized as a matrix equation:

$$\begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \tag{1.113}$$

Note that the matrix in this expression is the *transpose* of the matrix that consists of the coefficients of the components of the $\{R\}$ basis vectors in the $\{S\}$ basis.

1.5.2 Transformation from the Unit Cell Basis to an Orthonormal Basis

The strategy to formulate a method to transform vectors from a unit cell basis into an orthonormal basis should now be clear. We determine the components of the unit cell axes in the orthonormal basis and transpose the matrix of the components. This will provide a transformation matrix $\mathbf{B}$ such that $\mathbf{B}\mathbf{v_f} = \mathbf{v_c}$, where $\mathbf{v_f} = [x_f\ y_f\ z_f]$, in fractional coordinates in the unit cell basis, and $\mathbf{v_c} = [x_c\ y_c\ z_c]$, the Cartesian coordinates of the vector.

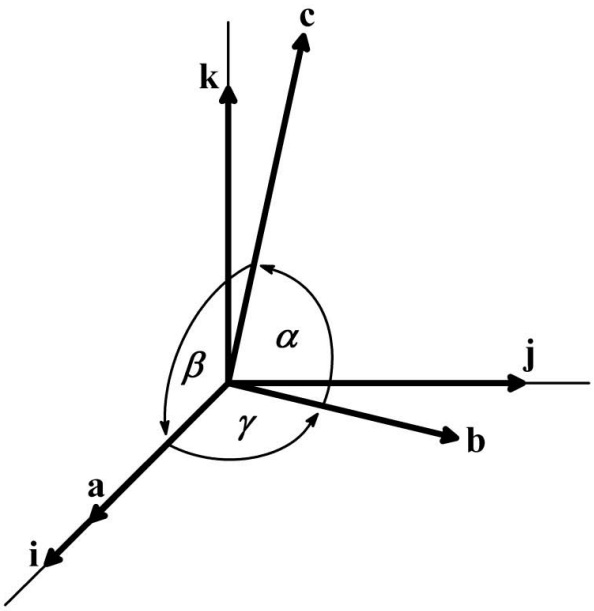

Figure 1.31 The standard orientation of a unit cell in a Cartesian coordinate system. The **a** axis is coincident with **i**, The **b** axis lies in the **ij** plane, and the z component of the **c** axis points in the same direction as **k**.

There is, however, one necessary step before the **B** matrix is determined. The orientation of the unit cell with respect to the Cartesian coordinate system is arbitrary. If we place the unit cell origin at the origin of the Cartesian system we can rotate the unit cell at any angle about the origin. Each orientation will, of course, create a different transformation matrix since the relation between the bases will change. We must therefore fix the orientation in some rational way. Fig. 1.31 illustrates the standard orientation of the unit cell with respect to the orthonormal basis. This orientation maintains right-handed coordinates for both systems. The **a** axis is coincident with the **i** axis, and points in the same direction (both are positive or negative simultaneously). The **b** axis is coplanar with **i** and **j** and points in a direction such that $\mathbf{a} \times \mathbf{b}$ is coincident with **i** and points in the same direction. The z component of the **c** axis points in the same direction as **k**.

We begin with a vector **v** with coordinates in the unit cell basis $\{\mathbf{a}\ \mathbf{b}\ \mathbf{c}\}$. Vectors will be considered column vectors in the discussion which follows. We will use the symbol $\mathbf{v_f} = [x_f\ y_f z_f]$ to denote a general vector with fractional coordinates in the unit cell basis:

$$\mathbf{v_f} = x_f\mathbf{a} + y_f\mathbf{b} + z_f\mathbf{c}. \tag{1.114}$$

We seek to determine the components *of the same vector* in Cartesian coordinates, which we will denote $\mathbf{v_c} = [x_c\ y_c\ z_c]$ such that

$$\mathbf{v_c} = x_c\mathbf{i} + y_c\mathbf{j} + z_c\mathbf{k}. \tag{1.115}$$

In the unit cell basis the components of the axial basis vectors are

$$\begin{aligned} \mathbf{a} &= 1\mathbf{a} + 0\mathbf{b} + 0\mathbf{c} = [1\ 0\ 0] \\ \mathbf{b} &= 0\mathbf{a} + 1\mathbf{b} + 0\mathbf{c} = [0\ 1\ 0] \\ \mathbf{c} &= 0\mathbf{a} + 0\mathbf{b} + 1\mathbf{c} = [0\ 0\ 1]. \end{aligned} \tag{1.116}$$

The lengths of the unit cell axes, a, b, and c, and the angles between the axes, α, β, and γ are collectively known as *the unit cell parameters.* In the orthonormal basis the unit cell basis vectors have components $\mathbf{a_c} = [a_x\ a_y\ a_z]$, $\mathbf{b_c} = [b_x\ b_y\ b_z]$, and $\mathbf{c_c} = [c_x\ c_y\ c_z]$. The unit cell lengths and the components are expressed in the same units as the unit vectors of the Cartesian system (e.g., Å):

$$\begin{aligned} \mathbf{a_c} &= a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k} \\ \mathbf{b_c} &= b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k} \\ \mathbf{c_c} &= c_x\mathbf{i} + c_y\mathbf{j} + c_z\mathbf{k}. \end{aligned} \tag{1.117}$$

