

## Ordre Logarithmique des Densités de Repartition. I

PAR E. F. BERTAUT

*Laboratoire d'Electrostatique et de Physique du Métal, Institut Fourier, Place du Doyen Gosse,  
Grenoble, France*

(Reçu le 17 juin 1959)

By logarithmic ordering one obtains the distribution function for the one—and multi-dimensional case in a form, positive everywhere and such that polynomials of the same degree are associated with coefficients of the same order of magnitude. The law of the hyperbolic tangent proposed by Cochran & Woolfson (1955) for sign determining relations is demonstrated rigorously. It is shown that only the odd part of the logarithm of the joint distribution function is significant.

### Introduction et plan

Dans un article récent, Klug (1958), a fourni une importante contribution à l'étude de la convergence du développement en série des densités de répartition  $P(A_1, \dots, A_m)$  des facteurs de structure

$$A_k (k=1, \dots, m).$$

Malgré la similitude des techniques mathématiques employées par Klug et par nous dans un mémoire antérieur (Bertaut, 1955a) en ce qui concerne la fonction caractéristique, les cumulants et les méthodes symboliques, il y a une différence essentielle dans la position du problème. Alors que nous avons rangé les termes de la fonction caractéristique  $C(u_1, \dots, u_m)$  selon leur degré polynomial en  $u_k (k=1, \dots, m)$  Klug les range selon leur ordre de grandeur, en définissant un paramètre d'ordre de grandeur  $t$ , égal au nombre d'atomes dans l'unité asymétrique.

Klug a montré qu'à une même puissance du paramètre  $t$  sont associés des polynômes de degrés différents en  $u_k$  dans la fonction caractéristique et par conséquent aussi de degrés différents en  $A_k$  dans la densité de répartition.

Dans ce mémoire nous poursuivons les idées de Klug plus loin. Mais au lieu d'ordonner les termes de la densité de répartition  $P(A_1, \dots, A_m)$ , nous ordonnons  $\log P(A_1, \dots, A_m)$ . Cette procédure présente de nombreux avantages. Nous parvenons d'abord à écrire la densité de répartition sous une forme positive partout. (Il n'en est pas ainsi des séries de Gram-Charlier, cf. Kendall, 1943, p. 150, 151). De plus, avantage par rapport à la présentation de Klug, dans

$$\log P(A_1, \dots, A_m)$$

les termes s'ordonnent de telle sorte qu'à un polynôme de degré  $s$  en  $A_k$  soit associé un paramètre d'ordre de grandeur  $\varepsilon^{s-2}$  ( $\varepsilon = N^{-\frac{1}{2}}$ ;  $N$  = nombre d'atomes dans la maille).

Enfin nous montrons que la probabilité  $P^+$  d'un signe positif de  $A_1$  peut s'écrire rigoureusement

$$P^+ = \frac{1}{2} + \frac{1}{2} \operatorname{th} Q_i,$$

$Q_i$  étant la partie impaire en  $A_1$  de  $\log P(A_1, \dots, A_m)$ .

Pour habituer le lecteur aux techniques utilisées, nous considérons dans une première partie le cas de la densité de répartition d'une seule variable  $P(A)$ .

Dans la deuxième partie nous considérons la fonction  $P(A_1, \dots, A_m)$  dans le cas général d'atomes différents et de positions atomiques quelconques (en position générale ou spéciale, mais dépendant de paramètres). Notre traitement tient implicitement compte des poids statistiques.

### Cas d'une variable

Nous exposerons d'abord aussi simplement que possible le principe. Soit une variable aléatoire  $E$  normalisée.

$$\langle E^2 \rangle = 1. \quad (\text{I-1})$$

On cherche la densité de répartition  $P(A)$  définie positive telle que  $P(A) \cdot dA$  est la probabilité pour que  $E$  soit compris entre les valeurs  $A$  et  $A + dA$ . On suppose que tous les moments  $\mu_q$  d'ordre  $q$  existent ( $\mu_1 = 0$ ;  $\mu_2 = 1$ )

$$\langle E^q \rangle = \mu_q. \quad (\text{I-2})$$

La fonction caractéristique est alors:

$$C(u) = \langle \exp iEu \rangle = \sum_q \mu_q v^q / q! \quad (\text{I-3})$$

où

$$v = iu$$

et la densité de répartition  $P(A)$  est sa transformée de Fourier.

$$P(A) = \int C(u) \exp(-iuA) du. \quad (\text{I-4})$$

On peut définir des cumulants  $k_q$  en fonction des  $\mu_q (k_1 = 0; k_2 = 1, \text{ cf. appendice A})$  par:

$$\begin{aligned} C(u) &= \exp \sum_q k_q v^q / q! \\ &= \exp -\frac{1}{2} u^2 \exp \sum_{q \geq 3} k_q v^q / q! \\ &= \exp -\frac{1}{2} u^2 (1 + v^3 k_3 / 3! + v^4 k_4 / 4! + (1/2!) \\ &\quad \times (v^3 k_3 / 3!)^2 + v^5 k_5 / 5! + v^3 v^4 k_3 k_4 / (3! 4!) \\ &\quad + (1/3!) (v^3 k_3 / 3!)^3 + \dots) \end{aligned} \quad (I-5)$$

Nous avons 'ordonnée'  $C(u)$  selon le principe suivant. Dans une grande classe de problèmes  $k_3/3!$ ,  $k_4/4!$  et  $k_5/5!$  sont de l'ordre de  $\varepsilon$ ,  $\varepsilon^2$  et de  $\varepsilon^3$  respectivement de sorte que  $(k_3/3!)^2$  est aussi de l'ordre de  $\varepsilon^2$ . De même  $k_3 k_4 / (3! 4!)$  et  $(k_3/3!)^3$  sont de l'ordre de  $\varepsilon^3$ . La transformée de Fourier de  $C(u)$  (I-5) est:

$$P(A) = (\sqrt{2\pi})^{-1} \exp(-\frac{1}{2} A^2) S \quad (I-6)$$

où:

$$\begin{aligned} S &= 1 + H_3(A) k_3 / 3! + H_4(A) k_4 / 4! \\ &\quad + (1/2!) H_6(A) (k_3 / 3!)^2 + H_5(A) k_5 / 5! \\ &\quad + H_7(A) k_3 k_4 / (3! 4!) + (1/3!) H_9(A) (k_3 / 3!)^3 \\ &\quad + \dots \end{aligned} \quad (I-7)$$

est formée simplement en remplaçant  $v^p$  par  $H_p(A)$  (cf. Bertaut, 1955; Klug, 1958), fonction d'Hermite d'ordre  $p$ .

