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

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

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

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

$\mathbf{A}$ and $\mathbf{M}$ are symmetric; the columns m of $\mathbf{M}$ are the direction cosines of, and the diagonal elements of $\mathbf{L}$ 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 $\mathbf{M}$ and $\mathbf{L}$ when the positions r_k of the atoms are incremented by the amounts $\delta\mathsf{r}_k \equiv \xi_k$ and the corresponding quantities $\mathsf{s}_k \equiv \mathsf{r}_k - \langle\mathsf{r}\rangle$ (the vectors from the centroid to the atoms) by the amounts $\eta_k, (\mathsf{s} \rightarrow \mathsf{s} + \eta)$, $\eta_k \equiv \xi_k - \langle\xi\rangle$. (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 $\mathbf{A}, \mathbf{M}$ and $\mathbf{L}$ are

references