$$0 = \frac{\partial \Phi}{\partial d} = \boldsymbol{\Lambda 1}.$$

As before, multiply the third equation by $\boldsymbol{M}$ and solve for $\mathbf{R}_a$. Then substitute the result into the first and fourth equations to obtain

$$\begin{aligned} \mathsf{mR} + \mathsf{m}\boldsymbol{M}\mathsf{m}\boldsymbol{\Lambda} &= d\boldsymbol{1}, \\ \mathsf{mm} &= 1, \\ \boldsymbol{\Lambda}\mathbf{R} + \boldsymbol{\Lambda M \Lambda}\mathsf{m} &= \mu\mathsf{m}, \\ \boldsymbol{\Lambda 1} &= 0 \end{aligned}$$

as the $n + 5$ mostly nonlinear equations to be solved for m, $\boldsymbol{\Lambda}$, d and μ by linearizing (differentiation), solving for increments, and iterating, in the pattern described more fully above. An approximate solution for m and d has first to be obtained somehow, perhaps by the method of SWMB (with isotropic uncorrelated weights), checked for suitability, and extended to a full complement of first approximations by

$$\begin{aligned} \boldsymbol{\Lambda} &= (\mathsf{m}\boldsymbol{M}\mathsf{m})^{-1}(d\boldsymbol{1} - \mathsf{mR}) \\ \mu &= \mathsf{m}\boldsymbol{\Lambda}\mathbf{R} + \mathsf{m}\boldsymbol{\Lambda M \Lambda}\mathsf{m}, \end{aligned}$$

which readily follow from the previous equations. As in the 'intercepts' solution the linearized expression for the increments in $\mu, \boldsymbol{\Lambda}, d$ and m can be used together with the equation for $\mathbf{R}_a$ to obtain all the covariances needed in the treatment described in Section 3.2.2.

3.2.3.2. *Concluding remarks*

Proper tests of statistical significance of this or that aspect of a least-squares plane can be made if the plane has been based on a proper weight matrix as discussed in Section 3.2.3; if it can be agreed that the random errors of observation are normally distributed; and if an agreeable test (*null*) hypothesis can be formulated. For example, one may ask for the probability that a degree of fit of the observed positions to the least-squares plane at least as poor as the fit that was found might occur if the atoms in truth lie precisely on a plane. The χ^2 test answers this question: a table of probabilities displayed as a function of χ^2 and ν provides the answer. Here χ^2 is just our minimized

$$S = \mathsf{b}\boldsymbol{\Lambda M P M \Lambda}\mathsf{b} = \mathsf{b}\boldsymbol{\Lambda M \Lambda}\mathsf{b},$$

and

$$\begin{aligned} \nu &= n_{\text{observations}} - n_{\text{adjusted parameters}} - n_{\text{constraints}} \\ &= 3n - (n + 3) - n = n - 3, \end{aligned}$$

is the number of degrees of freedom for the problem of the plane (erroneously cited in at least one widely used crystallographic system of programs as $3n - 3$). There will not usually be any reason to believe that the atoms are exactly coplanar in any case; nevertheless, this test may well give a satisfying indication of whether or not the atoms are, in the investigator's judgment, essentially coplanar. It must be emphasized that χ^2 as calculated in Section 3.2.3 will include proper allowance for uncertainties in the d and orientation of the plane with greater reliability than the estimates of Section 3.2.2, which are based on nominally arbitrary weights. Both, however, will allow for the large variations in d and *tilt* that can arise in either case if n is small. Some of the earlier, less complete discussions of this problem have been mentioned in Section 3.2.2.