$$\begin{aligned} H_p(x) &= \sum_{m=0}^{\leq p/2} (-1)^m x^{p-2m} p! / 2^m (p-2m)! m! \\ &= \exp \frac{1}{2} x^2 (-1)^p \frac{d^p}{dx^p} \exp -\frac{1}{2} x^2 \end{aligned} \quad (I-8)$$

De tels développements ont été proposés par Edgeworth (1904) et généralisés au cas de plusieurs variables par Klug. Ils possèdent cependant deux inconvénients. C'est que  $S$  (I-7) étant un développement limité,  $P(A)$  n'est pas forcément positif partout (cf. Kendall, 1943). Enfin au point de vue analytique  $P(A)$  est difficile à manipuler, car des polynômes de degrés différents (par exemple 5, 7 et 9) sont associés avec des coefficients de même ordre de grandeur (ici,  $k_5/5!$ ,  $k_3 k_4 / (3! 4!)$  et  $(k_3/3!)^3$ ).

C'est ici que nous introduisons un point de vue nouveau dont les avantages vont apparaître plus tard: au lieu d'ordonner  $P(A)$ , nous ordonnons  $\log P(A)$ . Pour cela nous appliquons de nouveau la procédure des cumulants à  $S$ , c'est-à-dire nous posons dans (I-6):

$$S = \exp Q \quad (I-9)$$

avec:

$$\begin{aligned} S &= 1 + m_3(1) k_3 / 3! + (1/2!) m_3(2) (k_3 / 3!)^2 \\ &\quad + (1/3!) m_3(3) (k_3 / 3!)^3 + \dots + m_4(1) (k_4 / 4!) \\ &\quad + m_{3,4}(1,1) k_3 k_4 / (3! 4!) + m_5(1) k_5 / 5! + \dots \end{aligned} \quad (I-10)$$

$$\begin{aligned} Q &= 1 + \beta_3(1) k_3 / 3! + (1/2!) \beta_3(2) (k_3 / 3!)^2 \\ &\quad + (1/3!) \beta_3(3) (k_3 / 3!)^3 + \dots + \beta_4(1) k_4 / 4! \\ &\quad + \beta_{3,4}(1,1) k_3 k_4 / (3! 4!) + \beta_5(1) k_5 / 5! + \dots \end{aligned} \quad (I-11)$$

Cette procédure revient à considérer  $(k_3/3!)$ ,  $(k_4/4!)$ ,

$(k_5/5!)$  etc. comme des 'variables' associées à des 'moments'

$$\begin{aligned} m_3(1) &= H_3(A); \quad m_3(2) = H_6(A); \\ m_{p_1, p_2, \dots}(q_1, q_2, \dots) &= H_{p_1 q_1 + p_2 q_2 + \dots}(A) \end{aligned} \quad (I-12)$$

et à chercher les 'cumulants' correspondants (cf. appendice A)

$$\begin{aligned} \beta_p(1) &= m_p(1); \quad \beta_p(2) = m_p(2) - (m_p(1))^2; \\ \beta_p(3) &= m_p(3) - 3m_p(2)m_p(1) + 2(m_p(1))^3; \\ \beta_{p_1, p_2}(1,1) &= m_{p_1, p_2}(1,1) - m_{p_1}(1)m_{p_2}(1) \end{aligned} \quad (I-13)$$

Ici l'indice inférieur  $p$  rappelle la 'variable'  $(k_p/p!)$  et l'indice en parenthèse ( $q$ ) se réfère à l'ordre du moment ou cumulant. On a ainsi:

$$\begin{aligned} \beta_p(1) &= H_p(A); \quad \beta_p(2) = H_{2p}(A) - (H_p(A))^2; \\ \beta_p(3) &= H_{3p}(A) - 3H_{2p}(A)H_p(A) + 2(H_p(A))^3; \\ \beta_{p_1, p_2}(1,1) &= H_{p_1 + p_2}(A) - H_{p_1}(A)H_{p_2}(A) \end{aligned} \quad (I-14)$$

Tous les  $\beta_{p_1, p_2, \dots}(q_1, q_2, \dots)$  peuvent donc s'exprimer à l'aide de polynômes d'Hermite.

Nous remarquons d'abord que  $P(A)$  (I-15) est maintenant sous une forme toujours positive, même si le développement de  $Q$  est limité, car  $Q$  est réel

$$P(A) = (\sqrt{2\pi})^{-1} \exp(-\frac{1}{2} A^2 + Q) \quad (I-15)$$

Ensuite, et c'est là le point capital, on peut montrer par induction que les  $\beta_p(q)$  sont des polynômes en  $A$  de degré  $s$

$$s = pq + 2(1 - q) \quad (I-16)$$

et plus généralement  $\beta_{p_1, p_2, \dots}(q_1, q_2, \dots)$  est de degré

$$s = p_1 q_1 + p_2 q_2 + \dots + 2(1 - q_1 - q_2 - \dots) \quad (I-17)$$

Par exemple  $\beta_3(2)$  et  $\beta_4(1)$ , associés à  $(k_3/3!)^2$  et  $(k_4/4!)$ , coefficients de même ordre  $\varepsilon^{s-2} = \varepsilon^2$ , sont de même degré  $s=4$ . De même aux coefficients  $(k_5/5!)$ ,  $k_3 k_4 / (3! 4!)$ ,  $(k_3/3!)^3$  qui sont du même ordre  $\varepsilon^{s-2} = \varepsilon^3$ , sont associés des polynômes  $\beta_5(1)$ ,  $\beta_{3,4}(1,1)$ ,  $\beta_3(3)$  de degré  $s=5$ . En effet

$$\begin{aligned} \beta_{3,4}(1,1) &= -12(A^5 - 7A^3 + 8A); \\ \beta_3(3) &= 6 \times 9(3A^5 - 16A^3 + 15A) \end{aligned} \quad (I-18)$$

