

La Statistique des Fonctions Complexes: Application à la Cristallographie

PAR E. F. BERTAUT

Laboratoire d'Électrostatique et de Physique du Métal, Institut Fourier, Grenoble, France

(Reçu le 19 septembre 1955)

The probability of a value of a given complex function and the joint probability of values of several given complex functions are written in a tractable form, and some applications to crystallography are discussed. The joint probability may be used in the determination of unknown phases of complex functions (e.g. structure factors) when the moduli (absolute values) are known by experiment. New criteria are developed which allow groups with and without centre of symmetry to be distinguished. A simple relation between phases, generalizing the Sayre-Cochran-Zachariasen relation in the centrosymmetrical case, is obtained. A method is indicated of finding the signs of invariant reflexions (not depending on the choice of the origin) in groups without centre of symmetry.

Introduction et plan

On rencontre des fonctions aléatoires complexes dans l'étude des phénomènes dépourvus d'un centre de symétrie. Leur importance en cristallographie réside dans le simple fait que les groupes sans centre de symétrie forment la grande majorité, en effet sur 230 groupes il n'y en a que 92 doués d'un centre de symétrie.

On évalue d'abord la probabilité pour qu'une fonction complexe

$$F(x_1, \dots, x_n) = R + iI$$

ait ses parties réelles R et imaginaires I comprises dans les intervalles respectifs $(A, A + dA)$ et $(B, B + dB)$. On en déduit aisément la fonction de répartition des modules, par exemple des facteurs de structure, dont l'équation proposée par Wilson (1949) constitue une première (et excellente) approximation.

On ajoute de nouveaux critères permettant la distinction entre groupes avec et sans centre de symétrie.

On forme ensuite l'expression de la probabilité composée pour que m fonctions complexes

$$F_k(x_1, \dots, x_n) = R_k + iI_k \quad (k = 1, \dots, m)$$

aient leurs parties réelles R_k et imaginaires I_k comprises dans les intervalles respectifs $(A_k, A_k + dA_k)$ et $(B_k, B_k + dB_k)$. Lorsque l'observation fournit non pas la mesure complexe

$$M_k = A_k + iB_k = |M_k| \exp i\varphi_k$$

mais seulement le module $|M_k|$, la probabilité composée peut servir à déterminer les phases les plus probables φ_k .

On donne deux applications. L'une montre que lorsque $\varphi(h)$, $\varphi(h')$ et $\varphi(h'')$ sont les phases des réflexions (h) , (h') et (h'') telles que $h + h' + h'' = 0$, alors $\cos[\varphi(h) + \varphi(h') + \varphi(h'')]$ est probablement positif. C'est une généralisation de la relation de Sayre-

Cochran-Zachariasen dans le cas centrosymétrique où $s(h)s(h')s(h'')$ est probablement positif.

La deuxième application concerne la détermination des signes des réflexions invariants. Elle présente un intérêt pratique évident.

Préliminaires mathématiques

Définitions

Soit une fonction complexe $F(x_1, \dots, x_n) = R + iI$, dépendant de n variables x_j ($j = 1, \dots, n$) et prenant toutes ses valeurs dans l'intervalle $0 < x_j < 1$. F est supposée intégrable ainsi que toutes ses puissances telles que $F^p(F^*)^q$. La probabilité élémentaire est définie par le volume élémentaire $dx_1 dx_2 \dots dx_n$. On a

$$\int_0^1 dx_1 \dots dx_n = 1. \quad (1)$$

On appellera mesure de F la valeur complexe $M = A + iB$.

La probabilité $P(A, B)dAdB$ pour que les parties réelles R et imaginaires I d'une fonction complexe $F = R + iI$, dépendant de n variables x_j ($j = 1, \dots, n$) soient comprises respectivement dans les intervalles $(A, A + dA)$ et $(B, B + dB)$ est telle que

$$P(A, B) = \int_0^1 \delta(R - A) \delta(I - B) \prod_{j=1}^n dx_j \quad (2)$$

(cf. Bertaut, 1955a, b).

