3.2. The least-squares plane

BY R. E. MARSH AND V. SCHOMAKER

3.2.1. Introduction

By way of introduction, we remark that in earlier days of crystal structure analysis, before the advent of high-speed computers and routine three-dimensional analyses, molecular planarity was often assumed so that atom coordinates along the direction of projection could be estimated from two-dimensional data [see, *e.g.*, Robertson (1948)]. Today, the usual aim in deriving the coefficients of a plane is to investigate the degree of planarity of a group of atoms as found in a full, three-dimensional structure determination. We further note that, for such purposes, a crystallographer will often be served just as well by establishing the plane in an almost arbitrary fashion as by resorting to the most elaborate, nit-picking and pretentious least-squares treatment. The approximate plane and the associated perpendicular distances of the atoms from it will be all he needs as scaffolding for his geometrical and structural imagination; reasonable common sense will take the place of explicit attention to error estimates.

Nevertheless, we think it appropriate to lay out in some detail the derivation of the 'best' plane, in a least-squares sense, through a group of atoms and of the standard uncertainties associated with this plane. We see two cases: (1) The weights of the atoms in question are considered to be isotropic and uncorrelated (*i.e.* the weight matrix for the positions of all the atoms is diagonal, when written in terms of Cartesian axes, and for each atom the three diagonal elements are equal). In such cases the weights may have little or nothing to do with estimates of random error in the atom positions (they may have been assigned merely for convenience or convention), and, therefore, no one should feel that the treatment is proper in respect to the theory of errors. Nevertheless, it may be desired to incorporate the error estimates (variances) of the atom positions into the *results* of such calculations, whereupon these variances (which may be anisotropic, with correlation between atoms) need to be propagated. In this case the distinction between *weights* (or their inverses) and *variances* must be kept very clear. (2) The weights are anisotropic and are presumably derived from a variance–covariance matrix, which may include correlation terms between different atoms; the objective is to achieve a truly proper Gaussian least-squares result.

3.2.2. Least-squares plane based on uncorrelated, isotropic weights

This is surely the most common situation; it is not often that one will wish to take the trouble, or be presumptive enough, to assign anisotropic or correlated weights to the various atoms. And one will sometimes, perhaps even often, not be genuinely interested in the hypothesis that the atoms actually are rigorously coplanar; for instance, one might be interested in examining the best plane through such a patently non-planar molecule as cyclohexane. Moreover, the calculation is simple enough, given the availability of computers and programs, as to be a practical realization of the off-the-cuff treatment suggested in our opening paragraph. The problem of deriving the plane's coefficients is intrinsically nonlinear in the way first discussed by Schomaker *et al.* (1959; SWMB). Any formulation other than as an eigenvalue–eigenvector problem (SWMB), as far as we can tell, will sometimes go astray. As to the propagation of errors, numerous treatments have been given, but none that we have seen is altogether satisfactory.

We refer all vectors and matrices to Cartesian axes, because that is the most convenient in calculation. However, a more elegant formulation can be written in terms of general axes [*e.g.*, as in Shmueli (1981)].

The notation is troublesome. Indices are needed for atom number and Cartesian direction, and the exponent 2 is needed as well, which is difficult if there are superscript indices. The best way seems to be to write all the indices as subscripts and distinguish among them by context – i, j, 1, 2, 3 for directions; k, l, p (and sometimes $K, \ldots$) for atoms. In any case, *atom* first then *direction* if there are two subscripts; *direction,* if only one index for a vector component, but *atom* (in this section at least) if for a weight or a vector. And σ_{d_1}, *e.g.*, for the standard uncertainty of the distance of atom 1 from a plane. For simplicity in practice, we use Cartesian coordinates throughout.