Aux coefficients d'un même ordre de grandeur  $\varepsilon^{s-2}$  sont associés des polynômes de degré  $s$

Exemple: Considérons le cas simple où  $E$  dépend de  $t$  variables  $x_j$  ( $j=1, \dots, t$ ) indépendantes par l'intermédiaire de  $t$  fonctions  $G(x_j)$  additives

$$E = \frac{1}{\sqrt{t}} \sum_{j=1}^t G(x_j) \quad (I-19)$$

Les fonctions  $G(x_j)$  sont toutes identiques (elles ne diffèrent que par la dénomination de la variable), aléatoires et normalisées. On admet que tous les

moments (I-20)  $\mu'_p$  et les cumulants correspondants  $k'_p$  existent ( $\mu'_1=0$ ;  $\mu'_2=1$ )

$$\mu'_p = \langle G^p \rangle. \quad (\text{I-20})$$

Dans ces conditions  $E$  (I-19) est aussi normalisé et la fonction caractéristique (I-4 à I-5) devient

$$\begin{aligned} \mathbf{C}(u) &= \left\langle \exp \frac{v}{\sqrt{t}} G(x_1) \dots \exp \frac{v}{\sqrt{t}} G(x_t) \right\rangle \\ &= \left( \left\langle \exp \frac{v}{\sqrt{t}} G(x) \right\rangle \right)^t \\ &= \left( \exp \sum_q v^q k'_q / (q! \sqrt{t}^q) \right)^t \\ &= \exp \left( \sum_q v^q k'_q / (q! (\sqrt{t})^{q-2}) \right). \quad (\text{I-21}) \end{aligned}$$

La comparaison avec (I-5) montre que

$$k_q = k'_q / (\sqrt{t})^{q-2}. \quad (\text{I-22})$$

Si  $t$  est assez grand et que les  $k'_q$  varient lentement avec  $q$ , c'est le terme  $(t^{-\frac{1}{2}})^{q-2} = \varepsilon^{q-2}$  qui fixera l'ordre de grandeur de  $k_q$ ;  $k_3$  sera de l'ordre de  $\varepsilon$ ,  $k_5$  de l'ordre de  $\varepsilon^3$  etc. ....

#### Application à la cristallographie

$E$  est un facteur de structure normalisé. Dans le cas d'atomes tous égaux et en position générale on peut écrire  $E$  sous la forme (I-19) avec

$$G(x_j) = \xi_j / \sqrt{\langle \xi_j^2 \rangle} \quad (\text{I-23})$$

où  $\xi$  est le facteur de structure trigonométrique. La probabilité d'une position atomique est supposée uniforme. Dans ce cas tous les moments et cumulants  $k'_q$  et  $\mu'_q$  d'ordre impair sont nuls et l'on trouve que

$$\begin{aligned} P(A) &= (\sqrt{2\pi})^{-1} \exp -\frac{1}{2} A^2 \exp \{ t^{-1} H_4(A) k'_4 / 4! \\ &+ t^{-2} [H_6(A) k'_6 / 6! + (1/2!) (H_8(A) - H_4^2(A)) (k'_4 / 4!)^2] \\ &+ \dots \} \quad (\text{I-24}) \end{aligned}$$

où

$$\begin{aligned} k'_4 &= \mu'_4 - 3(\mu'_2)^2 = \langle \xi^4 \rangle / \langle \xi^2 \rangle - 3; \\ k'_6 &= \mu'_6 - 15\mu'_4 \mu'_2 + 30(\mu'_2)^3. \quad (\text{I-25}) \end{aligned}$$

Exemple:  $P\bar{1}$ ;  $k'_4 = -3/2$ ;  $2t = N$  (= nombre d'atomes dans la maille).

On a:

$$\begin{aligned} P(A) &= (\sqrt{2\pi})^{-1} \exp \left[ -\frac{1}{2} A^2 - \frac{1}{8} N^{-1} (A^4 - 6A + 3) + \dots \right]. \quad (\text{I-26}) \end{aligned}$$

#### Ordre indiciel et ordre de grandeur

Si à la place de  $k_3/3!$ ,  $k_4/4!$  etc. on avait introduit  $\varepsilon = t^{-\frac{1}{2}}$  comme variable, on aurait obtenu le même résultat (I-15). Ce sont de tels paramètres d'ordre que nous introduirons dans le cas général (IIe partie).

Il faut cependant observer que la forme de  $Q$  (I-11) ainsi que le degré des polynômes  $\beta_{p_1, p_2, \dots} (q_1, q_2, \dots)$  ne dépend pas des hypothèses faites sur l'ordre de grandeur des cumulants  $k_q/q!$ . Pour éclairer ce point, envisageons le développement de  $P(A)$  en polynômes d'Hermitte (série de Gram-Charlier) donné par l'auteur (Bertaut, 1955b) sous la forme de l'espérance mathématique d'une fonction de Dirac.

$$P(A) = \langle \delta(E - A) \rangle = (\sqrt{2\pi})^{-1} \exp -\frac{1}{2} A^2 \sum_q B_q / q! \quad (\text{I-27})$$

$$B_q = H_q(A) \langle H_q(E) \rangle. \quad (\text{I-28})$$

Nous dirons que dans (I-27) les termes sont ordonnés selon l'ordre *indiciel*.

