On the Charge Density Refinement of Odd-Order Multipoles Invariant under Crystal Point Group Symmetry

Pietro Roversi 1,* and Riccardo Destro 2

1 Department of Biochemistry, Oxford Glycobiology Institute, Oxford University, South Parks Road, Oxford OX1 3QU, UK
2 Department of Chemistry, Università degli Studi di Milano, Via Golgi 19, 20133 Milano, Italy; riccardo.destro@unimi.it
* Correspondence: pietro.roversi@bioch.ox.ac.uk; Tel.: +44-1865-275339

Abstract: Charge density studies utilise a multipolar expansion of the atomic density (and the associated atomic scattering factor) in order to model asphericity. Contributions of the individual multipoles to the atomic density are then refined as multipole population coefficients. Refinement of these coefficients pertaining to odd-order multipoles that are invariant under the crystal point-group symmetry is often problematic, with ill-defined values and correlations plaguing the convergence to a good model. These difficulties have been discussed in generic terms in the literature, but never explicitly analysed in detail. In this communication, we show that the charge density multipolar atomic scattering factor can be partitioned in three contributions that differ in their behaviour under the point group symmetry of the crystal. This partitioning rationalises and predicts the conditions that give rise to ill-conditioning of the charge density refinement of these multipoles.

Keywords: charge-density; multipolar expansion; odd-order multipole; invariance under point symmetry

1. Introduction

Charge density studies based on ultra high-resolution X-ray diffraction data from crystals of biological macromolecules yield experimentally determined physico-chemical properties of proteins and nucleic acids [1], molecular systems often too large for simulations and quantum-mechanical calculations to be feasible or accurate [2–5]. X-ray scattering at ultra high-resolution carries information about bonding and non-bonding atomic valence electron density that departs from spherical symmetry around the nucleus, and therefore these charge density studies are based on multipolar expansions of atomic densities [6,7]. Crystals of biological macromolecules can only belong to non-centrosymmetric space groups because of the chirality of naturally occurring α-amino acids and nucleic acids [8]. Thus, ultra-high resolution charge density studies for crystals of biological macromolecules could be exposed to a long-known problem arising during multipolar refinement in non-centrosymmetric space groups: odd-order poles that are invariant under the crystal point-group symmetry can be ill-determined and/or plagued by correlations. The subject has always been discussed in generic terms in the literature [9–12], or in terms of individual atomic scattering factors [11], but never explicitly analysed in terms of the structure factor. In this communication, we derive formulae that make it easier to understand and predict the conditions giving rise to ill-conditioning of the refinement of these multipoles. We then proceed to discuss the topic in the context of charge density studies based on ultra high-resolution X-ray diffraction data from crystals of biological macromolecules.
2. Results

The formalism adopted here to discuss the interaction between atomic generalised scattering factors and crystal symmetry is the one introduced by Gérard Bricogne in Section §1.3.4. of the International Tables for Crystallography Volume B [13]. Moreover, the atomic densities are expanded in multipoles all oriented along the same axes, such as the crystal axes, irrespective of the atom [14]. Should local axes be chosen instead [15], the analysis presented here would still be valid, except that certain linear combinations of odd-order multipoles populations would be invariant—rather than their individual values—and it would be their refinement that would suffer from ill-determination. Different choices of axes will of course change the algebraic expression for the angular part of the multipoles and therefore the list of the odd-order invariant poles. For this reason, the list of the potentially troublesome multipoles should be carefully arrived at on the basis of the symmetry invariance, and not decided a posteriori on the basis of, say, high values of the correlation coefficients in the least-squares refinement.

2.1. The Structure Factor

We are interested in discussing the interplay between the symmetry-invariance and the resulting phases of certain atomic contributions to the structure factor. We begin by recalling that structure factor \( F_h \) can be written as the Fourier transform of the electron density in the unit cell, \( \rho_0(x) \):

\[
F_h = \mathcal{F} \left[ \rho_0(x) \right] (h) = \int_{R^3} \rho_0(x) \exp(2\pi i h \cdot x) \, d^3x
\]

(1)

The electron density in the unit cell, \( \rho_0(x) \), can be written as a sum over the atoms in the unit cell:

\[
\rho_0(x) = \sum_{k \in K} \rho_k(x-x_k) = \sum_{k \in K} \tau_{x_k} \rho_k(x)
\]

(2)

where \( K \) is the set of atoms in the unit cells, \( \rho_k(x) \) is the atomic density of atom \( k \) as a function centred around the origin, and \( \tau_{x_k} \) is the translational operator used to write the atomic density around the nucleus located at \( x_k \): \( \rho_k(x-x_k) = \tau_{x_k} \rho_k(x) \).