Choix de la représentation de la fonction de Dirac

Les fonctions δ de Dirac peuvent être représentées par des séries telles que (Bertaut, 1955c, d)

$$\delta(R - A) = G(\mathcal{A}) \sum_{p=0}^{\infty} \frac{1}{p!} H_p(\mathcal{A}) H_p(\mathcal{R}). \quad (3)$$

Ici \mathcal{A} et \mathcal{R} sont les grandeurs réduites

$$\mathcal{A} = A/\alpha; \quad \mathcal{R} = R/\alpha. \quad (4)$$

α^2 (5) est la moyenne quadratique de R ; $H_p(z)$ est le polynôme d'ordre p d'Hermite défini par (6); enfin $G(\mathcal{A})$ est la fonction de Gauss (7):

$$\alpha^2 = \int_0^1 R^2 \prod_{j=1}^n dx_j, \quad (5)$$

$$H_p(z) = (-1)^p \exp\left[\frac{1}{2}z^2\right] \frac{d^p}{dz^p} \exp\left[-\frac{1}{2}z^2\right], \quad (6)$$

$$G(\mathcal{A}) = (\alpha/2 \cdot \pi)^{-1} \exp\left[-\frac{1}{2}\mathcal{A}^2\right]. \quad (7)$$

Introduisant la définition de $H_p(z)$ (6) dans (3) on peut écrire aussi

$$\begin{aligned} \delta(R-A) \\ = (\alpha/2 \cdot \pi)^{-1} \exp\left[\frac{1}{2}\mathcal{R}^2\right] \exp[\mathbf{a}] \exp\left[-\frac{1}{2}(\mathcal{R}^2 + \mathcal{A}^2)\right], \end{aligned} \quad (8)$$

où $\exp[\mathbf{a}]$ est l'opérateur différentiel

$$\exp[\mathbf{a}] = \sum_{p=0}^{\infty} \frac{\mathbf{a}^p}{p!} \quad \text{avec} \quad \mathbf{a} = \frac{\delta}{\delta \mathcal{A}} \frac{\delta}{\delta \mathcal{R}}. \quad (9)$$

On a de même

$$\begin{aligned} \delta(I-B) = (\beta/2 \cdot \pi)^{-1} \exp\left[\frac{1}{2}\mathcal{I}^2\right] \\ \times \exp[\mathbf{b}] \exp\left[-\frac{1}{2}(\mathcal{I}^2 + \mathcal{B}^2)\right], \end{aligned} \quad (10)$$

avec les notations

$$\beta^2 = \int_0^1 I^2 \prod_{j=1}^n dx_j, \quad (11)$$

$$\mathcal{B} = B/\beta, \quad \mathcal{I} = I/\beta, \quad (12)$$

$$\mathbf{b} = \frac{\delta}{\delta \mathcal{B}} \frac{\delta}{\delta \mathcal{I}}. \quad (13)$$

Le produit des fonctions (8) et (10) peut alors se mettre sous la forme concise

$$\begin{aligned} \delta(R-A)\delta(I-B) = (2\pi\alpha\beta)^{-1} \exp\left[\frac{1}{2}|\mathcal{F}|^2\right] \exp[(\mathbf{a}+\mathbf{b})] \\ \times \exp\left[-\frac{1}{2}(|\mathcal{F}|^2 + |\mathcal{M}|^2)\right], \end{aligned} \quad (14)$$

où l'on a abrégé

$$\mathcal{F} = \mathcal{R} + i\mathcal{I}, \quad \mathcal{M} = \mathcal{A} + i\mathcal{B}, \quad (15)$$

$$\exp[(\mathbf{a}+\mathbf{b})] = \sum_{p=0}^{\infty} \frac{1}{p!} \left(\frac{\delta}{\delta \mathcal{A}} \frac{\delta}{\delta \mathcal{R}} + \frac{\delta}{\delta \mathcal{B}} \frac{\delta}{\delta \mathcal{I}} \right)^p. \quad (16)$$

Remarquons que la moyenne quadratique de $|\mathcal{F}|$ est 2. Il n'est pas avantageux, à ce stade du calcul, d'effectuer les opérations ce qui aboutirait à des expressions assez compliquées en \mathcal{A} , \mathcal{B} , \mathcal{R} et \mathcal{I} . Il vaut mieux garder la fonction \mathcal{F} et sa mesure complexe \mathcal{M} le plus longtemps possible. On parvient au but en introduisant les opérateurs suivants

$$S = \frac{\delta}{\delta \mathcal{A}} + i \frac{\delta}{\delta \mathcal{B}}; \quad T = \frac{\delta}{\delta \mathcal{R}} + i \frac{\delta}{\delta \mathcal{I}}. \quad (17)$$

En effet, on peut écrire

$$\mathbf{a} + \mathbf{b} = \frac{1}{2}(ST^* + S^*T). \quad (18)$$

L'avantage des opérateurs (17) vient de ce que l'on a par exemple

$$S\mathcal{M} = 0, \quad (19)$$

et

$$S^p \exp\left[-\frac{1}{2}|\mathcal{M}|^2\right] = (-\mathcal{M})^p \exp\left[-\frac{1}{2}|\mathcal{M}|^2\right].$$

