# On the Computation of the Fast Rotation Function

By J. Navaza

*Laboratoire de Physique (UPR 180 du CNRS), Faculté de Pharmacie, 92290 Chatenay-Malabry, France*

## Abstract

A new more efficient algorithm for the evaluation of the fast rotation function coefficients is derived. Its implementation in standard programs is straightforward.

## Introduction

The rotation function (RF) defined by Rossmann & Blow (1962) measures the overlap between one Patterson function and the rotated version of another, as a function of the rotation angle. The first codes to compute the RF were extremely slow. Later, Crowther (1972) showed the advantage of expanding the Patterson functions within a spherical domain, in terms of spherical harmonics. Indeed, $\beta$-sections of the RF are calculated with two-dimensional fast Fourier transforms. This fast rotation function (FRF) takes the form

$$R(\alpha, \beta, \gamma) = \sum_{l=0}^{\infty} \sum_{m,m'=-l}^{l} C_{mm'}^l \cdot D_{mm'}^l(\alpha, \beta, \gamma), \quad (1)$$

which is an expansion in terms of the orthonormal matrices $D_{mm'}^l$ of the irreducible representations of the rotation group. The coefficients $C_{mm'}^l$ of this expansion depend on the intensities of both crystals and on the definition of the spherical domain, but not on angular variables. Most of the computing time is spent in the calculation of these coefficients so that efficient codes are necessary. We will first review the existing algorithms, discuss their shortcomings, and eventually derive a more efficient one.

## The fast rotation function

The analysis follows the lines of a previous paper (Navaza, 1987; JN hereafter); the starting point is the quite general expansion of the Patterson function in terms of the spherical harmonics $Y_{lm}$,

$$P(\mathbf{r}) = \sum_{\mathbf{H}} |F_{\mathbf{H}}|^2 \exp(-2\pi i \mathbf{H}\mathbf{r})$$

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm}(r) Y_{lm}(\hat{\mathbf{r}}) \quad (2)$$

($r = |\mathbf{r}|$ and $\hat{\mathbf{r}} = \mathbf{r}/r$ stand for the radial and angular parts of vector $\mathbf{r}$, respectively). Replacing the exponential by its expansion in spherical harmonics (Landau & Lifschitz, 1972),

$$\exp(i\mathbf{p}\mathbf{r}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(pr) Y_{lm}(\hat{\mathbf{p}}) Y_{lm}(\hat{\mathbf{r}})^* \quad (3)$$

($j_l$ is the spherical Bessel function), we obtain the radial functions

$$c_{lm}(r) = 4\pi \sum_{\mathbf{H}} |F_{\mathbf{H}}|^2 (-i)^l j_l(2\pi Hr) Y_{lm}(\hat{\mathbf{H}})^*, \quad (4)$$

which are the basis of the FRF. In particular, the coefficients $C_{mm'}^l$ are given by the integral

$$C_{mm'}^l = (3/4\pi)(b^3 - a^3)^{-1} \int_a^b c_{lm}^{(t)}(r) c_{lm'}^{(s)}(r)^* r^2 \mathrm{d}r. \quad (5)$$

[$a$ and $b$ denote the inner and outer radii of the spherical domain, $(t)$ and $(s)$ stand for the target and search crystals, respectively.] Substituting (4) into (5) we obtain

$$C_{mm'}^l = 12\pi \sum_{\mathbf{H}} \sum_{\mathbf{K}} |F_{\mathbf{H}}^{(t)}|^2 Y_{lm}(\hat{\mathbf{H}}) T^l(H,K) |F_{\mathbf{K}}^{(s)}|^2 Y_{lm'}^*(\hat{\mathbf{K}}), \quad (6)$$

with

$$T^l(H,K) = (b^3 - a^3)^{-1} \int_a^b j_l(2\pi Hr) j_l(2\pi Kr) r^2 \mathrm{d}r. \quad (7)$$

Defining the dimensionless quantities $h_a = 2\pi Ha$, $h_b = 2\pi Hb$, $k_a = 2\pi Ka$, $k_b = 2\pi Kb$ and $\delta = b/a$, the array $T^l$ is conveniently written in terms of the elementary integral (Watson, 1958)

$$U^l(h,k) = \int_0^1 j_l(hx) j_l(kx) x^2 \mathrm{d}x$$

$$= \begin{cases} [j_{l-1}(h) j_l(k) h - j_{l-1}(k) j_l(h) k]/(k^2 - h^2), \\ \qquad\qquad\qquad\qquad\qquad \text{if } h \neq k, \\ \frac{1}{2}[j_l^2(h) - j_{l-1}(h) j_{l+1}(h)], \quad \text{if } h = k. \end{cases} \quad (8)$$

