notorious for its rather indeterminate mean and 'infinite' variance, resulting from the 'tail' of the denominator distributions extending through zero to negative values. The leading terms of the ratio distribution are given by Kendall & Stuart (1977, p. 288).

2.1.7. Non-ideal distributions: the correction-factor approach

2.1.7.1. *Introduction*

The probability density functions (p.d.f.'s) of the magnitude of the structure factor, presented in Section 2.1.5, are based on the central-limit theorem discussed above. In particular, the centric and acentric p.d.f.'s given by equations (2.1.5.11) and (2.1.5.8), respectively, are expected to account for the statistical properties of diffraction patterns obtained from crystals consisting of nearly equal atoms, which obey the fundamental assumptions of uniformity and independence of the atomic contributions and are not affected by noncrystallographic symmetry and dispersion. It is also assumed there that the number of atoms in the asymmetric unit is large. Distributions of structure-factor magnitudes which are based on the central-limit theorem, and thus obey the above assumptions, have been termed 'ideal', and the subjects of the following sections are those distributions for which some of the above assumptions/restrictions are not fulfilled; the latter distributions will be called 'non-ideal'.

We recall that the assumption of uniformity consists of the requirement that the fractional part of the scalar product $hx + ky + lz$ be uniformly distributed over the [0, 1] interval, which holds well if x, y, z are rationally independent (Hauptman & Karle, 1953), and permits one to regard the atomic contribution to the structure factor as a random variable. This is of course a necessary requirement for any statistical treatment. If, however, the atomic composition of the asymmetric unit is widely heterogeneous, the structure factor is then a sum of unequally distributed random variables and the Lindeberg–Lévy version of the central-limit theorem (*cf.* Section 2.1.4.4) cannot be expected to apply. Other versions of this theorem might still predict a normal p.d.f. of the sum, but at the expense of a correspondingly large number of terms/atoms. It is well known that atomic heterogeneity gives rise to severe deviations from ideal behaviour (*e.g.* Howells *et al.*, 1950) and one of the aims of crystallographic statistics has been the introduction of a correct dependence on the atomic composition into the non-ideal p.d.f.'s [for a review of the early work on non-ideal distributions see Srinivasan & Parthasarathy (1976)]. A somewhat less well known fact is that the dependence of the p.d.f.'s of $|E|$ on space-group symmetry becomes more conspicuous as the composition becomes more heterogeneous (*e.g.* Shmueli, 1979; Shmueli & Wilson, 1981). Hence both the composition and the symmetry dependence of the intensity statistics are of interest. Other problems, which likewise give rise to non-ideal p.d.f.'s, are the presence of heavy atoms in (variable) special positions, heterogeneous structures with complete or partial noncrystallographic symmetry, and the presence of outstandingly heavy dispersive scatterers.

The need for theoretical representations of non-ideal p.d.f.'s is exemplified in Fig. 2.1.7.1(*a*), which shows the ideal centric and acentric p.d.f.'s together with a frequency histogram of $|E|$ values, recalculated for a centrosymmetric structure containing a platinum atom in the asymmetric unit of $P\bar{1}$ (Faggiani *et al.*, 1980). Clearly, the deviation from the Gaussian p.d.f., predicted by the central-limit theorem, is here very large and a comparison with the possible ideal distributions can (in this case) lead to wrong conclusions.

Two general approaches have so far been employed in derivations of non-ideal p.d.f.'s which account for the above-mentioned problems: the correction-factor approach, to be dealt with in the following sections, and the more recently introduced Fourier method, to which Section 2.1.8 is dedicated. In what follows, we introduce briefly the mathematical background of the correction-factor approach, apply this formalism to centric and acentric non-ideal p.d.f.'s, and present the numerical values of the moments of the trigonometric structure factor which permit an approximate evaluation of such p.d.f.'s for all the three-dimensional space groups.