Determining these 9 components will provide the transformation matrix that we seek. Since $\mathbf{a}$ lies along the $\mathbf{i}$ axis,

$$a_x = a \tag{1.118}$$

$$b_x = 0 \tag{1.119}$$

$$c_x = 0. \tag{1.120}$$

Since $\mathbf{b_c}$ lies in the $\mathbf{ij}$ plane its z component is zero. The dot product of $\mathbf{a_c}$ and $\mathbf{b_c}$ gives us b_x:

$$\begin{aligned} \mathbf{a_c} \cdot \mathbf{b_c} &= ab\cos\gamma \\ \mathbf{a_c} \cdot \mathbf{b_c} &= \mathbf{a_c}^T\mathbf{b_c} = \begin{bmatrix} a & 0 & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ 0 \end{bmatrix} = ab_x + 0 + 0 \implies \\ ab_x &= ab\cos\gamma \\ b_x &= b\cos\gamma. \end{aligned} \tag{1.121}$$

The dot product of $\mathbf{b_c}$ with itself provides b_y:

$$\begin{aligned} \mathbf{b_c} \cdot \mathbf{b_c} &= bb\cos(0) = b^2 \\ \mathbf{b_c} \cdot \mathbf{b_c} &= \mathbf{b_c}^T\mathbf{b_c} = \begin{bmatrix} b\cos\gamma & b_y & 0 \end{bmatrix} \begin{bmatrix} b\cos\gamma \\ b_y \\ 0 \end{bmatrix} \\ &= b^2\cos^2\gamma + b_y^2 + 0 \implies \\ b^2 &= b^2\cos^2\gamma + b_y^2 \\ b_y^2 &= b^2(1 - \cos^2\gamma) \\ &= b^2\sin^2\gamma \\ b_y &= b\sin\gamma \end{aligned} \tag{1.122}$$

$$b_z = 0. \tag{1.123}$$

The x component of $\mathbf{c_c}$ is determined from the dot product of $\mathbf{a_c}$ and $\mathbf{c_c}$:

$$\begin{aligned}
\mathbf{a_c} \cdot \mathbf{c_c} &= ac\cos\beta \\
\mathbf{a_c} \cdot \mathbf{c_c} &= \mathbf{a_c}^T\mathbf{c_c} = \begin{bmatrix} a & 0 & 0 \end{bmatrix} \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} \\
&= ac_x + 0 + 0 \implies \\
ac_x &= ac\cos\beta \\
c_x &= c\cos\beta. \qquad (1.124)
\end{aligned}$$

The y component of $\mathbf{c_c}$ is determined from the dot product of $\mathbf{b_c}$ and $\mathbf{c_c}$:

$$\begin{aligned}
\mathbf{b_c} \cdot \mathbf{c_c} &= bc\cos\alpha \\
\mathbf{b_c} \cdot \mathbf{c_c} &= \mathbf{b_c}^T\mathbf{c_c} = \begin{bmatrix} b\cos\gamma & b\sin\gamma & 0 \end{bmatrix} \begin{bmatrix} c\cos\beta \\ c_y \\ c_z \end{bmatrix} \\
&= bc\cos\beta\cos\gamma + c_y b\sin\gamma + 0 \implies \\
bc\cos\alpha &= bc\cos\beta\cos\gamma + c_y b\sin\gamma \\
c_y b\sin\gamma &= bc(\cos\alpha - \cos\beta\cos\gamma) \\
c_y &= \frac{c(\cos\alpha - \cos\beta\cos\gamma)}{\sin\gamma} \qquad (1.125)
\end{aligned}$$

Finally, the z component of $\mathbf{c_c}$ is obtained from the length of the axis and the x and y components:

$$\begin{aligned}
c^2 &= c_x^2 + c_y^2 + c_z^2 \\
c_z^2 &= c^2 - c_x^2 - c_y^2 \\
&= c^2 - c^2\cos^2\beta - \frac{c^2(\cos\alpha - \cos\beta\cos\gamma)^2}{\sin^2\gamma} \\
&= c^2\left(\frac{\sin^2\gamma - \cos^2\beta\sin^2\gamma - (\cos\alpha - \cos\beta\cos\gamma)^2}{\sin^2\gamma}\right) \\
&= \frac{c^2}{\sin^2\gamma}(\sin^2\gamma - \cos^2\beta\sin^2\gamma - \cos^2\alpha - \cos^2\beta\cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)
\end{aligned}$$

Substituting $-\cos^2\beta\sin^2\gamma - \cos^2\beta\cos^2\gamma = -\cos^2\beta(\sin^2\gamma + \cos^2\gamma) = -\cos^2\beta$ and $\sin^2\gamma = 1 - \cos^2\gamma$,

$$\begin{aligned}
c_z^2 &= \frac{c^2}{\sin^2\gamma}(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma). \\
c_z &= \frac{c(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)^{\frac{1}{2}}}{\sin\gamma}. \qquad (1.126)
\end{aligned}$$

