Acta Cryst. (1958). **11**, 199

Diffraction by Helical Structures

By A. Klug*

Birkbeck College Crystallography Laboratory, University of London, England

F. H. C. Crick† and H. W. Wyckoff‡

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge, England

(*Received* 20 *June* 1957 *and in revised form* 30 *October* 1957)

The symmetry of helical structures and their diffraction patterns is discussed, and a list is given of the line groups for enantiomorphic helical structures. The main body of the paper concerns two special kinds of projection of a helical structure—the radial projection and the helical projection. It is shown that these projections provide a very convenient way of thinking about a helical structure and analysing its diffraction pattern. The theory of these projections is given in detail, and their uses are discussed.

Introduction

Since it was first developed some years ago by Cochran, Crick & Vand (1952; hereafter referred to as C.C.V. for short) and by Stokes (unpublished), the theory of diffraction by a helical structure has been of great value in the analysis of the structure of many substances, notably the synthetic polypeptides (for example Cochran & Crick, 1952), deoxyribonucleic acid (for example Wilkins, Stokes & Wilson, 1953; Franklin & Gosling, 1953), tobacco mosaic virus (for example Watson, 1954; Franklin & Klug, 1955) and collagen (for example Cohen & Bear, 1953; Cowan, North & Randall, 1953), to name only some of the most important.

In essence the results of C.C.V. contain the full diffraction theory, in much the same way as three-dimensional Fourier series analysis contains the theory of diffraction by a crystal lattice. But, just as in the latter there have been found many special techniques for considering the problem, so will there also be other ways of considering helical diffraction theory which are, or may prove to be, useful in practice. It is the purpose of this paper to give some of these ways of thinking about a helical structure.

The type of structure which we shall consider is one which is infinite in one special direction, which we shall call the 'fibre direction', and finite in all directions perpendicular to this. We shall consider that the structure repeats exactly after a distance c in the fibre direction, since any real structure can be made to approximate to this as closely as we please. Thus, the Fourier transform of such a structure will fall on to discrete layer planes, but will be continuous on each layer plane. Naturally, in actual specimens, the material may form a three-dimensional lattice, but the effect of this is easily allowed for—for example, by appropriate samplings of the semi-continuous Fourier transform.

Before embarking on the algebraical treatment we shall consider certain general properties of structures of this kind, including special projections and matters of symmetry.

I.1. Symmetry and special projections

Let us first consider the symmetry elements, which as usual must form a group. Then for a structure of this type—that is, infinite in one direction only—there will be at least one infinite straight line, the 'fibre axis', which every symmetry element of the group will turn into itself. We shall take this particular line (it is usually unique) as the z axis. Notice that the distance of any point from this line is unaltered by the operation of any symmetry element. Thus all the points produced by the operation of all the possible symmetry operations upon one chosen point will be equidistant from the z axis. That is, they will lie on a cylindrical sheet co-axial with the z axis.

Now consider two particular projections. The first is the projection of the structure parallel to the z axis. This will have the symmetry of one of the two-dimensional plane point groups, though naturally without the restrictions usually imposed by crystallographers. These point groups are of two kinds: the

* Work supported by the Nuffield Foundation.

† Part of this work is taken from a Ph.D. thesis accepted by the University of Cambridge in 1953.

‡ Some of the material presented herein has been described in a Ph.D. thesis by H. W. Wyckoff entitled *X-ray Diffraction Analysis of the Structure of Deoxyribonucleic Acid*, accepted by the Massachusetts Institute of Technology in 1955. This portion of the investigation was supported in part by Research Grant C-1780 from the National Cancer Institute of the National Institutes of Health, U.S. Public Health Service, under the supervision of Richard S. Bear. Further work by H. W. W. was carried out during the tenure of a U.S. Public Health Service, National Institute of Arthritis and Metabolic Diseases Fellowship at the Cavendish Laboratory, Cambridge, England.

cyclic point-groups, C_N, which have only an N-fold rotation axis, where N is *any* integer (not merely 1, 2, 3, 4 and 6); and the point groups, C_{Nv}, which have in addition mirror reflexion across radial lines, there being N such lines.

The second kind of projection is the more useful of the two for discussion of the symmetry. Consider a cylindrical surface which is co-axial with the z axis. Project the structure on to this surface along lines starting from the z axis and perpendicular to it—that is, along radial lines. We shall call the result a 'radial projection'. However, we shall usually think of it in a slightly modified form. Imagine such a cylindrical surface cut along a straight line, parallel to the z axis, and then opened out flat,* and imagine that identical sheets of this type are laid side by side, in register, till the plane pattern extends to infinity in a direction perpendicular to the z axis as well as parallel to it. It is this infinite two-dimensional plane pattern—the reiterated radial projection—that we shall refer to simply as the radial projection. It is clear that it will have the symmetry of one of the two-dimensional plane groups. These plane groups are listed in the *International Tables* (1952, pp. 58–72). There are 17 of them, but our rule that any symmetry element must leave the z axis unmoved eliminates the last eight of them. Of the remaining nine, the first two are enantiomorphous and the other seven are non-enantiomorphous.

We shall not consider the non-enantiomorphous ones further, since they are only rarely required. The vast majority of helical structures are necessarily enantiomorphous, since they are displayed by materials which contain only one optical isomer. Moreover, detailed examination of the non-enantiomorphous groups shows that a helical molecule is very unlikely to display such symmetries, for a variety of reasons.

Since the radial projections of the enantiomorphous groups are of only two kinds, namely $p1$ and $p211$, (*International Tables*, 1952, p. 58) the enumeration of the possible groups is simple. The only symmetry operations are:

t—a translation parallel to the z axis.
r—a rotation of $2\pi/N$ radians about the z axis, where $|N|$ is any positive integer greater than one.
s—a screw displacement; that is, a translation parallel to z together with a rotation of $2\pi/M$ radians about the z axis, where we have artificially restricted M to being a rational number written u/t in later sections (also $|M| \neq 1$).
2—a twofold rotation about a line passing through the z axis and perpendicular to it.

* If one wishes to distinguish between left-handed and right-handed helices then it is advisable to adhere to a fixed convention, say, that the cylindrical surface is opened out with the inside face upwards. The basic helix in Fig. 1 would then represent a right-handed helix.

It turns out that, omitting the case where there is no symmetry whatsoever (not even a translation), there are eight possible combinations of these symmetry elements. These are obtained by choosing

(*a*) either a translation (**t**) or a screw axis (**s**);
(*b*) either a parallel N-fold rotation axis (**r**), or not;
(*c*) either a perpendicular dyad axis (**2**), or not.

All combinations of these three choices are possible, giving the eight cases set out in Table 1. We tentatively suggest the nomenclature indicated above, putting the **2**, when it occurs, at the end. The symbol **t** could be omitted but is included for clarity.

Table 1. *The eight enantiomorphic line groups*

Proposed symbol	**t**	**t2**	**tr**	**tr2**	**s**	**s2**	**sr**	**sr2**
Position of z axis	Non-U	Non-U	U	U	U	U	U	U
Minimum number of asymmetric chains	1	2	N	$2N$	1	2	N	$2N$

U = unique. Non-U = not unique.
N refers to the N-fold parallel rotation axis.

Naturally the ones that will concern us most are the last four, since these apply to truly helical structures, whereas the first four could more aptly be called cylindrical. In place of the number of asymmetric units in the repeat, which depends on the particular screw axis involved, we may usefully specify the 'minimum number of asymmetric chains'. This can be defined as follows. Let the line group be represented by a discrete number of infinite continuous lines *in general positions*, that is, not passing through a symmetry axis. What is the minimum number possible? The answers are included in the table. Note that the actual number of chemical chains in a structure may be a multiple of this minimum number.

The first four line groups can usefully be regarded as special cases of the last four in which the angle of rotation of the screw axis has become zero. Thus in isolated fibrous molecules we should expect them to occur only very occasionally. In an assembly of fibrous molecules they may perhaps occur more often because of interactions between molecules, but even then they may be rather rare.

We note briefly that the symmetry suggested for the structure of deoxyribonucleic acid (DNA) is **s2** (Watson & Crick, 1953), while that of the proposed structure of polyadenylic acid (see Watson, 1957) is **sr**, $N = 2$. No case of **sr2** has so far been reported.

I.2. The effects of symmetry elements on the Fourier transform

We shall adhere as far as possible to the notation in C.C.V.; (r, φ, z) and (R, ψ, ζ) are cylindrical coordinates in real and reciprocal space respectively. Consider first the simplest helical line groups, **s**,

without rotation axes of either kind. We have restricted ourselves to cases where there is an exact repeat after a distance c in the z direction. Thus the fundamental parameters of the helix, the translation p and the rotation $2\pi p/P$ are such that the ratio P/p can be expressed as a rational fraction u/t, where u and t are integer. (For the effect of a departure from an exact integer ratio see C.C.V. and Franklin & Klug (1955).) Thus we shall have $up = tP = c$. In other words there are u asymmetric units of the structure distributed evenly along exactly t turns of the *basic* helix. Note that this is not necessarily a physical description of the structure; it corresponds to a particular choice of the asymmetric unit, and this is by no means unique, as will be familiar to crystallographers.