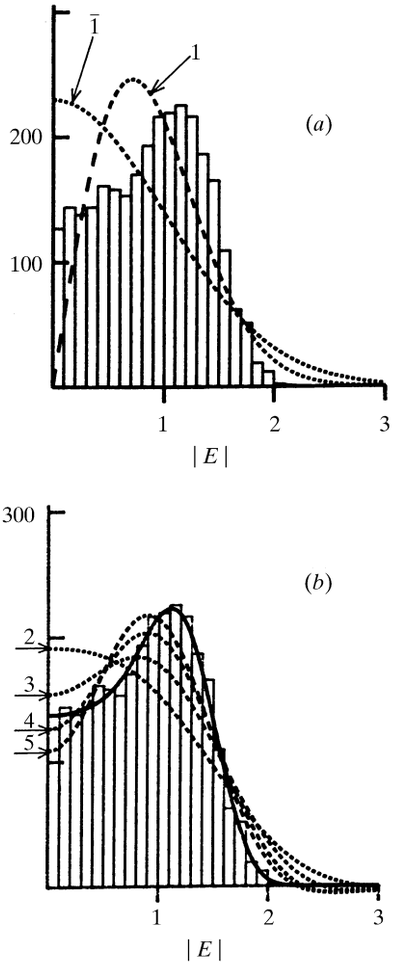

Fig. 2.1.7.1. Atomic heterogeneity and intensity statistics. The histogram appearing in (*a*) and (*b*) was constructed from $|E|$ values which were recalculated from atomic parameters published for the centrosymmetric structure of $C_6H_{18}Cl_2N_4O_4Pt$ (Faggiani *et al.*, 1980). The space group of the crystal is $P\bar{1}$, $Z = 2$, *i.e.* all the atoms are located in general positions. (*a*) A comparison of the recalculated distribution of $|E|$ with the ideal centric [equation (2.1.5.11)] and acentric [equation (2.1.5.8)] p.d.f.'s, denoted by $\bar{1}$ and 1, respectively. (*b*) The same recalculated histogram along with the centric correction-factor p.d.f. [equation (2.1.7.5)], truncated after two, three, four and five terms (dashed lines), and with that accurately computed for the correct space-group Fourier p.d.f. [equations (2.1.8.5) and (2.1.8.22)] (solid line).

2.1.7.2. *Mathematical background*

Suppose that $p(x)$ is a p.d.f. which accurately describes the experimental distribution of the random variable x, where x is related to a sum of random variables and can be assumed to obey (to some approximation) an ideal p.d.f., say $p^{(0)}(x)$, based on the central-limit theorem. In the correction-factor approach we seek to represent $p(x)$ as

$$p(x) = p^{(0)}(x) \sum_k d_k f_k(x), \qquad (2.1.7.1)$$

where d_k are coefficients which depend on the cause of the deviation of $p(x)$ from the central-limit theorem approximation and $f_k(x)$ are suitably chosen functions of x. A choice of the set $\{f_k\}$ is deemed suitable, if only from a practical point of view, if it allows the convenient introduction of the cause of the above deviation of $p(x)$ into the expansion coefficients d_k. This requirement is satisfied – also from a theoretical point of view – by taking $f_k(x)$ as a set of polynomials which are orthogonal with respect to the ideal p.d.f., taken as their weight function (*e.g.* Cramér, 1951). That is, the functions $f_k(x)$ so chosen have to obey the relationship

$$\int_a^b f_k(x) f_m(x) p^{(0)}(x)\, \mathrm{d}x = \delta_{km} = \begin{cases} 1, & \text{if} \quad 2k = m \\ 0, & \text{if} \quad 2k \neq m \end{cases}, \qquad (2.1.7.2)$$

where $[a, b]$ is the range of existence of all the functions involved. It can be readily shown that the coefficients d_k are given by

$$d_k = \int_a^b f_k(x) p(x)\, \mathrm{d}x = \langle f_k(x) \rangle = \sum_{n=0}^{k} c_n^{(k)} \langle x^n \rangle, \qquad (2.1.7.3)$$

where the brackets $\langle\ \rangle$ in equation (2.1.7.3) denote averaging with respect to the *unknown* p.d.f. $p(x)$ and $c_n^{(k)}$ is the coefficient of the nth power of x in the polynomial $f_k(x)$. The coefficients d_k are thus directly related to the moments of the non-ideal distribution and the coefficients of the powers of x in the orthogonal polynomials. The latter coefficients can be obtained by the Gram–Schmidt procedure (*e.g.* Spiegel, 1974), or by direct use of the Szegö determinants (*e.g.* Cramér, 1951), for any weight function that has finite moments. However, the feasibility of the present approach depends on our ability to obtain the moments $\langle x^n \rangle$ *without* the knowledge of the non-ideal p.d.f., $p(x)$.