Later in the chapter we will derive an expression for the unit cell volume (Eqn. 1.5.4): $V = abc(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)^{\frac{1}{2}}$. Multiplying the expression for c_z by ab/ab results in

$$c_z = \frac{V}{ab\sin\gamma}. \qquad (1.127)$$

We have now determined the components of the axial vectors in a Cartesian coordinate system, based on the unit cell parameters:

$$\begin{aligned}
\mathbf{a_c} &= a\ \mathbf{i} + 0\ \mathbf{j} + 0\ \mathbf{k} \\
\mathbf{b_c} &= b\cos\gamma\ \mathbf{i} + b\sin\gamma\ \mathbf{j} + 0\ \mathbf{k} \\
\mathbf{c_c} &= c\cos\beta\ \mathbf{i} + \left(\frac{c(\cos\alpha - \cos\beta\cos\gamma)}{\sin\gamma}\right)\ \mathbf{j} + \left(\frac{V}{ab\sin\gamma}\right)\ \mathbf{k}.
\end{aligned} \tag{1.128}$$

The matrix of coefficients is

$$\begin{aligned}
\mathbf{C} &= \begin{bmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{bmatrix} \\
&= \begin{bmatrix} a & 0 & 0 \\ b\cos\gamma & b\sin\gamma & 0 \\ c\cos\beta & \left(\frac{c(\cos\alpha - \cos\beta\cos\gamma)}{\sin\gamma}\right) & \frac{V}{ab\sin\gamma} \end{bmatrix}.
\end{aligned} \tag{1.129}$$

The matrix that will transform a vector in fractional coordinates based on a unit cell with cell parameters $a,\ b,\ c,\ \alpha,\ \beta\ \gamma$, and V into an orthonormal basis is therefore:

$$\mathbf{B} = \mathbf{C}^T = \begin{bmatrix} a & b\cos\gamma & c\cos\beta \\ 0 & b\sin\gamma & \left(\frac{c(\cos\alpha - \cos\beta\cos\gamma)}{\sin\gamma}\right) \\ 0 & 0 & \frac{V}{ab\sin\gamma} \end{bmatrix}. \tag{1.130}$$

$$\mathbf{B}\begin{bmatrix} x_f \\ y_f \\ z_f \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}. \tag{1.131}$$

In Chapter 3 a basis related to the diffraction pattern and known as a the *reciprocal basis* will be introduced. Each cell parameter in the unit cell basis is related to the cell parameters in the reciprocal basis, and we will later use these relationships to simplify **B**.

To transform a vector in Cartesian coordinates into unit cell coordinates we need only invert the **B** matrix:

$$\mathbf{B}\mathbf{v_f} = \mathbf{v_c} \tag{1.132}$$

$$\begin{aligned}
\mathbf{B^{-1}B}\mathbf{v_f} &= \mathbf{B^{-1}}\mathbf{v_c} \\
\mathbf{I}\mathbf{v_f} &= \mathbf{B^{-1}}\mathbf{v_c}
\end{aligned}$$

$$\mathbf{B^{-1}}\mathbf{v_c} = \mathbf{v_f} \tag{1.133}$$

1.5.3 Determining Distances and Angles In the Unit Cell

A virtual plethora of crystal structures are now available to the scientific investigator, either in the published literature or archived in databases[22]. These structures contain the unit cell parameters and the $[x\ y\ z]$ coordinates of the atoms inside the unit cell. It is unlikely that there is a crystallographer alive who has not been confronted by frustrated colleagues who have taken the atomic coordinates from

a published structure and attempted to calculate interatomic distances and angles from them. The reason, of course, is that they are treating the coordinates as Cartesian coordinates in an orthonormal basis, when the coordinates are actually listed as fractional coordinates in the unit cell basis. While there are software programs available which will accept fractional coordinates[2], allowing the user to determine molecular parameters, the use of the transform derived in the previous section is very straightforward, as an example using the 2-mercaptopyridine structure will illustrate.

The crystallographic data in the literature will ordinarily list the unit cell parameters, including the cell volume and the *space group* (the internal crystal symmetry – discussed in Sec. 2.4) along with other information related to data collection and refinement of the structure. Somewhere in the manuscript or archive* will be a list of the fractional coordinates of the atoms in the unit cell, labeled as x, y, and z (rather that x_f, etc.):

> The title compound, 2-mercaptopyridine, crystallized in the monoclinic space group $P2_1/n$, with $a = 6.112(5)$ Å, $b = 6.326(5)$ Å, $c = 14.314(5)$ Å, $\beta = 101.530(5)^\circ$, $V = 542.3(6)$ Å^3, $Z = 4$, T = 293 K.

The numbers in parentheses after each of the unit cell parameters are the estimated standard deviations of the parameters. They are a measure of the uncertainty in the last digit (e.g., for a, the standard deviation of the axial length is 0.005 Å). Standard deviations will be discussed in detail in Chapters 5 and 8.

