My sincere thanks are due to my teacher, Prof. G. Hägg, for having read my manuscript and to Dr W. Brown for linguistic assistance.

References

BERTAUT, E. F. (1955). *Acta Cryst.* **8**, 823.
COCHRAN, W. (1952). *Acta Cryst.* **5**, 65.
COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.
COCHRAN, W. (1958). *Acta Cryst.* **11**, 579.
COCHRAN, W. & DOUGLAS, A. S. (1955). *Proc. Roy. Soc.* A, **227**, 486.
COCHRAN, W. & DOUGLAS, A. S. (1957). *Proc. Roy. Soc.* A, **243**, 281.
COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1.
HARKER, D. & KASPER, J. S. (1948). *Acta Cryst.* **1**, 70.
KLUG, A. (1958). *Acta Cryst.* **11**, 515.
LÖFGREN, T. (1960). *Acta Cryst.* **13**, 429.
LÖFGREN, T. (1961). *Acta Cryst.* **14**, 436.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60.
WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.
WOOLFSON, M. M. (1958*a*). *Acta Cryst.* **11**, 277.
WOOLFSON, M. M. (1958*b*). *Acta Cryst.* **11**, 393.
ZACHARIASEN, W. H. (1952). *Acta Cryst.* **5**, 68.

Acta Cryst. (1961). **14**, 436

A Non-Statistical Derivation of some Intensity Laws. The Unpredictability of $\mathscr{P}(E)$, $N(z)$ etc.

BY TOR LÖFGREN

Institute of Chemistry, University of Uppsala, Uppsala, Sweden

(*Received* 3 *March* 1960)

The variance criterion, $\langle (I-\Sigma)^2\rangle$, is reconsidered in a slightly modified and generalized form. The approach is in terms of the Patterson function and permits a physical visualization of the criterion. The result is in agreement with standard statistical results in certain space groups only and with special assumptions.

It is maintained that because fundamental statistical quantities like $\mathscr{P}(E)$ critically depend on the unpredictable overlap in high order ϱ self-convolutes, they are quite generally useless.

(A summary of conclusions is given at the end of this article.)

As was stressed by Löfgren (1961), there is an advantage of avoiding probability distributions such as $\mathscr{P}(|{}_gF|)$ in deriving laws for reciprocal-space functions (for ${}_gF(\mathbf{h})$ see Löfgren (1960). We assume that ${}_gf(\mathbf{h})$, corresponding to ${}_gF(\mathbf{h})$ for correct phases, is such that ${}_gf(\mathbf{h})={}_gf(-\mathbf{h})$). Direct-space functions should be taken, instead, as a starting-point. We shall thus study some aspects of the Patterson function, ${}_gP$.

We start with:

$$\oint \{{}_gP(\mathbf{r}) - C_gP_{\text{origin}}(\mathbf{r})\}^2 dv(\mathbf{r}) = (1/V)\sum_{\mathbf{h}} \{|{}_gF(\mathbf{h})|^2 - C.\sum_{n=1}^{N} {}_gf^2_{(n)}(\mathbf{h})\}^2, \quad (1)$$

where (cf. Lipson & Cochran (1953), p. 152)

$${}_gP(\mathbf{r}) = \sum_{n,m=1}^{N} {}_gP_{(n,m)}(\mathbf{r}-(\mathbf{r}_{(n)}-\mathbf{r}_{(m)}))\ , \quad (2)$$

with:

$${}_gP_{(n,m)}(\mathbf{r}) = (1/V)\sum_{\mathbf{h}} {}_gf_{(n)}(\mathbf{h}).{}_gf_{(m)}(\mathbf{h}).\exp\,[-2\pi i\mathbf{h}.\mathbf{r}]\ , \quad (3)$$

and, further:

$${}_gP_{\text{origin}}(\mathbf{r}) = \sum_{n=1}^{N} {}_gP_{(n,n)}(\mathbf{r})\ , \quad (4)$$

and C is an arbitrary constant.

Let us suppose that for certain $n-m$ pairs,

$$\mathbf{r}_{(n)} - \mathbf{r}_{(m)} = \text{constant vector} = \mathbf{r}^{(\lambda)}; \quad (5)$$

i.e., $\mathbf{r}^{(\lambda)}$ are the different Patterson peak positions.