Let us now indicate the crystal space group as \( G \). It is advantageous to write the sum over all atoms in the unit cell as a sum over the orbits of the symmetry-unique atoms, i.e., the atoms in the crystal asymmetric unit \( J \).

We recall here the definitions of Isotropy group and Orbit of an element \( x \) of a set \( X \) with the \( g \in G \) that form a group, see section §1.3.4.2.2.2. of the International Tables for Crystallography Volume B [13]:

The Isotropy group of \( x \in X \) is the subgroup \( G_x \) of \( G \) containing the elements \( g \in G \) which map \( x \) onto itself:

\[
G_x = \{ g \in G \mid gx = x, x \in X \}
\]

The Orbit of \( x \in X \) under \( G \), is the subset of \( X \) consisting of all elements \( gx \) with \( g \) running through \( G \):

\[
Gx = \{ x' \in X \mid x' = gx, x \in X, g \in G \}
\]

Acting on the asymmetric unit atoms with the symmetry operators in \( G \):

\[
\rho_0^g(x) = \sum_{g \in G} \sum_{j \in J} S^g \left( \tau_{x_j} \rho_j(x) \right)
\]

(3)

where \( S^g \left( \tau_{x_j} \rho_j(x) \right) \) is the image of the \( j \)-th atomic density under the action of the \( g \)-th symmetry operator. The symmetry operator \( g \in G \), acting on coordinates \( x \), is defined as:

\[
g(x) = R_g x + \tau_g \text{ (mod } \Lambda)\]

(4)
where \( R \) is the rotation matrix and \( \tau \) the translation vector of the symmetry operator. \( \Delta \) is the set of crystal lattice translations.

The \( j \)-th atomic density is invariant under all its site-symmetry operations, i.e., the symmetry operations in the isotropy subgroup \( G_{x_j} \) of \( G \) for position \( x_j \). Thus, the sum over symmetry operations in (3) can be expressed without redundancies making use of the symmetry operators \( \gamma \in G/G_{x_j}' \), where \( G/G_{x_j} \) is the collection of representatives of the distinct left cosets of \( G_{x_j} \) (the isotropy group of position \( x_j \)) so that the sum in (3) runs over all those space group symmetry operations that produce a distinct copy of atom \( j \), and only on those:

\[
\rho^0(x) = \sum_{j \in J} \sum_{\gamma \in G/G_{x_j}} \mathcal{S}^\gamma_{j\gamma} \left( \tau_{x_j} \rho_j(x) \right)
\]  

(5)

Thanks to this notation, the \( j \)-th atomic density \( \rho_j(x) \) in the asymmetric unit only appears once and only once in the sum (5), in that the orbit contains all and only the distinct copies of \( \rho_j(x) \) under the symmetry in \( G \). Recalling now Equation (1), the structure factor can be written:

\[
F_h = \mathcal{F} \left[ \rho^0(x) \right](h) = \\
\sum_{j \in J} \sum_{\gamma \in G/G_{x_j}} \exp \left( 2\pi i h \cdot t_{x_j} \right) \times \exp \left( 2\pi i R_{j\gamma}^T h \cdot x_j \right) \\
\mathcal{F} [\rho_j](R_{j\gamma}^T h)
\]  

(6)

2.2. The Multipolar Structure Factor

The last factor in formula (6) is the transform of the atomic density for atom \( j \), evaluated at \( R_{j\gamma}^T h \); if we now recall the multipolar expansion of the atomic density, this transform (called the atomic generalised scattering factor \( Gsf_j(h) \) \([6,16]\)), is written:

\[
Gsf_j(h) = \mathcal{F} [\rho_j](h) = \\
\left[ C_{j\text{core}} \mathcal{F} \left[ \rho_{j\text{core}} \right](h) + \sum_{l=0}^{L} \sum_{m=-l}^{l} C_{jlm} \mathcal{F} \left[ R_{j\gamma} Y_{l}^{m} \right](h) \right] \times T_{j}(h)
\]  

(7)

where the \( C_{jlm} \) are the multipolar coefficients, the \( R_{j\gamma} Y_{l}^{m} \) are the products of atomic radial (\( R_{j\gamma} \)) and angular (\( Y_{l}^{m} \)) valence deformation functions, and \( T_{j}(h) \) is the Debye–Waller factor of atom \( j \).