The structure factor per asymmetric unit is for the lth layer-line (see C.C.V.)

$$F(R, \psi, l/c) = \sum_j \sum_n f_j J_n(2\pi R r_j) \times \exp\left[i\left\{n(\psi+\tfrac{1}{2}\pi) - n\varphi_j + \frac{2\pi l z_j}{c}\right\}\right] . \quad (1)^*$$

For any one atom there is a summation over the orders n of Bessel functions determined by the selection rule

$$l = tn + um , \quad (2)$$

and there is a further summation over all the atoms in the asymmetric unit, their co-ordinates being r_j, φ_j, z_j.

This result obtained by C.C.V. for helical structures is essentially a special case of the theory of Fourier transforms in cylindrical co-ordinates for (non-helical) structures periodic in z. The Fourier transform of such general structures will be finite only on a set of layer planes, on each of which the scattered amplitude will be of the form

$$F(R, \psi, \zeta) = \sum_{n=-\infty}^{\infty} A_n(R) \exp\left[in(\psi+\tfrac{1}{2}\pi)\right] ,$$

where

$$A_n(R) = \iiint \varrho(r, \varphi, z) J_n(2\pi R r) \times \exp\left[i(-n\varphi + 2\pi l z/c)\right] r\, dr\, d\varphi\, dz ,$$

or an analogous expression for the case of discrete atoms. This result is perfectly general. However, if the structure is helical there is a further rotational periodicity, linearly related to the translational periodicity in z. This has the effect of making many of the Bessel terms systematically zero, and only those obeying a relation between n and l can be finite. Thus we see that it is the selection rule that is the true characteristic of a helical structure, the appearance of Bessel functions in the theory being due to the use of cylindrical co-ordinates. This point of view may help to clarify the treatment of some of the problems considered later in this paper.

We must now consider the other helical line groups; in doing so we shall adhere to the notation already employed. This causes no difficulty except for those rare cases where the N-fold parallel rotation axis reduces the length of the true crystallographic repeat because N is a factor of u. For example if there were an 8-fold screw axis and a 2-fold parallel rotation axis the crystallographic repeat would be $\frac{1}{2}c$. For consistency in the algebra we shall, in such cases, use c to mean not the true repeat but the repeat the structure would have if the parallel rotation axis were absent.

The effect of the rotation axes on the general formula for the structure factors is easy to see. An N-fold parallel rotation axis makes all Bessel function contributions zero unless their order, n, is an integral multiple of N, so that this restriction becomes an additional selection rule.

A 2-fold perpendicular axis, parallel to the line $\varphi = 0$, causes equation (1) to be replaced by

$$F(R, \psi, l/c) = \sum_j \sum_n 2 f_j J_n(2\pi R r_j) \times \cos\left(-n\varphi_j + \frac{2\pi l z_j}{c}\right) \exp in(\psi+\tfrac{1}{2}\pi) , \quad (3)$$

where the sum goes over one asymmetric unit as before. Thus when $\psi = \pm\frac{1}{2}\pi$ all the structure factors are real, which is not surprising since this is the plane of the reciprocal lattice perpendicular to the dyad. Moreover, the phase of the contribution of any particular Bessel function varies in a predictable manner with ψ, since it depends only on the value of n, and not on the atomic parameters, apart from an ambiguity of π due to the ambiguity of the sign of the amplitude at $\psi = -\frac{1}{2}\pi$.

The symmetry of the Patterson function and of the intensity distribution in the reciprocal lattice (which are of course Fourier transforms of each other) follow the usual rules. That is, the u-fold screw axis in the real structure will be replaced in the Patterson by a u-fold rotation axis; the perpendicular dyads will, in the Patterson, all pass through the origin and a centre of symmetry will be added, so that mirror planes will be generated perpendicular to each even rotation axis. If U is the lowest common multiple of u (from the u-fold screw axis) and N (from the N-fold parallel rotation axis) then the net rotation axis of the Patterson is U-fold. The operation of the selection rules makes the difference between the orders of successive Bessel functions on the same layer line equal to U.

It is instructive to consider how the intensity varies as one travels round one layer plane at a constant distance from the ζ axis. This is most usefully discussed in terms of the number of Bessel functions contributing to that part of the reciprocal lattice. As pointed out in C.C.V., this number increases as R increases,

* This equation has been derived for a right-handed helix. In the case of a left-handed helix n must be replaced by $-n$ in equations (1) and (2).

though the effective number for any value of R varies from layer line to layer line.

If only one Bessel function contributes, as is usually the case near the meridian, it is easy to see from equation (1) that the intensity is *constant*, whereas the phase rotates at a uniform rate, making n complete revolutions for one circuit of the ζ axis (the Bessel function concerned being J_n).

If two Bessel functions contribute, the intensity will vary sinusoidally, going through U cycles in one complete traverse of the reciprocal lattice. In general, as pointed out by Stokes (1955), we may express the intensity as a Fourier series with respect to ψ:

$$I(R, \psi, l/c) = A_0+2A_1\cos(U\psi+\psi_1)+2A_2\cos(2U\psi+\psi_2)+\ldots . \quad (4)$$
(R constant, l/c constant).

If there are perpendicular dyads we can choose the origin so that $\psi_1 = \psi_2 = \psi_3 = \ldots = 0$.

It is easy to show that the highest term needed in this expansion at a particular value of R depends on the maximum *difference* in the orders of the contributing Bessel functions (with due allowance for sign of the orders), the number of cosine terms required being this difference divided by U. Thus usually the number of terms will equal the number of Bessel functions contributing. The only systematic exception to this occurs on the equator when U is odd, since in such a case, because of the centre of symmetry of the intensity distribution, the odd terms in the above expansion all vanish ($A_1 = A_3 = A_5 = \ldots = 0$).

I.3. Fibre diagrams

If crystal reflexions are present, a fibre diagram gives the average intensity for a finite set of values of ψ—those values at which reciprocal-lattice points occur. But, if the fibre diagram consists of layer-line streaks only, then it gives the intensity averaged over all values of ψ. It is a matter of some practical interest to obtain this average value for any proposed structure.

If the structure factors have been computed it is easy to show (Franklin & Klug, 1955) that the simplest procedure is to calculate the *intensity* due to each Bessel function, taken separately, and then add these intensities together (see equation (31) below). This gives correctly the average intensity over all values of ψ. However, if the structure is being studied by means of optical transforms the different Bessel functions on a given layer line cannot easily be separated, and an alternative procedure should be followed. The reciprocal lattice is sampled in a small number of planes of constant ψ (i.e. one makes optical transforms of the corresponding projections), and the *intensities* so obtained are averaged. This problem has been considered by Stokes (1955) but we shall treat it here in a slightly different manner by fixing attention on only one particular layer line at a time.

The way the intensity varies with ψ has already been set out in equation (4). It is clear that the number of samples, equally spaced in ψ, should equal the number of terms appearing in equation (4). Thus, normally, if three Bessel functions are contributing to the part of the layer line being considered, three samples, spaced at intervals of $\psi = \frac{1}{3}(2\pi/U)$, will average to the required average, A_0.

In the exceptional case mentioned earlier, which occurs on the equator when U is odd, it is possible to take rather fewer samples. It is easy to see from equation (1) that the number of samples can be equal to the number of Bessel functions of *positive* order (including zero) if the angular spacing between the samples is the appropriate sub-multiple of $2\pi/2U$ rather than of $2\pi/U$. However the general rule:

> 'a true average is obtained if the number of samples, uniformly spaced in the angular repeat of $2\pi/U$, equals the number of contributing Bessel functions',

is always true. In using this rule any Bessel function which is fortuitously absent within the sequence of Bessel functions should be counted as if it were present.

It is useful to remember that, since one normally sees both halves of an optical transform, a single optical transform will sample reciprocal space at points spaced $\psi = \pi$ apart. If U is odd the right- and left-hand sides of the optical transform may be different, except on the equator where the intensity at two such points is necessarily the same.

For structures with perpendicular dyads the intensity in any plane perpendicular to one of the dyads will be at a maximum or a minimum value as far as the variation with respect to ψ is concerned. In regions where only two Bessel functions (whose orders differ by U) are contributing, a single sample at an angular distance of $\pm\frac{1}{4}(2\pi/U)$ from one of these positions will give the average intensity.

II.1. Reciprocal relations between the (n, l) plot and the helix net

In this section we shall discuss the radial projection already described (§ I.1). We shall first consider the case in which there is only one atom in the asymmetric unit; then the case where there are several atoms in the asymmetric unit, but all at the same radius. This enables one to see in a simple way the relationship between the radial projection and the intensities of the diffraction pattern of the structure. Finally we shall discuss the general case in which atoms lie at several different radii.

(a) One atom in the asymmetric unit

An example of the reiterated radial projection (described in § I.1) is shown in Fig. 1, which represents

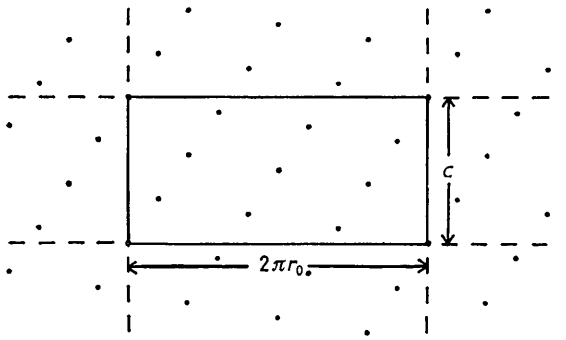

Fig. 1. The helix net (or radial projection) for a hypothetical case in which the axial repeat contains 10 units in 3 turns of the basic helix.

a helix with 10 asymmetric units in 3 turns. Bear (1955) has called a pattern of this kind a helix net, and we propose calling the points of the net *helical net points*.