Ecrivons alors:

$$\sum_q B_q / q! = \exp \left( \sum_q C_q / q! \right). \quad (\text{I-29})$$

La relation entre les  $B_q$  et  $C_q$  est encore exactement celle entre 'moments' ( $B_q$ ) et 'cumulants' ( $C_q$ ).  $E$  étant aléatoire et normalisé (d'où  $B_1 = B_2 = 0$ ), on a:

$$\begin{aligned} C_q &= B_q \quad \text{pour } q=1, 2, 3, 4, 5 \\ C_6/6! &= B_6/6! - (1/2!)(B_3/3!)^2; \\ C_7/7! &= B_7/7! - B_3 B_4 / (3!4!) \quad \text{etc.} \quad (\text{I-30}) \end{aligned}$$

Comme le développement (I-27) en fonctions orthogonales est complet, il doit contenir tous les termes. On doit donc avoir aussi:

$$Q = \sum_q C_q / q! \quad (\text{I-31})$$

c'est-à-dire les suites (I-11) et (I-31) ne peuvent différer que par l'ordre de succession des termes, ce que l'on peut d'ailleurs vérifier directement. (Exemple: on peut identifier  $B_3/3!$  avec  $\beta_3(1)k_3/3! = H_3(A)k_3/3!$  et chercher si (I-31) contient un terme

$$-(1/2!)(H_3(A)k_3/3!)^2$$

existant dans (I-11). Effectivement ce terme existe dans  $C_6/6!$  (I-30).

Klug (avec Edgeworth, 1904) a remarqué judicieusement que l'ordre indiciel ne coïncide pas avec l'ordre de grandeur. Il a ensuite objecté que le développement (I-27) ordonné selon l'ordre indiciel cache toute référence à l'ordre de grandeur des termes. Ceci n'est pas tout à fait correct et nous profitons de l'occasion pour énoncer une règle simple permettant encore d'ordonner les termes de (I-27). Cependant à cause de son intérêt tout académique nous avons relégué cette question dans l'appendice B.

#### Cas général

Le facteur de structure normalisé est:

$$E(h) = \sum_{j=1}^t \varphi_j(h) \xi_j(h) \quad (\text{II-1})$$

on a :

$$\langle \xi_j^2(h) \rangle = p_j(h) n_j. \quad (\text{II-2})$$

$\xi_j(h)$  est le facteur de structure trigonométrique des  $n_j$  atomes équivalents de l'espèce  $j$ ,  $t$  le nombre de positions indépendantes et  $p_j(h)$  le poids statistique\* de  $\xi_j(h)$ . Ecrivons :

$$\varphi_j(h) = \varphi_j / \sqrt{p_j(h)} \quad (\text{II-3})$$

et utilisons pour  $\varphi_j$  l'approximation :

$$\varphi_j = Z_j / \left( \sum_{i=1}^t (n_i Z_i^2) \right)^{\frac{1}{2}} \quad (\text{II-4})$$

où  $Z_j$  est le nombre d'électrons sur l'atome  $j$ . Les relations (II-1) à (II-4) entraînent que :

$$\langle |E(h)|^2 \rangle = 1. \quad (\text{II-5})$$

Considérons  $m$  facteurs de structure

$$E(h_k) = E_k(k=1, 2, \dots, m).$$

La fonction caractéristique  $C$  qui leur est associée est :

$$C(u_1, \dots, u_m) = \left\langle \exp \sum_{k=1}^m v_k E_k \right\rangle \quad (\text{II-6})$$

où :

$$v_k = i u_k. \quad (\text{II-7})$$

Grâce à (II-1) on l'écrira :

$$C = \left\langle \prod_{j=1}^t \exp \sum_{k=1}^m \varphi_j(h_k) v_k \xi_j(h_k) \right\rangle. \quad (\text{II-8})$$

Les atomes  $j=1, \dots, t$  étant supposés indépendants, on a :

$$C = \prod_{j=1}^t C_j \quad (\text{II-9})$$

où :

$$C_j = \left\langle \exp \sum_{k=1}^m \varphi_j(h_k) v_k \xi_j(h_k) \right\rangle. \quad (\text{II-10})$$

Les moments d'ordre  $r=r_1+r_2+\dots+r_m$  des facteurs de structure trigonométriques sont notés conventionnellement

$$m_{r_1 \dots r_m; j} = \langle \xi_j^{r_1}(h) \dots \xi_j^{r_m}(h) \rangle. \quad (\text{II-11})$$

A leur place, nous introduisons les moments réduits  $\mu_{r_1 \dots r_m; j}$  tels que :

$$\mu_{r_1 \dots r_m; j} = m_{r_1 \dots r_m; j} / (p_1^{r_1} \dots p_m^{r_m})^{\frac{1}{2}}. \quad (\text{II-12})$$

Leur introduction présente l'avantage que par la suite on n'aura plus à se préoccuper de poids statistiques. En effet, on pourra écrire :

$$C_j = \sum_{r=0}^{\infty} \sigma_{r; j} / r! \quad (\text{II-13})$$

\* Dans un groupe de translation d'ordre  $\tau$ , le poids statistique est  $\tau p_p$  où  $p_p$  est le poids statistique du groupe primitif correspondant (Bertaut, 1956).

où  $\sigma_{r; j}$  est un polynôme de degré  $r$ .

$$\sigma_{r; j} = \varphi_j^r \sum_{r_k} v_1^{r_1} \dots v_m^{r_m} \mu_{r_1 \dots r_m; j} / (r_1! \dots r_m!). \quad (\text{II-14})$$

Ici la sommation est faite sur toutes les partitions de  $r=r_1+\dots+r_m$ . On remarquera que dans (II-12) et (II-13) ne figure pas  $\varphi_j(h)$  (II-3), mais  $\varphi_j$  (II-4) qui ne dépend plus du poids statistique.

On introduit les cumulants en posant :

$$C_j = \exp \left( \sum_{r=0}^{\infty} \tau_{r; j} / r! \right) \quad (\text{II-15})$$

où  $\tau_{r; j}$  est un polynôme de degré  $r$

$$\tau_{r; j} = \varphi_j^r \sum_{r_k} v_1^{r_1} \dots v_m^{r_m} k_{r_1 \dots r_m; j} / (r_1! \dots r_m!). \quad (\text{II-16})^*$$

Les  $k_{r_1 \dots r_m; j}$  sont les cumulants calculés selon la méthode de l'appendice A en fonction des moments  $\mu_{r_1 \dots r_m; j}$ .

