

## 3. DUAL BASES IN CRYSTALLOGRAPHIC COMPUTING

the dihedral angle will behave like the standard dihedral angle or, instead, like the torsion angle, will be found to be whether or not two or more atoms are common to both planes.

All that we have tried to bring out about the covariances of derived quantities involving the plane requires that the covariances of the experimental atom positions (reduced in our formulations to Cartesian coordinates) be included. However, such covariances of derived quantities are often not available in practice, and are usually left unused even if they are. The need to use the covariances, not just the variances, has been obvious from the beginning. It has been emphasized in another context by Schomaker & Marsh (1983) and much more strongly and generally by Waser (1973), whose pleading seems to have been generally ignored, by now, for about thirty years.

## APPENDIX A3.2.1

Consider  $n$  atoms at observed vector positions  $r$  (expressed in Cartesians),  $n$  constraints (each adjusted position  $r_a$  – ‘a’ for ‘adjusted’ – must be on the plane) and  $3n + 3$  adjustable parameters ( $3n r_a$  components and the 3 components of the vector  $a$  of reciprocal intercepts of the plane), so that the problem has  $n - 3$  degrees of freedom. The weight matrices  $\mathbf{P}$  may be differently anisotropic for each atom, but there are no interatomic correlations. As before, square brackets, ‘[. . .]’, represent the Gaussian sum over all atoms, usually suppressing the atom indices. We also write  $\lambda$ , *not* the  $\lambda$  of Section 3.2.2, for the Lagrange multipliers (one for each atom);  $m$  for the direction cosines of the plane normal; and  $d$  for the perpendicular origin-to-plane distance.

As before,  $\mathbf{P}_k$  is the reciprocal of the atomic error matrix:  $\mathbf{P}_k = \mathbf{M}_k^{-1}$  (correspondingly,  $\mathbf{P} \equiv \mathbf{M}^{-1}$ ) but ‘ $\mathbf{M}$ ’ is no longer the ‘ $\mathbf{M}$ ’ of Section 3.2.2. The appropriate least-squares sum is

$$S = [(r - r_a)^T \mathbf{P} (r - r_a)]$$

and the augmented sum for applying the method of Lagrange multipliers is

$$\Phi = S/2 - [\lambda a^T r_a].$$

$\Phi$  is to be minimized with respect to variations of the adjusted atom positions  $r_{ka}$  and plane reciprocal intercepts  $b_i$ , leading to the equations

$$\begin{aligned} 0 &= \frac{-\partial\phi}{\partial r_a^T} = \mathbf{P}(r - r_a) + \lambda \mathbf{b} \text{ and} \\ 0 &= \frac{-\partial\Phi}{\partial \mathbf{b}^T} = [\lambda r_a], \end{aligned}$$

subject to the plane conditions  $\mathbf{b}^T r_a = 1$ , each atom, with  $d^2 = 1/(\mathbf{b}^T \mathbf{b})$ ,  $m = \mathbf{b}/\sqrt{\mathbf{b}^T \mathbf{b}}$ . These equations are nonlinear.

A convenient solution runs as follows: first multiply the first equation by  $\mathbf{M}$  and solve for the adjusted atom positions in terms of the Lagrange multipliers  $\lambda$  and the reciprocal intercepts  $b$  of the plane; then multiply *that* result by  $\mathbf{b}^T$  applying the plane conditions, and solve for the  $\lambda$ 's

$$\begin{aligned} r_a &= r + \lambda \mathbf{M} \mathbf{b}, \quad \mathbf{M} \equiv \mathbf{P}^{-1} \\ 1 &= \mathbf{b}^T r + \lambda \mathbf{b}^T \mathbf{M} \mathbf{b}, \\ \lambda &= (1 - \mathbf{b}^T r)/(\mathbf{b}^T \mathbf{M} \mathbf{b}). \end{aligned}$$

Next insert these values for the  $\lambda$ 's and  $r_a$ 's into the second equation:

$$0 = \frac{\partial\Phi}{\partial \mathbf{b}^T} = [\lambda r_a] = \left[ \frac{1 - \mathbf{b}^T r}{\mathbf{b}^T \mathbf{M} \mathbf{b}} \left( r + \frac{1 - \mathbf{b}^T r}{\mathbf{b}^T \mathbf{M} \mathbf{b}} \mathbf{M} \mathbf{b} \right) \right].$$

