Why the moments of E take the values they do

Norman Stein

CCP4, Daresbury Laboratory
Warrington WA4 4AD, United Kingdom

The values of the moments of the normalised structure amplitude E depend on whether or not the data is twinned and can therefore be used as a test for the presence of merohedral twinning. For example the CCP4 Truncate program [1] lists two sets of theoretical values for these moments, one for untwinned data and one for perfectly twinned data. Partially twinned data give rise to values for the moments of E which lie somewhere between these two extreme cases. This article explains how one can calculate these theoretical values. None of the results presented here are new, but it appears to be hard to find their derivation in the protein crystallography literature. Thus the calculations presented here can perhaps best be viewed as an extension of the last section of the 'Basic Maths for Protein Crystallographers' tutorial included in the CCP4 distribution.

The normalised structure factor amplitude E is defined in terms of the reduced intensity I by $E = \sqrt{I/\Sigma}$, where Σ is the mean intensity in a thin spherical resolution shell containing the reflection under consideration. (The reduced intensity is the measured intensity divided by the symmetry factor ε.) The intensities obey the Wilson distribution [1,2], which for acentric, untwinned reflections has probability density function

$$\begin{aligned} p_I(x) &= (1/\Sigma)\exp(-x/\Sigma) \qquad\qquad (x \geq 0) \\ p_I(x) &= 0 \qquad\qquad (x < 0) \end{aligned} \tag{1}$$

In other words, the probability that the intensity I lies in the range $[x, x+dx]$ is $p_I(x)dx$. Using angle brackets to denote an average over a resolution shell, the n'th moment of E is

$$\langle E^n \rangle = \left\langle I^{n/2} \right\rangle / \Sigma^{n/2} = \int_0^\infty (x/\Sigma)^{n/2} \exp(-x/\Sigma)\, dx \tag{2}$$

Making the change of variable $y = x/\Sigma$ gives

$$\langle E^n \rangle = \int_0^\infty y^{n/2} \exp(-y)\, dy = \Gamma(n/2+1) \tag{3}$$

Using the following properties of the Gamma function [3]

$$\begin{aligned} \Gamma(m+1) &= m! \qquad\qquad (m \text{ integer}) \\ \Gamma(z+1) &= z\Gamma(z) \\ \Gamma(1/2) &= \sqrt{\pi} \end{aligned} \tag{4}$$

it is straightforward to calculate the values shown in Table 1 for untwinned, acentric data. In order to do the same for perfectly twinned, acentric data, it is convenient to introduce the characteristic function

$$\tilde{p}_I(\theta) = \langle \exp(i\theta I) \rangle = \int_{-\infty}^{\infty} \exp(i\theta x)\, p_I(x)\, dx \tag{5}$$

which is just the Fourier transform of the probability density function. For the untwinned, acentric Wilson distribution

$$\tilde{p}_I(\theta) = \frac{1}{1 - i\theta\Sigma} \tag{6}$$

For perfectly twinned data, the observed intensities are given by $I = (J + K)/2$, where J and K represent the individual contributions of two reflections whose Miller indices are related by the twinning operator [4]. In order to obtain the characteristic function for such data, we make use of two results from probability theory. First, if I has probability density $p(x)$ then $I/2$ has probability density $2p\,(2x)$. Applying this to (1) gives

$$\begin{aligned} p_{J/2}(x) &= (2/\Sigma)\exp(-2x/\Sigma) \qquad\qquad (x \geq 0) \\ p_{J/2}(x) &= 0 \qquad\qquad (x < 0) \end{aligned} \tag{7}$$

with characteristic function

$$\tilde{p}_{J/2}(\theta) = \frac{1}{1 - i\theta\Sigma/2} \tag{8}$$

Secondly, if J are K are two independent random variables, then

$$\tilde{p}_{J+K}(\theta) = \langle \exp(i\theta(J + K)) \rangle = \langle \exp(i\theta J) \rangle \langle \exp(i\theta K) \rangle = \tilde{p}_J(\theta)\, \tilde{p}_K(\theta) \tag{9}$$