*Choix du paramètre d'ordre dans le cas général.* — On peut remarquer d'abord que tout moment  $\mu_j$  (II-12) est divisible par  $n_j$ . Il en est à fortiori de même des cumulants  $k_j$  (II-16). On pourrait donc mettre en évidence dans  $\tau_{r; j}$  (II-16) un facteur  $n_j \varphi_j^r$ . Cela suggère l'adoption des quantités suivantes comme paramètres d'ordre :

$$z_r = \sum_{j=1}^t n_j \varphi_j^r. \quad (\text{II-17})$$

Posons alors :

$$\Phi_{r; j} = \varphi_j^r / z_r \quad (\text{II-18})$$

et introduisons les abréviations suivantes :

$$\varkappa_{r_1 \dots r_m} = \sum_{j=1}^t \Phi_{r; j} k_{r_1 \dots r_m; j} \quad (\text{II-19})$$

$$T_r = \sum_{r_k} v_1^{r_1} \dots v_m^{r_m} \varkappa_{r_1 \dots r_m} / (r_1! \dots r_m!). \quad (\text{II-20})$$

La fonction caractéristique s'écrit maintenant :

$$C = \exp \sum_r z_r T_r. \quad (\text{II-21})$$

La raison des choix (II-17) et (II-18) est la suivante : lorsque tous les atomes sont égaux et en position générale  $\sum_j \Phi_{r; j}$  se réduit à  $(n^{\frac{1}{2}})^{r-2} = n^{\frac{1}{2}r-1}$  et  $z_r$  (17) à  $(N^{\frac{1}{2}})^{r-2}$  et le produit de ces quantités à  $(t^{\frac{1}{2}})^{r-2}$ .

De même  $T_r$  devient dans ce cas proportionnel au  $\mathcal{A}_r$  de Klug.

En vertu de la normalisation des facteurs de structure on a toujours

$$z_2 T_2 = -\frac{1}{2} \sum_{k=1}^m u_k^2 \quad (\text{II-22})$$

et

$$C = \exp -\frac{1}{2} \sum_{k=1}^m u_k^2 \exp \left( \sum_{r=3}^{\infty} z_r T_r \right). \quad (\text{II-23})$$

\* Ces notations  $\sigma_{r; j}$  et  $\tau_{r; j}$  figurent déjà dans (Bertaut, 1955a : formule (III-22)) au facteur  $(2\pi)^r$  près.

Sous la forme (II-23) la fonction caractéristique  $C(u_1, \dots, u_m)$  et plus précisément  $\log C(u_1, \dots, u_m)$  sont ordonnés.

La densité de répartition  $P(A_1, \dots, A_m)$  des facteurs de structure normalisés observés  $A_k (k=1, \dots, m)$  est la transformée de Fourier de la fonction caractéristique  $C(u_1, \dots, u_m)$  (cf. Bertaut, 1955a; Klug, 1958)

$$P(A_1, \dots, A_m) = \int C(u_1, \dots, u_m) \times \exp -i(A_1 u_1 + \dots + A_m u_m) du_1 \dots du_m. \quad (\text{II-24})$$

Malheureusement la forme ordonnée selon (II-22) n'est pas directement intégrable et l'on doit d'abord développer le deuxième facteur du second membre de (II-23) selon:

$$C(u_1, \dots, u_m) = \exp\left(-\frac{1}{2} \sum_k u_k^2\right) \{1 + z_3 T_3 + (\frac{1}{2} z_3^2 T_3^2 + z_4 T_4) + ((1/3!) z_3^3 T_3^3 + z_3 z_4 T_3 T_4 + z_5 T_5) + \dots\}. \quad (\text{II-25})$$

La série est écrite de façon que des termes de même ordre de grandeur se trouvent dans une même parenthèse. Dans le cas d'atomes tous égaux, on aurait:

$$z_3 = \varepsilon; \quad z_3^2 = z_4 = \varepsilon^2; \quad z_3^3 = z_3 z_4 = z_5 = \varepsilon^3; \quad \text{où } \varepsilon = N^{-\frac{1}{2}}. \quad (\text{II-26})$$

Comme dans la première partie, l'intégration sur  $u_k (k=1, \dots, m)$  peut être remplacée par un procédé purement symbolique (cf. Bertaut, 1955a; Klug, 1958) qui consiste à remplacer

$$\exp -\frac{1}{2} \sum_k u_k^2$$

par

$$(\sqrt{2\pi})^{-m} \exp -\frac{1}{2} \sum_k A_k^2$$

et chaque terme  $v_k^p = (iu_k)^p$  de l'accolade  $\{ \}$  de (II-25) par le polynôme d'Hermite d'ordre  $p$  et d'argument  $A_k$ , soit  $H_p(A_k)$ .

(Exemple: Un terme tel que  $v_1^3 v_2 v_5^2$  devient

$$H_3(A_1) H_1(A_2) H_2(A_5).$$

On obtient ainsi:

$$P(A_1, \dots, A_m) = (\sqrt{2\pi})^{-m} \exp\left(-\frac{1}{2} \sum_k A_k^2\right) \cdot S$$

où

$$S = 1 + z_3 [T_3] + (\frac{1}{2} z_3^2 [T_3^2] + z_4 [T_4]) + ((1/3!) z_3^3 [T_3^3] + z_3 z_4 [T_3 T_4] + z_5 [T_5]) + \dots \quad (\text{II-27})$$

Ici le crochet  $[ ]$  signifie simplement que le procédé symbolique (remplacement de  $v_k^p$  par  $H_p(A_k)$ ) a été effectué. Jusqu'ici le traitement suit étroitement celui donné dans (Bertaut, 1955a) et en ce qui concerne l'ordre des termes, celui de Klug. Remarquons notamment que dans la dernière parenthèse de (II-26) les paramètres  $z_3^2$ ,  $z_3 z_4$  et  $z_5$ , tous trois de l'ordre de  $\varepsilon^3 = N^{-3/2}$ , sont associés à des polynômes de degrés

différents (9, 7 et 5 respectivement) ce qui ne simplifie pas l'analyse de  $P(A_1, \dots, A_m)$ .