The Fourier transform of this two-dimensional pattern will clearly be the net pattern reciprocal to Fig. 1, as shown in Fig. 2(*a*). However, instead of labelling the axes of this reciprocal net in reciprocal-lattice units, we have marked the horizontal axis '*n*' and the vertical axis '*l*', so that each point in Fig. 2(*a*) can be characterized by a pair of integers (n, l). It is easy to see from the selection rule of equation (2) that if the point (n, l) is present in Fig. 2(*a*) it implies that the Bessel function, J_n, will occur in the diffraction pattern of this helical structure on the lth layer line. In other words, Fig. 2(*a*) expresses geometrically the selection rule for Bessel functions. To bring this out more clearly we have included Fig. 2(*b*), in which lines are drawn joining up all the points having the same value of m (see equation (2)).

Now consider a particular point (n, l) in the reciprocal lattice shown in Fig. 2(*a*), for example $(2, \bar{4})$. This corresponds in the real two-dimensional space of Fig. 1 to a set of sinusoidal density waves. A very simple and rapid way of finding the *direction* of these waves (defined, for example, as the line *parallel* to the zero lines of the wave) is shown in Fig. 3. This is simply Fig. 1, but with the horizontal axis labelled l (backwards) and the vertical axis labelled n. Alternatively, it can be regarded as Fig. 2(*a*) (on some suitable scale) turned 90° anticlockwise. Then the line joining the origin to, say, the point $(2, \bar{4})$ in Fig. 3 is parallel to the wave corresponding to the point $(2, \bar{4})$ in reciprocal space. The spacing of the wave can be found by noting that the (n, l) set of waves makes n intercepts along the horizontal edge and l intercepts along the vertical edge of the basic rectangle of Fig. 1. Also note that in three dimensions n gives the number of helical lines (chains) necessary to cover all the helix net points, except when n and l have a common factor,

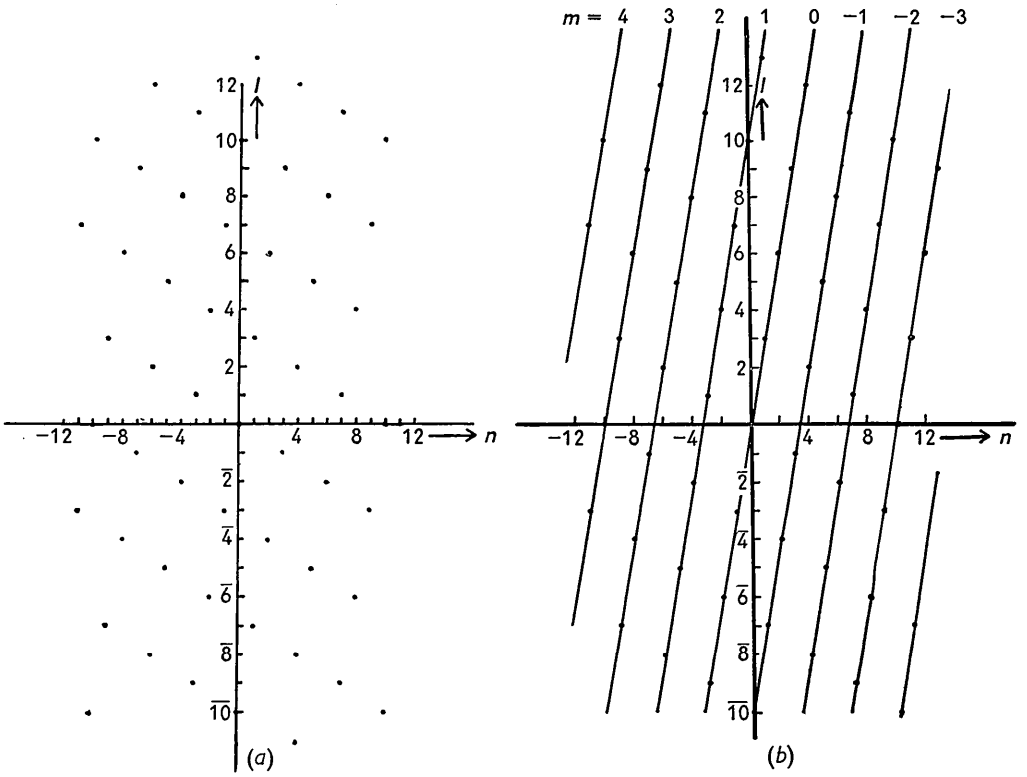

Fig. 2. (*a*) The net reciprocal to that of Fig. 1, read as an (n, l) plot.
(*b*) The (n, l) plot with the various 'branches' of the diffraction pattern labelled by the values of m.

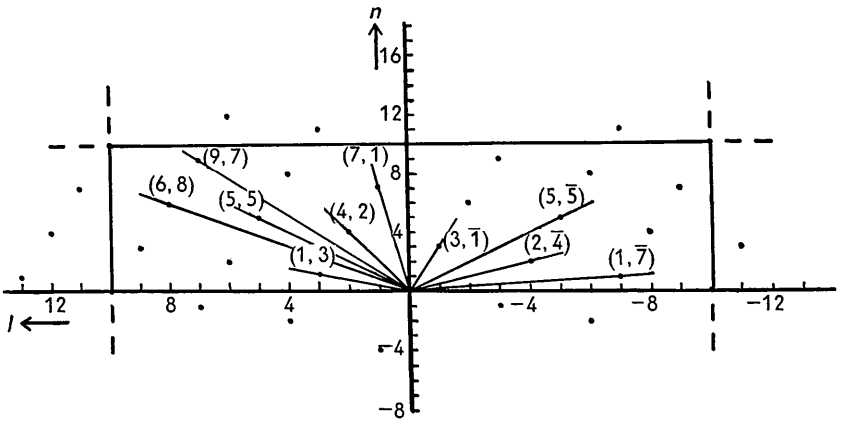

Fig. 3. Part of the radial projection of the example shown in Fig. 1. Each line is the projection of one of the continuous helices underlying the helix net, and is labelled with the appropriate pair of integers (n, l). (The discussion in the text shows that n and l play a role analogous to Miller indices.)

If we fix an origin and assign to each net point a horizontal co-ordinate l and a vertical co-ordinate n, as has been done in the figure, then the labelling of the continuous helices follows automatically.

when the number is equal to n divided by this factor. These rather obvious points are useful when considering the more complicated cases of radial projections (see below).

The other helical line groups can easily be dealt with. The existence of a parallel N-fold rotation axis means that the projection of the structure corresponds to N repeats in the horizontal direction of the infinite net shown in Fig. 1. The *scale* of n in Figs. 2 and 3 must correspondingly be multiplied by N, without in any way changing the pattern of points.

The existence of a perpendicular dyad implies that there will be two atoms, related by the dyad, at each net point in Fig. 1. However, we can retain the above description if we place the single atom *on* the dyad axis. The case when the atoms are not on the dyad is best considered as a special example of that treated in the next sub-section (*b*).

(*b*) *More than one atom in the asymmetric unit*

It is clear that there is a rather intimate *geometrical* relationship between the real net of Fig. 1 and the reciprocal net of Fig. 2. We shall now show that for the case in which all the atoms lie *at one radius* this relationship extends to the *intensity* distribution. As usual, we consider first the simplest helical line group, **s**.

Let all the atoms lie at radius r_0. We construct the radial projection as before, calling the co-ordinates of Fig. 1 (x, z) such that $x_j = -\varphi_j r_0$ and $z_j = z_j$. An example, with three atoms in the asymmetric unit, is shown in Fig. 4. The Fourier components of the radial projection will be

$$F(h, l) = \sum_j f_j \exp\left[2\pi i\left(\frac{hx_j}{a} + \frac{lz_j}{c}\right)\right], \qquad (5a)$$

where $a \equiv 2\pi r_0$. This has considerable similarity to equation (1). To bring this out more clearly we write

$$J_n(2\pi R r_0) \exp\left[in(\psi + \tfrac{1}{2}\pi)\right] \equiv B_n(R, \psi).$$

Then, since B_n is independent of j, we can write equation (1) for this case as

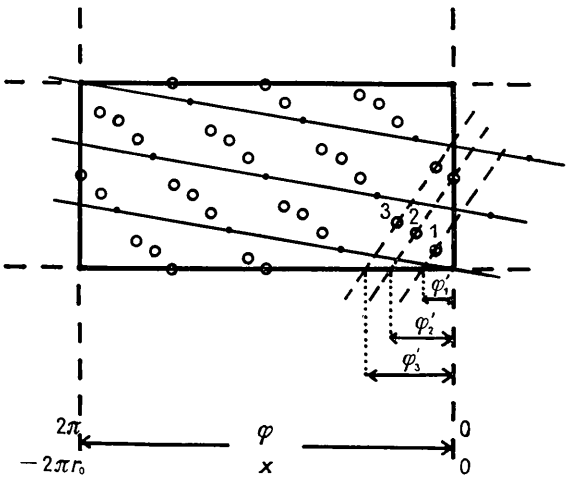

Fig. 4. Radial projection of a 10-unit 3-turn helix with 3 atoms in the asymmetric unit. The basic helix is shown as a full line and the 'helical net points' as dots. The atoms are shown as circles.