2.1.7.3. *Application to centric and acentric distributions*

We shall summarize here the non-ideal centric and acentric distributions of the magnitude of the normalized structure factor E (*e.g.* Shmueli & Wilson, 1981; Shmueli, 1982). We assume that (i) all the atoms are located in general positions and have rationally independent coordinates, (ii) all the scatterers are dispersionless, and (iii) there is no noncrystallographic symmetry. Arbitrary atomic composition and space-group symmetry are admitted. The appropriate weight functions and the corresonding orthogonal polynomials are

$p^{(0)}(\lvert E\rvert)$	$f_k(x)$	Non-ideal distribution
$(2/\pi)^{1/2} \exp(-\lvert E\rvert^2/2)$	$He_{2k}(\lvert E\rvert)/[(2k)!]^{1/2}$	Centric
$2\lvert E\rvert \exp(-\lvert E\rvert^2)$	$L_k(\lvert E\rvert^2)$	Acentric

(2.1.7.4)

where He_k and L_k are Hermite and Laguerre polynomials, respectively, as defined, for example, by Abramowitz & Stegun (1972). Equations (2.1.7.2), (2.1.7.3) and (2.1.7.4) suffice for the general formulation of the above non-ideal p.d.f.'s of $|E|$. Their full derivation entails (i) the expression of a sufficient number of moments of $|E|$ in terms of absolute moments of the trigonometric structure factor (*e.g.* Shmueli & Wilson, 1981; Shmueli, 1982) and (ii) calculation of the latter moments for the various symmetries (Wilson, 1978*b*; Shmueli & Kaldor, 1981, 1983). The notation below is similar to that employed by Shmueli (1982).

These non-ideal p.d.f.'s of $|E|$, for which the first five expansion terms are available, are given by

$$p_c(|E|) = p_c^{(0)}(|E|) \left[1 + \sum_{k=2}^{\infty} \frac{A_{2k}}{(2k)!} He_{2k}(|E|) \right] \qquad (2.1.7.5)$$

and

$$p_a(|E|) = p_a^{(0)}(|E|) \left[1 + \sum_{k=2}^{\infty} \frac{(-1)^k B_{2k}}{k!} L_k(|E|^2) \right] \qquad (2.1.7.6)$$

for centrosymmetric and noncentrosymmetric space groups, respectively, where $p_c^{(0)}(|E|)$ and $p_a^{(0)}(|E|)$ are the ideal centric and acentric p.d.f.'s [see (2.1.7.4)] and the unified form of the coefficients A_{2k} and B_{2k}, for $k = 2$, 3, 4 and 5, is

$$\begin{aligned} A_4 \quad &\text{or} \quad B_4 = a_4 Q_4 \\ A_6 \quad &\text{or} \quad B_6 = a_6 Q_6 \\ A_8 \quad &\text{or} \quad B_8 = a_8 Q_8 + U(a_4^2 Q_4^2 - \gamma_4^2) \\ A_{10} \quad &\text{or} \quad B_{10} = a_{10} Q_{10} + V(a_4 a_6 Q_4 Q_6 - \gamma_4 \gamma_6 Q_{10}) \\ &\qquad + W \gamma_4^2 Q_{10} \end{aligned} \qquad (2.1.7.7)$$

(Shmueli, 1982), where $U = 35$ or 18, $V = 210$ or 100 and $W = 3150$ or 900 according as A_{2k} or B_{2k} is required, respectively, and the other quantities in equation (2.1.7.7) are given below. The *composition-dependent* terms in equations (2.1.7.7) are

$$Q_{2k} = \frac{\sum_{j=1}^{m} f_j^{2k}}{\left(\sum_{n=1}^{m} f_n^2\right)^k}, \qquad (2.1.7.8)$$

where m is the number of atoms in the asymmetric unit, f_j, $j = 1, \ldots, m$ are their scattering factors, and the *symmetry* dependence is expressed by the coefficients a_{2k} in equation (2.1.7.7), as follows:

$$a_{2k} = (-1)^{k-1}(k-1)!\alpha_{k0} + \sum_{p=2}^{k} (-1)^{k-p}(k-p)!\alpha_{kp}\gamma_{2p}, \qquad (2.1.7.9)$$