With this expression for the atomic scattering factor, (6) is rewritten:

\[
F_h = \sum_{j \in J} \left[ C_{j\text{core}} \mathcal{F} \left[ \rho_{j\text{core}} \right](h) + C_{j\text{rad}} \mathcal{F} \left[ R_{j\gamma} \right](h) \right] \times T_{j}^0(h) \\
+ \sum_{\gamma \in G/G_{x_j}} T_{j\gamma}(R_{j\gamma}^T h) \\
\times \sum_{l=0}^{L} \sum_{m=-l}^{l} C_{jlm} \mathcal{F} \left[ R_{j\gamma} Y_{l}^{m} \right](R_{j\gamma}^T h)
\]  

(8)

with:

\[
T_{j\gamma}(R_{j\gamma}^T h) = \\
\exp \left( 2\pi i h \cdot t_{x_j} \right) \exp \left( 2\pi i R_{j\gamma}^T h \cdot x_j \right) \times T_{j}(R_{j\gamma}^T h)
\]  

(9)
and:

\[ T^\Phi_j(h) = \sum_{\gamma_j \in G_j/G} T_{\gamma_j}(R^\gamma_j h) \]  

\( T^\Phi_j(h) \) gathers phase factors and Debye–Waller factors multiplying the spherically symmetric poles (i.e., the monopoles).

2.3. Multipoles Invariant under Point-Group Symmetry

Let us now consider a multipole \( R^\Phi_j l^m_l Y^\Phi_l(x) \) (with \( l \geq 1 \)) that is invariant under all rotations \( R_g \) belonging to the crystal point-group:

\[ R^\Phi_g \left( R^\Phi_j l^m_l Y^\Phi_l(x) \right) = R^\Phi_j l^m_l Y^\Phi_l(x) \quad \forall g \in G \]  

The invariance of the multipole in real space is mirrored in reciprocal space by the invariance of its transform:

\[ \bar{F} \left[ R^\Phi_j l^m_l Y^\Phi_l \right](h) = F^\Phi_j l^m_l Y^\Phi_l(h) \]  

Thanks to this invariance, it is possible to factorise the invariant poles out of the the sum over symmetry operations in formula (8), as it was done with the spherically symmetrical poles. The structure factor is then a sum of three terms:

\[ F_h = F_{h}^{\text{Mono}} + F_{h}^{\text{Inv}} + F_{h}^{\text{Non Inv}} \]  

with:

1. **Monopoles** (\( l = 0 \)): the

\[ F_{h}^{\text{Mono}} = \sum_{j \in J} T^\Phi_j(h) \times \left[ C^\text{core}_j f^\text{core}_j(h) + C_{00} f_{j0}(h) \right] \]  

This is usually the largest contribution to the structure factor, especially at high resolution where diffraction is dominated by core electrons.

2. **Invariant poles** (\( l > 0 \)):

\[ F_{h}^{\text{Inv}} = \sum_{j \in J} T^\Phi_j(h) \times \left[ \sum_{I \neq 0} \sum_{(l,m) = (l,0)} 4\pi C_{jlm} l^l f_{jl}(h) Y^\Phi_l(h) \right] \]  

As \( T^\Phi_j(h) \) is the same factor appearing in the monopoles term, and because of the \( l^l \) phase factor, the invariant poles add contributions whose phase either is collinear with the phase line of the monopoles term (even-order invariant poles), or at \( \pm \frac{\pi}{2} \) from it (odd-order invariant poles). Notice that in centrosymmetric crystals no odd-order poles contribute to this term (because no odd-order pole is invariant under the center of inversion); only non-centrosymmetric crystals can have odd-order invariant poles, e.g., \( Y_{32} = q_x q_y q_z \) octupoles (\( l = 3 \)) in orthorhombic space groups of 222 point group symmetry; or \( Y_{10} = q_z \) dipoles (\( l = 1 \)) in monoxial tetragonal, trigonal and hexagonal space groups (point group symmetries 4, 3 and 6). The comprehensive list of such poles in each possible point group can be derived from the character tables and irreducible representations for the 32 crystallographic point groups in three dimensions, and the transformation properties of tensors, see [17,18].
3. **Non invariant poles** \( (l \neq \bar{l}, l > 0) \):