*Analyse de  $\log P(A_1, \dots, A_m)$ .* Au lieu d'ordonner la fonction  $P(A_1, \dots, A_m)$ , nous ordonnons

$$\log P(A_1, \dots, A_m).$$

Appliquons en effet, à nouveau la procédure des cumulants à  $S$  (II-26). Cela revient à considérer les paramètres  $z_3, z_4, \dots, z_p$  et leurs puissances comme des variables et leurs coefficients  $[T_3], [T_4], \dots, [T_p]$  et les produits et puissances  $[T_{p_1}^{q_1} T_{p_2}^{q_2} \dots]$  comme l'analogue de moments dont on cherche les cumulants correspondants. En d'autres mots, nous écrivons:

$$S = \exp Q \quad (\text{II-28})$$

avec

$$Q = z_3 K_3(1) + ((1/2!) z_3^2 K_3(2) + z_4 K_4(1)) + ((1/3!) z_3^3 K_3(3) + z_3 z_4 K_{3,4}(1,1) + z_5 K_5(1)) + \dots \quad (\text{II-29})$$

où  $K_p(q)$  est le 'cumulant d'ordre  $q$ ' correspondant au 'moment'  $[T_p^q]$  d'ordre  $q$  et plus généralement

$$K_{p_1, p_2, \dots}(q_1, q_2, \dots)$$

est le cumulant correspondant à  $[T_{p_1}^{q_1} T_{p_2}^{q_2} \dots]$ . On a les mêmes relations (cf. appendice A) qu'entre cumulants

$$\begin{aligned} K_p(1) &= [T_p]; & K_p(2) &= [T_p^2] - [T_p]^2; \\ K_p(3) &= [T_p^3] - 3[T_p^2][T_p] + 2[T_p]^3; \\ K_{p_1, p_2}(1, 1) &= [T_{p_1} T_{p_2}] - [T_{p_1}][T_{p_2}]. \end{aligned} \quad (\text{II-30})$$

L'introduction des nouveaux polynômes  $K_p(q)$  permet de mettre la densité de répartition des facteurs de structure sous la forme

$$P(A_1, \dots, A_m) = (\sqrt{2\pi})^{-m} \exp -\frac{1}{2} \sum_k A_k^2 \exp Q. \quad (\text{II-31})$$

$Q$  étant réel,  $P(A_1, \dots, A_m)$  est maintenant sous une forme *toujours positive* même si le développement de  $Q$  est limité.

Enfin, le principal avantage des polynômes  $K_p(q)$  et  $K_{p_1, p_2, \dots}(q_1, q_2, \dots)$  est que leur degré  $s$  est parfaitement défini par rapport à l'ordre de grandeur des paramètres d'ordre associés. En effet, on peut encore montrer que  $K_p(q)$  est de degré  $s$  (I-16) et que dans le cas d'atomes tous égaux, le paramètre d'ordre associé est (II-32)

$$z_p^q = \varepsilon^{s-2} \quad (\varepsilon = N^{-\frac{1}{2}}) \quad (\text{II-32})$$

avec la même valeur de  $s$ . De même

$$K_{p_1, p_2, \dots}(q_1, q_2, \dots)$$

est de degré  $s$  (I-17) et le paramètre d'ordre associé est:

$$z_{p_1}^{q_1} z_{p_2}^{q_2} \dots = \varepsilon^{s-2}. \quad (\text{II-33})$$

C'est ainsi que les coefficients  $z_5, z_3 z_4$  et  $z_3^3$  (de l'ordre de  $\varepsilon^{s-2} = \varepsilon^3$ ) accompagnent les polynômes  $K_5(1), K_{3,4}(1,1)$  et  $K_3(3)$  qui sont de même degré  $s=5$ .

Aux paramètres d'ordre de même ordre de grandeur  $\varepsilon^{s-2}$  sont donc associés des polynômes de même degré  $s$  en  $A_k$ . La forme de ces polynômes sera explicitée dans les exemples.

*La loi de la tangente hyperbolique.* Notons:

$$L(A_1, \dots, A_m) = \log P(A_1, \dots, A_m). \quad (\text{II-34})$$

Soit à déterminer le signe de  $A_1$ . Décomposons  $L$  en une partie paire  $L_p$  et une partie  $L_i$  impaire en  $A_1$ .

$$L = L_p + L_i. \quad (\text{II-35})$$

La probabilité pour que  $A_1$  soit positif,  $|A_1|$  étant connu, est:

$$P^+ = CP(+|A_1|, \dots, A_m) = C \exp L_p \exp L_i. \quad (\text{II-36})$$

De même la probabilité pour que  $A_1$  soit négatif est:

$$P^- = CP(-|A_1|, \dots, A_m) = C \exp L_p \exp -L_i. \quad (\text{II-37})$$

Eliminant la constante de normalisation  $C$  par la condition  $P^+ + P^- = 1$  on trouve que l'on a rigoureusement:

$$P^+ = \exp L_i / (\exp L_i + \exp -L_i) = \frac{1}{2} + \frac{1}{2} \text{th } L_i \quad (\text{II-38})$$

ce qui montre une fois de plus que c'est le logarithme de  $P(A_1, \dots, A_m)$  qu'il importe d'ordonner. En décomposant  $Q$  (II-29) en une partie paire  $Q_p$  et une partie impaire  $Q_i$

$$Q = Q_p + Q_i \quad (\text{II-39})$$

on voit en comparant (II-31, II-34, II-35 et II-39) que:

$$L_i = Q_i. \quad (\text{II-40})$$

En résumé, on a rigoureusement

$$P^+ = \frac{1}{2} + \frac{1}{2} \text{th } Q_i. \quad (\text{II-41})$$

Notons enfin que l'on peut toujours écrire  $Q_i$  sous la forme (II-42) où  $R$  est de nouveau une fonction paire en  $A_1$ .

$$Q_i = A_1 R \quad (\text{II-42})$$

Cochran & Woolfson (1955) ont indiqué les premiers une formule telle que (II-41) et dans laquelle l'approximation de  $A_1 R$  ne comprenait que le premier terme (d'ordre  $z_3$ ). Ici l'approximation peut-être poussée aussi loin que l'on veut.

Un fait remarquable est que  $P^+$  ne dépende que de la partie impaire de  $Q$ , soit de  $Q_i$ . (Il n'en est pas de même dans la formulation (II-27) de  $P(A_1, \dots, A_m)$ ).

