Acta Cryst. (1952). 5, 60

The Squaring Method: a New Method for Phase Determination

By D. Sayre*

Laboratory of Chemical Crystallography, University Museum, Oxford, England

(*Received* 31 *May* 1951)

A new set of relationships is given which exist among the structure factors of crystals composed only of like atoms and which hold approximately for ordinary organic crystals containing only C, N, O and H. They are the consequence of the similarity between the electron-density function $d(\bar{x})$ and its square, $d^2(\bar{x})$, for such crystals. In spite of considerable differences in form, the relationships found make the same general statement about the structure factors as do the Harker–Kasper–MacGillavry and Karle–Hauptmann–Goedkoop relationships, namely that the phase of $F(a+b)$ is closely related to that of $F(a)\,F(b)$. Because the present relationships apply to a special case (like atoms) and include the condition of a minimum separation of atoms, they should be a more powerful method for phase determination than the other methods, but this fact has not yet been demonstrated directly. The method has been applied to the structure of hydroxyproline, where it was possible in a few days to find the signs of thirty-one structure factors, of which thirty were correct, and to produce a Fourier projection in which the molecule was easily recognizable.

1. Essentials of the method

A crystal structure which is composed exclusively of like atoms will in three dimensions resemble in all essentials the one-dimensional example of Fig. 1, that is, its electron-density function $d(\bar{x})$ will be composed of identical and resolved peaks. For such a structure, and only for such a structure, the function $d(\bar{x})$ and its square $d^2(\bar{x})$ are very nearly alike, differing only in that the peaks in the latter function are sharper than in the former. Therefore, supposing that the atomic shape in $d(\bar{x})$ is known, the structure factors $F^{\text{sq.}}(h,k,l)$ of the squared structure may be calculated from the structure factors $F(h,k,l)$ of the original structure by the relations

$$F^{\text{sq.}}(h,k,l)=S(h,k,l)\,F(h,k,l) \text{ for all } h,k,l, \quad (1{\cdot}1)$$

where $S(h,k,l)$ is the appropriate function to account for the change in atomic shape.

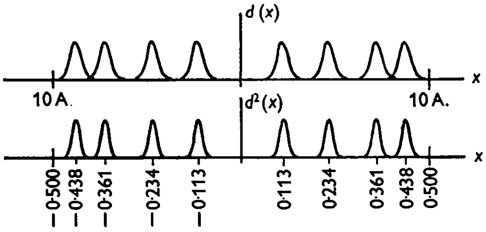

Fig. 1. The unit cell of a one-dimensional crystal composed of identical and non-overlapping atoms, showing the similarity between $d(\bar{x})$ and $d^2(\bar{x})$. This is the crystal used as an example in §2(*a*).

But it is also a general result of Fourier theory (see, for example, Bochner & Chandrasekharan, 1949) that the operation of squaring any function is equivalent to the operation of *self-convoluting* its array of structure factors; that is

$$F^{\text{sq.}}(h,k,l) = (1/v)\sum_p\sum_q\sum_r F(p,q,r)\,F(h-p,\,k-q,\,l-r) \quad \text{for all } h,k,l, \quad (1{\cdot}2)$$

where the right-hand side is an explicit statement of the self-convolution process. Here v is the volume of the crystal unit cell in Å^3, and the summations are to be made over all values of p, q, r.

Hence it follows that the structure factors of a crystal composed of identical atoms satisfy the relations

$$\sum_p\sum_q\sum_r F(p,q,r)\,F(h-p,k-q,l-r) = vS(h,k,l)\,F(h,k,l) \quad \text{for all } h,k,l. \quad (1{\cdot}3)$$

Consequently, for crystals of this single atomic species type, an array of phases or signs can be correct only if it causes the relations (1·3) *to be satisfied.* The present method for phase determination is therefore to seek an array of phases or signs which causes the relations to be satisfied.

Although strictly speaking only an insignificant number of crystal structures contain but a single atomic species, for practical purposes the method will be applicable to the entire class of ordinary organic crystals containing only carbon, nitrogen, oxygen and hydrogen. It therefore applies to precisely that class of structures not covered by the heavy-atom technique.