Positional Parameters of 2-Mercaptopyridine

Atom	x	y	z
S(1)	0.7403	0.0629	0.4073
H(1)	0.8714	0.0983	0.4799
N(1)	0.3705	0.2616	0.4294
C(1)	0.5501	0.2608	0.3860
C(2)	0.5630	0.4351	0.3256
H(2)	0.683	0.431	0.292
C(3)	0.4070	0.5920	0.3147
H(3)	0.423	0.693	0.271
C(4)	0.2289	0.5854	0.3617
H(4)	0.108	0.682	0.347
C(5)	0.2151	0.4162	0.4187
H(5)	0.083	0.368	0.453

Figure 1.32 displays a *displacement ellipsoid* plot[23] of the 2-mercaptopyridine molecule showing the atom labeling scheme. The displacement ellipsoids are informative in that they can tell us something about how the molecule vibrates (in good structures) or alternatively characterize the quality of the diffraction data (in not-so-good structures). For our purposes here we only note that this is typical of drawings in the literature and that the centers of the atoms lie at the centers of the ellipsoids. Note that not all of the unit cell parameters are listed. We will discover in Sec. 2.3.6 that some of the unit cell angles are fixed due to the lattice symmetry.

*Atomic positions were routinely published in the older literature. Current journals generally refer the reader to an archival database where the atomic parameters are stored.

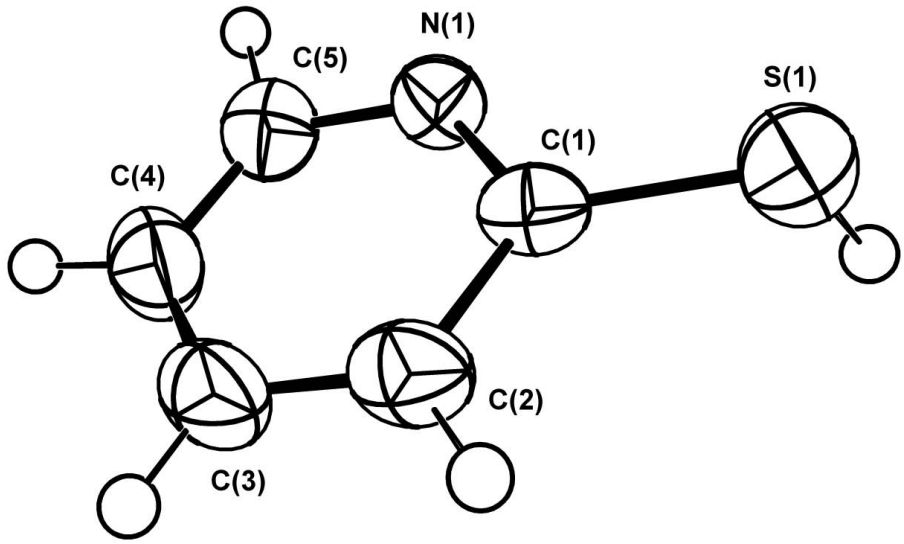

Figure 1.32 Displacement ellipsoid plot of the 2-mercaptopyridine molecule. Ellipsoids are plotted at the 50% level.

It is common practice to list only those parameters which are determined experimentally, rather than assigned by constraint. In this example these fixed angles are $\alpha = \gamma = 90°$. In a substantial majority of the structures found in the literature the unlisted angles will be 90°, *but not always*.

Suppose that we wish to know the interatomic distance (bond length) between S(1) and C(1), and the N(1)-C(1)-S(1) interatomic angle with C(1) at the vertex. To calculate these parameters we must first convert the fractional coordinates of the atom positions to Cartesian coordinates (in Å), requiring the **B** matrix:

$$\mathbf{B} = \begin{bmatrix} 6.112 & \{6.326\cos(90°)\} & \{14.314\cos(101.53°)\} \\ 0 & \{6.326\sin(90°)\} & \dfrac{14.314\{\cos(90°)-\cos(101.53°)\cos(90°)\}}{\sin(90°)} \\ 0 & 0 & \dfrac{542.3}{(6.112)(6.326)\sin(90°)} \end{bmatrix}.$$

$$= \begin{bmatrix} 6.112 & 0 & -2.861 \\ 0 & 6.326 & 0 \\ 0 & 0 & 14.025 \end{bmatrix}.$$

Using Eqn. 1.131 for the S(1) fractional coordinates transforms [(0.74025) (0.06290) (0.40731)] into Cartesian coordinates in Å:

$$\begin{bmatrix} 6.112 & 0 & -2.861 \\ 0 & 6.326 & 0 \\ 0 & 0 & 14.025 \end{bmatrix} \begin{bmatrix} 0.74025 \\ 0.06290 \\ 0.40731 \end{bmatrix} = \begin{bmatrix} 3.359 \\ 0.398 \\ 5.712 \end{bmatrix}.$$

The remainder of the coordinates, which we will call *crystal Cartesian coordinates*, allow for the calculation of any distance or angle in the molecule:

Crystal Cartesian Coordinates

Atom	x_c	y_c	z_c
S(1)	3.359	0.398	5.712
H(1)	3.953	0.622	6.731
N(1)	1.036	1.655	6.022
C(1)	2.258	1.650	5.414
C(5)	0.117	2.633	5.872
C(2)	2.510	2.752	4.567
C(4)	0.364	3.703	5.073
C(3)	1.587	3.745	4.414
H(5)	−0.789	2.328	6.353
H(2)	3.339	2.727	4.095
H(3)	1.810	4.384	3.801
H(4)	−0.333	4.31	4.867