Nous explicitons le développement de (14) jusqu'aux termes d'ordre 4:

$$\begin{aligned} \delta(R-A)\delta(I-B) = (2\pi\alpha\beta)^{-1} \exp\left[-\frac{1}{2}|\mathcal{M}|^2\right] \sum_{r=0}^{\infty} Q_r \quad (20) \\ = (2\pi\alpha\beta)^{-1} \exp\left[-\frac{1}{2}|\mathcal{M}|^2\right] \left\{ 1 + \frac{2^{-1}}{1!} (\mathcal{F}\mathcal{M}^* + \mathcal{F}^*\mathcal{M}) \right. \\ + \frac{2^{-2}}{2!} [(\mathcal{F}\mathcal{M}^*)^2 + (\mathcal{F}^*\mathcal{M})^2 + 2(|\mathcal{F}|^2 - 2)(|\mathcal{M}|^2 - 2)] \\ + \frac{2^{-3}}{3!} [(\mathcal{F}\mathcal{M}^*)^3 + (\mathcal{F}^*\mathcal{M})^3 \\ + 3(\mathcal{F}\mathcal{M}^* + \mathcal{F}^*\mathcal{M})(|\mathcal{F}|^2 - 4)(|\mathcal{M}|^2 - 4)] \\ + \frac{2^{-4}}{4!} [(\mathcal{F}\mathcal{M}^*)^4 + (\mathcal{F}^*\mathcal{M})^4 + 4(|\mathcal{F}\mathcal{M}^*|^2 + |\mathcal{F}^*\mathcal{M}|^2) \\ \times (|\mathcal{F}|^2 - 6)(|\mathcal{M}|^2 - 6) + 6(|\mathcal{F}|^4 - 8|\mathcal{F}|^2 + 8) \\ \times (|\mathcal{M}|^4 - 8|\mathcal{M}|^2 + 8)] + \dots \left. \right\}. \quad (21) \end{aligned}$$

Indiquons enfin pour être complet que Q_r (20) a la forme générale suivante, symétrique en \mathcal{F} et \mathcal{M} que l'on peut prouver par induction:

$$\begin{aligned} Q_r = 2^{-r} \sum_{p=0}^{p \leq r/2} p!(r-p)! [(\mathcal{F}\mathcal{M}^*)^{r-2p} + (\mathcal{F}^*\mathcal{M})^{r-2p}] \\ \times \sum_{q=0}^p (-1)^q 2^q [q!(r-p-q)!(p-q)!]^{-1} |\mathcal{F}|^{2p-2q} \\ \times \sum_{q'=0}^p (-1)^{q'} 2^{q'} [q'!(r-p-q')!(p-q')!]^{-1} |\mathcal{M}|^{2p-2q'}. \quad (22) \end{aligned}$$

Fonction de répartition des modules

L'intégration de (21) par rapport à toutes les variables x_j ($j = 1, \dots, n$) fournit la fonction de répartition $P(A, B)$ (2). Le premier terme de (21) a déjà été donné par Wilson (1949) avec les notations $\alpha^2 = \beta^2 = \frac{1}{2}\Sigma$.

Multiplions (21) par $dA dB$. On a

$$(\alpha\beta)^{-1} dA dB = d\mathcal{A} d\mathcal{B}. \quad (23)$$

Passons ensuite en coordonnées polaires $|\mathcal{M}|, \varphi$:

$$d\mathcal{A} d\mathcal{B} = |\mathcal{M}| d|\mathcal{M}| d\varphi. \quad (24)$$

Dans l'intégration par rapport à φ ne subsistent que les termes en $|\mathcal{M}|^{2n}$ de sorte que compte tenu de $|\overline{\mathcal{F}}|^2 = 2$ on a

$$\begin{aligned} P(|\mathcal{M}|) = |\mathcal{M}| \exp\left[-\frac{1}{2}|\mathcal{M}|^2\right] \\ \times [1 + 2^{-6}(|\mathcal{M}|^4 - 8|\mathcal{M}|^2 + 8)(|\mathcal{F}|^4 - 8) + \dots]. \quad (25) \end{aligned}$$