The projected co-ordinates φ' are for use in Fig. 5.

$$F(R, \psi, l/c) = \sum_n B_n(R, \psi) \sum_j f_j \exp\left[2\pi i\left(\frac{nx_j}{a} + \frac{lz_j}{c}\right)\right]$$
$$= \sum_n B_n(R, \psi) T_{n,l}, \qquad (5b)$$

where

$$T_{n,l} = \sum_j f_j \exp\left[2\pi i\left(\frac{nx_j}{a} + \frac{lz_j}{c}\right)\right]. \qquad (5c)$$

Thus the whole diffraction pattern *of the helical structure* can be characterized by a set of complex numbers, $T_{n,l}$, which we can plot at the points of a two-dimensional lattice, as in Fig. 2, with rectangular

co-ordinates n and l (strictly n/a and l/c). Identifying n and h in equations (5a) and (5c), we see that the $T_{n,l}$ array is nothing but the weighted reciprocal lattice of the two-dimensional *radial projection* of the helical structure. This relationship was first pointed out by Crick (1953).

The advantage of this way of looking at the structure is that we can quickly estimate from the two-dimensional radial projection (Fig. 4) which of its Fourier components are strong and which weak (by using Bragg–Lipson charts, for example, in which case Fig. 3 enables us to construct the correct chart for any particular Fourier component very quickly). This tells us immediately which Bessel terms will be strong and which weak. In short a radial projection is a convenient way of estimating the sum of the phase factors in equation (1).

The other helical line groups present no difficulty. The cases where there is an N-fold parallel rotation axis is exactly as in the last section. The effect of perpendicular dyads is to put a centre of symmetry into the radial projection, thus making the numbers $T_{n,l}$ real rather than complex.

(*c*) *Atoms at various radii*

When atoms are not all at the same radius, the simple result given above will not hold, since we cannot perform the factorization in equation (5b). The contribution of one Bessel function of order n to the scattering amplitude on the layer line l is now represented by a complex number $G_{n,l}(R)$ whose modulus and phase vary with R, where

$$G_{n,l}(R) = \sum_j f_j J_n(2\pi R r_j) \exp\left[i\left(-n\varphi_j + \frac{2\pi l z_j}{c}\right)\right]. \quad (6)$$

In this notation the scattered amplitude on a layer line l is given by

$$F(R, \psi, l/c) = \sum_n G_{n,l}(R) \exp\left[in(\psi + \tfrac{1}{2}\pi)\right], \quad (7)$$

where the sum is, of course, subject to the usual selection rule (2).

We should like to have a convenient way of estimating the contribution of the various atoms to $G_{n,l}(R)$. Now each atom contributes a Bessel function multiplied by a phase factor involving only its φ and z co-ordinates. The phase factor is the same form for all atoms, and is essentially scale-free, not depending on the radius. It can thus be conveniently estimated from a single diagram, namely the radial projection (on to a cylinder of radius r_0) and it is this which gives the radial projection its value. When considering any particular Bessel function, order n, at a radius in reciprocal space R, we regard each atom in the radial projection as having a weight $J_n(2\pi R r_j)$. We can then estimate the phase factor in exactly the way described for (unweighted) atoms in the last sub-section (*b*); for example, by using Bragg–Lipson charts. Moreover the same diagram will do for all values of n and R we wish to consider.

The radial projection, therefore, is the most convenient way of representing the structure in order to make quick estimates of its diffraction pattern. Its use was first pointed out by Wyckoff (1955).

II.2. Net ambiguities and connection ambiguities

Bear (1955) has discussed the relationship between the real and reciprocal nets and has described 'net ambiguities' and 'connection ambiguities'. His net ambiguities correspond to the ambiguity in deciding the correct line group from the diffraction data. His connection ambiguities correspond roughly to the usual ambiguity in drawing the unit cell of Fig. 1. In particular the helix net of a fibrous structure does not in itself tell how the chemical chains of the structure connect up the points of the helix net, let alone how many separate chemical chains the structure possesses.

However, if the atoms of a structure are concentrated along a chain direction it is sometimes possible, by using the methods just described, to determine the run of the chains, and hence the number of chains, from the order n of the strongest Bessel function observed in the diffraction pattern.

II.3. An interpretation of m

These considerations serve to show that the conventional description of a helical structure adopted in the first section as u units repeating in t turns of *one* helix is not necessarily the most illuminating. It is probably the simplest to imagine geometrically*. We have called it the basic helix since it leads to a very simple interpretation of the integer m, occurring in the selection rule (2) and characterizing the various 'branches' of the diffraction pattern.

The equation to the basic helix is

$$r = \text{constant}, \quad \varphi - 2\pi z/P = 0 .$$

This is a member of a space-filling family of helices all of the same pitch P, defined by

$$r = \text{constant}, \quad \varphi - 2\pi z/P = \text{constant} = \varphi', \text{ say.}$$

Now we can imagine a helix passing through each atom, and can fix the position of each atom by a new set of co-ordinates, namely the azimuth φ' of the helix passing through it, and the fraction of the helix screw measured along the helix. That is, we make the co-ordination transformation

* Note that our basic helix corresponds to one of the two 'genetic helices' of Bear (1955). The choice above is that having the greater pitch or smaller distance between equivalent points along the helix.

$$\left.\begin{aligned} r_j' &= r_j\,, \\ \varphi_j' &= \varphi_j - 2\pi z_j/P\,, \\ \alpha_j &= 2\pi z_j/p\,. \end{aligned}\right\} \tag{8}$$

Then the phase factor determining the contribution of the jth atom to J_n on the lth layer line becomes

$$\begin{aligned} &= \exp\left[i\{-n(\varphi_j'+2\pi z_j/P)+2\pi l z_j/c\}\right] \\ &= \exp\left[i\left\{-n\varphi_j'+\frac{2\pi z_j}{p}\left(\frac{-tn+l}{u}\right)\right\}\right], \\ &\qquad\qquad \text{and using equation (2),} \\ &= \exp\left[i\{-n\varphi_j'+m\alpha_j\}\right]. \end{aligned} \tag{9}$$

We thus see that m has the meaning of an index for translational periodicity along the basic helix, just as l is an index for translational periodicity along the z axis and n for rotational periodicity. The 'gearing' of translation and rotation in a helical screw is expressed by a linear relation between n, l and m, namely equation (2)*. We can use any two of three integers n, l and m to describe the diffraction pattern and the helix net; the pair (n, l) will usually be the most convenient.

III.1. Helical projections

In § II.1 we showed how the radial projection of a structure might be used to estimate for a particular value of R the value of the various Bessel function terms $G_{n,l}(R)$ corresponding to the sets of sinusoidal density waves (n, l). When the atoms are not all at one radius the weights $J_n(2\pi Rr)$ to be attached to the different atoms do not all vary at the same rate with R. The procedure would accordingly be rather heavy, and it may be much more convenient to fix our attention on one Bessel function term $G_{n,l}(R)$ at a time. In this section we shall show that it is then more useful to construct a special two-dimensional projection related to the particular (n, l) term being considered.

From the relation between $G_{n,l}(R)$ and the (n, l) set of waves in the radial projection, discussed in § II, it is clear that the translation of each atom in a direction parallel to the zero lines of the wave has no effect on $G_{n,l}(R)$. This suggests, that, in considering one Bessel function term $G_{n,l}(R)$, we may project the atoms parallel to the (n, l) zero lines; we shall denote this direction by (n, l). In three dimensions this procedure corresponds to projecting the structure down a corresponding set (n, l) of helices. The construction of such maps of the structure was first proposed by Crick (1953), who called them *helical projections* and pointed out that the (n, l) helical projection offered a quick method of estimating the effect of the phase factor in the mathematical expression for $G_{n,l}(R)$. Such a projection will be useful, when only one Bessel function effectively contributes to a layer line we wish to consider. This will happen when the other possible Bessel terms on the layer line are either of too high an order, or else lie slightly off the layer line, as in the case of a structure which does not repeat exactly in the distance c after a whole number of turns.

We may restate this idea mathematically as follows. When only one Bessel function is relevant, the scattered amplitude on the layer line l is given by

$$\begin{aligned} &F(R, \psi, l/c) \\ &\quad = \sum_j f_j J_n(2\pi R r_j) \exp\left[i\{n(\psi+\tfrac{1}{2}\pi)-n\varphi_j+2\pi l z_j/c\}\right] \\ &\quad = G_{n,l}(R) \exp\left[in(\psi+\tfrac{1}{2}\pi)\right], \end{aligned} \tag{10}$$

where $G_{n,l}(R)$ is defined by equation (6).

Let us for the moment restrict ourselves to the cases $n \neq 0$ and $l \neq 0$. Then if there are two atoms with cylindrical co-ordinates (r_1, φ_1, z_1) and (r_2, φ_2, z_2) such that

$$\left.\begin{aligned} -n\varphi_1+2\pi(l/c)z_1 &= -n\varphi_2+2\pi(l/c)z_2\,, \\ r_1 &= r_2\,, \end{aligned}\right\}$$

then, apart from the atomic scattering factor, their contributions to the above sum will be equal for all values of R and ψ. Thus if we project the entire structure along lines defined by the relations

$$-\varphi+2\pi(l/nc)z = \text{const.},\ r = \text{const.}\,, \tag{11}$$

the pattern we obtain is uniquely related to the structure factor given by equation (10).