The final result is

$$T^l(H,K) = (1 - \delta^3)^{-1}[U^l(h_b,k_b) - \delta^3 U^l(h_a,k_a)]. \quad (9)$$

Thus, by replacing (9) into (6), we get a closed analytical expression for the coefficients $C_{mm'}^l$. This involves a summation over a huge number of terms

[proportional to (number of vectors **H**) times (number of vectors **K**)] because the contributions of the target and search crystals are entangled by the arrays $T^l$. It is then crucial that the summations on **H** and **K** are performed independently of each other, which amounts to factorizing $U^l(h,k)$. The different implementations of the FRF differ precisely in the way the factorization is carried out.

The largest order $l_{max}$ that should be included in (1) is approximately related to the outer radius $b$ and the resolution of the data $d$, through

$$l_{max} \doteq h_{max} = 2\pi b/d. \qquad (10)$$

This criterion is usually obtained by physical considerations.

## Numerical methods

Two different procedures have so far been implemented in the available software.

### Fourier–Bessel expansion

This is Crowther's original formulation, in which $j_l(hx)$ is approximated by the truncated expansion

$$j_l(hx) \doteq \sum_{n=1}^{N} b_{ln}(h) j_l(\lambda_{ln}x) \qquad (11)$$

(strictly speaking, Crowther proceeded in a different but strictly equivalent way). The $\lambda_{ln}$ values are the zeros of $j_l$ and the coefficients are (Watson, 1958)

$$b_{ln}(h) = 2[\int_0^1 j_l(hx) j_l(\lambda_{ln}x)x^2 dx]/j_{l-1}(\lambda_{ln})$$

$$= \begin{cases} 2j_l(h)\lambda_{ln}/[j_{l-1}(\lambda_{ln})(h^2 - \lambda_{ln}^2)], & \text{if } h \neq \lambda_{ln} \\ 1, & \text{if } h = \lambda_{ln}. \end{cases}$$

$$(12)$$

The array $U^l$ is now

$$U^l(h,k) \doteq 2j_l(h)j_l(k) \sum_{n=1}^{N} \lambda_{ln}^2/[(h^2 - \lambda_{ln}^2)(k^2 - \lambda_{ln}^2)],$$

$$\text{if } h, k \neq \lambda_{ln}. \qquad (13)$$

The characteristics of this approximation have been discussed in JN. In particular, we have the possibility of choosing a special $N$ for each $l$ and $h$. The main shortcomings are that the series (13) converges every slowly and the roots $\lambda_{ln}$ are not available for high $l$. The largest allowed order in the *ALMN* code (Dodson, 1985) is $l = 60$, which puts a restriction on the radius that can be used with a given resolution.

### Numerical integration

This procedure consists in evaluating (8) by means of a quadrature formula

$$U^l(h,k) = \int_0^1 j_l(hx) j_l(kx)x^2 dx$$

$$\doteq \sum_{n=1}^{N} j_l(hx_n) j_l(kx_n)x_n^2 w_n, \qquad (14)$$

where $x_n$ and $w_n$ are the integration points and weights, respectively. Points and weights of high-order quadrature formulas may be easily obtained, so that situations with large values of the ratio $h_{max} = 2\pi b/d$ may now be considered. The main shortcoming of this method is that the number of integration points may be rather important for high values of $h_{max}$. The *ROTING* program written by the author is based on this procedure, where an optimal $N$ corresponding to a Gauss–Legendre quadrature formula was determined, as a function of $l$ and $h_{max}$, for $l \leq 120$. As a general rule it gives consistently better results than the Fourier–Bessel expansion – with a slightly greater computing effort (Alzari & Navaza, 1991). It is incorporated in the new molecular replacement package *AMoRe* (Navaza, 1992, 1993). The success of *AMoRe* in many cases where standard packages failed was because of the good quality of the FRF output (Strynadka, Adachi, Jensen, Johns, Sielecki, Betzel, Sutoh & James, 1992).