**APPENDICE A**

*Méthode des Cumulants*

Le problème est celui-ci: étant donné une suite de polynômes  $\sigma_q(u_1, \dots, u_m)$  de degré  $q$ , dépendant des

variables  $u_1, \dots, u_m$ , exprimer en fonction des  $\sigma_q$  les polynômes  $\tau_q(u_1, \dots, u_m)$  de degré  $q$  définis par:

$$\sum_q \sigma_q/q! = \exp \sum_q \tau_q/q! \quad (\text{A-1})$$

ou

$$\sum_q \tau_q/q! = \log \sum_q \sigma_q/q!. \quad (\text{A-2})$$

On supposera  $\sigma_0 = 1$ . Habituellement, on considère le cas simple d'une seule variable où:

$$\sigma_q = u^q \mu_q; \quad \tau_q = u^q k_q \quad (\text{A-3})$$

et l'on cherche les relations entre 'moments'  $\mu_q$  et 'cumulants'  $k_q$  (cf. Kendall, 1943) que l'on généralise ensuite pour deux et plusieurs variables. Nous traitons tout de suite le cas le plus général de plusieurs variables en déterminant d'abord les relations entre polynômes  $\sigma_q$  et  $\tau_q$  et ensuite celles entre moments et cumulants. Nous suivons une méthode déjà utilisée (Bertaut, 1955a).

Posons:

$$\sum_{q=0}^{\infty} \sigma_q/q! = 1 + \dots + \sigma_r/r! + \dots = 1 - \varepsilon. \quad (\text{A-4})$$

On a

$$\begin{aligned} \log(1 - \varepsilon) &= -\varepsilon - \varepsilon^2/2 - \dots - \varepsilon^r/r \\ &= \sum_{r=1}^{\infty} (-1)^{r-1} \left( \sum_{q=1}^{\infty} \sigma_q/q! \right)^r / r = \sum_{r=1}^{\infty} (r-1)! (-1)^{r-1} \\ &\quad \times \sum_{r_s} (\sigma_1/1!)^{r_1} \dots (\sigma_s/s!)^{r_s} / (r_1! \dots r_s!). \end{aligned} \quad (\text{A-5})$$

Ici la sommation sur  $r_s$  est faite sur toutes les partitions de  $r = \text{constant}$

$$r = r_1 + \dots + r_s. \quad (\text{A-6})$$

Le degré  $m$  d'un polynôme sous le signe  $\sum_{r_s}$  est:

$$m = 1.r_1 + 2.r_2 + \dots + s.r_s. \quad (\text{A-7})$$

Pour connaître un polynôme  $\tau_m/m!$  il faut collecter dans (A-5) tous les polynômes de même degré  $m$ .

$$\begin{aligned} \tau_m/m! &= \sum_{r=1}^m (-1)^{r-1} (r-1)! \\ &\quad \times \sum_{r_s} (\sigma_1/1!)^{r_1} \dots (\sigma_s/s!)^{r_s} / (r_1! \dots r_s!). \end{aligned} \quad (\text{A-8})$$

Exemple: Déterminer  $\tau_5/5!$ . L'équation (A-7) se réduit à:

$$5 = r_1 + 2r_2 + 3r_3 + 4r_4 + 5r_5. \quad (\text{A-9})$$

On commence par l'indice le plus élevé ( $s=5$ ) en posant  $r_5 = 1$ . (A-6) fournit  $r = 1$  et (A-8) fournit une contribution  $\sigma_5/5!$ . On pose ensuite  $r_5 = 0; r_4 = 1$ . La seule solution de (A-9) est  $r_1 = 1$  et de (A-6)  $r = 2$ , d'où d'après (A-8) une contribution  $-\sigma_4 \sigma_1 / (4!1!)$ . On fait ensuite  $r_5 = r_4 = 0; r_3 = 1$  etc. On a finalement:

$$\begin{aligned} \tau_5/5! &= \sigma_5/5! - \sigma_4 \sigma_1 / (4!1!) - \sigma_3 \sigma_2 / (3!2!) + 2 \sigma_3 \sigma_1^2 / 3! \\ &\quad - \sigma_2 \sigma_1^2 / 2! + (\sigma_2/2!)^2 \sigma_1 + \sigma_1^5 / 5. \end{aligned} \quad (\text{A-10})$$

On trouve de même:

$$\begin{aligned}\tau_1 &= \sigma_1 \\ \tau_2/2! &= \sigma_2/2! - \frac{1}{2}\sigma_1^2 \\ \tau_3/3! &= \sigma_3/3! - \sigma_2\sigma_1/2! + 2!\sigma_1^3/3! \\ \tau_4/4! &= \sigma_4/4! - \sigma_3\sigma_1/3! + \sigma_1^2\sigma_2/2! - \sigma_1^4/4! - (1/2!)(\sigma_2/2!)^2.\end{aligned}\quad (\text{A-11})$$

Cas d'une seule variable

On n'a qu'à remplacer  $\tau_q$  par  $k_q$  et  $\sigma_q$  par  $\mu_q$  pour retrouver les relations classiques entre cumulants et moments.

$$k_1 = \mu_1; \quad k_2 = \mu_2 - \mu_1^2; \quad k_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3. \quad (\text{A-12})$$

En somme les relations entre moments  $\mu_q$  et cumulants  $k_q$  sont les mêmes qu'entre polynômes  $\sigma_q$  et  $\tau_q$  de degré  $q$ . Ces relations sont tabulées (cf. Kendall, 1943). Elles peuvent donc servir dans le

Cas de plusieurs variables

Soit à calculer  $k_{311}$  à l'aide de (A-10). On a:

$$\tau_5/5! = \sum_{r_s} k_{r_1 \dots r_s} u_1^{r_1} \dots u_s^{r_s} / (r_1! \dots r_s!) \quad (\text{A-13})$$

$$\sigma_4/4! = \sum_{r_s} \mu_{r_1 \dots r_s} u_1^{r_1} \dots u_s^{r_s} / (r_1! \dots r_s!) \text{ etc.} \quad (\text{A-14})$$