The first task is to find the plane, which we write as

$$0 = \mathbf{m} \cdot \mathbf{r} - d \equiv \mathsf{m}^T \mathsf{r} - d,$$

where $\mathbf{r}$ is here the vector from the origin to any point on the plane (but usually represents the measured position of an atom), $\mathbf{m}$ is a unit vector parallel to the normal from the origin to the plane, d is the length of the normal, and m and r are the column representations of $\mathbf{m}$ and $\mathbf{r}$. The least-squares condition is to find the stationary values of $S \equiv [w_k(\mathsf{m}^T \mathsf{r}_k - d)^2]$ subject to $\mathsf{m}^T \mathsf{m} = 1$, with $\mathsf{r}_k, k = 1, \ldots, n$, the vector from the origin to atom k and with weights, w_k, isotropic and without interatomic correlations for the n atoms of the plane. We also write S as $S \equiv [w(\mathsf{m}^T \mathsf{r} - d)^2]$, the subscript for atom number being implicit in the Gaussian summations ($[\ldots]$) over all atoms, as it is also in the angle-bracket notation for the weighted average over all atoms, for example in $\langle \mathsf{r} \rangle$ – the weighted centroid of the groups of atoms – just below.

First solve for d, the origin-to-plane distance.

$$0 = -\frac{1}{2}\frac{\partial S}{\partial d} = [w(\mathsf{m}^T \mathsf{r} - d)] = 0,$$

$$d = [w\mathsf{m}^T \mathsf{r}]/[w] \equiv \mathsf{m}^T \langle \mathsf{r} \rangle.$$

Then

$$S \equiv [w(\mathsf{m}^T \mathsf{r} - d)^2] = [w\{\mathsf{m}^T(\mathsf{r} - \langle \mathsf{r} \rangle)\}^2]$$

$$\equiv [w(\mathsf{m}^T \mathsf{s})^2] \equiv \mathsf{m}^T[w\mathsf{s}\mathsf{s}^T]\mathsf{m} \equiv \mathsf{m}^T \mathsf{A}\mathsf{m}.$$

Here $\mathbf{s}_k \equiv \mathsf{r}_k - \langle \mathsf{r} \rangle$ is the vector from the centroid to atom k. Then solve for $\mathbf{m}$. This is the eigenvalue problem – to diagonalize A (bear in mind that A_{ij} is just $[ws_i s_j]$) by rotating the coordinate axes, *i.e.*, to find the 3×3 arrays M and L, L diagonal, to satisfy

$$\mathsf{M}^T \mathsf{A}\mathsf{M} = \mathsf{L}, \qquad \mathsf{M}^T \mathsf{M} = \mathsf{I}.$$

A and M are symmetric; the columns m of M are the direction cosines of, and the diagonal elements of L are the sums of weighted squares of residuals from, the best, worst and intermediate planes, as discussed by SWMB.

3.2.2.1. *Error propagation*

Waser *et al.* (1973; WMC) carefully discussed how the random errors of measurement of the atom positions propagate into the derived quantities in the foregoing determination of a least-squares plane. This section presents an extension of their discussion. To begin, however, we first show how standard first-order perturbation theory conveniently describes the propagation of error into M and L when the positions r_k of the atoms are incremented by the amounts $\delta \mathsf{r}_k \equiv \xi_k$ and the corresponding quantities $\mathbf{s}_k \equiv \mathsf{r}_k - \langle \mathsf{r} \rangle$ (the vectors from the centroid to the atoms) by the amounts $\eta_k, (\mathbf{s} \rightarrow \mathbf{s} + \eta)$, $\eta_k \equiv \xi_k - \langle \xi \rangle$. (The need to account for the variation in position of the centroid, *i.e.* to distinguish between η_k and ξ_k, was overlooked by WMC.) The consequent increments in A, M and L are

$$\delta \mathbf{A} = [w\eta \mathbf{s}^T] + [w\mathbf{s}\eta^T] \equiv \alpha,$$

$$\delta \mathbf{M} = \mathbf{M}\mu,$$

$$\delta \mathbf{L} \equiv \lambda.$$

Here the columns of μ are expressed as linear combinations of the columns of $\mathbf{M}$. Note also that both perturbations, μ and λ, which are the adjustments to the orientations and associated eigenvalues of the principal planes, will depend on the reduced coordinates $\mathbf{s}$ and the perturbing influences ξ by way of α, which in turn depends only on the reduced coordinates and the reduced shifts η_k. In contrast,

