2.2. Direct methods

BY C. GIACOVAZZO

2.2.1. List of symbols and abbreviations

f_j atomic scattering factor of jth atom
Z_j atomic number of jth atom
N number of atoms in the unit cell
m order of the point group

$$[\sigma_r]_p, [\sigma_r]_q, [\sigma_r]_N, \ldots = \sum_{j=1}^{p} Z_j^r, \sum_{j=1}^{q} Z_j^r, \sum_{j=1}^{N} Z_j^r, \ldots$$

$[\sigma_r]_N$ is always abbreviated to σ_r when N is the number of atoms in the cell

$$\sum_p, \sum_q, \sum_N, \ldots = \sum_{j=1}^{p} f_j^2, \sum_{j=1}^{q} f_j^2, \sum_{j=1}^{N} f_j^2, \ldots$$

s.f.   structure factor
n.s.f.   normalized structure factor
cs.   centrosymmetric
ncs.   noncentrosymmetric
s.i.   structure invariant
s.s.   structure seminvariant
$\mathbf{C} = (\mathbf{R}, \mathbf{T})$ symmetry operator; $\mathbf{R}$ is the rotational part, $\mathbf{T}$ the translational part
$\varphi_{\mathbf{h}}$ phase of the structure factor $F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\varphi_{\mathbf{h}})$

2.2.2. Introduction

Direct methods are today the most widely used tool for solving small crystal structures. They work well both for equal-atom molecules and when a few heavy atoms exist in the structure. In recent years the theoretical background of direct methods has been improved to take into account a large variety of prior information (the form of the molecule, its orientation, a partial structure, the presence of pseudosymmetry or of a superstructure, the availability of isomorphous data or of data affected by anomalous-dispersion effects, ...). Owing to this progress and to the increasing availability of powerful computers, a number of effective, highly automated packages for the practical solution of the phase problem are today available to the scientific community.

The *ab initio* crystal structure solution of macromolecules seems not to exceed the potential of direct methods. Many efforts will certainly be devoted to this task in the near future: a report of the first achievements is given in Section 2.2.10.

This chapter describes both the traditional direct methods tools and the most recent and revolutionary techniques suitable for macromolecules.

The theoretical background and tables useful for origin specification are given in Section 2.2.3; in Section 2.2.4 the procedures for normalizing structure factors are summarized. Phase-determining formulae (inequalities, probabilistic formulae for triplet, quartet and quintet invariants, and for one- and two-phase s.s.'s, determinantal formulae) are given in Section 2.2.5. In Section 2.2.6 the connection between direct methods and related techniques in real space is discussed. Practical procedures for solving crystal structures are described in Sections 2.2.7 and 2.2.8, and references to the most extensively used packages are given in Section 2.2.9. The techniques suitable for the *ab initio* crystal structure solution of macromolecules are described in Section 2.2.10.2. The integration of direct methods with isomorphous-replacement and anomalous-dispersion techniques is briefly described in Sections 2.2.10.3 and 2.2.10.4.

The reader will find full coverage of the most important aspects of direct methods in the recent books by Giacovazzo (1998) and Woolfson & Fan (1995).

2.2.3. Origin specification

(*a*) Once the origin has been chosen, the symmetry operators $\mathbf{C}_s \equiv (\mathbf{R}_s, \mathbf{T}_s)$ and, through them, the algebraic form of the s.f. remain fixed.

A shift of the origin through a vector with coordinates $\mathbf{X}_0$ transforms $\varphi_{\mathbf{h}}$ into

$$\varphi_{\mathbf{h}}' = \varphi_{\mathbf{h}} - 2\pi\mathbf{h} \cdot \mathbf{X}_0 \tag{2.2.3.1}$$

and the symmetry operators $\mathbf{C}_s$ into $\mathbf{C}_s' = (\mathbf{R}_s', \mathbf{T}_s')$, where

$$\mathbf{R}_s' = \mathbf{R}_s; \mathbf{T}_s' = \mathbf{T}_s + (\mathbf{R}_s - \mathbf{I})\mathbf{X}_0 \quad s = 1, 2, \ldots, m. \tag{2.2.3.2}$$

(*b*) *Allowed* or *permissible origins* (Hauptman & Karle, 1953, 1959) for a given algebraic form of the s.f. are all those points in direct space which, when taken as origin, maintain the same symmetry operators $\mathbf{C}_s$. The allowed origins will therefore correspond to those points having the same symmetry environment in the sense that they are related to the symmetry elements in the same way. For instance, if $\mathbf{T}_s = 0$ for $s = 1, \ldots, 8$, then the allowed origins in *Pmmm* are the eight inversion centres.

To each functional form of the s.f. a set of permissible origins will correspond.

(*c*) A translation between permissible origins will be called a *permissible* or *allowed translation*. Trivial allowed translations correspond to the lattice periods or to their multiples. A change of origin by an allowed translation does not change the algebraic form of the s.f. Thus, according to (2.2.3.2), all origins allowed by a fixed functional form of the s.f. will be connected by translational vectors $\mathbf{X}_p$ such that

$$(\mathbf{R}_s - \mathbf{I})\mathbf{X}_p = \mathbf{V}, \quad s = 1, 2, \ldots, m, \tag{2.2.3.3}$$

where $\mathbf{V}$ is a vector with zero or integer components.

In centred space groups, an origin translation corresponding to a centring vector $\mathbf{B}_v$ does not change the functional form of the s.f. Therefore all vectors $\mathbf{B}_v$ represent permissible translations. $\mathbf{X}_p$ will then be an allowed translation (Giacovazzo, 1974) not only when, as imposed by (2.2.3.3), the difference $\mathbf{T}_s' - \mathbf{T}_s$ is equal to one or more lattice units, but also when, for any s, the condition

$$(\mathbf{R}_s - \mathbf{I})\mathbf{X}_p = \mathbf{V} + \alpha\mathbf{B}_v, \quad s = 1, 2, \ldots, m; \quad \alpha = 0, 1 \tag{2.2.3.4}$$

is satisfied.

We will call any set of cs. or ncs. space groups having the same allowed origin translations a Hauptman–Karle group (H–K group). The 94 ncs. primitive space groups, the 62 primitive cs. groups, the 44 ncs. centred space groups and the 30 cs. centred space groups can be collected into 13, 4, 14 and 5 H–K groups, respectively (Hauptman & Karle, 1953, 1956; Karle & Hauptman, 1961; Lessinger & Wondratschek, 1975). In Tables 2.2.3.1–2.2.3.4 the H–K groups are given together with the allowed origin translations.

(*d*) Let us consider a product of structure factors

$$F_{\mathbf{h}_1}^{A_1} \times F_{\mathbf{h}_2}^{A_2} \times \ldots \times F_{\mathbf{h}_n}^{A_n} = \prod_{j=1}^{n} F_{\mathbf{h}_j}^{A_j}$$

$$= \exp\left(i \sum_{j=1}^{n} A_j \varphi_{\mathbf{h}_j}\right) \prod_{j=1}^{n} |F_{\mathbf{h}_j}|^{A_j},$$

$$\tag{2.2.3.5}$$

A_j being integer numbers.

The factor $\sum_{j=1}^{n} A_j \varphi_{\mathbf{h}_j}$ is the phase of the product (2.2.3.5). A *structure invariant* (s.i.) is a product (2.2.3.5) such that