Equation (11) defines a space-filling family of helices each having a pitch nc/l. If one projects all the atoms by moving them along the appropriate helices of this set, one obtains the helical projection along (n, l). We may draw the projection either on the horizontal plane $z = 0$ or the vertical plane $\varphi = 0$. In Fig. 5 we have drawn the helical projection along $(3, -1)$ of an example similar to that discussed in the last section and shown in Fig. 4, but where the atoms are now to be considered as lying at various radii. For definiteness, we shall henceforth consider only the projection on to $z = 0$, which has the form of a two-dimensional map in which the jth atom of the first asymmetric unit of structure has the polar co-ordinates:

$$r_j,\ \varphi_j' = \varphi_j - 2\pi(l/nc)z_j\,. \tag{12}$$

Now we have shown in § II that in a radial projection the (n, l) lines make n intercepts on the horizontal axis. Accordingly, if one travels round the (n, l) *helical projection*, keeping at a constant radius, the variations in projected electron density will, in general, repeat n times in one complete revolution, i.e. the pattern will have n-fold rotational symmetry. This pattern can then be considered as made up of a number of angular Fourier components of frequency $0, \pm n, \pm 2n, \ldots, \pm hn, \ldots$, where h is an integer. Now consider one of these angular components, the (hn)th, say.

* From the point of view of § II.1, m gives the number of intercepts that the set of lines (n, l) makes on the set of lines $(1, t)$ corresponding to the basic helix.

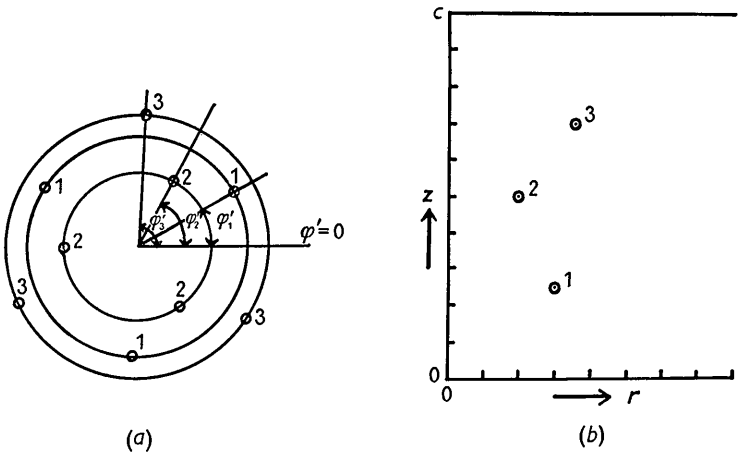

Fig. 5. The helical projection in the (3, −1) direction of the hypothetical structure of Fig. 4 plotted (*a*) in the $z = 0$ plane, (*b*) in the $\varphi = 0$ plane. The direction of projection is indicated by dotted lines in Fig. 4.

It is a two-dimensional pattern with (hn)-fold rotational symmetry, and it can be shown (see § IV.2 below) that its two-dimensional Fourier transform is identical with the term $G_{hn,hl}(R) \exp[inh(\psi+\frac{1}{2}\pi)]$, that is, with the structure factor on the (hl)th layer line.

We thus see that a Fourier analysis of this single projection would lead to the whole set of terms $G_{hn,hl}(R)$, h being any integer as before. We might well have expected all these terms to be involved since the direction of the (n, l) helical projection is determined only by the ratio of n and l, the pitch of the helix being nc/l. It is in this sense that we may think of the term $G_{hn,hl}(R)$ as the hth order of $G_{n,l}(R)$. If the helical projection has a large (hn)-fold lumpiness, we can see at once that the layer line hl will have strong X-ray intensities on it. Alternatively, if its (hn)th angular Fourier component is weak, the X-ray intensities will be weak or absent.

Note that it is not necessary to project the whole structure. It is clear from the manner of derivation that one need only project one asymmetric unit along the relevant helix on to plane $z = 0$ and then operate on this limited projection with an n-fold rotation to produce the entire helical projection. It is the fact that only one asymmetric unit needs to be projected that makes the method useful.

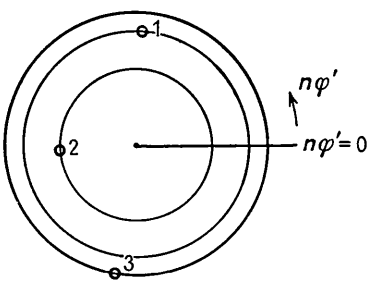

Fig. 6. The $r, n\varphi'$ plot corresponding to Fig. 5(*a*). Note that in this example atoms 1 and 3 have phases differing nearly by π.

The fact that we need only project one unit also leads to another useful kind of plot if we wish to consider each of the $G_{hn,hl}$ terms separately. Consider, for instance, the Bessel term $G_{n,l}$. The phase angle for each atom is $n\varphi'_j$, where φ'_j is given by equation (12) and is the angular co-ordinate in the helical projection. Hence if we project one asymmetric unit and plot it on n times the angular scale we shall have a plot showing the radius and the phase angle of each atom. An example is shown in Fig. 6, where we have constructed the plot corresponding to the term $G_{3,-1}(R)$ for the hypothetical structure of Fig. 4. Lumping together atoms at approximately the same radius, we see that in the example the phase angles of atoms 1 and 3 differ by a value close to π, and we can thus expect that their contributions to the Bessel term will almost cancel. Thus this kind of plot might be useful when one is moving atoms about in building a trial structure for a substance. Future experience may show whether this direct phase plot is any more useful than the actual helical projection, since the latter in principle contains not only $G_{n,l}$ but all the higher order $G_{hn,hl}$.

For completeness, we should add that so far we have considered only cases where $n \neq 0$ and $l \neq 0$. When $l = 0$ we get the usual projection parallel to the z axis. When $n = 0$, $l \neq 0$ the helices degenerate into circles perpendicular to the z axis, in which case it would be more useful to consider the projection on to the $\varphi = 0$ plane. Indeed, for a cylindrical, as distinct from a helical structure, these are the only projections of this kind possible (see Whittaker, 1955)*.

We must also deal briefly with the other helical line

* By a cylindrical structure is usually meant one in which the helical screw has degenerated into a pure translation, but an N-fold rotation axis remains (**tr** in the notation of § I). The general diffraction theory given here and in C.C.V. is still valid, but in this case the same orders of Bessel functions occur on all layer lines, i.e. the helix net and (n, l) plot are rectangular arrays. Cylindrical lattices are discussed briefly in § VI at the end.

groups. When there is an N-fold parallel rotation axis, the helical projection along (n, l) will have nN-fold rotational symmetry, instead of the straightforward n-fold symmetry of the simplest helical line group **s**. The effect of perpendicular dyads on the (n, l) helical projection is to add radial mirror reflexion lines, i.e. to increase the symmetry from C_N to C_{Nv}. The $G_{n,l}(R)$ terms will then be real rather than complex (see equations (3) and (7)), but the helical projection itself will have a centre of symmetry only when n is even.

III.2. An application of helical projections to disordered structures

It sometimes happens that we have helical molecules or polymers arranged as in a crystal lattice but with a disorder consisting of a variable displacement of the molecules along their axes or a variable rotation about their axes. It is also possible to have—although this has not been often recognized—a combination of both, i.e. a variable screw. In such cases the overall crystallinity of the arrangement will be destroyed and the diffraction pattern will correspond to the continuous transform of the molecule, but some parts of the X-ray diagram will still correspond to diffraction by a crystal. For if the disorder consists of a screw (taking the most general case) in the direction determined by the family of helices (n, l) (using our previous notation), the helical projection of the structure along (n, l) will be unchanged. Hence in reciprocal space the corresponding set of Bessel terms $G_{hn,hl}(R)$, h an integer, will be unaffected by disordering, and so, as far as these parts of the diffraction pattern are concerned, the scattering from different molecules is coherent. Thus the corresponding parts of the X-ray diagram will resemble diffraction by a crystal and show sharp reflexions.

An analysis of this kind, but without using the terminology of helical projections, has been given by Franklin & Klug (1956) for diffraction by dry tobacco mosaic virus, where the rod-shaped virus particles pack together in an hexagonal arrangement and are interlocked by helical grooves. The original paper should be consulted for further details.

Another example of disorder where this type of analysis should be applicable is to be found in the transition at 20° C. reported by Bunn & Howells (1954) in their X-ray study of the polymer, polytetrafluoroethylene. The fact that the reflexions corresponding to the J_0 and J_1 parts of the molecular transform remain sharp suggests that the disordering involves a screw motion along the basic helices.

A simpler type of disorder involving only axial displacements, has been found in the case of deoxyribonucleic acid by Wyckoff (1955).

IV. Fourier syntheses of electron density

In general the X-ray photographs that can be obtained from a helical structure are fibre diagrams. These will often be of poor quality compared with X-ray photographs of single crystals, and the only method we have of testing a proposed structure is to compare observed and calculated intensities (see Crick (1954) for further discussion of these points). However, there are instances, notably in the cases of tobacco mosaic virus and deoxyribonucleic acid, where the quality of the photographs is high, and accordingly, by analogy with single-crystal work, the question arises of calculating the electron density from the scattered amplitude. We shall, therefore, now discuss the problem of Fourier inversion in three-dimensional cylindrical co-ordinates. The further question of whether there is a useful analogue of the Patterson function will be dealt with in a later section.

IV.1. The three-dimensional Fourier synthesis

Suppose that the phases have been determined for at least some of the accessible structure-factor terms $G_{n,l}(R)$ either by trial or otherwise (e.g. isomorphous replacement, as in tobacco mosaic virus (Franklin, 1957)), and we wish to use them with the observed moduli to calculate the electron density.