where

$$\alpha_{kp} = \binom{k}{p} \frac{(2k-1)!!}{(2p-1)!!} \quad \text{or} \quad \binom{k}{p} \frac{k!}{p!} \qquad (2.1.7.10)$$

according as the space group is centrosymmetric or noncentrosymmetric, respectively, and γ_{2p} in equation (2.1.7.9) is given by

$$\gamma_{2p} = \frac{\langle |T|^{2p} \rangle}{\langle |T|^2 \rangle^p}, \qquad (2.1.7.11)$$

where $\langle |T|^k \rangle$ is the kth absolute moment of the trigonometric structure factor

$$T(\mathbf{h}) = \sum_{s=1}^{g} \exp[2\pi i \mathbf{h}^T (\mathbf{P}_s \mathbf{r} + \mathbf{t}_s)] \equiv \xi(\mathbf{h}) + i\eta(\mathbf{h}). \qquad (2.1.7.12)$$

In equation (2.1.7.12), g is the number of general equivalent positions listed in *IT* A (1983) for the space group in question, times the multiplicity of the Bravais lattice, $(\mathbf{P}_s, \mathbf{t}_s)$ is the sth space-group operator and $\mathbf{r}$ is an atomic position vector.

Table 2.1.7.1. *Some even absolute moments of the trigonometric structure factor*

The symbols p, q, r and s denote the second, fourth, sixth and eighth absolute moments of the trigonometric structure factor T [equation (2.1.7.12)], respectively, and the columns of the table contain (for some conciseness) $p, q, r/p$ and s/p^2. The numbers in parentheses, appearing beside some space-group entries, refer to *hkl* subsets which are defined in the note at the end of the table. These subset references are identical with those given by Shmueli & Kaldor (1981, 1983). The symbols q, r and s are also equivalent to $\gamma_4 P^2$, $\gamma_6 P^3$ and $\gamma_8 P^4$, respectively, where γ_{2n} are the normalized absolute moments given by equation (2.1.7.11).

Space groups(s)	p	q	r/p	s/p^2
Point group: 1				
$P1$	1	1	1	1
Point group: $\bar{1}$				
$P\bar{1}$	2	6	10	17
Point groups: 2, *m*				
All P	2	6	10	17
All C	4	48	160	560
Point group: 2/*m*				
All P	4	36	100	$306\frac{1}{4}$
All C	8	288	1600	9800
Point group: 222				
All P	4	28	64	$169\frac{3}{4}$
All C and I	8	224	1024	5432
$F222$	16	1792	16384	173824
Point group: *mm*2				
All P	4	36	100	$306\frac{1}{4}$
All A, C and I	8	288	1600	9800
$Fmm2$	16	2304	25600	313600
$Fdd2$ (1)	16	2304	25600	313600
$Fdd2$ (2)	16	1280	7168	43264
Point group: *mmm*				
All P	8	216	1000	$5359\frac{3}{8}$
All C and I	16	1728	16000	171500
$Fmmm$	32	13824	256000	5488000
$Fddd$ (1)	32	13824	256000	5488000
$Fddd$ (2)	32	7680	71680	757120
Point group: 4				
$P4, P4_2$	4	36	100	$306\frac{1}{4}$
$P4_1$* (3)	4	36	100	$306\frac{1}{4}$
$P4_1$* (4)	4	20	28	$42\frac{1}{4}$
$I4$	8	288	1600	9800
$I4_1$ (5)	8	288	1600	9800
$I4_1$ (6)	8	160	448	1352
Point group: $\bar{4}$				
$P\bar{4}$	4	28	64	$169\frac{3}{4}$
$I\bar{4}$	8	224	1024	5432
Point group: 4/*m*				
All P	8	216	1000	$5359\frac{3}{8}$
$I4/m$	16	1728	16000	171500
$I4_1/a$ (7)	16	1728	16000	171500
$I4_1/a$ (8)	16	960	4480	23660
Point group: 422				
$P422, P42_12, P4_222, P4_22_12$	8	136	424	$1682\frac{1}{8}$
$P122$,* $P4_12_12$* (3)	8	136	424	$1682\frac{1}{8}$
$P4_122$,* $P4_12_12$* (4)	8	104	208	$470\frac{1}{8}$
$I422$	16	1088	6784	53828
$I4_122$ (7)	16	1088	6784	53828
$I4_122$ (8)	16	832	3328	15044