Application aux facteurs de structure

Nous désignerons par E le facteur de structure normalisé (26) et par V sa valeur complexe. $\xi(h)$ est la partie trigonométrique usuelle du facteur de structure. (Pour les définitions cf. Bertaut, 1955*d*). On a

$$E(h) = \sum_{j=1}^l \varphi_j \xi_j(h) \quad (26)$$

et

$$E = \mathcal{F}/\sqrt{2}; \quad V = \mathcal{M}/\sqrt{2}. \quad (27)$$

Il n'y a aucune difficulté à transcrire (21) dans ces variables. Nous n'explicitons que la relation (25). On a

$$P(|V|) = 2|V| \exp[-|V|^2] \times [1 + 2^{-2}(|V|^4 - 4|V|^2 + 2)(\overline{|E|^4} - 2) + \dots]. \quad (28)$$

Le calcul direct montre d'autre part que pour une réflexion générale (hkl)

$$\begin{aligned} \overline{|E|^4} &= 2\overline{|E|^2} + \sum_{j=1}^l \varphi_j^4 [|\xi_j|^4 - 2(|\xi_j|^2)^2] \\ &= 1 \times 2 + \sum_{j=1}^l \varphi_j^4 (|\xi_j|^4 - 2n_j^2), \end{aligned} \quad (29)$$

et que plus généralement la partie principale de $\overline{|E|^{2n}}$ est $n!$. Dans le cas de N atomes, tous égaux, dans la maille on a

$$\overline{|E|^{2n}} = n! + O(N^{-1}). \quad (30)$$

Or les moments d'ordre $2n$ correspondant à la fonction de répartition $2|V| \exp[-|V|^2]$ sont précisément $n!$. Cela prouve que la statistique des modules suit bien la formule proposée par Wilson (1949), aux termes en $1/N$ près. Au même degré d'approximation on a dans un groupe centrosymétrique

$$\overline{|E|^{2n}} = (2n-1)!/2^n(n-1)!. \quad (31)$$

De la comparaison de (30) et de (31) on déduit des critères sensibles quant à la présence ou absence d'un centre de symétrie.

On a par exemple $\overline{|E|^4} = 3$; $\overline{|E|^6} = 15$ dans un groupe centrosymétrique et $\overline{|E|^4} = 2$; $\overline{|E|^6} = 6$ dans un groupe sans centre de symétrie.

(Pour les réflexions spéciales, cf. Bertaut, 1955*d*.)

La probabilité composée

La probabilité composée

$$P(A_1, B_1, \dots, A_m, B_m) dA_1 dB_1 \dots dA_m dB_m$$

pour que m fonctions complexes

$$F_k = R_k + iI_k, \quad (k = 1, \dots, m)$$

dépendant chacune de n variables réelles x_j ($j = 1, \dots, n$), aient leurs parties réelles R_k et imaginaires I_k com-

prises dans les intervalles respectifs ($A_k, A_k + dA_k$) et ($B_k, B_k + dB_k$) est telle que (Bertaut, 1955*a, b*)

$$P(A_1, B_1, \dots, A_m, B_m) = \int_0^1 \prod_{k=1}^m \delta(R_k - A_k) \delta(I_k - B_k) \prod_{j=1}^n dx_j. \quad (32)$$

Il suffit de former le produit de formes (21) et d'intégrer par rapport à toutes les variables x_j . Nous écrivons le résultat avec les fonctions normalisées (27):

$$\begin{aligned} P(A_1, B_1, \dots, A_m, B_m) &= (2\pi)^{-m} (\alpha_1 \beta_1 \dots \alpha_m \beta_m)^{-1} \exp[-(|V_1|^2 + \dots + |V_m|^2)] \\ &\times \int_0^1 \prod_{j=1}^n dx_j \left[1 + \sum_k \alpha_k + \sum_{k,l} \alpha_k \alpha_l + \sum_{k,l,p} \alpha_k \alpha_l \alpha_p + \frac{1}{2!} \sum_{k,l} \mathfrak{b}_k \alpha_l \right. \\ &\left. + \sum_{k,l,p,q} \alpha_k \alpha_l \alpha_p \alpha_q + \frac{1}{2!} \sum_{k,l,p} \mathfrak{b}_k \alpha_l \alpha_p + \frac{1}{3!} \sum_{k,l} \mathfrak{c}_k \alpha_l + \dots \right], \end{aligned} \quad (33)$$

où l'on a abrégé†

$$\left. \begin{aligned} \alpha_k &= E_k V_k^* + E_k^* V_k; \\ \mathfrak{b}_k &= (E_k V_k^*)^2 + (E_k^* V_k)^2 + 2(|E_k|^2 - 1)(|V_k|^2 - 1); \\ \mathfrak{c}_k &= (E_k V_k^*)^3 + (E_k^* V_k)^3 + 3\alpha_k (|E_k|^2 - 2)(|V_k|^2 - 2). \end{aligned} \right\} \quad (34)$$

Lorsque les fonctions E_k sont aléatoires et orthogonales, le premier terme non nul dans (33) est d'ordre trois en V .