$$\delta \mathsf{d} = \delta(\mathsf{m}^T \langle \mathsf{r} \rangle) = (\delta \mathsf{m}^T) \langle \mathsf{r} \rangle + \mathsf{m}^T \langle \xi \rangle;$$

the change in the origin-to-plane distance for the plane defined by the column vectors $\mathbf{m}$ of $\mathbf{M}$, depends on the $\langle \mathsf{r} \rangle$ and $\langle \xi \rangle$ directly as well as on the $\mathbf{s}$ and η by way of the $\delta \mathbf{m}$.

The first-order results arising from the standard conditions, $\mathbf{M}^T \mathbf{M} = \mathbf{I}, \mathbf{L}$ diagonal, and $\mathbf{M}^T \mathbf{A} \mathbf{M} = \mathbf{L}$, are

$$\mu^T + \mu = 0, \lambda \text{ diagonal},$$

and

$$\mu^T \mathbf{M}^T \mathbf{A} \mathbf{M} + \mathbf{M}^T \alpha \mathbf{M} + \mathbf{M}^T \mathbf{A} \mathbf{M} \mu = \mu^T \mathbf{L} + \mathbf{L}\mu + \mathbf{M}^T \alpha \mathbf{M} = \lambda.$$

Stated in terms of the matrix components λ_{ij} and μ_{ij}, the first condition is $\mu_{ij} = -\mu_{ji}$, hence $\mu_{ii} = 0$, $i,j = 1,2,3$, and the second is $\lambda_{ij} = 0$, $i \neq j$. With these results the third condition then reads

$$\lambda_{jj} = (\mathbf{M}^T \alpha \mathbf{M})_{jj}, \qquad j = 1,2,3$$

$$\mu_{ij} = (\mathbf{M}^T \alpha \mathbf{M})_{ij}/(L_{jj} - L_{ii}), \quad i \neq j, \quad i,j = 1,2,3.$$

All this is analogous to the usual first-order perturbation theory, as, for example, in elementary quantum mechanics.

Now rotate to the coordinates defined by WMC, with axes parallel to the original eigenvectors $[\mathbf{M} = \mathbf{I}, A_{ij} = L_{ij}\delta_{ij}, (\mathbf{M}^T \alpha \mathbf{M})_{ij} = \alpha_{ij}]$, restrict attention to the best plane $(M_{13} \equiv m_1 = 0, M_{23} \equiv m_2 = 0, M_{33} \equiv m_3 = 1)$, and define ε^T as $(\delta m_1, \delta m_2, \delta d_c)$, keeping in mind $\delta m_3 = \mu_{33} = 0$; d_c itself, the original plane-to-centroid distance, of course vanishes. One then finds

$$\delta m_i \equiv \varepsilon_i = \alpha_{i3}/(L_{33} - L_{ii})$$

$$= [w(s_i \eta_3 + s_3 \eta_i)]/([ws_3^2] - [ws_i^2]), \quad i = 1,2,$$

$$\delta d_c \equiv \varepsilon_3 = [w\xi_3]/[w] \equiv \langle \xi_3 \rangle, \qquad \xi_k \equiv \delta r_k,$$

and also

$$\delta d = \varepsilon_1 \langle r_1 \rangle + \varepsilon_2 \langle r_2 \rangle + \varepsilon_3.$$