Applying this to (8) gives the characteristic function for perfectly twinned, acentric data

$$\tilde{p}_I(\theta) = \frac{1}{(1 - i\theta\Sigma/2)^2} \tag{10}$$

A useful property of the characteristic function is that the moments of the intensity distribution can be expressed in terms of the derivatives of the characteristic function evaluated at $\theta = 0$

$$\langle I^n \rangle = \tilde{p}_I^{(n)}(0)/i^n \tag{11}$$

This gives a quick way of calculating the even moments of E. For example, to find $\langle E^4 \rangle$, we have

$$\tilde{p}_I''(\theta) = \frac{3\Sigma^2}{2(1 - i\theta\Sigma/2)^4} \tag{12}$$

so that

$$\left\langle E^4 \right\rangle = \left\langle I^2 \right\rangle / \Sigma^2 = -\tilde{p}_I''(0)/\Sigma^2 = 1.5 \tag{13}$$

In order to calculate odd moments of E, we need to invert the Fourier Transform using contour integration or by consulting tables of Fourier transforms. The result is

$$\begin{aligned} p(x) &= (4x/\Sigma^2)\exp(-2x/\Sigma) \qquad (x \geq 0) \\ p(x) &= 0 \qquad (x < 0) \end{aligned} \tag{14}$$

The n'th moment of E is therefore

$$\left\langle I^{n/2} \right\rangle / \Sigma^{n/2} = 4 \int_0^\infty x^{n/2+1} \exp(-2x/\Sigma)\, dx / \Sigma^{n/2+2} \tag{15}$$

The substitution $y = 2x/\Sigma$ then gives

$$\langle E^n \rangle = \int_0^\infty y^{n/2+1} \exp(-y)\, dy / 2^{n/2} = \Gamma(n/2 + 2)/2^{n/2} \tag{16}$$

For centric reflections, the Wilson distribution is

$$\begin{aligned} p_I(x) &= (1/\sqrt{2\pi\Sigma\, x})\exp(-x/2\Sigma) \qquad (x \geq 0) \\ p_I(x) &= 0 \qquad (x < 0) \end{aligned} \tag{17}$$

with characteristic function

$$\tilde{p}_I(\theta) = \frac{1}{\sqrt{1 - 2\Sigma i\theta}} \tag{18}$$

The moments are given by

$$\langle E^n \rangle = \left(1/\sqrt{2\pi\Sigma^{n+1}}\right) \int_0^\infty x^{(n-1)/2} \exp(-x/2\Sigma)\, dx \tag{19}$$

The substitution $y = x/2\Sigma$ yields

$$\langle E^n \rangle = \left(2^{n/2}/\sqrt{\pi}\right) \int_0^\infty y^{(n-1)/2} \exp(-y)\, dy = \frac{2^{n/2}}{\sqrt{\pi}}\, \Gamma\left(\frac{n+1}{2}\right) \tag{20}$$

Applying the same arguments as for the acentric case to (18), the characteristic function for perfect twinning is

$$\tilde{p}_I(\theta) = \frac{1}{1 - \Sigma i\theta} \tag{21}$$

Note that this is exactly the same as for the untwinned acentric case. This means that the moments for the perfectly twinned centric case are the same as for the untwinned centric case.

	$\langle E\rangle$	$\langle E^3\rangle$	$\langle E^4\rangle$	$\langle E^5\rangle$	$\langle E^6\rangle$	$\langle E^7\rangle$	$\langle E^8\rangle$
acentric untwinned	0.886	1.329	2	3.323	6	11.632	24
acentric perfect twinning	0.940	1.175	1.5	2.056	3	4.626	7.5
centric untwinned	0.798	1.596	3	6.383	15	38.30	105
centric perfect twinning	0.886	1.329	2	0.000	6	11.632	24

Table 1. Numerical values for the first eight moments of E. ($\langle E^2\rangle = 1$ by definition.)

References

1. S. French & K. S. Wilson *Acta Cryst.* A**34**, 517–525 (1978).
2. A. J. C. Wilson *Acta Cryst.* **2**, 318–321 (1949).
3. M. S. Abramovitz & I. A. Stegun *Handbook of Mathematical Functions* (Dover, New York, 1965).
4. T. O. Yeates *Acta Cryst.* **44**, 142–144 (1988).