Supposons que les intégrations dans (33) sur les fonctions E_k puissent être effectuées et que, expérimentalement, on connaisse seulement les modules $|V_k|$. Remplaçons alors les V_k par

$$V_k = |V_k| \exp[i\varphi_k]. \quad (35)$$

On obtient ainsi pour la probabilité composée une expression où figurent comme seules inconnues les phases φ_k . Leurs valeurs les plus probables sont celles qui rendent (33) maximum:

$$\delta P(\varphi_1, \dots, \varphi_m) / \delta \varphi_k = 0 \quad (\text{avec } \delta^2 P / \delta \varphi_k^2 < 0). \quad (36)$$

La résolution du système d'équations (36) par rapport aux phases fournit donc les phases les plus probables, vue l'information donnée. Signalons à ce propos une indétermination. La probabilité (33) ne change pas si à chaque grandeur E_k et V_k on substitue la grandeur conjuguée. Cela signifie qu'à toute solution $\varphi_1, \varphi_2, \dots, \varphi_m$ de (33) correspond aussi la solution $-\varphi_1, -\varphi_2, \dots, -\varphi_m$.

Application à la cristallographie

L'expression (33) constitue la probabilité composée des facteurs de structure complexes de tous les groupes non centrosymétriques. Les moyennes

$$\overline{E_1^a (E_1^*)^{a'} \dots E_m^a (E_m^*)^{a'}} \quad (37)$$

† Les abréviations α_k et \mathfrak{b}_k dans (34) ne doivent pas être confondues avec les opérateurs \mathbf{a} (9) et \mathbf{b} (13).

que l'on doit calculer ne diffèrent de zéro que lorsque certaines conditions sont imposées aux h_1, \dots, h_m (cf. Bertaut, 1955*b*). Les énoncés sur la linéarisation des produits et des puissances des parties trigonométriques des facteurs de structure (Bertaut, 1955*d*) restent valables dans le cas non centrosymétrique. Voici le principe du calcul. On formera d'abord le produit

$$\xi_1^a (\xi_1^*)^{a'} \dots \xi_m^q (\xi_m^*)^{q'} \varphi_j^{a+a'+\dots+q+q'} \quad (38)$$

analogue à (37), mais correspondant à une seule position atomique d'indice j . On le linéarise sous la forme

$$a_{1j} \xi(H_1) + a_{2j} \xi(H_2) + \dots + a_{pj} \xi(H_p), \quad (39)$$

où les H_q sont des combinaisons linéaires des h_1, \dots, h_m . On posera ensuite par exemple $H_1 = 0$. Cela impose une condition (linéaire) entre les h_1, \dots, h_m . Lorsqu'elle est remplie, H_2, \dots, H_p seront différents de zéro et la moyenne de (38) se réduira à $a_{1j} \xi(0)$. On sommerà sur toutes les positions atomiques j ($j = 1, \dots, t$). Cela fournit la moyenne de (37) dans la condition $H_1 = 0$.

On mène les calculs de la même façon dans la condition $H_2 = 0$ et ainsi de suite, jusqu'à ce qu'on ait épuisé les p conditions possibles. (Des applications simples sont rencontrées dans la section suivante.) Mentionnons que les moyennes (37) sont des quantités réelles.

Remarque.—Le cas où la condition $H_1 = 0$ peut se décomposer en deux conditions indépendantes $H_1' = 0$ et $H_1'' = 0$ mérite une mention spéciale. Il correspond au cas où le produit (37) peut se décomposer en un produit de deux facteurs, la moyenne de chacun étant différent de zéro. Dans ce cas la moyenne du produit est égale au produit des moyennes des facteurs. Cette éventualité se rencontre pour les produits (37) d'ordre supérieur ou égal à six. (Le cas le plus simple est celui de

$$\overline{E_1 E_2 E_3 \times E_4 E_5 E_6}$$

avec la condition $h_1 + h_2 + h_3 = 0$ et $h_4 + h_5 + h_6 = 0$. Le résultat est, dans nos notations (Bertaut, 1955*d*) égal à z_3^2 .)

Généralisation

Dans tout groupe sans centre de symétrie (à l'exception de $P1$) on peut partager les facteurs de structure en deux catégories principales: (a) catégorie complexe ($R \neq 0, I \neq 0$) pour laquelle la répartition (33) est valable, (b) catégorie simple (ou $R = 0$ ou $I = 0$) pour laquelle la répartition centrosymétrique (cf. Bertaut, 1955*c, d*) est valable.