In the event that there exist projections of the structure in which there is no overlapping of atoms, the appropriate two- or one-dimensional forms

$$\sum_p\sum_q F(p,q)\,F(h-p,k-q)=aS(h,k)\,F(h,k) \quad \text{for all } h,k \quad (1{\cdot}3')$$

and $$\sum_p F(p)\,F(h-p)=lS(h)\,F(h) \quad \text{for all } h \quad (1{\cdot}3'')$$

---

* Present address: The Johnson Foundation for Medical Physics, University of Pennsylvania, Philadelphia 4, Pa., U.S.A.

will be applicable. Here a and l are the area or length of the unit cell of the projected structure, in Å^2 or Å respectively.

Although the relations (1·3) are valid whether or not the crystal structure possesses a centre of symmetry, the method has so far been used only for centrosymmetrical problems, and in this paper only that case will be discussed further.

2. Examples of the use of the method

(a) *A hypothetical one-dimensional example*

Consider the structure shown in Fig. 1, which is of line-group $P\overline{1}$, with unit cell of length $l = 20$ Å, and with eight atoms at positions $x = \pm 0{\cdot}113$, $\pm 0{\cdot}234$, $\pm 0{\cdot}361$, and $\pm 0{\cdot}438$. With X-rays of $\lambda = 1{\cdot}54$ Å structure factors out to index 26 could be observed. Supposing that the atoms are Gaussian, of form $\exp(-2\pi r^2)$, their scattering factor will be

$$(1/\sqrt{2}) \exp(-\tfrac{1}{2}\pi R^2),$$

and the squared atoms will be of form $\exp(-4\pi r^2)$ with scattering factor $\tfrac{1}{2} \exp(-\tfrac{1}{4}\pi R^2)$. This information permits the $F(h)$'s and the $F^{\text{sq.}}(h)$'s to be calculated; they are given in Table 1.

Table 1. *Structure factors of the crystal of Fig.* 1

h	$F(h)$	$F^{\text{sq.}}(h)$	$lF^{\text{sq.}}(h)$	$\Sigma F(p)\,F(h-p)$
0	5·657	4·000	80·00	79·98
1	−0·999	−0·708	−14·16	−14·16
2	−0·408	−0·291	− 5·82	− 5·85
3	−0·481	−0·346	− 6·92	− 6·87
4	−1·276	−0·931	−18·62	−18·56
5	0·347	0·258	5·16	5·12
6	−1·782	−1·352	−27·04	−26·98
7	−0·540	−0·421	− 8·42	− 8·36
8	1·408	1·129	22·58	22·54
9	2·798	2·320	46·40	46·29
10	−1·290	−1·110	−22·20	−22·28
11	0·481	0·431	8·62	8·52
12	−0·637	−0·598	−11·96	−11·63
13	−0·519	−0·511	−10·22	−10·14
14	0·383	0·398	7·96	7·80
15	−1·813	−1·995	−39·90	−39·34
16	0·766	0·895	17·90	17·61
17	0·718	0·895	17·90	17·91
18	0·381	0·509	10·18	9·79
19	−0·049	−0·071	− 1·42	− 1·96
20	−0·070	−0·109	− 2·18	− 1·68
21	−0·098	−0·165	− 3·30	− 3·17
22	−0·026	−0·047	− 0·94	− 0·96
23	−0·174	−0·348	− 6·96	− 6·59
24	−0·367	−0·805	−16·10	−15·49
25	0·362	0·874	17·48	16·92
26	0·027	0·072	1·44	1·50

Knowing the correct array of signs we may first check the validity of equations (1·3″). By these equations the last two columns of the table should be equal, and it is evident that they are very nearly so, the conventional disagreement factor R being only 1·2 %. The slight discrepancies which do appear are undoubtedly due to the facts (*a*) that the atoms are not entirely free of overlap, and (*b*) that the unobservable structure factors of index higher than 26 are not entirely negligible. These considerations emphasize the necessity of using atoms which are simultaneously confined as nearly as possible to a radius of about 0·65 Å (half the minimum separation of atoms) in direct space and to a radius of 1·3 Å^{-1} (the limit of observation with Cu $K\alpha$ radiation) in reciprocal space. The Gaussian atom $\exp(-2\pi r^2)$ comes as near as possible to this ideal, falling to about 0·07 of its central values by these limits.

