3. DUAL BASES IN CRYSTALLOGRAPHIC COMPUTING

3.3.2.2.1. *Methods based on conformational variables*

Suppose that $\mathbf{t}$ represents a vector from the current position of an atom in the model to a target position then (see Section 3.3.1.1.3), to first order, the observational equations are

$$t_{IA} = D_{IpA}\theta_p + v_{IA}$$

in which $\boldsymbol{\theta}$ represents changes to conformational variables which may include dihedral angles, bond angles, bond lengths, and parameters determining overall position and orientation of the molecule as a whole. If every such parameter is included the model acquires $3n$ degrees of freedom for n atoms, in which case the methods of the next section are more appropriate, but if bond lengths and some or all bond angles are being treated as constants then the above equation becomes the basis of the treatment.

$$D_{IPA} = \frac{\partial r_{IA}}{\partial \theta_P} = \varepsilon_{Ijk} n_{jP}(r_{kA} - r_{kP})$$

in which $\mathbf{n}_P$ is a unit vector defining the axis of rotation for an angular variable θ_P, $\mathbf{r}_A$ and $\mathbf{r}_P$ are position vectors of the atom A and the site of the parameter P, and $\mathbf{v}_A$ represents a residual vector.

$\sigma = v_{ia}v_{ia}$ is minimized by

$$\boldsymbol{\theta} = \boldsymbol{M}^{-1}\mathbf{V}$$

in which

$$M_{PQ} = D_{iPa}D_{iQa}$$
$$V_P = D_{iPa}t_{ia}.$$

More generally, if σ represents any scalar quantity which is to be minimized, *e.g.* an energy, then

$$V_P = -\frac{1}{2}\frac{\partial \sigma}{\partial \theta_P}$$

$$M_{PQ} = \frac{1}{2}\frac{\partial^2 \sigma}{\partial \theta_P \partial \theta_Q}.$$

It is beyond the scope of this chapter to review the methods available for evaluating $\boldsymbol{\theta}$ from these equations. Difficulties may arise from two sources:

(i) Inversion of $\boldsymbol{M}$ may be difficult if $\boldsymbol{M}$ is large or ill conditioned and impossible if $\boldsymbol{M}$ is singular.

(ii) Successful evaluation of $\boldsymbol{M}^{-1}\mathbf{V}$ will not minimize σ in one step if $\mathbf{t}$ is not linearly dependent on $\boldsymbol{\theta}$ or, equivalently, $\partial^2\sigma/\partial\theta_P\partial\theta_Q$ is not constant, and substantial changes $\boldsymbol{\theta}$ are involved. Iteration is then necessary.

Difficulties of the first kind may be overcome by gradient methods, for example the conjugate gradient method without searches if $\boldsymbol{M}$ is available or with searches if it is not available, or they may be overcome by methods based on eigenvalue decompositions. If non-linearity is serious less dependence should be placed on $\boldsymbol{M}$ and gradient methods using searches are more valuable. In this connection Diamond (1966) introduced a sliding filter technique which produced rapid convergence in extreme conditions of non-linearity. These topics have been reviewed elsewhere (Diamond, 1981, 1984*b*) and are the subject of many textbooks (Walsh, 1975; Gill *et al.*, 1981; Luenberger, 1984).

Warme *et al.* (1972) have developed a similar system using dihedral angles as variables and a variety of alternative optimization algorithms.

The modelling of flexible rings or lengths of chain with two or more fixed parts is sometimes held to be a difficulty in methods using conformational variables, although a simple two-stage solution does exist. The principle involved is the sectioning of the space of the variables into two orthogonal subspaces of which the first is used to satisfy the constraints and the second is used to perform the optimization subject to those constraints.