Pour fixer les idées, supposons qu'il y ait

(a) des fonctions normalisées complexes

$$E_k = R_k + iI_k \quad (k = 1, \dots, m),$$

(b₁) des fonctions normalisées réelles

$$E_r = R_r \quad (r = 1, \dots, s)$$

et

(b₂) des fonctions normalisées purement imaginaires

$$E_j = iI_j \quad (j = 1, \dots, t).$$

Pour construire la densité de probabilité composée

$$\begin{aligned} P(A_k, B_k, \dots; A_r, \dots, B_j, \dots) \\ = \int_0^1 \prod_k \delta(R_k - A_k) \delta(I_k - B_k) \times \prod_r \delta(R_r - A_r) \\ \times \prod_j \delta(I_j - B_j) dx_1 \dots dx_n \end{aligned} \quad (40)$$

on utilisera pour la catégorie complexe (a) les produits de formes (21); pour la catégorie réelle (b₁), on utilisera le développement (3) et un développement analogue pour la catégorie purement imaginaire (b₂).

On fait le produit de toutes les fonctions de Dirac et l'on intègre par rapport à toutes les coordonnées x_1, \dots, x_n . Le résultat est la densité de probabilité composée *complète* des facteurs de structure du groupe considéré. Nous n'explicitons pas le résultat qu'il est d'ailleurs facile d'écrire. Les considérations sur les phases les plus probables restent encore valables.

Sans aborder le problème difficile de la résolution pratique du système (36) nous sommes maintenant en mesure d'envisager deux applications.

Application I: une relation simple entre phases

Soient trois facteurs de structure complexes et normalisés $E(h), E(h'), E(h'')$ tels que

$$h + h' + h'' = 0. \quad (41)$$

Soient $V(h), V(h'), V(h'')$ leurs valeurs, $\varphi(h), \varphi(h'), \varphi(h'')$ leurs phases. La probabilité composée est de la forme

$$\begin{aligned} P[V(h), V(h'), V(h'')] \\ = C[1 + \overline{a(h)a(h')a(h'')} + \dots] \\ = C\{1 + z_3[V(h)V(h')V(h'')] \\ + (V(h)V(h')V(h''))^* + \dots\}, \end{aligned} \quad (42)$$

où l'on a abrégé

$$a(h) = E(h)V^*(h) + E^*(h)V(h), \quad (43)$$

$$z_3 = \overline{E(h)E(h')E(h'')} = \sum_{j=1}^t n_j \varphi_j^3. \quad (44)$$

Posons

$$\varphi = \varphi(h) + \varphi(h') + \varphi(h'') \quad (45)$$

et envisageons l'intervalle $-\frac{1}{2}\pi < \varphi < \frac{3}{2}\pi$. Dans l'intervalle $-\frac{1}{2}\pi < \varphi < \frac{1}{2}\pi$, $\cos \varphi$ est positif. La probabilité pour que $\cos \varphi$ soit dans cet intervalle est de la forme

$$P^+ = C \int_{-\pi/2}^{\pi/2} d\varphi (1 + 2z_3 \cos \varphi |V(h)V(h')V(h'')|). \quad (46)$$

La probabilité P^- pour que $\cos \varphi$ soit négatif se

calculé de la même façon. La constante de normalisation C est déterminé par la condition $P^+ + P^- = 1$. On a

$$C = (2\pi)^{-1} \quad (47)$$

et

$$P^+ = \frac{1}{2} + (2/\pi)z_3 |V(h)V(h')V(h'')|. \quad (48)$$

Partant de facteurs de structure unitaires $U(h)$, $U(h')$, $U(h'')$ on aurait obtenu

$$P^+ = \frac{1}{2} + (2/\pi)\alpha |U(h)U(h')U(h'')|, \quad (49)$$

où

$$\alpha = \sum n_j^3 / (\sum n_j^2)^3 \quad (50)$$

se réduit à N dans le cas d'atomes tous égaux.

La probabilité pour que $\cos [\varphi(h) + \varphi(h') + \varphi(h'')]$ soit positif, dépasse donc toujours $\frac{1}{2}$. Cela constitue une généralisation de la relation de Sayre-Cochran-Zachariasen.

Remarque 1. — Il semble qu'il faille envisager un domaine de variation de 2π pour chacune des phases $\varphi(h)$, $\varphi(h')$, $\varphi(h'')$ et une probabilité élémentaire $d\varphi(h)d\varphi(h')d\varphi(h'')$. On obtient cependant exactement le même résultat, mais après des calculs plus longs.

