COOPER, M. J. & ROUSE, K. D. (1972). *Z. Kristallogr.* **135**, 316–317.
CORLISS, L. M., HASTINGS, J. M., KUNNMANN, W. & BANKS, E. (1966). *Acta Cryst.* **21**, A95.
DAVIS, H. H. (1960). *Phys. Rev.* **120**, 789–801.
FRIEDMAN, Z., SHAKAD, H. & SHTRIKMAN, S. (1967). *Phys. Lett.* **25A**, 9–10.
GALLAGHER, P. K., MACCHESNEY, J. B. & BUCHANAN, D. N. E. (1964). *J. Chem. Phys.* **41**, 2429–2434.
GALLAGHER, P. K., MACCHESNEY, J. B. & BUCHANAN, D. N. E. (1965). *J. Chem. Phys.* **43**, 516–520.
GELLER, S., GRANT, R. W. & GONSER, U. (1971). *Prog. Solid State Chem.* **5**, 1–26.
HUBBARD, J. & MARSHALL, W. (1965). *Proc. Phys. Soc.* **86**, 561–572.
NEUTRON DIFFRACTION COMMISSION (1972). *Acta Cryst.* A**28**, 357–358.
RIETVELD, H. M. (1967). *Acta Cryst.* **22**, 151–152.
SMITH, D. K. (1962). *Acta Cryst.* **15**, 1146–1152.
STRENZWILK, D. F. & ANDERSON, E. E. (1968). *Phys. Rev.* **175**, 654–659.
TAKEDA, T., YAMAGUCHI, Y., TOMIYOSHI, S., FUKASE, M., SUGIMOTO, M. & WATANABE, H. (1968). *J. Phys. Soc. Japan*, **24**, 446–452.
TOFIELD, B. C. & FENDER, B. E. F. (1970). *J. Phys. Chem. Solids*, **31**, 2741–2749.
TOFIELD, B. C., GREAVES, C. G. & FENDER, B. E. F. (1975). To be published.
WATSON, R. E. & FREEMAN, A. J. (1961). *Acta Cryst.* **14**, 27–37.

Acta Cryst. (1975). B**31**, 646

Scaling of Intensities

BY T. R. LOMER AND A. J. C. WILSON

Department of Physics, University of Birmingham, Birmingham B15 2TT, England

(*Received* 13 *September* 1974; *accepted* 27 *September* 1974)

The supposed object of a scaling factor is to make the calculated intensity of reflexion, on the average, equal to the observed intensity. Refinement of an adjustable scaling factor by least-squares procedures does not achieve this aim; there is a systematic error whose magnitude depends on the residual refined.

Ideally the calculated and observed values for the intensity of an X-ray reflexion should agree exactly. In practice, there are always differences between the observed intensities I and the calculated intensities H; the various residuals used in least-squares refinement are different average measures of the differences. Only if certain conditions are satisfied will refinement of different residuals lead to the same values of the parameters refined, and still other conditions must be satisfied if the values of the parameters are to be unbiased (see, for example, Wilson, 1973). Observed intensities are normally on a relative scale, and the calculated intensities must therefore be multiplied by a scaling factor E before making the comparison. [The reasons for including the scaling factor in H rather than in I have been given by, for example, Cruickshank (1970).] Appropriate values of the scaling factor can be obtained at each stage of the refinement by making the average values of H and I agree, an extension of the approximate procedure proposed by Wilson (1942). Least-squares adjustment of the scaling parameter will normally give significantly different values of E. Wilson (1974*a*) has shown that if

$$R_1=\langle(F-G)^2\rangle/\langle F^2\rangle \qquad (1)$$

is the residual based on the magnitude of the observed structure factor F and of the calculated structure factor G, scaled so that

$$\langle F^2\rangle\equiv\langle I\rangle=\langle H\rangle\equiv\langle G^2\rangle\,, \qquad (2)$$

then refinement of the scaling factor E_1 in the modified residual

$$S_1=\langle(F-E_1G)^2\rangle/\langle F^2\rangle \qquad (3)$$

leads to

$$E_1=1-\tfrac{1}{2}R_1\,, \qquad (4)$$

$$(S_1)_{\min}=R_1(1-\tfrac{1}{4}R_1)\,. \qquad (5)$$

In typical cases E_1 will differ from the value 1 [expected from (2)] by some per cent. The present note discusses scaling in the case of the residual R_2. Ordinarily this gives simpler calculations than R_1 does (Wilson, 1969; Lenstra, 1974), but here the results are more complex and less straightforward in interpretation.

The residual R_2 is defined by

$$R_2=\langle(I-H)^2\rangle/\langle I^2\rangle\,, \qquad (6)$$

where H has been scaled in accordance with (2). The residual modified by including an adjustable scaling factor is

$$S_2=\langle(I-E_2H)^2\rangle/\langle I^2\rangle\,. \qquad (7)$$

Multiplying out and minimizing with respect to E_2 gives

$$R_2=1-2\langle HI\rangle/\langle I^2\rangle+\langle H^2\rangle/\langle I^2\rangle\,, \qquad (8)$$

$$E_2=\langle HI\rangle/\langle H^2\rangle\,, \qquad (9)$$

$$(S_2)_{\min}=1-\langle HI\rangle^2/\langle H^2\rangle\langle I^2\rangle\,. \qquad (10)$$

These equations contain two unknown ratios, $\langle H^2\rangle/\langle I^2\rangle$ and $\langle HI\rangle/\langle I^2\rangle$, so that E_2 and $(S_2)_{\min}$ cannot be expressed in terms of R_2 alone. Eliminating $\langle HI\rangle$ between (8) and (9) gives

$$E_2=1+[(1-R_2)\langle I^2\rangle-\langle H^2\rangle]/2\langle H^2\rangle\,. \qquad (11)$$

The scaling factor E_2 is thus greater than (less than) unity if $(1-R_2)\langle I^2\rangle$ is greater than (less than) $\langle H^2\rangle$.