This last equation,  $F(\mathbf{b}) = 0$ , is to be solved for  $\mathbf{b}$ . It is highly nonlinear:  $F(\mathbf{b}) = O(\mathbf{b}^3)/O(\mathbf{b}^4)$ . One can proceed to a first approximation by writing  $0 = [(1 - \mathbf{b}^T r)r]$ ; *i.e.*,  $[\mathbf{r}\mathbf{r}] \cdot \mathbf{b} = [\mathbf{r}]$ , in dyadic notation. [ $\mathbf{M} = \mathbf{I}$ , all atoms;  $1 - \mathbf{b}^T r = 0$  in the multiplier of  $\mathbf{M}\mathbf{b}/(\mathbf{b}^T \mathbf{M}\mathbf{b})$ .] A linear equation in  $\mathbf{b}$ , this approximation usually works well.<sup>6</sup> We have also used the iterative Frazer, Duncan & Collar eigenvalue solution as described by SWMB (1959), which works even when the plane passes exactly through the origin. To continue the solution of the nonlinear equations, by linearizing and iterating, write  $F(\mathbf{b}) = 0$  in the form

$$\begin{aligned} 0 &= [\lambda r + \lambda^2 \mathbf{M} \mathbf{b}] \\ &\approx \left[ r \frac{\partial\lambda}{\partial \mathbf{b}} + 2\mathbf{M} \mathbf{b} \lambda \frac{\partial\lambda}{\partial \mathbf{b}} + \lambda^2 \mathbf{M} \right] \delta \mathbf{b} + [\lambda(r + \lambda \mathbf{M} \mathbf{b})]_0, \end{aligned}$$

solve for  $\delta \mathbf{b}$ , reset  $\mathbf{b}$  to  $\mathbf{b} + \delta \mathbf{b}$ , *etc.*, until the desired degree of convergence of  $|\delta \mathbf{b}|/|\mathbf{b}|$  toward zero has been attained. By  $\partial\lambda/\partial \mathbf{b}$  is meant the partial derivative of the above expression for  $\lambda$  with respect to  $\mathbf{b}$ , as detailed in the next paragraph.

In the Fortran program *DDLELSP* (double precision Deming Lagrange, with error estimates, least-squares plane, written to explore this solution) the preceding equation is recast as

$$\begin{aligned} \mathbf{B} \delta \mathbf{b} &\equiv \left[ \lambda^2 \mathbf{M} + (r + 2\lambda \mathbf{M} \mathbf{b}) \frac{\partial\lambda}{\partial \mathbf{b}} \right] \delta \mathbf{b} \\ &\equiv \mathbf{Y} = -[\lambda(r + \lambda \mathbf{M} \mathbf{b})]_0, \end{aligned}$$

with

$$\begin{aligned} \frac{\partial\lambda}{\partial \mathbf{b}} &= r^T / \mathbf{b}^T \mathbf{M} \mathbf{b} - [2(1 - \mathbf{b}^T r)/(\mathbf{b}^T \mathbf{M} \mathbf{b})^2] \\ &= -(\mathbf{r}^T + 2\lambda \mathbf{b}^T \mathbf{M}) / \mathbf{b}^T \mathbf{M} \mathbf{b}. \end{aligned}$$

The usual goodness of fit, GOF2 in *DDLELSP*, evaluates to

$$\begin{aligned} \text{GOF2} &= \left( \frac{S_{\min}}{n-3} \right)^{1/2} = \left( \frac{1}{n-3} [\lambda^2 \mathbf{b}^T \mathbf{M} \mathbf{P} \mathbf{M} \mathbf{b}] \right)^{1/2} \\ &= \left( \frac{1}{n-3} [\lambda^2 \mathbf{b}^T \mathbf{M} \mathbf{b}] \right)^{1/2} \\ &= \left( \frac{1}{n-3} \left[ \frac{(1 - \mathbf{b}^T r)^2}{\mathbf{b}^T \mathbf{M} \mathbf{b}} \right] \right)^{1/2}. \end{aligned}$$

<sup>6</sup> We do not fully understand the curious situation of this equation. It arises immediately if the isotropic problem is formulated as one of minimizing  $[(1 - \mathbf{b}^T r)^2]$  by varying  $\mathbf{b}$ , and it fails then [SWMB (1959) referred to it as ‘an incorrect method’], as it obviously must – observe the denominator – if the plane passes too close to the origin. However, it fails in other circumstances also. The main point about it is perhaps that it is linear in  $\mathbf{b}$  and is obtained as the supposedly exact and *unique* solution of the isotropic problem, whereas the problem has no unique solution but three solutions instead (SWMB, 1959). From the point of view of Gaussian least squares, the essential fault in minimizing  $S_{\min} = [(1 - \mathbf{b}^T r)^2]$  may be that the apparently simple weighting function in it, *i.e.* the identity, is actually complicated and unreasonable. In terms of distance deviations from the plane, we have  $S_{\min} = [w(d - m^T r)^2]$ , with  $w = \mathbf{b}^T \mathbf{b} = d^{-2}$ . Prudence requires that the origin be shifted to a point sufficiently far from the plane and close enough to the centroid normal to avoid the difficulties discussed by SWMB. Note that for the one-dimensional problem of fitting a constant to a set of measurements of a single entity the Deming–Lagrange treatment with the condition  $1 = c x_a$  and weights  $w$  reduces immediately to the standard result  $1/c = [w x]/[w]$ .