Next, pretending the signs to be unknown, it is possible to deduce every one of them by use of the relations (1·3″). The process is facilitated by preparing Table 2, which contains all the interproducts appearing in the left-hand sides of the relations. The indices of the two terms whose product appears in any particular position in the table are given by the numbers along the top and left-hand side. All of the products which enter into the equation for any particular h will be found along a diagonal (lower left to upper right). Notice that the form of the equations causes each product to occur twice in an equation; in Table 2 the products are listed only once, but with double weight. The only exceptions are the products along the lower left-hand margin, which occur only once and are listed with single weight. These products, like those along the upper left-hand margin, are certainly positive. The products $2F(0)\,F(h)$ will certainly have the same signs as the expected totals $20F^{\text{sq.}}(h)$; hence they are subtracted (omitted) from the table and from the expected totals, and the revised expected totals

$$20F^{\text{sq.}}(h) - 2F(0)\,F(h)$$

are listed at the extreme right of the table, at the ends of the corresponding diagonals. The problem now is to find a set of twenty-six signs which, when applied consistently throughout the entire table, will cause the sums along the diagonals to equal as nearly as possible the expected totals.

A few of the products (indicated in the table by asterisks) are so large that they enforce sign relationships. Thus, for example, the term $2F(\overline{6})\,F(15)$ in the equation for $F(9)$ is so large that its sign and that of $F(9)$ must be the same.* Using these first relationships, and choosing $F(9)$ to be positive (the sign of any one of the odd structure factors may be chosen arbitrarily), it is easy to show that $F(0)$, $F(9)$, $F(16)$, $F(18)$ and $F(25)$ are all positive, and that $F(1)$, $F(6)$, $F(10)$, $F(15)$ and $F(24)$ all have the same sign, whether positive or negative. When all this information is inserted into the table it becomes possible to show that these five structure factors must in fact be negative. The reasoning is typical and will be outlined. Suppose that they are positive. Then it is easy to show that $F(7)$, $F(8)$, $F(17)$ and $F(23)$ must all have the same sign. But

---

* Supposing that the magnitude of the expected total at the right of Table 2 is $E(h)$, and that the sum of the magnitudes of the terms in the corresponding diagonal (that is, the maximum possible total) is $M(h)$, then it is easy to show that any term or terms whose sum has a magnitude equal to or greater than $\tfrac{1}{2}[M(h) - E(h)]$ are sufficient to determine the sign of $F(h)$.