These results have simple interpretations. The changes in direction of the plane normal (the δm_i) are rotations, described by ε_1 and ε_2, in response to changes in moments acting against effective torsion force constants. For ε_2, for example, the contribution of atom k to the total relevant moment, about direction 1, is $-w_k s_{k3} s_{k2}$ ($w_k s_{k3}$ the 'force' and s_{k2} the lever arm), and its nominally first-order change has two parts, $-w_k s_{k2} \eta_3$ from the change in force and $-w_k s_{k3} \eta_2$ from the change in lever arm; the resisting torsion constant is $[ws_2^2] - [ws_3^2]$, which, reflection will show, is qualitatively reasonable if not quantitatively obvious. The perpendicular displacement of the plane from the original centroid $\langle r \rangle$ is ε_3, but there are two further contributions to δd, the change in distance from origin to plane along the plane normal, that arise from the two components of out-of-plane rotation of the plane about its centroid. Note that ε_3 is not given by $[w\eta_3]/[w] = [w(\xi_3 - \langle \xi_3 \rangle)]/[w]$, which vanishes identically.

There is a further, somewhat delicate point: If the group of atoms is indeed essentially coplanar, the s_{k3} are of the same order of magnitude as the η_{ki}, unlike the $s_{ki}, i \neq 3$, which are in general about as big as the lateral extent of the group. It is then appropriate to drop all terms in η_i or ξ_i, $i \neq 3$, and, in the denominators, the terms in s_{k3}^2.

The covariances of the derived quantities (by *covariances* we mean here both variances and covariances) can now be written out rather compactly by extending the implicit designation of atom numbers to double sums, the first of each of two similar factors referring to the first atom index and the second to the second, *e.g.*, $\sum_{kl} ww(s_i s_j) \ldots \equiv \sum_{kl} w_k w_l (s_{ki} s_{lj}) \ldots$. Note that the various covariances, *i.e.* the averages over the presumed population of random errors of replicated measurements, are indicated by overlines, angle brackets having been pre-empted for averages over sets of atoms.

$$\mathrm{cov}(m_i, m_j) \equiv \overline{\varepsilon_i \varepsilon_j}$$

$$= \frac{\sum_{kl} ww(s_i s_j \overline{\eta_3 \eta_3} + s_3 s_3 \overline{\eta_i \eta_j} + s_i s_3 \overline{\eta_3 \eta_j} + s_3 s_j \overline{\eta_i \eta_3})}{\{[w(s_3^2 - s_i^2)]\}\{[w(s_3^2 - s_j^2)]\}}, i,j = 1,2$$

$$\mathrm{cov}(m_i, d_c) \equiv \overline{\varepsilon_i \varepsilon_3} = \frac{\sum_{kl} ww(s_{ki} \overline{\eta_3 \xi_3} + s_{k3} \overline{\eta_i \xi_3})}{\{[w(s_3^2 - s_i^2)]\}[w]}, \quad i,j = 1,2$$

$$\sigma^2(d_c) \equiv \overline{\varepsilon_3^2} = \frac{\sum_{kl} ww \overline{\xi_3 \xi_3}}{[w]^2}$$

$$\sigma^2(d) \equiv \langle (\delta d)^2 \rangle = \langle r_1 \rangle^2 \overline{\varepsilon_1^2} + \langle r_2 \rangle^2 \overline{\varepsilon_2^2} + \overline{\varepsilon_3^2} + 2\langle r_1 \rangle \langle r_2 \rangle \overline{\varepsilon_1 \varepsilon_2}$$
$$+ 2\langle r_1 \rangle \overline{\varepsilon_1 \varepsilon_3} + 2\langle r_2 \rangle \overline{\varepsilon_2 \varepsilon_3}.$$

Interatomic covariance (*e.g.*, $\overline{\eta_{k3} \eta_{l3}}$, $k \neq l$) thus presents no formal difficulty, although actual computation may be tedious. Nonzero covariance for the η's *may* arise explicitly from interatomic covariance (*e.g.*, $\overline{\xi_{ki} \xi_{lj}}$, $k \neq l$) of the errors in the atomic positions r_k, and it will *always* arise implicitly because $\langle \xi \rangle$ in $\eta_k = \xi_k - \langle \xi \rangle$ includes all the ξ_k and therefore has nonzero covariance with all of them and with itself, even if there is no interatomic covariance among the ξ_i's.