The algebra of the method may be outlined as follows, and is given in more detail by Diamond (1971, 1980*a,b*). Parametric shifts $\boldsymbol{\theta}_1$ which satisfy the constraints are solutions of

$$\mathbf{V}_1 = \boldsymbol{M}_1\boldsymbol{\theta}_1$$

in which $\mathbf{V}_1$ and $\boldsymbol{M}_1$ depend only on the target vectors, $\mathbf{t}_1$, of the atoms with constrained positions and on the corresponding derivatives. We then find a partitioned orthogonal matrix $(\boldsymbol{A}_1\boldsymbol{B}_1)$ satisfying

$$(\boldsymbol{A}_1\boldsymbol{B}_1) = (\boldsymbol{E}_\lambda\boldsymbol{E}_0)\begin{pmatrix} \boldsymbol{R}_\lambda & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_0 \end{pmatrix}$$

in which $\boldsymbol{E}_\lambda$ are the eigenvectors of $\boldsymbol{M}_1$ having positive eigenvalues, $\boldsymbol{E}_0$ are those having zero eigenvalues, and $\boldsymbol{R}_\lambda$ and $\boldsymbol{R}_0$ are arbitrary orthogonal matrices. Then

$$\begin{pmatrix} \boldsymbol{A}_1^T \\ \boldsymbol{B}_1^T \end{pmatrix}\mathbf{V}_1 = \begin{pmatrix} \boldsymbol{A}_1^T \\ \boldsymbol{B}_1^T \end{pmatrix}\boldsymbol{M}_1(\boldsymbol{A}_1\boldsymbol{B}_1)\begin{pmatrix} \boldsymbol{A}_1^T \\ \boldsymbol{B}_1^T \end{pmatrix}\boldsymbol{\theta}_1$$

$$= \begin{pmatrix} \boldsymbol{A}_1^T\boldsymbol{M}_1\boldsymbol{A}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_1^T \\ \boldsymbol{B}_1^T \end{pmatrix}\boldsymbol{\theta}_1$$

$$\boldsymbol{A}_1^T\boldsymbol{\theta}_1 = (\boldsymbol{A}_1^T\boldsymbol{M}_1\boldsymbol{A}_1)^{-1}\boldsymbol{A}_1^T\mathbf{V}_1$$

in which the matrix to be inverted is positive definite. $\boldsymbol{A}_1$, however, is rectangular so that multiplying on the left by $\boldsymbol{A}_1$ does not necessarily serve to determine $\boldsymbol{\theta}_1$, but we may write

$$\boldsymbol{\theta}_1 = (\boldsymbol{A}_1\boldsymbol{B}_1)\begin{pmatrix} \boldsymbol{\varphi}_1 \\ \boldsymbol{\psi}_1 \end{pmatrix}$$

giving

$$\begin{pmatrix} \boldsymbol{A}_1^T \\ \boldsymbol{B}_1^T \end{pmatrix}\mathbf{V}_1 = \begin{pmatrix} \boldsymbol{A}_1^T\boldsymbol{M}_1\boldsymbol{A}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{\varphi}_1 \\ \boldsymbol{\psi}_1 \end{pmatrix}$$

$$\boldsymbol{\varphi}_1 = (\boldsymbol{A}_1^T\boldsymbol{M}_1\boldsymbol{A}_1)^{-1}\boldsymbol{A}_1^T\mathbf{V}_1$$

and $\boldsymbol{\psi}_1$ is indeterminate and free to adopt any value. We therefore adopt

$$\boldsymbol{\theta}_1 = \boldsymbol{A}_1\boldsymbol{\varphi}_1 = \boldsymbol{A}_1(\boldsymbol{A}_1^T\boldsymbol{M}_1\boldsymbol{A}_1)^{-1}\boldsymbol{A}_1^T\mathbf{V},$$

which is the smallest vector of parametric shifts which will satisfy the constraints, and allow $\boldsymbol{\psi}_1$ to be determined by the remaining observational equations in which the target vectors, $\mathbf{t}$, are now modified to $\mathbf{t}_2$ according to

$$\mathbf{t}_2 = \mathbf{t} - \boldsymbol{D}_2\boldsymbol{\theta}_1,$$

$\boldsymbol{D}_2$ and $\mathbf{t}_2$ being the derivatives and effective target vectors for the unconstrained atoms. We then solve