It is convenient to assume a continuous electron density $\varrho(r, \varphi, z)$, find its Fourier transform, and then show how this may be inverted again. In so doing we shall re-derive in a very general way the basic result of C.C.V.

Since the electron density ϱ is periodic in φ and in z we may expand it in the form of a double Fourier series

$$\varrho(r, \varphi, z) = \frac{1}{c} \sum_{l=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g_{n,l}(r) \exp\left[i(n\varphi - 2\pi lz/c)\right], \quad (13)$$

where the $g_{n,l}(r)$ are, in general, complex functions of r. They are the obvious generalizations of ordinary two-dimensional Fourier coefficients and are given by

$$g_{n,l}(r) = \frac{1}{2\pi} \int_0^c \int_0^{2\pi} \varrho(r, \varphi, z) \times \exp\left[-i(n\varphi - 2\pi lz/c)\right] d\varphi\, dz\,. \quad (13a)$$

The Fourier transform of ϱ is

$$F(R, \psi, \zeta) = \int_0^c \int_0^{2\pi} \int_0^{\infty} \varrho(r, \varphi, z) \times \exp\left[2\pi i\{Rr\cos(\varphi - \psi) + \zeta z\}\right] r\, dr\, d\varphi\, dz\,.$$

Substituting from (13), we find, as we expect, that the integral over z will be non-zero only when $\zeta = l/c$, i.e. that the scattering is confined to layer planes. On any one layer plane l we then have

$$F(R, \psi, l/c) = \iint \sum_n g_{n,l}(r) \exp(in\varphi) \exp\left[2\pi iRr\cos(\varphi - \psi)\right] r\, dr\, d\varphi$$

$$= \sum_n \exp\left[in(\psi + \tfrac{1}{2}\pi)\right] \int_0^{\infty} g_{n,l}(r) J_n(2\pi Rr) 2\pi r\, dr\,, \quad (14)$$

where we have used in the last step the integral representation of the Bessel function

$$2\pi i^n J_n(x) = \int_0^{2\pi} \exp\left[i\{x\cos\varphi + n\varphi\}\right] d\varphi \,. \qquad (15)$$

In keeping with our previous notation, we write (14) as

$$F(R, \psi, l/c) = \sum_n G_{n,l}(R) \exp\left[in(\psi + \tfrac{1}{2}\pi)\right], \quad (14a)$$

where, now,

$$G_{n,l}(R) = \int_0^\infty g_{n,l}(r) J_n(2\pi Rr) 2\pi r\, dr \qquad (16)$$

$$= \iiint \varrho(r, \varphi, z) J_n(2\pi Rr) \times \exp\left[i(-n\varphi + 2\pi lz/c)\right] r\, dr\, d\varphi\, dz \qquad (16a)$$

(cf. (6)).

Note that this analysis is perfectly general and we have nowhere used the fact that we are dealing with a helix; in general, all the $g_{n,l}(r)$ terms will occur in equation (13), and hence, of course, all $G_{n,l}(R)$ will also occur. However, when the structure has helical or cylindrical symmetry, some of the $g_{n,l}(r)$ terms will vanish. The values of n and l for which the $g_{n,l}(r)$ remain non-zero may be found very simply by considering the radial projection of the structure (see the remarks at the end of § IV.3 below). The effect of the symmetry is to reduce the size of the asymmetric unit, and indeed it is easily shown that only those $g_{n,l}(r)$ occur which obey the selection rule (2).* In the case of the helical line groups with rotation axes and perpendicular dyads there will, of course, be additional restrictions on n and l, as already discussed in § I.

With these restrictions on n and l, we now recognize equations (14a) and (16) as the rather obvious extension of equations (1) and (6) to the case of a continuous density distribution. Our derivation has, however, been such that we can easily find the inverse transform. By the Fourier–Bessel inversion theorem (Titchmarsh, 1937) it follows from equation (16) that

$$g_{n,l}(r) = \int_0^\infty G_{n,l}(R) J_n(2\pi Rr) 2\pi R\, dR \,. \qquad (17)$$

We use this equation (17) to obtain the $g_{n,l}(r)$, and then finally evaluate the Fourier synthesis, using equation (13).

It will now be clear that to make the Fourier synthesis of electron density we need to know the *individual* $G_{n,l}(R)$. If one were using, for instance, the method of isomorphous replacement one would have to separate the various terms in a region of R where two or more $G_{n,l}$ overlapped on the same layer line.

Although it seems impossible to use in practice, we might note, for completeness, that if we knew only the resultant modulus and phase along a layer line, that is, if we have $F(R, 0, l) = \sum_n G_{n,l}(R)$, we would be able to determine the individual $G_{n,l}(R)$ by Fourier–Bessel analysis. One would use the property of the orthogonality of Bessel functions of different order

$$\int J_n(2\pi Rr) J_p(2\pi Rr') 2\pi R\, dR = 0, \quad \text{if} \quad n \neq p \,. \qquad (18)$$

Further, if $n = p$, the integral vanishes unless $r = r'$. Hence an individual $g_{n,l}(r)$ is given by the J_n transform of the scattered amplitude on a layer line. Formally

$$g_{n,l}(r) = \int F(R, 0, l) J_n(2\pi Rr) 2\pi R\, dR \,. \qquad (19)$$

Ideally, the J_n transform would be non-zero only for those n obeying the selection rule (2). Indeed, if we knew the phase along a layer line, we could discover the helix net, i.e. the values of u and t in equation (2), by evaluating Bessel transforms of different orders for a number of layer lines in turn and noting those that do not vanish.

IV.2. Fourier synthesis of a helical projection

When we construct a helical projection along (n, l) we obtain in the (r, φ') plane (see equation (12)) a two-dimensional density distribution having n-fold symmetry, which we shall denote by $\varrho(r, \varphi')$. Then the expression (16a) for the corresponding terms $G_{hn,hl}(R)$, where h is, as before, any positive or negative integer, may be written as

$$G_{hn,hl}(R) = \int_0^\infty \int_0^{2\pi} \sigma(r, \varphi') J_n(2\pi Rr) \times \exp\left(-ihn\varphi'\right) r\, dr\, d\varphi' \,. \qquad (20)$$

Now we can expand $\sigma(r, \varphi')$ in terms of its angular components

$$\sigma(r, \varphi') = \sum_k \gamma_k(r) \exp\left(ik\varphi'\right), \quad k = 0, \pm n, \pm 2n, \ldots, \pm hn, \ldots, \qquad (21)$$

where

$$\gamma_k(r) = \int \sigma(r, \varphi') \exp\left(-ik\varphi'\right) d\varphi' \,. \qquad (22)$$

The two-dimensional transform of one of these angular components $\gamma_{hn} \exp\left(ihn\varphi'\right)$ is

$$F_{hn}(R, \psi) = \int_0^\infty \int_0^{2\pi} \gamma_{hn}(r) \exp\left(ihn\varphi'\right) \times \exp\left[2\pi iRr\cos\left(\varphi' - \psi\right)\right] r\, dr\, d\varphi'$$

$$= \exp\left[ihn(\psi + \tfrac{1}{2}\pi)\right] \int \gamma_{hn}(r) J_n(2\pi Rr) 2\pi r\, dr$$

$$= G_{hn,hl}(R) \exp\left[ihn(\psi + \tfrac{1}{2}\pi)\right], \qquad (23)$$

* The diffraction pattern of a *continuous* helix may be derived from this point of view, by considering it as the limiting case of a discrete helix with an infinite number of units per turn. It is then obvious that only the branch $m = 0$ occurs.

since it follows from (22) and (20) that

$$G_{hn,hl}(R) = \int \gamma_{hn}(r) J_n(2\pi Rr) 2\pi r\, dr\,. \qquad (24)$$

Thus we have proved the statement made in § III that the Bessel function term $G_{n,l}$ on the lth layer line is the transform of the nth angular component of the helical projection down (n, l), $G_{2n,2l}$ on the $2l$th layer line is the transform of the $2n$th component, and so on.

Note that, by comparing the right-hand sides of equations (24) and (16), we can identify $\gamma_n(r)$ (for a particular pair (n, l)) with $g_{n,l}(r)$, thus establishing the connection between the full three-dimensional Fourier synthesis and the limited synthesis of a helical projection. Indeed, the γ_n required for the helical projection are to be found in exactly the same way as in the three-dimensional case, since by inversion of equation (24):

$$\gamma_{hn}(r) = \int G_{hn,hl}(R) J_n(2\pi Rr) 2\pi R\, dR\,. \qquad (25)$$

Equations (25) and (21) between them show us how to construct the helically projected electron density, if we know the relevant set of $G_{n,l}(R)$. A synthesis of this kind is being carried out for tobacco mosaic virus by Dr R. E. Franklin (private communication), using a limited number of $G_{n,l}$ whose phases have been determined by isomorphous replacement. The particular value of such a procedure in a substance as complex as tobacco mosaic virus is that a single, appropriately chosen, helical projection can be used to show up certain specific features of the structure.