The components of the $\overrightarrow{\mathrm{C(1)S(1)}}$ vector translated to the origin are

$$[(3.359)\ (0.398)(5.712)] - [(2.258)\ (1.650)\ (5.414)] = [(1.101)\ (-1.252)\ (0.298)].$$

The carbon-sulfur bond length is $|\overrightarrow{\mathrm{C(1)S(1)}}| = (1.101^2 + (-1.252)^2 + 0.298^2)^{\frac{1}{2}} = 1.694$ Å. The components of the $\overrightarrow{\mathrm{C(1)N(1)}}$ vector translated to the origin are [(-1.222) (0.005) (0.608)], resulting in a carbon-nitrogen bond length of $|\overrightarrow{\mathrm{C(1)N(1)}}| = 1.365$ Å. The N(1)-C(1)-S(1) angle, ν, is determined from the dot product of these two vectors: $\overrightarrow{\mathrm{C(1)S(1)}} \cdot \overrightarrow{\mathrm{C(1)N(1)}} = [(1.101)(-1.222) + (-1.252)(0.005) + (0.298)(0.608)] = $ -1.170.

$$\overrightarrow{\mathrm{C(1)S(1)}} \cdot \overrightarrow{\mathrm{C(1)N(1)}} = \left(|\overrightarrow{\mathrm{C(1)S(1)}}|\right)\left(|\overrightarrow{\mathrm{C(1)N(1)}}|\right)\cos\nu = -1.170.$$

$$\cos\nu = \frac{-1.170}{\left(|\overrightarrow{\mathrm{C(1)S(1)}}|\right)\left(|\overrightarrow{\mathrm{C(1)N(1)}}|\right)} = \frac{-1.170}{(1.694)(1.365)} = -0.506.$$

$$\nu = 120.4^\circ.$$

1.5.4 Determining the Volume of the Unit Cell

The unit cell volume, which we have already used to simplify the **B** matrix, is also a useful parameter for the crystallographer during the early stages of structural investigation. With the cell volume in hand, the density of the crystal under investigation can be calculated and tested for consistency with the putative contents of the unit cell. A strange density usually means that the cell does not contain what it is thought to contain, or that the unit cell is incorrect. The general unit cell, shown in Fig. 1.33 is a parallelepiped, based on axes **a**, **b**, and **c**. The volume of a parallelepiped is determined as the area of its base, A_{bc}, times its height, h_a. The base, in turn, is a parallelogram with an area determined as the length of its

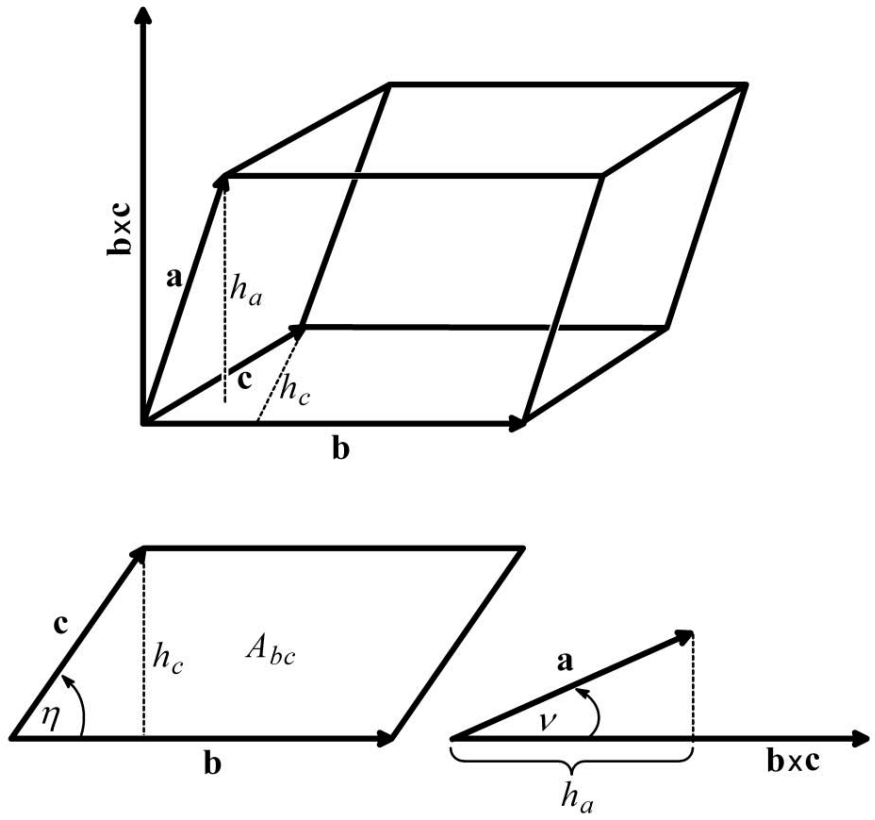

Figure 1.33 Unit cell parallelepiped. The **bc** vectors determine the base, with the perpendicular h_a the height. The b axis is the base of the parallelogram that makes up the base of the parallelepiped. The height of the parallelogram is the perpendicular, h_c.