Table 2. *Table of interproducts for sign determination*

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	Expected totals $E(h)$
−26																											+0·00	48·00
−25																										+0·26	0·02	2·86
−24																									+0·27	0·27	0·02	1·20
−23																								+0·06	0·13	0·13	0·01	1·48
−22																							+0·00	0·01	0·02	0·02	0·00	4·18
−21																						+0·02	0·01	0·03	0·07	0·07	0·01	1·23
−20																					+0·01	0·01	0·00	0·02	0·05	0·05	0·00	6·88
−19																				+0·01	0·01	0·01	0·00	0·02	0·04	0·04	0·00	2·31
−18																			+0·29	0·04	0·05	0·08	0·02	0·13	0·28	0·28	0·02	6·65
−17																		+1·03	0·55	0·07	0·10	0·14	0·04	0·25	0·53	0·52	0·04	14·74
−16																	+1·17	1·10	0·58	0·08	0·11	0·15	0·04	0·27	0·56	0·56	0·04	7·61
−15																+6·57	2·78	2·60	1·38	0·18	0·25	0·36	0·09	0·63	1·33	1·31	0·10	3·18
−14															+0·29	1·39	0·59	0·55	0·29	0·04	0·05	0·08	0·02	0·13	0·28	0·28	0·02	4·75
−13														+0·54	0·40	1·88	0·80	0·75	0·40	0·05	0·07	0·10	0·03	0·18	0·38	0·38	0·03	4·35
−12													+0·81	0·66	0·49	2·31	0·98	0·92	0·49	0·06	0·09	0·13	0·03	0·22	0·47	0·46	0·03	3·63
−11												+0·46	0·61	0·50	0·37	1·74	0·74	0·69	0·37	0·05	0·07	0·09	0·03	0·17	0·35	0·35	0·03	19·39
−10											+3·33	1·24	1·64	1·34	0·99	4·68	1·98	1·85	0·98	0·13	0·18	0·25	0·07	0·45	0·95	0·93	0·07	9·23
− 9										+15·66	7·22	2·69	3·57	2·90	2·14	10·15	4·29	4·02	2·13	0·27	0·39	0·55	0·15	0·97	2·05*	2·03	0·15	9·78
− 8									+3·97	7·88	3·63	1·35	1·79	1·46	1·08	5·11	2·16	2·02	1·07	0·14	0·20	0·28	0·07	0·49	1·03	1·02	0·08	5·87
− 7								+0·58	1·52	3·02	1·39	0·52	0·69	0·56	0·41	1·96	0·83	0·78	0·41	0·05	0·08	0·11	0·03	0·19	0·40	0·39	0·03	0·87
− 6							+6·35	1·93	5·02	9·97	4·60	1·71	2·27	1·85	1·37	6·46*	2·73	2·56	1·36	0·18	0·25	0·35	0·09	0·62	1·31	1·29	0·10	1·39
− 5						+0·24	1·24	0·38	0·98	1·94	0·90	0·33	0·44	0·36	0·27	1·26	0·53	0·50	0·26	0·03	0·05	0·07	0·02	0·12	0·26	0·25	0·02	2·19
− 4					+3·26	0·89	4·55	1·38	3·59	7·14	3·29	1·23	1·63	1·32	0·98	4·63	1·96	1·83	0·97	0·13	0·18	0·25	0·07	0·44	0·94	0·92	0·07	0·65
− 3				+0·46	1·23	0·33	1·71	0·52	1·35	2·69	1·24	0·46	0·61	0·50	0·37	1·74	0·74	0·69	0·37	0·05	0·07	0·09	0·03	0·17	0·35	0·35	0·03	4·99
− 2			+0·33	0·39	1·04	0·28	1·45	0·44	1·15	2·28	1·05	0·39	0·52	0·42	0·31	1·48	0·63	0·59	0·31	0·04	0·06	0·08	0·02	0·14	0·30	0·30	0·02	11·95
− 1		+2·00	0·82	0·96	2·55	0·69	3·56	1·08	2·81	5·59	2·58	0·96	1·27	1·04	0·77	3·62	1·53*	1·44	0·76	0·10	0·14	0·20	0·05	0·35	0·73	0·72	0·05	13·38
0																												1·14
1		+1·00	0·82	0·96	2·55	0·69	3·56	1·08	2·81	5·59	2·58	0·96	1·27	1·04	0·77	3·62	1·53	1·44	0·76	0·10	0·14	0·20	0·05	0·35	0·73*	0·72		
2			+0·17	0·39	1·04	0·28	1·45	0·44	1·15	2·28	1·05	0·39	0·52	0·42	0·31	1·48	0·63	0·59	0·31	0·04	0·06	0·08	0·02	0·14	0·30			
3				+0·23	1·23	0·33	1·71	0·52	1·35	2·69	1·24	0·46	0·61	0·50	0·37	1·74	0·74	0·69	0·37	0·05	0·07	0·09	0·03	0·17				
4					+1·63	0·89	4·55	1·38	3·59	7·14	3·29	1·23	1·63	1·32	0·98	4·63	1·96	1·83	0·97	0·13	0·18	0·25	0·07					
5						+0·12	1·24	0·38	0·98	1·94	0·90	0·33	0·44	0·36	0·27	1·26	0·53	0·50	0·26	0·03	0·05	0·07						
6							+3·18	1·93	5·02	9·97*	4·60	1·71	2·27	1·85	1·37	6·46	2·73	2·56	1·36	0·18	0·25							
7								+0·29	1·52*	3·02	1·39	0·52	0·69	0·56	0·41	1·96	0·83	0·78	0·41	0·05								
8									+1·98	7·88*	3·63	1·35	1·79	1·46	1·08	5·11*	2·16	2·02*	1·07									
9										+7·83	7·22	2·69	3·57	2·90	2·14	10·15*	4·29*	4·02										
10											+1·66	1·24	1·64	1·34	0·99	4·68*	1·98											
11												+0·23	0·61	0·50	0·37	1·74												
12													+0·41	0·66*	0·49													
13														+0·27														