It might be helpful to give a concrete example of the form taken by the $g_{n,l}(r)$ in a simple case. Since $g_{n,l}(r) \equiv \gamma_n(r)$ of the helical projection (n, l), we see that, in our example in Fig. 5, we would have, for the case of point atoms,

$$g_{3,\bar{1}}(r) = \delta(r-r_1)\exp[-i3\varphi_1'] + \delta(r-r_2)\exp[-i3\varphi_2'] + \delta(r-r_3)\exp[-i3\varphi_3']\,,$$

$$g_{6,\bar{2}}(r) = \delta(r-r_1)\exp[-i6\varphi_1'] + \delta(r-r_2)\exp[-i6\varphi_2'] + \delta(r-r_3)\exp[-i6\varphi_3']\,,$$

and so on, where $\delta(x)$ is a cylindrical δ function defined by

$$\int f(x)\delta(x-x_0)2\pi x\, dx = f(x_0)\,.$$

This form for $g_{n,l}(r)$ may be checked by substituting it in equation (16), to obtain, for example,

$$G_{3,\bar{1}}(R) = \sum_j J_n(2\pi Rr_j)\exp[-i3\varphi_j'] = \sum_j J_n(2\pi Rr_j)\exp[i(-3\varphi_j + 2\pi z_j/c)]\,, \quad \text{using (12)}\,,$$

which is the correct result.

For real atoms, the δ functions must be replaced by say, Gaussians, and the expression for $G_{n,l}(R)$ will correspondingly contain the atomic scattering factors.

IV.3. The radial projection

It is possible to derive equations for constructing a Fourier synthesis of the radial projection of a structure, as defined in § I. However, they turn out to be very cumbersome and not particularly useful and we shall not reproduce them here. In dealing with continuous distributions of density, it is much more illuminating mathematically to derive results, not for the true radial projection

$$\varrho_r(\varphi, z) = \int \varrho(r, \varphi, z) r dr\,, \qquad (26a)$$

but for a related function

$$\varrho_c(\varphi, z) = \int \varrho(r, \varphi, z) dr\,, \qquad (26b)$$

which is a radial projection of the density weighted by $1/r$. Because of its physical significance and the ensuing mathematical simplicity, it is this 'de-weighted' projection which we shall refer to in this section simply as the radial projection. Note that it corresponds to a superposition of the density in various cylindrical shells of the structure, all drawn out to the same linear dimensions without regard to radius.

Substituting from (13) into (26b), we obtain

$$\varrho_c(\varphi, z) = \int \sum_l \sum_n \exp[in\varphi - 2\pi ilz/c] g_{n,l}(r) dr = \sum_l \sum_n A_{n,l} \exp[in\varphi - 2\pi ilz/c]\,. \qquad (27)$$

This is an ordinary two-dimensional Fourier series with constant coefficients $A_{n,l}$ given by

$$A_{n,l} = \int_0^\infty g_{n,l}(r) dr\,. \qquad (28)$$

The $A_{n,l}$ can also be expressed in terms of reciprocal-space quantities by using (17) to obtain

$$A_{n,l} = \iint G_{n,l}(R) J_n(2\pi Rr) 2\pi R dR dr\,.$$

The double integral may be reduced by interchanging the order of integration and using the result

$$\int_0^\infty J_n(2\pi Rr) dr = \frac{1}{2\pi R}\,, \qquad (29)$$

and hence the coefficients $A_{n,l}$ are given by the following integral in reciprocal space:

$$A_{n,l} = \int_0^\infty G_{n,l}(R) dR\,. \qquad (30)$$

The similarity between equations (28) and (30) is very striking.

As a concrete example, we note that if we have two point atoms of weights Z_1 and Z_2 at different radii r_1 and r_2, but with the same φ and z co-ordinates, φ_j and z_j, then

$$A_{n,l} = (Z_1/2\pi r_1 + Z_2/2\pi r_2) \exp\left[i(-n\varphi_j + 2\pi l z_j/c)\right] .$$

Equations (26*b*) and (27) show how to make a Fourier synthesis of a radial projection from the diffraction data. But since the determination of the Fourier coefficients $A_{n,l}$ would seem to require prior knowledge of the whole course of $G_{n,l}(R)$ as a function of R, it would be more realistic to construct the full three-dimensional electron density. If there were some simple means of estimating, say, the average value of $G_{n,l}(R)$, then this synthesis might prove very valuable indeed.

We have included it here chiefly because it brings out, in an instructive way, how the $g_{n,l}(r)$ of the full three-dimensional structure reduce to simple Fourier coefficients $A_{n,l}$ in the radial projection. The selection rule (2) relating n and l—which, as we have seen, is what characterizes a helical structure—will be unchanged by the projection. We can thus conveniently use the radial projection for discussion of helical symmetry, as has been done in § IV.1.

V. The Patterson function of a helical projection

We shall now consider what information may be obtained directly from the intensities, when we have no knowledge of the phases. If we had three-dimensional data we could of course construct the Patterson function, but, as mentioned earlier, the only data we usually have is from a fibre diagram. From this the cylindrically averaged Patterson function may be calculated (MacGillavry & Bruins, 1948) in a perfectly general way that makes no use of any helical features that might be present in the structure.

The question naturally arises: if we know the helical symmetry, as expressed, say, by the selection rule (2), can we do anything better? We shall now show that if one can determine the separate intensity contributions $G^2_{n,l}(R)$* it is possible to construct certain maps related to the Patterson of a helical projection. To avoid the cumbersome convolution integrals that would arise for a continuous density distribution, we revert to the case of discrete atoms.

Now it has been shown by Franklin & Klug (1955) that the cylindrically averaged intensity along a layer line is given by

$$\begin{aligned}\langle F^2(R, \psi, l/c)\rangle_\psi &= \sum_n \sum_{i,j} J_n(2\pi R r_i) J_n(2\pi R r_j) \\ &\qquad \times \cos\left\{n(\varphi_i - \varphi_j) - \frac{2\pi l}{c}(z_i - z_j)\right\} , \\ &= \sum_n G^2_{n,l}(R) \qquad (31)\end{aligned}$$

* Strictly, this should be written $|G_{n,l}(R)|^2$, since G is, in general, complex.

in our notation. Essentially this means that, in a fibre diagram, Bessel functions of different order effectively do not interfere, and indeed they may even lie at slightly different levels, as has been found in the case of tobacco mosaic virus. Let us fix our attention on one term $G^2_{n,l}(R)$ (or rather—as it will turn out—the set of terms $G^2_{hn,hl}(R)$) and consider how the corresponding helical projection is related to it. We shall now derive a relation between $G^2_{n,l}(R)$, and the Patterson of the helical projection along (n, l).

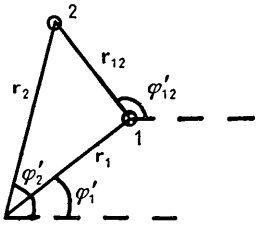

Fig. 7.

Fig. 7 is intended to represent two non-equivalent atoms 1 and 2 in the projection and the vector $\mathbf{r}_{12}$ between them. To construct the Patterson we simply place the vector $\mathbf{r}_{12}$ at the origin of r, φ' co-ordinates in Patterson space. To relate $|\mathbf{r}_{12}|$ to r_1 and r_2 we require the addition theorem for Bessel functions (see, for instance, Stratton, 1941) which states that

$$J_0(\lambda r_{12}) = \sum_{k=-\infty}^{\infty} J_k(\lambda r_1) J_k(\lambda r_2) \exp\left[ik(\varphi'_2 - \varphi'_1)\right] . \qquad (32)$$

Now, in the helical projection, for any one atom there will be another $(n-1)$ equivalent ones at the same radius, successive atoms being separated by an angle $2\pi/n$. Hence if we sum equation (32) over the set 2 of atoms, we obtain

$$\begin{aligned}&\sum_{\text{set 2}} J_0(\lambda r_{12}) \\ &\quad = \sum_{k=-\infty}^{\infty} J_k(\lambda r_1)\, J_k(\lambda r_2) \sum_{p=0}^{n-1} \exp\left[ik\left(\varphi'_2 + p\frac{2\pi}{n} - \varphi'_1\right)\right] \\ &\quad = \sum_{k=0,\,\pm n,\,\pm 2n,\,\ldots} J_k(\lambda r_1) J_k(\lambda r_2) \exp\left[ik(\varphi'_2 - \varphi'_1)\right] .\end{aligned}$$

This equation will hold for any choice of atom out of the set 1. Hence, summing over the set 1, we get

$$\begin{aligned}&\sum_{\text{set 1}} \sum_{\text{set 2}} J_0(\lambda r_{12}) \\ &\quad = n \sum_{k=0,\,\pm n,\,\pm 2n,\,\ldots} J_k(\lambda r_1) J_k(\lambda r_2) \exp\left[ik(\varphi'_2 - \varphi'_1)\right] .\end{aligned}$$

Finally, summing over all the pairs of atoms in the asymmetric unit, we obtain

$$\begin{aligned}&\frac{1}{n} \sum_i \sum_j J_0(\lambda r_{ij}) \\ &\quad = \sum_{k=0,\,\pm n,\,\ldots} \sum_{i,j} J_k(\lambda r_i) J_k(\lambda r_j) \exp\left[ik(\varphi'_j - \varphi'_i)\right] . \qquad (33)\end{aligned}$$

Now, in terms of the co-ordinates r, φ' (see equation

(12)) of the helical projection along (n, l), $G_{n,l}$ as defined by equation (6) is to be written as

$$G_{n,l} = \sum_j f_j J_n(2\pi R r_j) \exp[-in\varphi'] .$$

Hence, if we put $\lambda = 2\pi R$ and insert factors $f_i f_j$, equation (33) can be written as

$$\frac{1}{n} \sum_i \sum_j f_i f_j J_0(2\pi R r_{ij}) = G^2_{0,0}(R) + 2G^2_{n,l}(R) + 2G^2_{2n,2l}(R) + \ldots , \quad (34)$$

which shows that the J_0 transform of the Patterson of a helical projection is equal to the sum of the squares of the set of $G_{hn,hl}$ corresponding to that projection. We thus have proved an analogue of Parseval's theorem. Since a J_0 transform has cylindrical symmetry, we can express our result formally as follows. If $P_{n,l}(r, \varphi')$ represents the Patterson density of the helical projection, and

$$\bar{P}_{n,l}(r) = \frac{1}{2\pi r} \int_0^{2\pi} P_{n,l}(r, \varphi')\, d\varphi' = \frac{n}{2\pi r} \int_0^{2\pi/n} P_{n,l}(r, \varphi')\, d\varphi' \quad (35)$$

is its rotational average, then

$$\int_0^\infty \bar{P}_{n,l}(r) J_0(2\pi R r) 2\pi r\, dr = \sum_{h=0, \pm 1, \pm 2, \ldots} G^2_{hn,hl}(R) . \quad (36)$$

Inverting, we obtain

$$\bar{P}_{n,l}(r) = \int_0^\infty \left\{ \sum_{h=0, \pm 1, \ldots} G^2_{hn,hl}(R) \right\} J_0(2\pi R r) 2\pi R\, dR , \quad (37)$$

which is the desired relation.