Space groups(s)	p	q	r/p	s/p^2
Point group: 4*mm*				
All P	8	168	640	$2970\frac{5}{5}$
$I4mm$, $I4cm$	16	1344	10240	95060
$I4_1md, I4_1cd$ (7)	16	1344	10240	95060
$I4_1md, I4_1cd$ (8)	16	832	3328	15188
Point groups: $\bar{4}2m, \bar{4}m2$				
All P	8	136	424	$1682\frac{1}{8}$
$I\bar{4}m2, I\bar{4}2m, I\bar{4}c2$	16	1088	6784	53828
$I\bar{4}2d$ (5)	16	1088	6784	53828
$I\bar{4}2d$ (6)	16	832	3328	15044
Point group: 4/*mmm*				
All P	16	1008	6400	$51985\frac{15}{16}$
$I4/mmm, I4/mcm$	32	8064	102400	1663550
$I4_1/amd, I4_1/acd$ (5)	32	8064	102400	1663550
$I4_1/amd, I4_1/acd$ (6)	32	4992	33280	265790
Point group: 3				
All P and R	3	15	31	71
Point group: $\bar{3}$				
All P and R	6	90	310	1242
Point group: 32				
All P and R	6	66	166	508
Point group: 3*m*				
$P3m1, P31m, R3m$	6	66	178	604
$P3c1, P31c$, (3); $R3c$ (1)	6	66	178	604
$P3c1, P31c$, (4); $R3c$ (2)	6	66	154	412
Point group: $\bar{3}m$				
$P\bar{3}1m, P\bar{3}m1, R\bar{3}m$	12	396	1780	$10578\frac{3}{4}$
$P\bar{3}1c, P\bar{3}c1$ (3); $R\bar{3}c$ (1)	12	396	1780	$10578\frac{3}{4}$
$P\bar{3}1c, P\bar{3}c1$ (4); $R\bar{3}c$ (2)	12	396	1540	$7218\frac{3}{4}$
Point group: 6				
$P6$	6	90	340	1522
$P6_1$* (9)	6	90	340	1522
$P6_1$* (10)	6	54	91	161
$P6_1$* (11)	6	54	97	193
$P6_1$* (12)	6	90	280	962
$P6_2$* (13)	6	90	340	1522
$P6_2$* (14)	6	54	97	193
$P6_3$ (3)	6	90	340	1522
$P6_3$ (4)	6	90	280	962
Point group: $\bar{6}$				
$P\bar{6}$	6	90	310	1242
Point group: 6/*m*				
$P6/m$	12	540	3400	$26643\frac{3}{4}$
$P6_3/m$ (3)	12	540	3400	$26643\frac{3}{4}$

Table 2.1.7.1. *Some even absolute moments of the trigonometric structure factor* (*cont.*)