If both types of interatomic covariance (explicit and implicit) are negligible, the ε covariances simplify a great deal, the double summations reducing to single summations. [The formal expression for $\sigma^2(d)$ does not change, so it will not be repeated.]

$$\mathrm{cov}(m_i, m_j) \equiv \overline{\varepsilon_i \varepsilon_j}$$

$$= \frac{[w^2(s_i s_j \overline{\eta_3^2} + s_3^2 \overline{\eta_i \eta_j} + s_i s_3 \overline{\eta_3 \eta_j} + s_3 s_j \overline{\eta_i \eta_3}]}{\{[w(s_3^2 - s_i^2)]\}\{[w(s_3^2 - s_j^2)]\}}, i,j = 1,2$$

$$\mathrm{cov}(m_i, d_c) \equiv \overline{\varepsilon_i \varepsilon_3} = \frac{[w^2(s_i \overline{\eta_3 \xi_3} + s_3 \overline{\eta_i \xi_3})]}{\{[w(s_3^2 - s_i^2)]\}[w]}, \quad i,j = 1,2$$

$$\sigma^2(d_c) \equiv \overline{\varepsilon_3^2} = \frac{[w^2 \overline{\xi_3^2}]}{[w]^2}.$$

When the variances are the same for ξ as for η (i.e. $\overline{\xi_i \xi_j} = \overline{\eta_i \eta_j}$, all i,j) and the covariances all vanish ($\overline{\xi_i \xi_j} = 0$, $i \neq j$), the $\overline{\varepsilon_i \varepsilon_j}$ simplify further. If the variances are also isotropic ($\overline{\xi_i^2} = \overline{\xi_j^2} = \sigma^2$, all i, j), there is still further simplification to

$$\sigma^2(m_i) \equiv \overline{\varepsilon_i^2} = \frac{[w^2 \sigma^2(s_i^2 + s_3^2)]}{\{[w(s_3^2 - s_i^2)]\}^2}, \quad i = 1,2$$

$$\sigma^2(d_c) \equiv \overline{\varepsilon_3^2} = [w^2 \sigma^2]/[w]^2$$

$$\mathrm{cov}(m_1, m_2) \equiv \overline{\varepsilon_1 \varepsilon_2} = \frac{[w^2 \sigma^2 s_1 s_2]}{\{[w(s_3^2 - s_1^2)]\}\{[w(s_3^2 - s_2^2)]\}}$$

$$\mathrm{cov}(m_i, d_c) \equiv \overline{\varepsilon_i \varepsilon_3} = \frac{[w^2 \sigma^2 s_i]}{\{[w(s_3^2 - s_i^2)]\}[w]}, \quad i = 1,2.$$

If allowance is made for the difference in definition between ε_3 and

δd, these expressions are equivalent to the ones (equations 7–9) given by WMC, who, however, do not appear to have been aware of the distinction between η and ξ and the possible consequences thereof.

If, finally, w^{-1} for each atom is taken equal to its $\overline{\eta_{lj}^2} = \sigma^2$, all j, there is still further simplification.

$$\sigma^2(m_i) \equiv \overline{\varepsilon_i^2} = \frac{[w(s_3^2 + s_i^2)]}{\{[w(s_3^2 - s_i^2)]\}^2}, \quad i = 1, 2$$

$$\sigma^2(d_c) \equiv \overline{\varepsilon_3^2} = [w]/[w]^2 = 1/[w]$$

$$\text{cov}(m_1, m_2) \equiv \overline{\varepsilon_1 \varepsilon_2} = \frac{[w s_1 s_2]}{\{[w(s_3^2 - s_1^2)]\}\{[w(s_3^2 - s_2^2)]\}}$$

$$\text{cov}(m_i, d_c) \equiv \overline{\varepsilon_i \varepsilon_3} = \frac{[w s_i]}{\{[w(s_3^2 - s_i^2)]\}[w]}, \quad i = 1, 2.$$