base, b, times its height, h_c. The area of the base can be determined from the vector product of **b** and **c**:

$$\begin{aligned} A_{bc} &= bh_c \\ h_c &= c\sin\eta \\ |\mathbf{b}\times\mathbf{c}| &= bc\sin\eta = bh_c \\ A_{bc} &= |\mathbf{b}\times\mathbf{c}|. \end{aligned} \tag{1.134}$$

The height of the parallelepiped, h_a, is a perpendicular from the base, and is therefore parallel to $(\mathbf{b}\times\mathbf{c})$. The dot product of $\mathbf{a}$ and $(\mathbf{b}\times\mathbf{c})$ determines h_a, which is the projection of $\mathbf{a}$ onto $(\mathbf{b}\times\mathbf{c})$ (Eqn. 1.6):

$$\begin{aligned} V &= A_{bc}h_a \\ h_a &= \frac{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})}{|\mathbf{b}\times\mathbf{c}|} \\ &= \frac{\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})}{A_{bc}} \\ V &= \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}). \end{aligned} \tag{1.135}$$

The scalar, $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$,* is known as a *scalar triple product*; the volume of the unit cell is the scalar triple product of unit cell vectors. The scalar triple product provides us with the means to determine the cell volume without having to transform the unit cell parameters into orthonormal coordinates. We begin with Eqn. 1.71:

$$\begin{aligned} \mathbf{b}\times\mathbf{c} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} = \mathbf{i}\begin{vmatrix} b_y & b_z \\ c_y & c_z \end{vmatrix} - \mathbf{j}\begin{vmatrix} b_x & b_z \\ c_x & c_z \end{vmatrix} + \mathbf{k}\begin{vmatrix} b_x & b_y \\ c_x & c_y \end{vmatrix} \\ &= \left[\begin{vmatrix} b_y & b_z \\ c_y & c_z \end{vmatrix}\left(-\begin{vmatrix} b_x & b_z \\ c_x & c_z \end{vmatrix}\right)\begin{vmatrix} b_x & b_y \\ c_x & c_y \end{vmatrix}\right] \\ \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= a_x\begin{vmatrix} b_y & b_z \\ c_y & c_z \end{vmatrix} - a_y\begin{vmatrix} b_x & b_z \\ c_x & c_z \end{vmatrix} + a_z\begin{vmatrix} b_x & b_y \\ c_x & c_y \end{vmatrix} \\ &= \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}. \end{aligned} \tag{1.136}$$

The volume of the unit cell is the determinant of the matrix consisting of row vectors composed of the Cartesian coordinates of the unit cell axes. Thus,

$$\begin{aligned} V &= \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} \\ V^2 &= \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} \\ &= \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}\begin{vmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{vmatrix} \quad \text{(Determinant Property 3)} \end{aligned} \tag{1.137}$$

*Note that $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) \equiv \mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$ is unambiguous, since $(\mathbf{a}\cdot\mathbf{b})\times\mathbf{c}$ is meaningless.

$$
\begin{aligned}
V^2 &= \left| \begin{bmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{bmatrix} \begin{bmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{bmatrix} \right| \quad \text{(Determinant Property 4)} \\
&= \begin{vmatrix} \mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c} \\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot\mathbf{b} \\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot\mathbf{c} \end{vmatrix} \\
&= \begin{vmatrix} a^2 & ab\cos\gamma & ac\cos\beta \\ ab\cos\gamma & b^2 & bc\cos\alpha \\ ac\cos\beta & bc\cos\alpha & c^2 \end{vmatrix}. \qquad (1.138)
\end{aligned}
$$

Expansion of the determinant gives

$$
\begin{aligned}
V^2 &= (a^2b^2c^2) + (a^2b^2c^2\cos\alpha\cos\beta\cos\gamma) + (a^2b^2c^2\cos\alpha\cos\beta\cos\gamma) \\
&\quad -(a^2b^2c^2\cos^2\beta) - (a^2b^2c^2\cos^2\alpha) - (a^2b^2c^2\cos^2\gamma). \\
&= a^2b^2c^2(1-\cos^2\alpha-\cos^2\beta-\cos^2\gamma+2\cos\alpha\cos\beta\cos\gamma). \\
V &= abc(1-\cos^2\alpha-\cos^2\beta-\cos^2\gamma+2\cos\alpha\cos\beta\cos\gamma)^{\frac{1}{2}}. \qquad (1.139)
\end{aligned}
$$

This is the expression used to simplify the **B** matrix (Eqn. 1.126). The matrix of the determinant in Eqn. 1.138 contains all of the metrics of the unit cell and is known as *the metric tensor** of the lattice. We will encounter the metric tensor later on, as it is generally obtained experimentally without prior knowledge of the unit cell parameters, and provides a means for obtaining them!

1.5.5 Important Identities

In Chapter 3 a new lattice will be introduced which is reciprocal to the crystal lattice. The determination of the relationships between the basis vectors in this new lattice and those in the crystal lattice will require identities relating the basis vectors. These identities involve combinations of scalar and vector products, and will be developed here for later use.