If we further suppose that the sums

$$\sum_{\substack{n,m\\ \mathbf{r}^{(\lambda)}=\text{constant}\neq 0}} {}_gP_{(n,m)}(\mathbf{r}-\mathbf{r}^{(\lambda)})$$

are completely resolved for different λ and completely resolved from $(1-C).{}_gP_{\text{origin}}(\mathbf{r})$, we have, in the same manner as in Löfgren (1961):

$$\begin{aligned}\oint \{{}_gP(\mathbf{r}) - C.{}_gP_{\text{origin}}(\mathbf{r})\}^2 dv(\mathbf{r}) &= \oint (1-C)^2{}_gP^2{}_{\text{origin}}(\mathbf{r})\,dv(\mathbf{r}) \\ &\quad + \sum_{\lambda} \oint \{\sum_{\substack{n,m\\ \mathbf{r}^{(\lambda)}=\text{constant}\neq 0}} {}_gP_{(n,m)}(\mathbf{r})\}^2 dv(\mathbf{r}) \\ &= (1-C)^2(1/V)\sum_{\mathbf{h}}\{\sum_{n} {}_gf^2_{(n)}(\mathbf{h})\}^2 \\ &\quad + (1/V)\sum_{\lambda}\sum_{\mathbf{h}}\{\sum_{\substack{n,m\\ \mathbf{r}^{(\lambda)}=\text{constant}\neq 0}} {}_gf_{(n)}(\mathbf{h}).{}_gf_{(m)}(\mathbf{h})\}^2. \quad (6)\end{aligned}$$

If we assume for example that the only causes of overlap in $\mathbf{r}^{(\lambda)}$ are the operations of the (primitive) group, equations (1) and (6) give ($\Sigma = \sum_{n} {}_gf^2_{(n)}(\mathbf{h})$):

$$\sum_{\mathbf{h}} \{|{}_gF(\mathbf{h})|^2 - C\,.\,\Sigma\}^2 = (1-C)^2 \sum_{\mathbf{h}} \Sigma^2$$

$$+ \sum_{\mathbf{h}} \sum_{\lambda'} \left\{ \begin{pmatrix} \text{multiplicity of} \\ \text{non-Harker } n-m \\ \text{peak in } \mathbf{r}^{(\lambda')} \neq 0 \end{pmatrix} \times {}_gf_{(n)}(\mathbf{h})\,.\,{}_gf_{(m)}(\mathbf{h}) \right\}^2$$

$$+ \sum_{\mathbf{h}} \sum_{\lambda''} \left\{ \begin{pmatrix} \text{multiplicity of} \\ \text{Harker } n \text{ peak in} \\ \mathbf{r}^{(\lambda'')} \neq 0 \end{pmatrix} \times {}_gf^2_{(n)}(\mathbf{h}) \right\}^2 . \qquad (7)$$

Any more-or-less exact overlap which has not been considered will render the right-hand side of (6) and (7) a lower limit to the left-hand side, at least in so far as the ${}_gP_{(n,m)}$'s are ≥ 0.

In (7), we agree to define a Harker peak as any peak between equivalent atoms (i.e., for centrosymmetry we include peaks in $\pm 2\mathbf{r}_{(n)}$). For a primitive cell and only general atomic positions, we have the second sum, i.e., the 'non-Harker contribution', as*

for non-centrosymmetry:

$$\sum_{\mathbf{h}} \{\Sigma^2 - s \sum_{n=1}^{N} {}_gf^4_{(n)}(\mathbf{h})\}\,, \qquad (8)$$

for centrosymmetry:

$$2 \sum_{\mathbf{h}} \{\Sigma^2 - s \sum_{n=1}^{N} {}_gf^4_{(n)}(\mathbf{h})\}\,, \qquad (9)$$

where s is the symmetry number.

The Harker contribution for general positions only, takes the form:

$$s' \sum_{\mathbf{h}} \sum_{n=1}^{N} {}_gf^4_{(n)}(\mathbf{h})\,, \qquad (10)$$

where s' is an integer easily found for any space group.

For $C=1$ and under the conditions for (8), (9), and (10), equation (7) reads:

$$\sum_{\mathbf{h}} \{|{}_gF(\mathbf{h})|^2 - \Sigma\}^2 =$$

$$\begin{cases} \sum_{\mathbf{h}} \{\Sigma^2 + (s'-s) \sum_{n=1}^{N} {}_gf^4_{(n)}(\mathbf{h})\} & \text{for non-centro-symmetry,} \qquad (11) \\ \sum_{\mathbf{h}} \{2\,\Sigma^2 + (s'-2s) \sum_{n=1}^{N} {}_gf^4_{(n)}(\mathbf{h})\} & \text{for centro-symmetry.} \qquad (12) \end{cases}$$

For non-centrosymmetry, one finds that $s'-s \geq -1$, for centrosymmetry that $s'-2s \geq -3$.