this sign cannot be positive, for if it is the equation for $F(24)$ will have too many positive terms to permit its being satisfied. Nor can it be negative, because then the equation for $F(8)$ necessitates $F(2)$ and $F(14)$ being negative, in which case the equation for $F(24)$ will force $F(24)$ to be negative, contrary to the initial assumption. Therefore the initial assumption is untenable, and the list of signs determined becomes

+	−
0	1
9	6
16	10
18	15
25	24

(2·1)

whereupon still more signs follow, until all 26 have been found. To carry out the sign deductions in detail is an interesting logical problem for those who care to spend an evening or two on it.

(*b*) *The crystal structure of hydroxyproline*

Hydroxyproline, which occurs in crystals of space group $P2_12_12_1$, with 4 molecules of $C_5H_9NO_3$ per unit cell of dimensions $a=5{\cdot}01$, $b=8{\cdot}35$, $c=14{\cdot}1$ Å (Zussman, 1951), appeared, because of its short a axis, to be a suitable real problem for attack with the two-dimensional form of the method, equations (1·3′). As it turned out, however, there was considerable overlapping of atoms in the projection, so that in fact there could exist no array of signs which would cause equations (1·3′) to be properly satisfied. Inconsistencies therefore developed which made it impossible to determine signs with anything like the completeness of the preceding example. Nevertheless, it was possible to establish with reasonable certainty enough signs to give a Fourier projection in which the molecule was easily recognizable and which could certainly have been refined by the usual methods.

The X-ray data for this attempt were very kindly supplied by Mr J. Zussman of the Cavendish Laboratory, Cambridge. By Wilson's method (Wilson, 1949) it was first found that the atoms in this structure had scattering factors approximately that of nitrogen (Viervoll & Ögrim, 1949) with a temperature factor

$$B=2{\cdot}0\times10^{-16}\ \text{cm.}^2.$$

The data were divided through by this function to render them the data for a point atom crystal, and then multiplied by $\frac{1}{2}\exp(-\frac{1}{2}\pi R^2)$, the scattering curve of the two-dimensional Gaussian atom $\exp(-2\pi r^2)$, yielding the $F(k,l)$'s on a relative scale, and by $\frac{1}{4}\exp(-\frac{1}{4}\pi R^2)$, the scattering curve of the squared atoms $\exp(-4\pi r^2)$, yielding the $F^{\text{sq.}}(k,l)$'s on the same relative scale. Next, these structure factors were put on an absolute scale by using the equation (1·3′) for (0, 0); that is,

$$\sum_p\sum_q|F(p,q)|^2=aF^{\text{sq.}}(0,0). \tag{2·2}$$

Since there are 36 atoms in the unit cell, and each atom contains $\frac{1}{2}$ electron and each squared atom contains $\frac{1}{4}$ electron, $F(0,0)=18$ and $F^{\text{sq.}}(0,0)=9$. Thus equation (2·2) becomes

$$\sum_{p^2+q^2>0}\sum|F(p,q)|^2=9a-18^2\doteq736, \tag{2·2′}$$

which suffices to determine the scale of the F's. (It is possible, incidentally, to account by Wilson's method for the terms which lie beyond the observable limit R_0; they can easily be shown to contribute to the left-hand side of equation (2·2) an amount $\frac{1}{4}Na\exp(-\pi R_0^2)$, where N, the number of atoms in the unit cell, here is 36, and R_0 here is 1·3.)