Space groups(s)	p	q	r/p	s/p^2
$P6_3/m$ (4)	12	540	2800	$16843\frac{3}{4}$
Point group: 622				
$P622$	12	324	1150	$5506\frac{1}{4}$
$P6_122$* (9)	12	324	1150	$5506\frac{1}{4}$
$P6_122$* (10)	12	252	577	$1537\frac{3}{4}$
$P6_122$* (11)	12	252	583	$1601\frac{3}{4}$
$P6_122$* (12)	12	324	1090	$4746\frac{1}{4}$
$P6_222$* (13)	12	324	1150	$5506\frac{1}{4}$
$P6_222$* (14)	12	252	583	$1601\frac{3}{4}$
$P6_322$ (3)	12	324	1150	$5506\frac{1}{4}$
$P6_322$ (4)	12	324	1090	$4746\frac{1}{4}$
Point group: 6*mm*				
$P6mm$	12	396	1930	$12818\frac{3}{4}$
$P6cc$ (3)	12	396	1930	$12818\frac{3}{4}$
$P6cc$ (4)	12	396	1450	$6098\frac{3}{4}$
$P6_3cm, P6_3mc$ (3)	12	396	1930	$12818\frac{3}{4}$
$P6_3cm, P6_3mc$ (4)	12	396	1630	$8338\frac{3}{4}$
Point groups: $\bar{6}m2, \bar{6}2m$				
$P\bar{6}m2, P\bar{6}2m$	12	396	1780	$10578\frac{3}{4}$
$P\bar{6}c2, P\bar{6}2c$ (3)	12	396	1780	$10578\frac{3}{4}$
$P\bar{6}c2, P\bar{6}2c$ (4)	12	396	1540	$7218\frac{3}{4}$
Point group: 6/*mmm*				
$P6/mmm$	24	2376	19300	$224328\frac{1}{8}$
$P6/mcc$ (3)	24	2376	19300	$224328\frac{1}{8}$
$P6/mcc$ (4)	24	2376	14500	$106728\frac{1}{8}$
$P6/mcm$, $P6/mmc$ (3)	24	2376	19300	$224328\frac{1}{8}$
$P6/mcm$, $P6/mmc$ (4)	24	2376	16300	$145928\frac{1}{8}$
Point group: 23				
$P23$, $P2_13$	12	276	760	$2695\frac{1}{4}$
$I23$, $I2_13$	24	2208	12160	86248
$F23$	48	17664	194560	2759936
Point group: $m\bar{3}$				
$Pm\bar{3}, Pn\bar{3}, Pa3$	24	1800	9400	$67703\frac{1}{8}$
$Im\bar{3}, Ia\bar{3}$	48	14400	150400	2166500
$Fm\bar{3}$	96	115200	2406400	69328000
$Fd\bar{3}$ (1)	96	115200	2406400	69328000
$Fd\bar{3}$ (2)	96	96768	1484800	28183680

Space groups(s)	p	q	r/p	s/p^2
Point group: 432				
$P432, P4_232$	24	1272	4648	$25216\frac{7}{8}$
$P4_132$* (15)	24	1272	4648	$25216\frac{7}{8}$
$P4_132$* (16)	24	1176	3568	$13916\frac{7}{8}$
$P4_132$* (17)	24	1080	2776	$8664\frac{7}{8}$
$P4_132$* (18)	24	984	2272	$6580\frac{7}{8}$
$I432$	48	10176	74368	806940
$I4_132$ (15)	48	10176	74368	806940
$I4_132$ (17)	48	8640	44416	277276
$F432$	96	81408	1189888	25822080
$F4_132$ (15)	96	81408	1189888	25822080
$F4_132$ (18)	96	62976	581632	6738816
Point group: $\bar{4}3m$				
$P\bar{4}3m$	24	1272	5128	$32896\frac{7}{8}$
$P\bar{4}3n$ (1)	24	1272	5128	$32896\frac{7}{8}$
$P\bar{4}3n$ (2)	24	1272	4168	$17536\frac{7}{8}$
$I\bar{4}3m$	48	10176	82048	1052700
$I\bar{4}3d$ (15); (20)	48	10176	82048	1052700
$I\bar{4}3d$ (15); (21)	48	10176	66688	561180
$I\bar{4}3d$ (17)	48	8640	44416	277276
$F\bar{4}3m$	96	81408	1312768	33686400
$F\bar{4}3c$ (15)	96	81408	1312768	33686400
$F\bar{4}3c$ (18)	96	81408	1067008	17957760
Point group: $m\bar{3}m$				
$Pm\bar{3}m, Pn\bar{3}m$	48	8784	72160	$972717\frac{13}{16}$
$Pn\bar{3}n, Pm\bar{3}n$ (1)	48	8784	72160	$972717\frac{13}{16}$
$Pn\bar{3}n, Pm\bar{3}n$ (2)	48	8784	56800	$488877\frac{13}{16}$
$Im\bar{3}m$	96	70272	1154560	31126970
$Ia\bar{3}d$ (15); (20)	96	70272	1154560	31126970
$Ia\bar{3}d$ (15); (21)	96	51840	432640	4497850
$Ia\bar{3}d$ (17)	96	70272	908800	15644090
$Fm\bar{3}m$	192	562176	18472960	996063040
$Fm\bar{3}c$ (1)	192	562176	18472960	996063040
$Fm\bar{3}c$ (2)	192	562176	14540800	500610880
$Fd\bar{3}m$ (1)	192	562176	18472960	996063040
$Fd\bar{3}m$ (2)	192	414720	7782400	205432640
$Fd\bar{3}c$ (1)	192	562176	18472960	996063040
$Fd\bar{3}c$ (2)	192	414720	6799360	136619840