The *scalar triple product*, $\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$ has already been introduced (Eqn. 1.136) and shown to be equal to the unit cell volume. It is useful to generate the remaining two scalar triple products from the first in order to determine the vector products that will yield positive cell volumes. We do this by switching rows in the determinant representation of $\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$, recalling that each switch changes the sign of the determinant (and therefore the volume):

$$
\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} = -\begin{vmatrix} b_x & b_y & b_z \\ a_x & a_y & a_z \\ c_x & c_y & c_z \end{vmatrix} = \begin{vmatrix} b_x & b_y & b_z \\ c_x & c_y & c_z \\ a_x & a_y & a_z \end{vmatrix}
$$

$$
= \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) \qquad (1.140)
$$

*The word *tensor* will be used here in its "physics" context, as an entity that characterizes the properties of a physical system. All of the tensors that we will encounter will be represented by 3×3 matrices and we will often use *tensor* and *matrix* interchangeably.

and

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} = -\begin{vmatrix} c_x & c_y & c_z \\ b_x & b_y & b_z \\ a_x & a_y & a_z \end{vmatrix} = \begin{vmatrix} c_x & c_y & c_z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}$$

$$= \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}). \tag{1.141}$$

It follows that $V = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a})$ and $V = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$. Note that reversal of any of the vector products produces a negative volume for the unit cell. This is a useful diagnostic during the course of the structural investigation. If a negative unit cell volume is determined, one of the axes is pointing in a direction that renders the coordinate system left-handed. Reversing the direction of one of the axes in the vector product will correct this.

The *vector triple product* is the vector analog of the scalar triple product: $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. The vector $\mathbf{b}\times\mathbf{c}$ is perpendicular to the *bc* plane. The vector triple product produces a vector that is perpendicular to that vector, and therefore lies in the *bc* plane. Its components are determined from the scalar products of $\mathbf{a}$ with $\mathbf{b}$ and $\mathbf{c}$:

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}. \tag{1.142}$$

We prove this by showing that the expression on the right reduces to the vector triple product:

$$\begin{aligned}(\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c} &= (a_xc_x + a_yc_y + a_zc_z)(b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k}) \\ &- (a_xb_x + a_yb_y + a_zb_z)(c_x\mathbf{i} + c_y\mathbf{j} + c_z\mathbf{k}).\end{aligned}$$

Expanding and collecting terms gives

$$\begin{aligned}(\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c} &= [a_y(b_xc_y - b_yc_x) - a_z(b_zc_x - b_xc_z)]\,\mathbf{i} \\ &+ [a_z(b_yc_z - b_zc_y) - a_x(b_xc_y - b_yc_x)]\,\mathbf{j} \\ &+ [a_x(b_zc_x - b_xc_z) - a_y(b_yc_z - b_zc_y)]\,\mathbf{k}.\end{aligned}$$

We define $\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k}$ such that $v_x = b_yc_z - b_zc_y$, $v_y = b_zc_x - b_xc_z$, and $v_z = b_xc_y - b_yc_x$, that is, $\mathbf{v} = \mathbf{b}\times\mathbf{c}$. Substituting these components:

$$\begin{aligned}(\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c} &= (a_yv_z - a_zv_y)\,\mathbf{i} \\ &+ (a_zv_x - a_xv_z)\,\mathbf{j} \\ &+ (a_xv_y - a_yv_x)\,\mathbf{k} \\ &= \mathbf{a}\times\mathbf{v} = \mathbf{a}\times(\mathbf{b}\times\mathbf{c}).\end{aligned}$$

The *scalar quadruple product* is the scalar product of two vector products, $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})$:

$$(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}). \tag{1.143}$$

The scalar triple product and the vector triple product are employed to prove this:

$$
\begin{aligned}
&\text{Let } \mathbf{v} = \mathbf{c} \times \mathbf{d}. \\
&(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{v} = \mathbf{v} \cdot (\mathbf{a} \times \mathbf{b}) \quad (\textit{scalar triple product}) \\
&= \begin{vmatrix} v_x & v_y & v_z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = - \begin{vmatrix} a_x & a_y & a_z \\ v_x & v_y & v_z \\ b_x & b_y & b_z \end{vmatrix} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ v_x & v_y & v_z \end{vmatrix} \\
&= \mathbf{a} \cdot (\mathbf{b} \times \mathbf{v}) = \mathbf{a} \cdot [\mathbf{b} \times (\mathbf{c} \times \mathbf{d})] \quad (\textit{vector triple product}) \\
&= \mathbf{a} \cdot [(\mathbf{b} \cdot \mathbf{d})\,\mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\,\mathbf{d}] \\
&= (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}).
\end{aligned}
$$