With the data thus prepared the interproducts among the 153 observable structure factors were calculated and arranged into 153 lists corresponding to the 153 equations (1·3′). This process took about a week. There were no terms so large as to compel any sign relationships beyond doubt, but it nevertheless was obvious, upon inspection of the equations for the largest terms, that for a certain group of 19 terms, which included the ten strongest terms, only eight different sign arrangements were at all likely. Further development of these eight arrangements showed that six of them were very unlikely to be correct, and that of the remaining two arrangements, one was superior to the other in the sense that it caused the equations to hold rather better. This arrangement could quite confidently be extended to include thirty-one terms in all, and much less confidently to include fifty-three terms. By this time the inconsistencies caused by the overlapping were so severe that no further progress could be made.

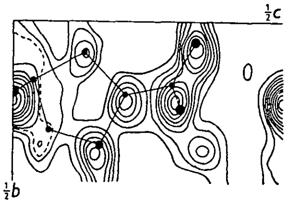

Fig. 2. The Fourier projection down the a axis of hydroxyproline, based on the first 19 signs found. Contours are at intervals of 0·66 e.Å². The correct atomic positions, as found by Zussman (1951), are indicated.

At this point the test was over. The signs found were compared with those which had been found independently by Zussman by usual methods, and it was found that at the 19-term stage all the signs were correct, at the 31-term stage one small term had incorrect sign, and that at the 53-term stage nine signs were incorrect. The Fourier projection based on the nineteen terms is shown in Fig. 2.

3. Discussion of the method

This method for phase determination has the merit of inserting into the phase problem more of the information initially known about a crystal structure than do the other comparable methods, namely, the Harker–Kasper inequalities (Harker & Kasper, 1947, 1948; Gillis, 1948*a*, *b*; MacGillavry, 1950) and the relations of Karle & Hauptman (1950) or Goedkoop (1950), for it includes the facts that all the atoms are identical and that there is a minimum distance between atoms, which the other methods do not. It should, therefore, be a more powerful method than they in the sense that it would give a larger number of correct phase relationships in any structure problem to which it is applicable. Unfortunately, that this is the case has not yet been proved directly. What is at the bottom of the trouble is that it is not yet known how to realize the full power inherent in the equations (1·3). This section will be devoted to a discussion of this matter. For simplicity a centrosymmetrical one-dimensional structure will be assumed.

The first thing to be noted is the strong qualitative similarity between the present method and the Karle–Hauptmann method (which will be taken as the prototype of the other methods, since according to its authors it includes all of the Harker–Kasper type of inequalities and also is identical with Goedkoop's relations except for certain determinantal equalities of very high order). Both methods essentially state a tendency for the sign of $F(a+b)$ to be the same as that of $F(a)F(b)$, the tendency being the stronger the larger $F(a)$, $F(b)$ and $F(a+b)$. In the Karle–Hauptmann treatment this statement takes the form of the inequality

$$\left|F(a+b)-\frac{F(a)F(b)}{F(0)}\right| \leqslant \frac{\sqrt{\{|F(0)|^2-|F(a)|^2\}}\,\sqrt{\{|F(0)|^2-|F(b)|^2\}}}{F(0)}, \quad (3{\cdot}1)$$

whereas in the present treatment it comes about because in the equation for $F(a+b)$ there appears the term $2F(a)F(b)$; if this term is large enough, it will dominate the equation and the sign of $F(a+b)$ will necessarily be the same as its sign. This was the basis upon which signs were determined in the examples of the preceding section.