The Wilson (1951) statistical result for a non-centrosymmetric crystal and $C=1$, leads to (11) with $s'-s=-1$. This holds true for the space group $P1$, but might be seriously wrong; e.g. in space group $P6$, $s'-s=3$. The reason for the discrepancy is evident: no phase conditions besides $\alpha(\mathbf{h}) = -\alpha(-\mathbf{h})$ were considered by Wilson.

For a centrosymmetric space group, Wilson (1951) gave an expression that leads to (12) with $s'-2s=-3$. This holds true in the space group $P\bar{1}$, for example. Only the phase relations of this centrosymmetric space group were considered by Wilson. But this might again be severely wrong. For example in space group *Pmmm*, $s'-2s=3$.

The terms in $\sum_{n} {}_gf^4_{(n)}(\mathbf{h})$ might be of the same order of magnitude as those in Σ^2, and therefore it is necessary to consider *all* symmetry operations. For specialized positions (not necessarily hypersymmetry) even the right hand side of (11) or (12) might be much too small.

If the $\mathscr{P}$'s given by Wilson (1949) are used for deriving $\langle(|{}_gF|^2-\Sigma)^2\rangle$, all terms in $\sum_{n} {}_gf^4_{(n)}(\mathbf{h})$ drop out. To derive the general expressions given in this paper (e.g., (6)) by means of more elaborate probability distribution functions, it would be *necessary* to consider several types of atomic distributions for each space group. As we shall see presently, it is doubtful if $\mathscr{P}$ is a practically useful concept at all.

We generalize (1) for $C=0$, and have:

$$\oint ({}_g\varrho^{*\nu})^2 dv(\mathbf{r}) = (1/V) \sum_{\mathbf{h}} |{}_gF(\mathbf{h})|^{2\nu}\,, \qquad (13)$$

where $*\nu$ indicates the $(\nu-1)$th self-convolution (an origin inversion might be included for every second 'factor'). It has been made clear that even for $\nu=2$, the integral (13) will depend in a complicated manner on systematic and fortuitous overlap. For $\nu \geq 3$ the more-or-less perfect 'random' overlap at least, cannot be predicted for an unknown structure and any limited $|{}_gF(\mathbf{h})|^2$ material. (Bertaut (1955), however, explicitely gives $\langle|{}_gF|^{2\nu}\rangle$ for a special case.)

It is a mathematical fact that a knowledge of $\mathscr{P}({}_gF)$—and thus of $\mathscr{P}(|{}_gF|)$, $N(z)$ etc., with which one might start—leads to a knowledge of $\langle|{}_gF|^{2\nu}\rangle$ and thus of the integral (13) for all ν's. The integral (13) is, however, unpredictable for general ν's. In other words: if, for example, $\mathscr{P}(|{}_gF|)$ is thought of in terms of its moments $\langle|{}_gF|^{2\nu}\rangle$ (which are, demonstrably, sufficient for a complete definition of $\mathscr{P}(|{}_gF|)$) it cannot be predicted any further than $\nu=2$, if that far.

We conclude:

(i) Expression (1) has a minimum for $C \approx 1$, but irrespective of C, the right hand side is only a measure of the degree of overlap in ${}_gP(\mathbf{r})$ (cf. Rogers & Wilson (1953), § 5·2). It is generally practical to chose $C=0$.

(ii) Starting with (13), no symmetry criterion more general than that for $\nu=2$ exists in practice. Statistical considerations do not lead any further because $\mathscr{P}(|{}_gF|)$, etc., will depend on assumptions about the atomic distribution in a manner that is too complicated.

---

* It can be proved that parallelism for vectors between *non-equivalent* atoms, within one cell and in general positions, cannot occur through space-group operations except centering and inversion. For vectors between *equivalent* atoms, such parallelism also occurs in certain non-centrosymmetric, primitive space groups, e.g., $P6$. Cf. Rogers & Wilson (1953), § 5·1.

(We assume, as usual, that $\mathscr{P}$ versus $|\mathbf{h}|$—defining $\mathbf{h}$ intervals—is independent of structure, i.e., that $\sum_{\mathbf{h}}|_gF(\mathbf{h})|^{2\nu}$ is as informative as $\langle|_gF(\mathbf{h})|^{2\nu}\rangle$. In a coming paper $\langle|_gF(\mathbf{h})|^{2\nu}\rangle$ versus $|\mathbf{h}|$ will be studied.) The direct-space interpretation of (13) (which for $\nu=2$ reduces to (1) for $C=0$) shows physically why traditional, statistical $|_gF|$ studies have met with complications, for example by heavy atoms, low multiplicity—or otherwise specialized—positions, hypersymmetry etc. (A wide literature is revealed if one begins with the entry 'Probability distribution' in the general Index of *Acta Cryst.*)

(iii) Under all circumstances, only a *lower* limit can be put to $\sum_{\mathbf{h}}|_gF(\mathbf{h})|^4$.