### 3.2. THE LEAST-SQUARES PLANE

This is only an approximation, because the residuals  $1 - \mathbf{b}^T \mathbf{r}$  are not the differences between the observations and appropriate linear functions of the parameters, nor are their variances (the  $\mathbf{b}^T \mathbf{M} \mathbf{b}$ 's) independent of the parameters (or, in turn, the errors in the observations).

We ask also about the perpendicular distances,  $e$ , of atoms to plane and the mean-square deviation  $(\delta e)^2$  to be expected in  $e$ .

$$e = (1 - \mathbf{b}^T \mathbf{r}) / \sqrt{\mathbf{b}^T \mathbf{b}} = d(1 - \mathbf{b}^T \mathbf{r})$$

$$\delta e = -d(\mathbf{b}^T \eta + \mathbf{r}^T \varepsilon) + O(\eta^2).$$

Here  $\eta$  and  $\varepsilon$  are the errors in  $\mathbf{r}$  and  $\mathbf{b}$ . Neglecting ' $O(\eta^2)$ ' then leads to

$$\overline{(\delta e)^2} = d^2(\mathbf{b}^T \overline{\eta \eta^T} \mathbf{b} + 2\mathbf{b}^T \overline{\eta \varepsilon^T} \mathbf{r} + \mathbf{r}^T \overline{\varepsilon \varepsilon^T} \mathbf{r}).$$

We have  $\overline{\eta \eta^T} = \mathbf{M} = \mathbf{P}^{-1}$ , but  $\overline{\eta \varepsilon^T}$  and  $\overline{\varepsilon \varepsilon^T}$  perhaps still need to be evaluated.

#### References

- Arley, N. & Buch, K. R. (1950). *Introduction to the Theory of Probability and Statistics*. New York: John Wiley and London: Chapman and Hall.
- Clews, C. J. B. & Cochran, W. (1949). *The structures of pyrimidines and purines. III. An X-ray investigation of hydrogen bonding in aminopyrimidines*. *Acta Cryst.* **2**, 46–57.

- Deming, W. E. (1943). *Statistical Adjustment of Data*. New York: John Wiley.
- Hamilton, W. C. (1961). *On the least-squares plane through a set of points*. *Acta Cryst.* **14**, 185–189.
- Hamilton, W. C. (1964). *Statistics in Physical Science*. New York: Ronald Press.
- Ito, T. (1981a). *Least-squares refinement of the best-plane parameters*. *Acta Cryst.* **A37**, 621–624.
- Ito, T. (1981b). *On the least-squares plane through a group of atoms*. *Sci. Pap. Inst. Phys. Chem. Res. Saitama*, **75**, 55–58.
- Ito, T. (1982). *On the estimated standard deviation of the atom-to-plane distance*. *Acta Cryst.* **A38**, 869–870.
- Kalantar, A. H. (1987). *Slopes of straight lines when neither axis is error-free*. *J. Chem. Educ.* **64**, 28–29.
- Lybanon, M. (1984). *A better least-squares method when both variables have uncertainties*. *Am. J. Phys.* **52**, 22–26.
- Robertson, J. M. (1948). *Bond-length variations in aromatic systems*. *Acta Cryst.* **1**, 101–109.
- Schomaker, V. & Marsh, R. E. (1983). *On evaluating the standard deviation of  $U_{eq}$* . *Acta Cryst.* **A39**, 819–820.
- Schomaker, V., Waser, J., Marsh, R. E. & Bergman, G. (1959). *To fit a plane or a line to a set of points by least squares*. *Acta Cryst.* **12**, 600–604.
- Shmueli, U. (1981). *On the statistics of atomic deviations from the 'best' molecular plane*. *Acta Cryst.* **A37**, 249–251.
- Waser, J. (1973). *Dyadics and variances and covariances of molecular parameters, including those of best planes*. *Acta Cryst.* **A29**, 621–631.
- Waser, J., Marsh, R. E. & Cordes, A. W. (1973). *Variances and covariances for best-plane parameters including dihedral angles*. *Acta Cryst.* **B29**, 2703–2708.
- Whittaker, E. T. & Robinson, G. (1929). *The Calculus of Observations*. London: Blackie.