Note. hkl subsets: (1) $h+k+l=2n$; (2) $h+k+l=2n+1$; (3) $l=2n$; (4) $l=2n+1$; (5) $2h+l=2n$; (6) $2h+l=2n+1$; (7) $2k+l=2n$; (8) $2k+l=2n+1$; (9) $l=6n$; (10) $l=6n+1, 6n+5$; (11) $l=6n+2, 6n+4$; (12) $l=6n+3$; (13) $l=3n$; (14) $l=3n+1, 3n+2$; (15) hkl all even; (16) only one index odd; (17) only one index even; (18) hkl all odd; (19) two indices odd; (20) $h+k+l=4n$; (21) $h+k+l=4n+2$.
* And the enantiomorphous space group.

The cumulative distribution functions, obtained by integrating equations (2.1.7.5) and (2.1.7.6), are given by

$$N_c(|E|) = \operatorname{erf}\left(\frac{|E|}{\sqrt{2}}\right) - \frac{2}{\sqrt{\pi}}\exp\left(-\frac{|E|^2}{2}\right) \times \left[\sum_{k=2}^{\infty}\frac{A_{2k}}{(2k)!}He_{2k-1}(|E|)\right] \quad (2.1.7.13)$$

and

$$N_a(|E|) = 1 - \exp(-|E|^2) + \exp(-|E|^2) \times \left\{\sum_{k=2}^{\infty}\frac{(-1)^k B_{2k}}{k!}[L_{k-1}(|E|^2) - L_k(|E|^2)]\right\} \quad (2.1.7.14)$$

for centrosymmetric and noncentrosymmetric space groups, respectively, where the coefficients are defined in equations (2.1.7.7)–(2.1.7.12). Note that the first term on the right-hand side

Table 2.1.7.2. *Closed expressions for γ_{2k} [equation (2.1.7.11)] for space groups of low symmetry*

The normalized moments γ_{2k} are expressed in terms of M_k, where

$$M_k = \frac{(2k)!}{2^k(k!)^2} = \frac{(2k-1)!!}{k!},$$

and l', which takes on the values 1, 2 or 4 according as the Bravais lattice is of type P, one of the types A, B, C or I, or type F, respectively. The expressions for γ_{2k} are identical for all the space groups based on a given point group, except *Fdd*2 and *Fddd*. The expressions are valid for general reflections and under the restrictions given in the text.

Point group(s)	Expression for γ_{2k}
1	1
$\bar{1}, 2, m$	$l'^{k-1}M_k$
$2/m, mm2$	$l'^{k-1}M_k^2$
mmm	$l'^{k-1}M_k^3$
222	$\dfrac{l'^{k-1}}{2^k(k!)^2}\sum_{p=0}^{k}(M_pM_{k-p})^3[p!(k-p)!]^2$

of equation (2.1.7.13) and the first two terms on the right-hand side of equation (2.1.7.14) are just the cumulative distributions derived from the ideal centric and acentric p.d.f.'s in Section 2.1.5.6.

The moments $\langle|T|^{2k}\rangle$ were compiled for all the space groups by Wilson (1978*b*) for $k = 1$ and 2, and by Shmueli & Kaldor (1981, 1983) for $k = 1$, 2, 3 and 4. These results are presented in Table 2.1.7.1. Closed expressions for the normalized moments γ_{2p} were obtained by Shmueli (1982) for the triclinic, monoclinic and orthorhombic space groups except *Fdd*2 and *Fddd* (see Table 2.1.7.2). The composition-dependent terms, Q_{2k}, are most conveniently computed as weighted averages over the ranges of $(\sin\theta)/\lambda$ which were used in the construction of the Wilson plot for the computation of the $|E|$ values.