Le coefficient de  $u_1^3 u_2^1 u_3^1$  dans (A-13), est  $k_{311}/(3!1!1!)$ . On cherchera de même le coefficient de  $u_1^3 u_2^1 u_3^1$  dans le second membre de (A-10). Dans  $\sigma_5/5!$  ce coefficient est  $\mu_{311}/(3!1!1!)$ . Dans  $-\sigma_4\sigma_1/(4!1!)$ , il est d'après (A-14):

$$-(\mu_{211}\mu_{100}/2! + \mu_{301}\mu_{010}/3! + \mu_{310}\mu_{001}/3!).$$

On a finalement:

$$\begin{aligned}k_{311}/3! &= \mu_{311}/3! - (\mu_{211}\mu_{100}/2! + \mu_{301}\mu_{010}/3! + \mu_{310}\mu_{001}/3!) \\ &\quad - (\mu_{111}\mu_{200} + \mu_{201}\mu_{110} + \mu_{210}\mu_{101})/2! + 2\mu_{111}\mu_{100}^2 \\ &\quad + 4\mu_{300}\mu_{010}\mu_{001}/3! - \mu_{011}\mu_{100}^3 + 2\mu_{110}\mu_{101}\mu_{100} \\ &\quad + 4\mu_{100}^3\mu_{010}\mu_{001}.\end{aligned}\quad (\text{A-15})$$

Selon les hypothèses de ce mémoire, les moments à indices mixtes, sont nuls et (A-15) se réduit à:

$$k_{311}/3! = \mu_{311}/3! - \mu_{111}\mu_{200}/2!.$$

## APPENDICE B

Considérons l'exemple cité par Klug (1958; rel. (3.4)) de la densité de répartition de deux facteurs de structure  $E_1 = E(2h)$ ,  $E_2 = E(h)$  que nous écrivons tout de suite sous la forme ordonnée:

$$\begin{aligned}P(A_1, A_2) &= (2\pi)^{-1} \exp(-\frac{1}{2}A_1^2 - \frac{1}{2}A_2^2) \sum_{pq} H_p(A_1)H_q(A_2) \\ &\quad \langle H_p(E_1)H_q(E_2) \rangle / (p!q!) = (2\pi)^{-1} \exp(-\frac{1}{2}A_1^2 - \frac{1}{2}A_2^2) \cdot \\ &\quad \{1 + H_1(A_1)H_2(A_2)\langle H_1(E_1)H_2(E_2) \rangle / 2! \\ &\quad + H_4(A_1)\langle H_4(E_1) \rangle / 4! + H_4(A_2)\langle H_4(E_2) \rangle / 4! \\ &\quad + H_2(A_1)H_4(A_2)\langle H_2(E_1)H_4(E_2) \rangle / (2!4!) \\ &\quad + H_3(A_1)H_2(A_2)\langle H_1(E_1)H_2(E_2) \rangle / (3!2!) \\ &\quad + H_1(A_1)H_4(A_2)\langle H_1(E_1)H_4(E_2) \rangle / (4!) \\ &\quad + H_5(A_1)H_2(A_2)\langle H_5(E_1)H_2(E_2) \rangle / (5!2!) \\ &\quad + H_1(A_1)H_6(A_2)\langle H_1(E_1)H_6(E_2) \rangle / 6! \\ &\quad + H_3(A_1)H_6(A_2)\langle H_3(E_1)H_6(E_2) \rangle / (3!6!) \}.\end{aligned}\quad (\text{B-1})$$

Dans l'accolade  $\{ \}$  nous distinguerons deux sortes de moyennes. Une moyenne  $\langle H_p(E_1)H_q(E_2) \rangle$  d'ordre indiciel  $p+q$  est dite 'fondamentale' si elle est de l'ordre de  $\varepsilon^{p+q-2}$  ( $\varepsilon = N^{-\frac{1}{2}}$ ). Par exemple

$$\begin{aligned}\langle H_1(E_1)H_2(E_2) \rangle &= 0(\varepsilon); \quad \langle H_4(E) \rangle = 0(\varepsilon^2); \\ \langle H_3(E_1)H_2(E_2) \rangle &= 0(\varepsilon^3).\end{aligned}$$

Une moyenne  $\langle H_p(E_1)H_q(E_2) \rangle$  est dite 'composée' si le couple d'indices  $(p, q)$  admet une partition

$$(p, q) = (p_1, q_1) + (p_2, q_2) + \dots \quad (\text{B-2})$$

telle que  $\langle H_{p_1}(E_1)H_{q_1}(E_2) \rangle, \langle H_{p_2}(E_1)H_{q_2}(E_2) \rangle, \dots$  sont des moyennes 'fondamentales'. L'ordre de grandeur d'une moyenne 'composée' est  $0(\varepsilon^{p_1+q_1-2} \cdot \varepsilon^{p_2+q_2-2} \dots)$ .

Par exemple  $(2,4) = (1,2) + (1,2)$ ;  $(5,2) = (1,2) + (4,0)$ ;  $(1,6) = (1,2) + (0,4)$  de sorte que  $\langle H_2(E_1)H_4(E_2) \rangle = 0(\varepsilon^2)$ ;  $\langle H_5(E_1)H_2(E_2) \rangle = 0(\varepsilon^3)$ ;  $\langle H_1(E_1)H_6(E_2) \rangle = 0(\varepsilon^3)$ .

Cette règle simple peut encore se généraliser au cas de plusieurs variables  $A_k (k=1, \dots, m)$ . Cependant si d'après la règle énoncée il est encore possible de ranger les termes de la série de Charlier-Gram selon leur ordre de grandeur (contrairement à l'objection de Klug) nous croyons être parfaitement d'accord avec Klug en disant que le calcul des cumulants est plus facile à mener que celui des moyennes  $\langle H_p(E_1)H_q(E_2) \rangle$  pour des indices  $p, q$  élevés.

## Références

- BERTAUT, E. F. (1955a). *Acta Cryst.* **8**, 537.  
 BERTAUT, E. F. (1955b). *Acta Cryst.* **8**, 823.  
 BERTAUT, E. F. (1956). *Acta Cryst.* **9**, 322.  
 COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1.  
 EDGEWORTH, F. Y. (1904). *Camb. Phil. Trans.* **20**, 36, 113.  
 KENDALL, M. G. (1943). *The Advanced Theory of Statistics*. London: Griffin.  
 KLUG, A. (1958). *Acta Cryst.* **11**, 515.