$$\mathbf{V}_2 = \boldsymbol{M}_2\boldsymbol{\theta}_2$$

in which $\boldsymbol{\theta}_2$ is required to be of the form

$$\boldsymbol{\theta}_2 = \boldsymbol{B}_1\boldsymbol{\psi}_1$$

giving

$$\boldsymbol{\theta}_2 = \boldsymbol{B}_1(\boldsymbol{B}_1^T\boldsymbol{M}_2\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^T\mathbf{V}_2$$

and apply the total shifts

$$\boldsymbol{\Theta} = \boldsymbol{\theta}_1 + \boldsymbol{\theta}_2$$

to obtain a structure which is optimized within the restrictions imposed by the constraints.

It may happen that $\boldsymbol{B}_1^T \boldsymbol{M}_2 \boldsymbol{B}_1$ is itself singular because there are insufficient data in the vector $\mathbf{t}_2$ to control the structure and the parametric shifts contained in $\boldsymbol{\theta}_2$ fully. In this event the same process may be applied again, basing the solution for $\boldsymbol{\theta}_2$ on

$$\begin{pmatrix} A_2^T \\ B_2^T \end{pmatrix} \boldsymbol{B}_1^T \boldsymbol{M}_2 \boldsymbol{B}_1 (A_2 B_2)$$

so that the vectors in $\boldsymbol{B}_2$ represent the degrees of freedom which remain uncommitted. This method of application of constraints by subspace sectioning may be nested to any depth and is completely general.

A valid matrix A_1 may be found from $\boldsymbol{M}_1$ by using the fact that the columns of $\boldsymbol{M}_1$ are all linear combinations of the columns of $\boldsymbol{E}_\lambda$ and are void of any contribution from $\boldsymbol{E}_0$. It follows that A_1 may be found by using the columns of $\boldsymbol{M}_1$ as priming vectors in the Gram–Schmidt process [Section 3.3.1.2.3 (i)] until the normalizing step involves division by zero. A_1 is then complete if all the columns of $\boldsymbol{M}_1$ have been tried. $(A_1 B_1)$ may then be completed by using arbitrary vectors as primers.

Manipulation of a ring of n atoms may be achieved by treating it as a chain of $(n + 2)$ atoms [having $(n + 1)$ bond lengths, n bond angles and $(n - 1)$ dihedral angles] in which atom 1 is required to coincide with atom $(n + 1)$ and atom 2 with $(n + 2)$. $\mathbf{t}_1$ then contains two vectors, namely the lack-of-closure vectors at these points, and is typically zero. A_1 is then found to have five columns corresponding to the five degrees of freedom of two points of fixed separation; $\boldsymbol{\theta}_1$ contains only zeros if the ring is initially closed, and contains ring-closure corrections if, through non-linearity or otherwise, the ring has opened. $\boldsymbol{B}_1$ contains $(p - 5)$ columns if the chain of $(n + 2)$ points has p variable parameters. It follows, if bond lengths and bond angles are treated as constants, that the seven-membered ring is the smallest ring which is flexible, that the six-membered ring (if it can be closed with the given bond angles) has no flexibility (though it may have discrete alternatives) and that it may be impossible to close a five-membered ring. Therefore some variation of bond angles and/or bond lengths is essential for the modelling of flexible five- and six-membered rings. Treating the ring as a chain of $(n + 1)$ atoms is less satisfactory as there is then no control over the bond angle at the point of ring closure.

A useful concept for the modelling of flexible five-membered rings with near-constant bond angles is the concept of the pseudorotation angle P, and amplitude θ_m, for which the jth dihedral angle is given by

$$\theta_j = \theta_m \cos\left(P + \frac{4\pi j}{5}\right).$$

This formulation has the property $\sum_{j=0}^4 \theta_j = 0$, which is not exactly true; nevertheless, θ_j values measured from observed conformations comply with this formulation within a degree or so (Altona & Sundaralingam, 1972).

Software specialized to the handling of condensed ring systems has been developed by van der Lieth *et al.* (1984) (Section 3.3.3.3.1) and by Cohen *et al.* (1981) (Section 3.3.3.3.2).