The differences between the observed and calculated intensities arise from three sources: (i) random errors in the observed intensities (statistical fluctuations in counting rates or the equivalent, irregular backlash, *etc.*); (ii) defects in the model from which the intensities are calculated (incorrect scattering factors, lack of proper allowance for extinction, failure to allow for anharmonic temperature motion, *etc.*); and (iii) incomplete refinement of adjustable positional *etc.* parameters in the model. This division has been discussed many times, for example by Shoemaker (1968) and Wilson (1973). It will be noted that what are sometimes regarded as 'systematic errors' in the observations are here classified as defects in the model. From one point of view the true intensity, say K, is

$$K=I+\sigma\,, \qquad (12)$$

where σ is the random error, and from another

$$K=H+\delta\,, \qquad (13)$$

where δ is the effect of defects in the model. Obviously, from the nature of random error,

$$\langle\sigma\rangle=\langle K\sigma\rangle=0\,, \qquad (14)$$

but there is no reason to suppose that

$$\langle K\delta\rangle=0\,, \qquad (15)$$

though (2) requires that

$$\langle\delta\rangle=0\,. \qquad (16)$$

The scaling factor E_2 derived from (9) will have the 'expected' value unity if the only error present is random, as in this case $H=K$ and $I=K-\sigma$. At the other extreme, if the random error is negligible but the model is defective, the scaling factor becomes

$$E_2=\langle H(H-\delta)\rangle/\langle H^2\rangle \qquad (17)$$

$$=1-\langle H\delta\rangle/\langle H^2\rangle\,. \qquad (18)$$

Although $\langle\delta\rangle=0$, there is no reason to suppose that $\langle H\delta\rangle=0$; one knows that if extinction is neglected δ is likely to be large and negative for large H, but small for small H, and many other defects in the model may vary systematically with the magnitude of H.

If both types of error are present the scaling factor is, from (9), (12) and (13),

$$E_2=1-\langle H\delta\rangle/\langle H^2\rangle+\langle H\sigma\rangle/\langle H^2\rangle\,, \qquad (19)$$

or, on using (13) and (14),

$$E_2=1-\langle H\delta\rangle/\langle H^2\rangle-\langle\delta\sigma\rangle/\langle H^2\rangle\,. \qquad (20)$$

The third term in (20) depends on the covariance of the defects in the model and the statistical error, and at first sight one is tempted to put it equal to zero. However, δ includes the effects of incorrect values of the adjustable parameters (positional, thermal, ...), and the refinement program will attempt to adjust these so that δ becomes as nearly as possible equal to σ, rather than as nearly as possible equal to zero. The covariance will presumably be small when the number of observed intensities greatly exceeds the number of adjustable parameters, but will become important if the difference is small; in the extreme case when the numbers are equal it might be possible to obtain an exact match, so that all calculated intensities agreed exactly with those observed, however large the statistical errors and however defective the model.

Any method of scaling is, of course, subject to defects in the model; even so-called absolute methods depend on measurements of crystal shape, volume, absorption coefficient, and density, as well as on chemical analysis. Equalization of observed and calculated intensities depends on, in addition, the values adopted for the scattering factors, and on effects like extinction (though this could largely be avoided by omitting reflexions of high intensity); it is insensitive to errors in positional parameters, and can be carried out in a fashion that makes it insensitive to assumptions about thermal parameters. Minimization of a residual is sensitive, in addition, to errors in the placement of atoms (Wilson, 1969, 1974*b*), to errors in the model of thermal motion (*e.g.* neglect of libration or anharmonic oscillation), and to incomplete refinement. It is likely that there will be difficulties in determining, for example, occupancy factors if it is attempted to treat both scaling factor and the occupancy factors as parameters to be refined by least-squares calculations.

We are greatly indebted to Professor D. W. J. Cruickshank for helpful correspondence.

References

CRUICKSHANK, D. W. J. (1970). In *Crystallographic Computing*, edited by F. R. AHMED, pp. 187–197. Copenhagen: Munksgaard.

LENSTRA, A. T. H. (1974). *Acta Cryst.* A**30**, 363–369.

SHOEMAKER, D. P. (1968). *Acta Cryst.* A**24**, 136–142.

WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 152.

WILSON, A. J. C. (1969). *Acta Cryst.* B**25**, 1288–1293.

WILSON, A. J. C. (1973). *Acta Cryst.* B**29**, 1488–1490.

WILSON, A. J. C. (1974*a*). Paper presented at the IUCr Conference on Anomalous Scattering, Madrid, 25 April 1974.

WILSON, A. J. C. (1974*b*). *Acta Cryst.* A**30**, 836–838.