Remarque 2. — On peut montrer que la relation (48) reste encore vraie lorsque l'un des trois facteurs de structure est réel.

Application II: réflexions invariantes

Notant toujours par $\xi(h, k, l)$ la partie trigonométrique du facteur de structure, $\xi(h, k, l)\xi(\bar{h}, \bar{k}, \bar{l}) = |\xi(h, k, l)|^2$ est évidemment une quantité réelle. La linéarisation (Bertaut, 1955d) de $|\xi(h, k, l)|^2$ fournira automatiquement soit des facteurs $\xi(H, K, L)$ réels, soit des parties réelles de ces facteurs, c'est-à-dire les combinaisons telles que $\xi(H, K, L) + \xi^*(H, K, L)$ ou $[\xi(H, K, L) - \xi^*(H, K, L)]/i$.

De plus la valeur du produit $\xi(h, k, l)\xi(\bar{h}, \bar{k}, \bar{l})$ est indépendante du choix de l'origine. Il en sera de même des $\xi(H, K, L)$ réels ou de leurs parties réelles qui figurent dans la somme linéarisée. On obtient ainsi toutes les réflexions invariantes dans un changement d'origine (comme dans le cas centrosymétrique). De la corrélation entre les réflexions invariantes (H, K, L) et les réflexions générales (hkl) on déduit l'application suivante.

Lorsque dans un groupe non centrosymétrique il y a une projection centrosymétrique, les signes des réflexions invariantes réelles dépendent non seulement des réflexions réelles de la zone centrosymétrique, mais aussi des réflexions générales. La corrélation fait intervenir uniquement les modules des facteurs de structure des réflexions générales.

Nous donnons deux exemples pour fixer les idées.

(1°) Groupe $D_{2d}^2 - P42c$. — On a ici

$$|\xi(h, k, l)|^2 = \xi(0, 0, 0) + \xi(2h, 2k, 0) + [\xi(0, 2k, 2l) + \xi(2h, 0, 2l)] \exp[\pi il] + \xi(h', k', 2l) + \xi^*(h', k', 2l) + [\xi(h', h', 0) + \xi(k', k', 0)] \exp[\pi il], \quad (51)$$

où

$$h' = h + k \quad \text{et} \quad k' = k - h. \quad (52)$$

Les réflexions qui figurent dans le second membre de (51) sont invariantes; en particulier les signes des parties réelles des réflexions $\xi(h', k', 2l)$ où h' et k' ont la même parité (52) ne dépendent pas du choix de l'origine.

(2°) Groupe $D_2^4 - P2_12_12_1$. — On a ici

$$|\xi(h, k, l)|^2 = \xi(0, 0, 0) + \xi(2h, 2k, 0)s + \xi(0, 2k, 2l)t + \xi(2h, 0, 2l)v \quad (53)$$

avec

$$s = \exp[\pi i(h+l)], \quad t = \exp[\pi i(h+k)], \quad v = \exp[\pi i(k+l)]. \quad (54)$$

On a par exemple

$$\overline{|\xi(h, k, l)|^2 \xi(2h, 2k, 0)} = \overline{s \xi^2(2h, 2k, 0)} = s \xi(0, 0, 0), \quad (55)$$

d'où l'on déduit que

$$\overline{|E(h, k, l)|^2 E(2h, 2k, 0)} = s \sum_{j=1}^t \overline{\varphi_j^3 \xi_j^2(2h, 2k, 0)} = s \sum_{j=1}^t n_j \varphi_j^3 = s z_3. \quad (56)$$

La densité de probabilité composée, formée avec les réflexions $(2h, 2k, 0)$ et (h, k, l) , comme il a été dit plus haut, s'écrit

$$P = C [1 + A(2h, 2k, 0) \sum (|V(h, k, l)|^2 - 1) \times (|E(h, k, l)|^2 - 1) E(2h, 2k, 0) + \dots], \quad (57)$$

et l'on trouve aisément que la probabilité P^+ pour que $A(2h, 2k, 0)$ soit positif est donnée par

$$P^+ = \frac{1}{2} [1 + |A(2h, 2k, 0)| z_3 \sum_t (-1)^{h+l} (|V(hkl)|^2 - 1) + \dots]. \quad (58)$$

Le signe probable de $A(2h, 2k, 0)$ est donc celui de la somme

$$(-1)^h \sum_t (|V(h, k, l)|^2 - 1) (-1)^l \quad (59)$$

où l'indice de sommation est l .