It is too early to say whether (37) will prove to be useful. It is easy to show that the $r\bar{P}_{n,l}(r)$ curve will contain a rather broad peak near $2r$ when there is a concentration of density at the radius r in the helical projection. It will thus reveal the radial limits of a structure, but it is not superior in this respect to the cylindrical Patterson. We have tried out equation (37) for the α-helix, using the intensities calculated by Pauling, Corey, Yakel & Marsh (1955) to evaluate* the rotationally averaged Patterson of the helical projection down the basic helix $(1, \bar{5})$. The resulting curve was, however, not easily interpretable in terms of the individual Patterson vectors.

It is perhaps worth while noting that equation (32), which is the basis of our derivation, is a special case of the formula (Stratton, 1941) relating the vector $\mathbf{r}_{12}$ to $\mathbf{r}_1$ and $\mathbf{r}_2$:

$$J_n(\lambda r_{12}) \exp[in\varphi'_{12}] = \sum_{k=-\infty}^{\infty} J_k(\lambda r_1) J_{n+k}(\lambda r_2) \exp[in\varphi'_2] \exp[ik(\varphi'_2 - \varphi'_1)] , \quad (38)$$

and we might hope to use the latter in the same way to obtain results involving $P_{n,l}(r, \varphi)$ rather than $\bar{P}_{n,l}(r)$. But we see at once that (38) involves products of Bessel functions of different order and such terms occur only in the expression for the full diffracted intensity $F^2(R, \psi, l/c)$ before the cylindrical averaging. Indeed, from the point of view of diffraction theory, equation (38) is nothing but the expression (in cylindrical co-ordinates) of the perfectly general result that the Fourier transform of the Patterson is the intensity distribution.

It thus appears that equation (37) represents all we can hope to learn from the cylindrically averaged intensity alone.

VI. Cylindrical lattices

It is interesting to note that the theory given in this paper bears a certain resemblance to that in recent papers on cylindrical *lattices* by Whittaker, Waser (1955) and Kunze & Jagodzinski (for references, see Whittaker, 1955). There is, however, very little in common apart from the mathematics of Bessel functions. A cylindrical lattice is made up of concentric cylindrical sheets of crystal-like unit cells, and the angular separation between asymmetric units is very small compared with the radius of the cylinder. In other words, Bessel functions of very high order are involved, and the approximation

$$J_n(x) \sim \sqrt{\left(\frac{2}{\pi x}\right)} \cdot \cos(x - \tfrac{1}{2}n\pi - \tfrac{1}{4}\pi)$$

may be introduced, resulting in structure factors of the usual trigonometric form applicable to three-dimensional lattices. The analogues of our $G_{n,l}(R)$ extend over very short distances in reciprocal space and may be considered as broadened crystal reflexions.

* We are indebted to Mr K. C. Holmes for carrying out the computation.

References

Bear, R. S. (1955). *Fibrous Proteins and their Biological Significance*, p. 97. Symposium No. 9 of the Society of Experimental Biology. Cambridge: University Press.

Bunn, C. W. & Howells, E. R. (1954). *Nature, Lond.* **174**, 549.

Cochran, W. & Crick, F. H. C. (1952). *Nature, Lond.* **169**, 234.

Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581. [Referred to in the text as C.C.V.]

Cohen, C. & Bear, R. S. (1953). *J. Amer. Chem. Soc.* **75**, 2783.

Cowan, P. M., North, A. C. T. & Randall, J. T. (1953). *Nature and Structure of Collagen*, p. 241. London: Butterworths.

Crick, F. H. C. (1953). Thesis, Cambridge.

Crick, F. H. C. (1954). *Science Progress*, p. 205.

Franklin, R. E. (1957). Private communication.

Franklin, R. E. & Gosling, R. G. (1953). *Nature, Lond.* **171**, 742.

Franklin, R. E. & Klug, A. (1955). *Acta Cryst.* **8**, 777.

Franklin, R. E. & Klug, A. (1956). *Biochim. Biophys. Acta*, **19**, 403.

International Tables for X-ray Crystallography (1952), vol. 1. Birmingham: Kynoch Press.
MACGILLAVRY, C. H. & BRUINS, E. M. (1948). *Acta Cryst.* **1**, 156.
PAULING, L., COREY, R. B., YAKEL, H. L. JR. & MARSH, R. E. (1955). *Acta Cryst.* **8**, 853.
STOKES, A. R. (1955). *Acta Cryst.* **8**, 27.
STRATTON, J. A. (1941). *Electromagnetic Theory*, chap. 6. New York: McGraw-Hill.
TITCHMARSH, E. C. (1937). *Theory of the Fourier Integral.* Oxford: Clarendon Press.
WASER, J. (1955). *Acta Cryst.* **8**, 142.
WATSON, J. D. (1954). *Biochim. Biophys. Acta*, **13**, 10.
WATSON, J. D. (1957). *The Chemical Basis of Heredity.* Baltimore: Johns Hopkins Press.
WATSON, J. D. & CRICK, F. H. C. (1953). *Nature, Lond.* **171**, 737.
WHITTAKER, E. J. W. (1955). *Acta Cryst.* **8**, 571.
WILKINS, M. H. F., STOKES, A. R. & WILSON, H. R. (1953). *Nature, Lond.* **171**, 737.
WYCKOFF, H. W. (1955). Thesis, Massachusetts Institute of Technology.

Acta Cryst. (1958). **11**, 213

Tobacco Mosaic Virus: Application of the Method of Isomorphous Replacement to the Determination of the Helical Parameters and Radial Density Distribution

BY ROSALIND E. FRANKLIN AND K. C. HOLMES

Birkbeck College Crystallography Laboratory (University of London), 21 Torrington Square, London W. C. 1, England

(*Received* 8 *August* 1957)

It is known that in tobacco mosaic virus (TMV) structurally equivalent protein sub-units lie in helical array about the long axis of the rod-shaped particle. It has not, however, proved possible to make a reliable determination of the helical parameters by the usual methods of direct measurement on X-ray fibre diagrams.

A quantitative comparison has now been made of the equatorial intensities in fibre diagrams of TMV and of a mercury-substituted TMV, TMV-Hg. This has led to the determination, first of the radial distance of the substituted mercury from the particle axis in TMV-Hg, and then of the parameters of the helical arrangement of protein sub-units in TMV. It also made possible the calculation of the radial density distribution in TMV, the result obtained being in good agreement with the earlier work of Caspar.

From a knowledge of the helical parameters and other physical and chemical data it can be shown that the symmetry of TMV is that of the simplest possible helical line group, there being no symmetry elements other than the helical axis.

Introduction

The X-ray fibre-diagrams obtained from orientated preparations of tobacco mosaic virus (TMV) show that the particles have a highly ordered internal structure (Bernal & Fankuchen, 1941). The best fibre diagrams are obtained from concentrated aqueous gel preparations in thin-walled glass capillary tubes (Bernal & Fankuchen, 1941), in which the virus particles, of length about 3000 Å (see Williams & Steere, 1951) and mean diameter about 150 Å (Bernal & Fankuchen, 1941; Franklin & Klug, 1956) lie with their long axes parallel to the axis of the tube. Since the virus particles are in random rotation about their long axes, and, moreover, do not lie strictly on a lattice, these fibre diagrams record, effectively, the cylindrically averaged squared structure factor of a *single* virus particle; interparticle interference effects are appreciable only on the equator at spacings larger than 100 Å.

Certain features of the structure of the virus can be deduced directly from measurements made on TMV fibre diagrams alone. In this way it has been shown (Watson, 1954; Franklin, 1955) that the virus protein (which comprises about 94% of the virus, the remainder being ribonucleic acid (RNA)) consists of equivalent sub-units set in helical array about the particle axis, the axial repeat period of 69 Å containing $3n+1$ such sub-units on 3 turns of the helix. It was also shown (Franklin & Klug, 1956) from a direct study of TMV fibre diagrams that the surface of the TMV particle is not smooth, but bears a rather deep helical groove, between the turns of which lies a helical array of knobs.

While it is possible to push somewhat further this type of direct interpretation of the TMV diagrams, more substantial progress can be made by comparing these diagrams with those of related substances. It has been shown by Green, Ingram & Perutz (1954)