For the earlier, more general expressions for the components of $\overline{\varepsilon \varepsilon^T}$ it is still necessary to find $\overline{\eta_{ki}\eta_{lj}}$ and $\eta_{ki}\xi_{l3}$ in terms of $\xi_{ki}\xi_{lj}$, with $\eta_{ki} \equiv \delta s_{ki} = \delta(r_{ki} - \langle r_{ki} \rangle) = \xi_{ki} - \langle \xi_i \rangle = \sum_l w_l (\xi_{ki} - \xi_{li})/[w]$.

$$\overline{\eta_{ki}\eta_{pj}} = \sum_{l,q} w_l w_q \overline{(\xi_{ki} - \xi_{li})(\xi_{pj} - \xi_{qj})}/[w]^2$$

$$= \overline{(\xi_{ki} - \langle \xi_i \rangle)(\xi_{pj} - \langle \xi_j \rangle)}$$

$$\overline{\eta_{ki}\xi_{p3}} = \sum_l w_l \overline{(\xi_{ki} - \xi_{li})\xi_{p3}}/[w] = \overline{\xi_{ki}\xi_{p3}} - \overline{\langle \xi_i \rangle \xi_{p3}}.$$

In the isotropic, 'no-correlation' case, for example, these reduce to

$$\overline{\eta_{ki}\eta_{pi}} = -w_k \overline{\xi_{ki}^2}/[w] - w_p \overline{\xi_{pi}^2}/[w]$$
$$+ [w^2 \overline{\xi_i^2}]/[w]^2, \quad k \neq p; i = 1, 2$$

$$\overline{\eta_{ki}^2} = (1 - 2w_k/[w])\overline{\xi_{ki}^2} + [w^2 \overline{\xi_i^2}]/[w]^2, \quad i = 1, 2$$

$$\overline{\eta_{k3}\eta_{p3}} = -w_p \overline{\xi_{p3}^2}/[w],$$

and

$$\overline{\eta_{k3}^2} = \overline{\xi_{k3}^2} - w_k \overline{\xi_{k3}^2}/[w] = \overline{\xi_{k3}^2}(1 - w_k/[w]).$$

Here the difference between the correct covariance values and the values obtained on ignoring the variation in $\langle r \rangle$ may be important if the number of defining atoms is small, say, 5 or 4 or, in the extreme, 3.

3.2.2.2. *The standard uncertainty of the distance from an atom to the plane*

There are two cases, as has been pointed out, *e.g.*, by Ito (1982).

(1) The atom (atom K) was *not* included in the specification of the plane.

$$d_K = \mathsf{m}^T(\mathsf{r}_K - \langle \mathsf{r} \rangle) = r_{K3} - \langle r_3 \rangle$$
$$\delta d_K = \xi_{K3} + s_{K1}\varepsilon_1 + s_{K2}\varepsilon_2 - \varepsilon_3$$
$$\sigma_{d_K}^2 = \overline{\xi_{K3}^2} + s_{K1}^2 \overline{\varepsilon_1^2} + s_{K2}^2 \overline{\varepsilon_2^2} + \overline{\varepsilon_3^2}$$
$$+ 2s_{K1}s_{K2}\overline{\varepsilon_1 \varepsilon_2} - 2s_{K1}\overline{\varepsilon_1 \varepsilon_3} - 2s_{K2}\overline{\varepsilon_2 \varepsilon_3}$$
$$+ 2s_{K1}\overline{\xi_{K3}\varepsilon_1} + 2s_{K2}\overline{\xi_{K3}\varepsilon_2} - 2\overline{\xi_{K3}\varepsilon_3}.$$

In the isotropic, 'no-correlation' case the last three terms, *i.e.* the terms in $\overline{\varepsilon_i \xi_{K3}}$, are all negligible or zero.

In either case the value for $\overline{\xi_{K3}^2}$ and the appropriate $\overline{\varepsilon_i \varepsilon_j}$ values from the least-squares-plane calculation need to be inserted.

(2) Atom K *was* included in the specification of the plane. The expression for $\sigma_{d_K}^2$ remains the same, but the averages in it may be importantly different.