3.3.2.2.2. *Methods based on positional coordinates*

Modelling methods in which atomic coordinates are the independent variables are mathematically simpler than those using angular variables especially if the function to be minimized is a quadratic function of interatomic distances or of distances between atoms and fixed points. The method of Dodson *et al.* (1976) is representative of this class and it may be outlined as follows. If $\mathbf{d}$ is a column vector containing ideal values of the scalar distances from atoms to fixed target points or to other atoms, and if $\mathbf{l}$

is a column vector containing the prevailing values of these quantities obtained from the model, then

$$\mathbf{d} = \mathbf{l} + \boldsymbol{D}\delta\mathbf{x} + \boldsymbol{\varepsilon}$$

in which the column matrix $\boldsymbol{\delta}\mathbf{x}$ contains alterations to the atomic coordinates, $\boldsymbol{\varepsilon}$ contains residual discrepancies and $\boldsymbol{D}$ is a large sparse rectangular matrix containing values of $\partial l / \partial x$, of which there are not more than six non-zero values on any row, consisting of direction cosines of the line of which $\mathbf{l}$ is the length. $\boldsymbol{\varepsilon}^T \boldsymbol{W} \boldsymbol{\varepsilon}$ is then minimized by setting

$$\boldsymbol{D}^T \boldsymbol{W}(\mathbf{d} - \mathbf{l}) = \boldsymbol{D}^T \boldsymbol{W} \boldsymbol{D} \delta\mathbf{x},$$

which they solve by the method of conjugate gradients without searches. This places reliance on the linearity of the observational equations (Diamond, 1984*b*). It also works entirely with the sparse matrix $\boldsymbol{W}^{1/2}\boldsymbol{D}$, the dense matrix $\boldsymbol{D}^T \boldsymbol{W} \boldsymbol{D}$, and its inverse, being never calculated.

The method is extremely efficient in annealing a model structure for which an initial position for every atom is available, especially if the required shifts are within the quasi-linear region, but is less effective when large dihedral-angle changes are involved or when many atoms are to be placed purely by interpolation between a few others for which target positions are available. Interbond angles are controlled by assigning d values to second-nearest-neighbour distances; this is effective except for bond angles near 180° so that, in particular, planar groups require an out-of-plane dummy atom to be included which has no target position of its own but does have target values of distances between itself and atoms in the planar group. The method requires a d value to be supplied for every type of nearest- and next-nearest-neighbour distance in the structure, of which there are many, together with W values which are the inverse variances of the distances concerned as assessed by surveys of the corresponding distances in small-molecule structures, or from estimates of their accuracy, or from estimates of accuracy of the target positions.

Hermans & McQueen (1974) published a similar method which differs in that it moves only one atom at a time, in the environment of its neighbours, these being considered fixed while the central atom is under consideration. This is inefficient in the sense that in any one cycle one atom moves only a small fraction ($\sim$3%) of the distance it will ultimately be required to move, but individual atom cycles are so cheap and simple that many cycles can be afforded. The method was selected for inclusion in *Frodo* by Jones (1978) (Section 3.3.3.2.7) and is an integral part of the *GRIP* system (Tsernoglou *et al.*, 1977; Girling *et al.*, 1980) (Section 3.3.3.2.2) for which it was designed.

3.3.2.2.3. *Approaches to the problem of multiple minima*

Modelling methods which operate by minimizing an objective function of the coordinates (whether conformational or positional) suffer from the fact that any realistic objective function representing the potential energy of the structure is likely to have many minima in the space of the variables for any but the simplest problems. No general system has yet been devised that can ensure that the global minimum is always found in such cases, but we indicate here two approaches to this problem.

The first approach uses dynamics to escape from potential-energy minima. Molecular-mechanics simulations allow each atom to possess momentum as well as position and integrate the equations of motion, conserving the total energy. By progressively removing energy from the simulation by scaling down the momentum vectors some potential-energy minimum may be found. Conversely, a minimization of potential energy which has led to a minimum thought not to be the global minimum may be continued by introducing atomic momenta sufficient to overcome potential-