Des relations similaires ont été indiquées par Vand & Pepinsky (1953) pour ce même groupe. La méthode de linéarisation nous paraît à la fois plus générale et plus élégante.

Nous avons rédigé, en collaboration avec Monsieur Jean Dulac, des tables contenant les carrés des modules $|\xi(h, k, l)|^2$ linéarisés pour tout groupe non centrosymétrique. Ces tables qui contiennent également les carrés, troisièmes et quatrièmes puissances de $\xi(h, k, l)$ ainsi que les produits $\xi(h, k, l)\xi(h', k', l')$ linéarisés des

groupes centrosymétriques, seront publiés prochainement.

Conclusion

La statistique dans le cas du cristal non centrosymétrique, abordée par Wilson (1949), Hauptmann & Karle (1953) et Vand & Pepinsky (1953), est présentée sous une forme nouvelle qui, espérons-le, nous approchera de la solution définitive du problème des phases.

Références

- BERTAUT, E. F. (1955a). *C. R. Acad. Sci., Paris*, **240**, 938.
 BERTAUT, E. F. (1955b). *Acta Cryst.* **8**, 537.
 BERTAUT, E. F. (1955c). *C. R. Acad. Sci., Paris*, **240**, 152.
 BERTAUT, E. F. (1955d). *Acta Cryst.* **8**, 823.
 HAUPTMANN, H. & KARLE, J. (1953). *Acta Cryst.* **6**, 136.
 VAND, V. & PEPINSKY, R. (1953). *The Statistical Approach to X-ray Structure Analysis*. State College: Pennsylvania State University.
 WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.

Acta Cryst. (1956). **9**, 460

Crystalline Forms of Bovine Pancreatic Ribonuclease: Techniques of Preparation, Unit Cells, and Space Groups*

BY MURRAY VERNON KING, BEATRICE S. MAGDOFF†, MYRA BURT ADELMAN AND DAVID HARKER

Protein Structure Project, Polytechnic Institute of Brooklyn, 55 Johnson Street, 4th Floor, Brooklyn 1, N.Y., U.S.A.

(Received 25 October 1955)

Seven crystalline forms of bovine pancreatic ribonuclease have been prepared and characterized by X-ray diffraction. Two of these forms (I and VII) are complexes with nickel, and one (VI) is a complex with a dye, iodophenol blue. Methods of preparing each of the polymorphs are given. The observed lattice parameters are found to vary by small amounts on changing the composition (pH, etc.) of the medium from which the crystals are obtained. Some of the polymorphs were dyed by various members of the sulfonphthalein indicator series, while other polymorphs could not be dyed under similar conditions.

Techniques of crystallization

Methods of crystallizing bovine pancreatic ribonuclease from salt solutions at room temperature, or from aqueous alcohol at lower temperature, have been reported by Kunitz (1940). In the present work, techniques were developed for crystallizing ribonuclease at room temperature from aqueous solutions of various organic liquids. The crystals, which were required for an X-ray diffraction study of the crystal structure of ribonuclease, appeared in a number of distinct modifications, of which seven were characterized by X-ray diffraction data. The techniques reported here have made possible the reproducible production of these modifications, as their occurrence was found to depend on factors, such as the pH, the identity and concentration of the organic liquid used as a precipitant, and in some cases on the presence of small amounts of heavy metals or dyes.

The sources of ribonuclease for this work were: (1) A crude fraction from bovine pancreas precipitated

from 0.7–0.8 saturated ammonium sulfate, purchased from Worthington Biochemical Company, and purified in this laboratory by either the method of McDonald (1948) or that of Kunitz (1940). (2) Lyophilized recrystallized salt-free ribonuclease purchased from Armour Laboratories. Different lots of the Armour ribonuclease differed in pH, but material giving a solution of a required pH could be obtained by dissolving the ribonuclease in water, adjusting the pH by addition of 0.1M NaOH or HCl, and lyophilizing again.

The methods of crystallization were:

(A) An aqueous solution of ribonuclease is frozen and then partially thawed. The organic liquid is added with gentle shaking, and the solution is then stored in a constant-temperature bath at 24–25° C. By use of this method of addition, both the effects of evolution of heat of mixing and of contact of nearly anhydrous organic liquid with the ribonuclease are avoided as much as possible.

(B) An aqueous solution of the organic liquid is chilled and poured on to solid lyophilized ribonuclease. The suspension is stored in the constant-temperature water bath at 24–25° C. The particles of ribonuclease

* Contribution No. 2 from the Protein Structure Project, Polytechnic Institute of Brooklyn

† Present address: Boyce Thompson Institute for Plant Research, Yonkers 3, New York, U.S.A.