Among the problems not considered here are ones of fitting more than one plane to a set of observed positions, *e.g.* of two planes fitted to three sets of atoms associated, respectively, with the first plane, the second plane, and both planes, and of the angle between the two planes. For the atoms common to both planes there will be a fundamental point of difference between existing programs (in which, in effect, the positions of the atoms in common are inconsistently adjusted to one position on the first plane and, in general, a different position on the second) and what we would advocate as the proper procedure of requiring the adjusted positions of such atoms to lie on the line of intersection of the two planes. As to the dihedral angle there is a difficulty, noted by WMC (1973, p. 2705), that the usual formulation of $\sigma^2(\theta_0)$ in terms of the cosine of the dihedral angle reduces to 0/0 at $\theta_0 = 0$. However, this variance is obviously well defined if the plane normals and their covariances are well defined. The essential difficulty lies with the ambiguity in the direction of the line of intersection of the planes in the limit of zero dihedral angle. For the torsion angle about a line defined by two atoms, there should be no such difficulty. It seems likely that for the two-plane problem proposed above, the issue that decides whether 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 3.2.1.

Consider n atoms at observed vector positions r (expressed in Cartesians), n constraints (each adjusted position $\mathsf{r}_a - a$ 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 λ, *not* the λ 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, $\boldsymbol{P} \equiv \boldsymbol{M}^{-1}$) but '$\mathbf{M}$' is no longer the '$\mathbf{M}$' of Section 3.2.2. The appropriate least-squares sum is

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

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

$$\Phi = S/2 - [\lambda \mathsf{a}^T \mathsf{r}_a].$$

Φ 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

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

subject to the plane conditions $\mathsf{b}^T \mathsf{r}_a = 1$, each atom, with $d^2 = 1/(\mathsf{b}^T \mathsf{b})$, $\mathsf{m} = \mathsf{b}/\sqrt{\mathsf{b}^T \mathsf{b}}$. 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 λ and the reciprocal intercepts b of the plane; then multiply *that* result by b^T applying the plane conditions, and solve for the λ's

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

Next insert these values for the λ's and r_a's into the second equation:

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

This last equation, $F(\mathsf{b}) = 0$, is to be solved for b. It is highly nonlinear: $F(\mathsf{b}) = O(\mathsf{b}^3)/O(\mathsf{b}^4)$. One can proceed to a first approximation by writing $0 = [(1 - \mathsf{b}^T\mathsf{r})\mathsf{r}]$; *i.e.*, $[\mathbf{rr}] \cdot \mathsf{b} = [\mathsf{r}]$, in dyadic notation. [$\mathbf{M} = \mathbf{I}$, all atoms; $1 - \mathsf{b}^T\mathsf{r} = 0$ in the multiplier of $\mathbf{M}\mathsf{b}/(\mathsf{b}^T\mathbf{M}\mathsf{b})$.] A linear equation in b, this approximation usually works well.* 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 $F(\mathsf{b}) = 0$ in the form

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

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

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

with

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

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

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

This is only an approximation, because the residuals $1 - \mathsf{b}^T\mathsf{r}$ are not the differences between the observations and appropriate linear functions of the parameters, nor are their variances (the $\mathsf{b}^T\mathbf{M}\mathsf{b}$'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 $\overline{(\delta e)^2}$ to be expected in e.

$$\begin{aligned} e &= (1 - \mathsf{b}^T\mathsf{r})/\sqrt{\mathsf{b}^T\mathsf{b}} = d(1 - \mathsf{b}^T\mathsf{r}) \\ \delta e &= -d(\mathsf{b}^T\eta + \mathsf{r}^T\varepsilon) + O(\eta^2). \end{aligned}$$

Here η and ε are the errors in r and b. Neglecting '$O(\eta^2)$' then leads to

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

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.

* We do not fully understand the curious situation of this equation. It arises immediately if the isotropic problem is formulated as one of minimizing $[(1 - \mathsf{b}^T\mathsf{r})^2]$ by varying 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 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 $S_{\mathrm{lin}} = [(1 - \mathsf{b}^T\mathsf{r})^2]$ 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 $S_{\mathrm{lin}} = [w(d - \mathsf{m}^T\mathsf{r})^2]$, with $w = \mathsf{b}^T\mathsf{b} = d^{-2}$. 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 $1 = cx_a$ and weights w reduces immediately to the standard result $1/c = [wx]/[w]$.

references