The scalar quadruple product is conveniently represented as a 2×2 determinant:

$$
(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}. \tag{1.144}
$$

The *vector quadruple product* is the vector product of two vector products, $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})$:

$$
\begin{aligned}
&\text{Let } \mathbf{v} = \mathbf{a} \times \mathbf{b}. \\
&\mathbf{v} \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{v} \cdot \mathbf{d})\mathbf{c} - (\mathbf{v} \cdot \mathbf{c})\mathbf{d} \quad (\textit{vector triple product}) \\
&= (\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b}))\mathbf{c} - (\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}))\mathbf{d} \\
&= (\mathbf{a} \cdot (\mathbf{b} \times \mathbf{d}))\mathbf{c} - (\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}))\mathbf{d} \\
&\quad \Longrightarrow (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{b} \times \mathbf{d})\mathbf{c} - (\mathbf{a} \cdot \mathbf{b} \times \mathbf{c})\mathbf{d}
\end{aligned} \tag{1.145}
$$

Exercises

1. The copper atoms depicted in Fig. 1.1 are arranged in a cubic unit cell. Each edge of the cell has the same length: $a = b = c = 3.6147$. Determine the distances between the planes with (a) (1 1 1), (b) (2 2 2), and (c) (3 3 3) indices.

2. Consider two 2-dimensional unit cells, each with the same axial lengths: $a = 2.40$ and $b = 3.20$. For unit cell A, $\gamma = 90°$; for unit cell B, $\gamma = 117°$. A point p is located in each unit cell at the end of the sum of a vector of magnitude 1.80 , parallel to the **a** axis and a vector of magnitude 2.40 , parallel to the **b** axis. (a) Determine the fractional coordinates of point p in each unit cell. (b) Determine the distance from the origin to point p in each unit cell.

3. (a) Derive a formula for the inverse of a 2×2 matrix and use matrix multiplication to demonstrate that your formula is correct ($\mathbf{D}\mathbf{D}^{-1} = \mathbf{I}$). (b) Compute the inverse of the following matrix:

$$\mathbf{D} = \begin{bmatrix} 1.000 & 2.000 & 3.000 \\ 2.000 & 1.000 & 3.000 \\ 3.000 & 2.000 & 1.000 \end{bmatrix}.$$

   (c) Demonstrate that the matrix calculated in part (b) is $\mathbf{D}^{-1}$.

4. Show that (a) the inverse of a matrix for the rotation of angle φ about a coordinate axis (e.g., the x axis) is the rotation matrix for the $-\varphi$ rotation about the same axis, (b) the matrix for a reflection across a coordinate plane (e.g., the xz plane) is its own inverse, and (c) the inversion matrix is its own inverse.

5. Using matrices, show that sequential rotations about a coordinate axis of φ_1 followed by φ_2 is equivalent to a single rotation of $(\varphi_1 + \varphi_2)$ about the same axis.

6. Show that (a) the rotation matrices for rotation about the coordinate axes are orthonormal matrices, (b) the inverses of these matrices are their transposes, and (c) the rotation of a general vector effected by any of these matrices does not alter the length of the vector.

7. The monoclinic unit cell of CuO has the following parameters: $a = 4.6837(5)$ Å, $b = 3.4226(6)$ Å, $c = 5.1288(6)$ Å, $\alpha = 90.00°$, $\beta = 99.54°(1)$ and $\gamma = 90.00°$.* The fractional coordinates of the contents of the unit cell are

atom	x_f	y_f	z_f	atom	x_f	y_f	z_f
Cu1	0.2500	0.2500	0.0000	O1	0.0000	0.4184	0.2500
Cu2	0.7500	0.7500	0.5000	O2	0.5000	0.9184	0.2500
Cu3	0.2500	0.7500	0.5000	O3	0.0000	0.5816	0.7500
Cu4	0.7500	0.2500	0.5000	O4	0.5000	0.0816	0.7500

*Åsbrink, S. and Norrby, L.-J., *Acta. Cryst.*, B26, 8(1970).

(a) Determine the shortest (contact) distance between the copper(II) ions and the oxide ions in the unit cell. (b) Determine the volume of the unit cell. (c) Determine the mass of the unit cell in grams. (d) Determine the density of solid copper(II) oxide in g/cm^3.

8. The orthorhombic unit cell of $CuSO_4$ has the following parameters:* $a = 8.39$ Å, $b = 6.89$ Å $c = 4.83$ Å, $\alpha = 90.00°$, $\beta = 90.00°$ and $\gamma = 90.00°$

   The fractional coordinates of the basic unit in the unit cell (repeated in order to fill the cell) are

atom	x_f	y_f	z_f	atom	x_f	y_f	z_f
Cu1	0.000	0.000	0.893	O2	0.375	0.250	0.439
S1	0.185	0.250	0.445	O3	0.129	0.069	0.307
O1	0.141	0.250	0.755				

   (a) Determine the average sulfur-oxygen distance and the average O-S-O angle in the sulfate ion. (b) The experimentally measured density of anhydrous copper sulfate is 3.6 g/cm^3. How Many $CuSO_4$ units are in the unit cell? (c) The basic unit in the unit cell seems to be missing an oxygen atom. How can this be if the stoichiometry in the crystal is $CuSO_4$? Hint: The basic unit in the unit cell is called the *asymmetric* unit. You may have to read ahead in Chapter 2 to answer this question. The space group of the crystal is *Pnma*.

*Rao, B.R., *Acta. Cryst.*, 14, 321(1961).