But this cannot be the real basis of the relationship among these structure factors. For imagine what would be the case if atoms were very sharp and very short wavelength X-rays could be used. Then a very large array of structure factors could be observed. Each equation (1·3) would contain a very large number of terms, of which none could possibly be large enough to override all the others, and the present method could not even make a start at sign determination. Yet obviously this imaginary situation is in fact much more favourable to structure analysis than the actual one. Clearly, then, the relationships among signs cannot depend merely upon the presence of single large terms in the equations. There must be some other, more fundamental, basis which has so far been overlooked.

The author believes that he has found this basis, although only in qualitative form, and that it can be expressed as follows: that $F(a+b)$ is related to $F(a)$ and $F(b)$ not merely by the presence of the term $2F(a)F(b)$ in its equation, but because if $F(a)$ and $F(b)$ are strong each imposes a characteristic pattern upon the entire array of signs, no matter how large the array, the result of which is strongly to affect $F(a+b)$.

Suppose, for example, that $F^{\mathrm{sq.}}(a)=0{\cdot}6F^{\mathrm{sq.}}(0)$. Then $lF^{\mathrm{sq.}}(a)$ must be at least 0·6 of the maximum possible total of the terms in its equation, because this total cannot exceed $lF^{\mathrm{sq.}}(0)$.* Hence the sum of the terms in the $F(a)$ equation which have the same sign as $F(a)$ must be at least 80 % of the total, while that of the opposing terms cannot exceed 20 %. Having $F(a)$ large, therefore, imposes upon the entire array of signs a pattern which can be described by saying that $F(p)F(a-p)$ has the same sign as $F(a)$ for most p. Similarly with $F(b)$. Now these patterns are such as strongly to affect $F(a+b)$. For

$$\begin{aligned} lF^{\mathrm{sq.}}(a+b) &= \sum_p F(p)F(a+b-p) \\ &= \sum_p [F(p)F(a-p)] \\ &\quad \times [F(p-a)F(a+b-p)]/|F(a-p)|^2. \quad (3{\cdot}2) \end{aligned}$$

Most of the terms in the first bracket will have the same sign as $F(a)$, and most of those in the second will have the sign of $F(b)$; hence $F(a+b)$ will have the sign of $F(a)F(b)$.

If this argument could be cast into appropriate mathematical form it would doubtless be possible to define precisely how large $F(a)$, $F(b)$ and $F(a+b)$ must be in order that a sign relationship necessarily exist among them, and the full power of the relations would thus be realized. Unfortunately, the problem has proved to be more difficult than was expected, and this analysis has not been done. It may nevertheless be supposed that, for the reasons given at the beginning of this section, the result would be a stronger set of relationships than those afforded by the other methods, and that this increased strength would in some way be connected with the fact referred to in the previous footnote, that the maximum possible total of the terms in any equation is often considerably smaller than $lF^{\mathrm{sq.}}(0)$, and that therefore a term considerably smaller than $F(0)$ may nevertheless impose a strong pattern. Advantage would thus be taken of the particular distribution of structure-factor magnitudes. (In the Karle–Hauptmann inequalities this is not so, equation (3·1) indicating that the sign-determining power of terms depends only upon their own magnitudes.)

---

* The maximum possible total

$$\Sigma\,|F(p)F(a-p)| = \Sigma\,|F(p)|\,|F(a-p)| \leqslant \Sigma\,|F(p)|^2 = lF^{\mathrm{sq.}}(0),$$

an inequality which is valid for any set of numbers $F(p)$. For any particular structure, of course, the actual values of the maximum possible totals will be known, so that this inequality may be considerably strengthened.

The author wishes to express his thanks first to Dr W. Cochran of the Cavendish Laboratory, Cambridge, for conversations which suggested this method, to Mr J. Zussman for very kindly supplying the data on hydroxyproline, to Mrs D. Crowfoot Hodgkin for many helpful conversations, and to the Rockefeller Foundation for a grant of research funds.

References

BOCHNER, S. & CHANDRASEKHARAN, K. (1949). *Fourier Transforms*. Princeton: Princeton University Press.