## A new recurrence relation

A new factorization of $U^l$ may be obtained by using the well known recurrence relation between spherical Bessel functions,

$$j_{l-1}(h) = (2l + 1)j_l(h)/h - j_{l+1}(h), \qquad (15)$$

to systematically increase the lowest $l$ order in expression (8). Two substitutions give the simple relation

$$U^l(h,k) = [j_{l-1}(h) j_l(k)h - j_{l-1}(k) j_l(h)k]/(k^2 - h^2)$$

$$= [-j_{l+1}(h) j_l(k)h + j_{l+1}(k) j_l(h)k]/(k^2 - h^2)$$

$$= (2l+3) j_{l+1}(h) j_{l+1}(k)/(hk) + [j_{l+1}(h) j_{l+2}(k)h$$

$$- j_{l+1}(k) j_{l+2}(h)k]/(k^2 - h^2)$$

$$= (2l+3) j_{l+1}(h) j_{l+1}(k)/(hk) + U^{l+2}(h,k), \qquad (16)$$

from which we obtain the desired approximation

$$U^l(h,k) \doteq \sum_{n=1}^{N} [2(l + 2n) - 1] j_{l+2n-1}(h) j_{l+2n-1}(k)/(hk), \qquad (17)$$

and its associated error, $U^{l+2N}$. In practice one is forced to limit the number of terms in the summation. The same criterion (10) that was used to neglect all coefficients (6) with $l > l_{max}$ leads us to the condition $l + 2N = l_{max} + 2$, or

$$N \doteq (l_{max} - l + 2)/2 \doteq (h_{max} - l + 2)/2. \qquad (18)$$

A better estimation of $N$ may be obtained by quantitatively assessing the errors of the approximation; for a given accuracy $\varepsilon$, we can find the smallest $N$ such that

$$|U^{l+2N}(h,k)|/\tilde{u}^l \le \varepsilon. \qquad (19)$$

($\tilde{u}^l$, the maximum value of $U^l$, is introduced in order to work with relative quantities). However, since the aim of the factorization is to perform the summations on $h$ and $k$ independently of each other, the best $N$ we can determine is the smallest one that satisfies (19) for every $k$ in $[0,h_{\max}]$. This gives $N$ as a function of $l$, $h$ and $h_{\max}$.

There is another way of estimating the optimal $N$. Let us call $h_\varepsilon^l$ the smallest argument of the equation $U^l(h,h)/\tilde{u}^l = \varepsilon$ (see Fig. 1), and choose the smallest $N$ such that $h_\varepsilon^{l+2N} \ge h_{\max}$. Then, using Schwartz's inequality [(8) is in fact a scalar product],

$$|U^{l+2N}(h,k)|/\tilde{u}^l \le |U^{l+2N}(h,h)|^{1/2}|U^{l+2N}(k,k)|^{1/2}/\tilde{u}^l$$

$$\le U^{l+2N}(h_\varepsilon^{l+2N},h_\varepsilon^{l+2N})/\tilde{u}^l = \varepsilon\tilde{u}^{l+2N}/\tilde{u}^l$$

$$\le \varepsilon, \quad \text{for } 0 \le h,k \le h_{\max} \le h_\varepsilon^{l+2N} \qquad (20)$$

[$\tilde{u}^{l+2N} \le \tilde{u}^l$ because of (16)].

Inspection of Fig. 1 shows that: (18) is a sensible estimation; the accuracy is a sharp function of the number of terms (adding two more terms makes the errors drop by two orders of magnitude); $l_{\max}$ may be determined as the largest $l$ such that $|U^l(h,k)|$ is smaller than a fraction of its maximum value $\tilde{u}^l$, for
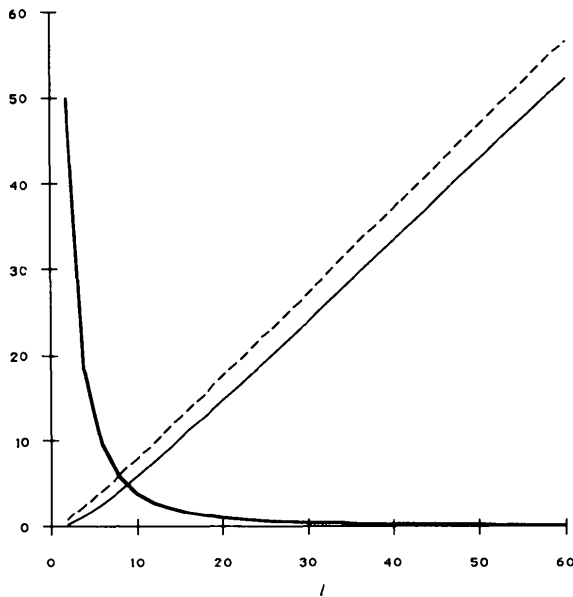
every $(h,k)$ in $[0,h_{\max}; 0,h_{\max}]$. This criterion gives, in general, $l_{\max} > h_{\max}$.