For example, consider a plane defined by only three atoms, one of overwhelmingly great w at $(0, 0, 0)$, one at $(1, 0, 0)$ and one at $(0, 1, 0)$. The centroid is at $(0, 0, 0)$ and we take $K = 2$, *i.e.* σ_{d_2} is the item of interest. Of course, it is obvious without calculation that the standard uncertainties vanish for the distances of the three atoms from the plane they alone define; the purpose here is only to show, in one case for one of the atoms, that the calculation gives the same result, partly, it will be seen, because the change in orientation of the plane is taken into account. If the only variation in the atom positions is described by $\overline{\xi_{23}^2} = \sigma^2$, one has $s_{21} = 1, \varepsilon_3 = \varepsilon_2 = 0$, $\varepsilon_1 = -\xi_{23}$, and $\overline{\xi_{K3}\varepsilon_1} = \sigma^2$. The non-vanishing terms in the desired variance are then

$$\sigma_{d_2}^2 = \overline{\xi_{23}^2} + s_{21}^2 \overline{\varepsilon_1^2} + 2s_{21}\overline{\xi_{23}\varepsilon_1}$$
$$= (1 + 1 - 2)\sigma^2 = 0.$$

If, however, the problem concerns the same plane and a fourth atom at position $(1, 0, r_{43})$, *not* included in the specification of the plane and uncertain only in respect to r_{43} (which is arbitrary) with $\overline{\xi_{43}^2} = \sigma^2$ (the same mean-square variation in direction 3 as for atom 2) and $\overline{\xi_{43}\xi_{23}} = 0$, the calculation for $\sigma_{d_4}^2$ runs the same as before, except for the third term:

$$\sigma_{d_4}^2 = (1 + 1 - 0)\sigma^2 = 2\sigma^2.$$

Extreme examples of this kind show clearly enough that variation in the direction of the plane normal or in the normal component of the centroid position will sometimes be important, the remarks to the contrary by Shmueli (1981) and, for the centroid, the omission by WMC notwithstanding. If only a few atoms are used to define the plane (*e.g.*, three or, as is often the case, a very few more), both the covariance with the centroid position and uncertainty in the direction of the normal are likely to be important. The uncertainty in the normal may still be important, even if a goodly number of atoms are used to define the plane, whenever the test atom lies near or beyond the edge of the lateral domain defined by the other atoms.

3.2.3. The proper least-squares plane, with Gaussian weights

If it is desired to weight the points to be fitted by a plane in the sense of Gaussian least squares, then two items different from what we have seen in the crystallographic literature have to be brought into view: (1) the weights may be anisotropic and include interatomic correlations, because the error matrix of the atom coordinates may in general be anisotropic and include interatomic correlations; and (2) it has to be considered that the atoms truly lie on a plane and that their observed positions are to be adjusted to lie precisely on that plane, whatever its precise position may turn out to be and no matter what the direction, in response to the anisotropic weighting, of their approach to the plane.

An important consequence of (1), the non-diagonal character of the weight matrix, even with Cartesian coordinates, is that the problem is no longer an ordinary eigenvalue problem as treated by SWMB (1959),* not even if there is no interatomic correlation and

---

* A simple two-dimensional problem illustrates the point. A regular polygon of n atoms is to define a 'best' line (always a *central* line). If the error matrix (the same for each atom) is isotropic, the weighted sum of squares of deviations from the line is independent of its orientation for $n > 2$, *i.e.* the problem is a degenerate eigenvalue problem, with two equal eigenvalues. However, if the error ellipsoids are not isotropic and are all oriented radially or all tangentially (these are merely the two orientations tried), the sum has $n/2$ equal minima for even n and 2 equal minima for odd n, in the one-π range of possible orientations of the line.

Possibly similar peculiarities might be imagined if the anisotropic weights were more complicated (*e.g.*, 'star' shaped) than can be described by a non-singular matrix, or by any matrix. Such are of course excluded here.