2.1.7.4. *Fourier versus Hermite approximations*

As noted in Section 2.1.8.7 below, the Fourier representation of the probability distribution of $|F|$ is usually much better than the particular orthogonal-function representation discussed in Section 2.1.7.3. Many, perhaps most, non-ideal centric distributions look like slight distortions of the ideal (Gaussian) distribution and have no resemblance to a cosine function. The empirical observation thus seems paradoxical. The probable explanation has been pointed out by Wilson (1986*b*). A truncated Fourier series is a best approximation, in the least-squares sense, to the function represented. The particular orthogonal-function approach used in equation (2.1.7.5), on the other hand, is not a least-squares approximation to $p_c(|E|)$, but is a least-squares approximation to

$$p_c(|E|)\exp(|E|^2/4). \qquad (2.1.7.15)$$

The usual expansions (often known as Gram–Charlier or Edgeworth) thus give great weight to fitting the distribution of the (compararively few) strong reflections, at the expense of a poor fit for the (much more numerous) weak-to-medium ones. Presumably, a similar situation exists for the representation of acentric distributions, but this has not been investigated in detail. Since the centric distributions $p_c(|E|)$ often look nearly Gaussian, one is led to ask if there is an expansion in orthogonal functions that (i) has the leading term $p_c(|E|)$ and (ii) is a least-squares (as well as an orthogonal-function)* fit to $p_c(|E|)$. One does exist, based on the orthogonal functions

$$f_k = n(x)He_k(2^{1/2}x), \qquad (2.1.7.16)$$

where $n(x)$ is the Gaussian distribution (Myller-Lebedeff, 1907). Unfortunately, no reasonably simple relationship between the coefficients d_k and readily evaluated properties of $p_c(|E|)$ has been found, and the Myller-Lebedeff expansion has not, as yet, been applied in crystallography. Although Stuart & Ord (1994, p. 112) dismiss it in a three-line footnote, it does have important applications in astronomy (van der Marel & Franx, 1993; Gerhard, 1993).

2.1.8. Non-ideal distributions: the Fourier method

The starting point of the method described in the previous section is the central-limit theorem approximation, and the method consists of finding correction factors which result in better approximations to the actual p.d.f. Conceptually, this is equivalent to improving the approximation of the characteristic function [*cf.* equation (2.1.4.10)] over that which led to the central-limit theorem result.

The method to be described in this section does not depend on any initial approximation and will be shown to utilize the dependence of the exact value of the characteristic function on the space-group symmetry, atomic composition and other factors. This approach has its origin in a simple but ingenious observation by Barakat (1974), who noted that if a random variable has lower and upper bounds then the corresponding p.d.f. can be non-zero only within these bounds and can therefore be expanded in an ordinary Fourier series and set to zero (identically) outside the bounded interval. Barakat's (1974) work dealt with intensity statistics of laser speckle, where sinusoidal waves are involved, as in the present problem. This method was applied by Weiss & Kiefer (1983) to testing the accuracy of a steepest-descents approximation to the exact solution of the problem of random walk, and its first application to crystallographic intensity statistics soon followed (Shmueli *et al.*, 1984). Crystallographic (*e.g.* Shmueli & Weiss, 1987; Rabinovich *et al.*, 1991*a*,*b*) and noncrystallographic (Shmueli *et al.*, 1985; Shmueli & Weiss, 1985*a*; Shmueli, Weiss & Wilson, 1989; Shmueli *et al.*, 1990) symmetry was found to be tractable by this approach, as well as joint conditional p.d.f.'s of several structure factors (Shmueli & Weiss, 1985*b*, 1986; Shmueli, Rabinovich & Weiss, 1989). The Fourier method is illustrated below by deriving the exact counterparts of equations (2.1.7.5) and (2.1.7.6) and specifying them for some simple symmetries. We shall then indicate a method of treating higher symmetries and present results which will suffice for evaluation of Fourier p.d.f.'s of $|E|$ for a wide range of space groups.

2.1.8.1. *General representations of p.d.f.'s of $|E|$ by Fourier series*

We assume, as before, that (i) the atomic phase factors $\vartheta_j = 2\pi\mathbf{h}^T\mathbf{r}_j$ [*cf.* equation (2.1.1.2)] are uniformly distributed on (0–2π) and (ii) the atomic contributions to the structure factor are independent. For a centrosymmetric space group, with the origin chosen at a centre of symmetry, the random variable is the (real) normalized structure factor E and its bounds are $-E_M$ and E_M, where

* The condition for this desirable property seems to be that the weight function in equation (2.1.7.2) should be unity.