The numerical tests showed that, for a given accuracy, the number of terms required by (17) is about 30% smaller than that required by a Gauss–Legendre numerical integration formula. Comparison with the truncated Fourier–Bessel expansion was not performed, because (13) is a poor approximation of $U^l$, as demonstrated in JN.

## Concluding remarks

The advantage of the formula (17) is threefold: (a) no quantities other than spherical Bessel functions are needed; (b) the number of terms can be easily and precisely determined as a function of $l$ and the accuracy; (c) it is faster than Gauss–Legendre quadratures. Moreover, formula (17) may be easily put into the existing FRF programs. Indeed, substitution of the approximation (17) into (9), and the latter into (6), gives

$$C_{mm'}^l \doteq 12\pi(1 - \delta^3)^{-1} \sum_{n=1}^{N} [e_{lmn}^{(l)}(b)e_{lm'n}^{(s)}{}^*(b)$$

$$- \delta^3 e_{lmn}^{(l)}(a)e_{lm'm}^{(s)}{}^*(a)], \qquad (21)$$

where

$$e_{lmn}(x) = [2(l + 2n) - 1]^{1/2}\sum_{\mathbf{H}}|F_{\mathbf{H}}|^2 Y_{lm}(\hat{\mathbf{H}})$$

$$\times j_{l+2n-1}(2\pi Hx)/(2\pi Hx), \qquad (22)$$

with $x$ denoting any radius. Thus, we have simply to replace the radial part of the $c_{lmn}$ coefficients in the *ROTING* program by that of the $e_{lmn}$. The optimal number of terms is now obtained using the criterion (20), which is calculated at each run [this gives a finer estimation than criterion (18)], so that there is no limit for the ratio $h_{\max} = 2\pi b/d$ (the determination of the optimal $N$ for numerical integration formulas is painful work!). The code has been tested for $l_{\max} \le 200$. The results are sensibly the same as those obtained with the ancient program (for $l_{\max} < 100$), but the computing time is about 30% shorter.

Similarly, the *ALMN* program may be easily modified in order to compute expression (21): the radial part of the $a_{lmn}$ coefficients

$$a_{lmn}(x) = 2^{1/2}\sum_{\mathbf{H}}|F_{\mathbf{H}}|^2 Y_{lm}(\hat{\mathbf{H}})$$

$$\times j_l(2\pi Hx)\lambda_{ln}/[(2\pi Hx)^2 - \lambda_{ln}^2], \qquad (23)$$

must be replaced by that of the corresponding $e_{lmn}$,

$$2^{1/2}j_l(2\pi Hx)\lambda_{ln}/[(2\pi Hx)^2 - \lambda_{ln}^2]$$

$$\rightarrow [2(l + 2n) - 1]^{1/2}j_{l+2n-1}(2\pi Hx)/(2\pi Hx), \qquad (24)$$

the subindex $n$ now labelling Bessel functions instead of roots. The number of terms are now determined by (18).



Fig. 1. Smallest argument $h_\varepsilon^l$ satisfying $U^l(h_\varepsilon^l,h_\varepsilon^l)/\tilde{u}^l = \varepsilon$ (—$\varepsilon = 10^{-4}$; – – –$\varepsilon = 10^{-2}$) and $\tilde{u}^l$, the maximum value of $U^l$ (—$\tilde{u}^l \times 2 \times 10^3$).

**References**

ALZARI, P. & NAVAZA, J. (1991). *Crystallographic Computing* 5, edited by D. MORAS, A. D. PODJARNY & J. C. THIERRY, pp. 348–360. IUCr/Oxford Univ. Press.

CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN. pp. 173–178. New York: Gordon & Breach.

DODSON, E. (1985). In *Molecular Replacement*, edited by P. A. MACHIN. Daresbury, England: SERC.

LANDAU, L. D. & LIFSCHITZ, E. M. (1972). In *Théorie Quantique Relativiste*. Moscow: Mir.

NAVAZA, J. (1987). *Acta Cryst.* A43, 645–653.

NAVAZA, J. (1992). In *Molecular Replacement*, edited by E. J. DODSON, S. GOVER & W. WOLF. Daresbury, England: SERC.

NAVAZA, J. (1993). *Acta Cryst.* A49. In the press.

ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* 15, 24–31.

STRYNADKA, C. J., ADACHI, H., JENSEN, S. E., JOHNS, K., SIELECKI, A., BETZEL, C., SUTOH, K. & JAMES, M. N. G. (1992). *Nature (London)*, 359, 700–705.

WATSON, G. N. (1958). In *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge Univ. Press.