GILLIS, J. (1948*a*). *Acta Cryst.* **1**, 76.

GILLIS, J. (1948*b*). *Acta Cryst.* **1**, 174.

GOEDKOOP, J. A. (1950). *Acta Cryst.* **3**, 374.

HARKER, D. & KASPER, J. S. (1947). *J. Chem. Phys.* **15**, 882.

HARKER, D. & KASPER, J. S. (1948). *Acta Cryst.* **1**, 70.

KARLE, J. & HAUPTMANN, H. (1950). *Acta Cryst.* **3**, 181.

MACGILLAVRY, C. H. (1950). *Acta Cryst.* **3**, 214.

VIERVOLL, H. & ÖGRIM, O. (1949). *Acta Cryst.* **2**, 277.

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

ZUSSMAN, J. (1951). *Acta Cryst.* **4**, 72.

Acta Cryst. (1952). **5**, 65

A Relation between the Signs of Structure Factors*

BY W. COCHRAN

*Physics Department, State College, Pa., U.S.A. and Crystallographic Laboratory,
Cavendish Laboratory, Cambridge, England*

(*Received* 14 *June* 1951)

By considering the extent to which a partial Fourier series, which contains as coefficients only a few of the larger structure factors, will represent the electron density in a crystal structure, and using the relations between the structure factors of an electron distribution and those of the square of this distribution, a probable relation between the signs of structure factors is deduced. The relation is suggested by the form of Sayre's equality, and is also related to Harker–Kasper inequalities.

1. Introduction

Relations between the magnitudes and phases of structure factors have been given by Harker & Kasper (1948), Karle & Hauptmann (1950) and Goedkoop (1950). All are based on the fact that the electron density in crystals is everywhere positive, and is approximately a superposition of spherically symmetric atoms of the same shape. An equality relation between structure factors, derived by Sayre (1952), is based on the fact that when a crystal contains equal atoms which are resolved from one another, the operation of squaring the electron density leaves it unchanged, in the sense that it is still composed of equal atoms in the same positions as before.

In this paper we shall use one of Sayre's results to show that the conditions fulfilled by a partial Fourier series lead to a very simple probable relation between the signs of structure factors. The relation is suggested by the form of Sayre's equality, and is also related to Harker–Kasper inequalities.

2. An exact condition satisfied by a partial series

Corresponding to the electron density

$$\rho(\mathbf{r})=\frac{1}{V}\sum_{-\infty}^{+\infty}F(\mathbf{h})\cos 2\pi\mathbf{h}.\mathbf{r}, \qquad (1)$$

let

$$\rho_p(\mathbf{r})=\frac{1}{V}\sum_{1}^{p}F(\mathbf{h})\cos 2\pi\mathbf{h}.\mathbf{r}, \qquad (2)$$

where the partial series (2) involves only p coefficients, which are not chosen from any particular region of the reciprocal lattice. By dividing the F's by the factor $\hat{f}$ (Harker & Kasper, 1948), (1) may be made to represent a distribution composed of point atoms.

Putting $F_s=F/\hat{f}$,

$$\sigma(\mathbf{r})=\frac{1}{V}\sum_{-\infty}^{+\infty}F_s(\mathbf{h})\cos 2\pi\mathbf{h}.\mathbf{r}, \qquad (3)$$

and correspondingly

$$\sigma_p(\mathbf{r})=\frac{1}{V}\sum_{1}^{p}F_s(\mathbf{h})\cos 2\pi\mathbf{h}.\mathbf{r}. \qquad (4)$$

Hence

$$\int_V \sigma\sigma_p\,dV=\frac{1}{V}\sum_{1}^{p}F_s^2. \qquad (5)$$

Let $(\sigma_p)_j$ denote the value of σ_p at the centre of the jth atom. The value of $\int\sigma\,dV$, where the integral is now

---

* A preliminary account of this work has already been given (Pepinsky & Cochran, 1951). Computations on X-RAC were carried out under contract N6o.n.r.–26916, T.O. 16 with the Office of Naval Research.