My sincere thanks are due to my teacher, Prof. G. Hägg, for having read my manuscript and to Dr W. Brown for linguistic assistance.

References

BERTAUT, E. F. (1955). *Acta Cryst.* **8**, 823.

LIPSON, H. & COCHRAN, W. (1953). *The Determination of Crystal Structures.* London: Bell.

LÖFGREN, T. (1960). *Acta Cryst.* **13**, 429.

LÖFGREN, T. (1961). *Acta Cryst.* **14**, 434.

ROGERS, D. & WILSON, A. J. C. (1953). *Acta Cryst.* **6**, 439.

WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.

WILSON, A. J. C. (1951). *Research, Lond.* **4**, 141.

Short Communications

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 500 words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible; and proofs will not generally be submitted to authors. Publication will be quicker if the contributions are without illustrations.

Acta Cryst. (1961). **14**, 438

Facteur d'echelle et coefficient d'agitation thermique.

Par J. GAULTIER et C. HAUW, *Laboratoire de Minéralogie et de Rayons X, Faculté des Sciences, Université de Bordeaux, France*

(*Reçu le* 15 *juillet* 1960)

Le premier travail, avant d'entreprendre la détermination d'une structure cristalline, consiste à déterminer le rapport d'échelle C des facteurs de structure calculés aux facteurs de structure observés et le coefficient d'agitation thermique moyen B des atomes.

La méthode de Wilson donne souvent des résultats incertains lorsque les points du diagramme

$$\log \langle F_o^2\rangle/\langle\Sigma f^2\rangle$$

en fonction de $(\sin\theta/\lambda)^2$ ne se trouvent pas groupés autour d'une droite moyenne.

Remarquons que ses résultats dépendent de la manière dont on effectue l'échantillonnage des $|F_o|^2$ par zones concentriques de l'espace réciproque. L'ambiguité persiste quel que soit ce mode d'échantillonnage, fut-il celui utilisant le maximum d'informations, c'est-à-dire résultant de tous les échantillonnages possibles.

Une autre méthode nous a paru excellente pour déterminer C et B dans des cas ou la méthode de Wilson était inappliquable: cette méthode consiste à évaluer la densité à l'intérieur du pic origine de la fonction de Patterson.

$$p_o(r)=C^2/V\sum_H F_o(H)^2\cos(2\pi H.r)$$

en limitant le calcul à une sphère de rayon r (environ 1,5 Å) autour de l'origine, et à comparer $p_o(r)$ à la densité théorique calculée dans le même domaine.

$$p_c(r)=1/V\sum_H\left\{\sum_{j=1}^{N}f_j^2(H)\exp\left[-2B(2\sin\theta/\lambda)^2\right]\right\}\cos(2\pi H.r).$$

Théoriquement ces deux fonctions devraient être identiques, et la différence $\delta(r)=p_o(r)-p_c(r)$ devrait être nulle dans tout le domaine sphérique envisagé. En fait il n'en est pas ainsi par suite:

— des erreurs expérimentales sur les facteurs de structure observés.

— de l'adoption d'une agitation thermique isotrope moyenne pour tous les atomes.

— de la valeur nulle donnée aux termes H non observés.

— de l'approximation initiale introduite dans le calcul de $p_c(r)$.

Nous supposerons que les constantes B et C sont correctement déterminées lorsque $\delta(r)$ présentera de faibles variations par rapport au zéro (inférieures à $\pm\varepsilon$).*

Cette seule condition est insuffisante pour définir de façon satisfaisante les paramètres B et C.

Nous avons pensé que l'introduction de l'équation

$$C^2\sum_H F_o(H)^2=\sum_H\left\{\sum_{j=1}^{N}f_j^2(H)\exp\left[-2B(\sin\theta/\lambda)^2\right]\right\}$$

qui lie de façon univoque ces deux constantes nous conduirait, en quelques itérations, très près de la solution cherchée.

Ecrire cette relation suppose que $\delta(r)=0$ pour $r=0$ c'est-à-dire que l'on admet arbitrairement que les facteurs d'erreur n'interviennent pas à l'origine.

La méthode pratique employée est la suivante:

---

* ε de l'ordre de 2% de $p_o(r)$ pour $r=0$.