\[
F_{h}^{\text{Non Inv}} = \sum_{j \in J} \left[ \sum_{\gamma_j \in G/G_{\bar{h}}} T_{\gamma_j} (R_{\gamma_j}^T h) \times \sum_{(l,j,m) \neq (l,\bar{h})} 4\pi C_{ljm}^i f_{jl} (h) Y_{m}^{l} \left( \frac{R_{\gamma_j}^T h}{\bar{h}} \right) \right]
\]  

The contributions to this third term have phases that bear no special relationship to the phase of the monopoles term, and depend in a combined way on the parity of the pole and the action of the symmetry operators.

When, for a significant number of structure factors, the odd-order non-invariant poles contribution in (16) is by chance aligned to the monopolar contribution in (14), a situation arises in which odd-order invariant poles add a contribution to those structure factors which is at \( \pm \frac{\pi}{2} \) from the contributions to structure factor computed from all the other poles. If in addition, the contribution from the odd-order invariant poles to the first order is small, then the population coefficients for these odd-order invariant multipoles only alter the phase and not the amplitude of the structure factor, and least-squares refinement of these population coefficients against amplitude data is ill-determined.

It is noted that—should one analyse the contributions to the structure factor from atom-centered nuclear probability functions (for example, the well known Gram–Charlier expansion, see Section §1.2.12. of the International Tables for Crystallography Volume B [19])—similar ambiguities would arise, for example, when attempting refinement of the third order cumulants invariant under point-group symmetry in non-centrosymmetric crystals.

For crystals whose asymmetric unit only contains a few atoms, most of which sit on a special position, so that only a few multipoles contribute to the expansion of the atomic density, for example, cubic hexamethylenetetramine [9,20–22], it is often the case that odd-order invariant poles add a contribution to those structure factors which is at \( \pm \frac{\pi}{2} \) from the contributions to structure factor computed from all the other poles, and refinement of the odd-order invariant multipolar populations is plagued by correlations. The same problems can still arise in multipolar refinements for crystals with several atoms in the asymmetric unit, none of which sits on a special position, when the conditions discussed above are met by a significant fraction of structure factors. To give a sense of the relative size and phase of the contributions from the odd-order invariant poles in (15) to all the other ones in a standard charge density study, we analysed the calculated structure factors from the multipolar study of DL-alanine (DL-Ala) [23–25] that appeared in [26].

In that study, a group of 2748 reflexions up to a resolution of 0.43 Å \( (\sin \theta / \lambda)_{\text{max}} = 1.15 \, \text{Å}^{-1} \) were measured from a Pna2₁ crystal which has one molecule in the asymmetric unit (chemical composition \( C_3 NO_2 H_7 \)), and the electron density modelled with a multipolar model up to \( l = 4 \) (hexadecapoles). Under the \( mm2 \) point group, and with a standard choice of axes along the orthorhombic cell edges, the odd-order invariant multipoles are the \( Y_{10} \) dipoles \( (Y_{10} = \bar{q}_z) \), and the \( Y_{32}^+ \) and \( Y_{30} \) octupoles \( (Y_{32}^+ = (q_x^2 - q_y^2)q_z, Y_{30} = (5q_z^2 - 3)q_z) \). Figure 1 illustrates the scatter plot of the log of the ratio of amplitudes \( F_h^{\text{InvOdd}} / F_h^{\text{Mono+InvEven+NonInv}} \) vs. the value of the angle between \( F_h^{\text{InvOdd}} \) and \( F_h^{\text{Mono+InvEven+NonInv}} \). All contributions of \( F_h^{\text{InvOdd}} \) to the structure factor are rather small; for a group of 493 structure factors, the absolute value of the angle between the lines of \( F_h^{\text{InvOdd}} \) and \( F_h^{\text{Mono+InvEven+NonInv}} \) is between 80 and 90 degrees, i.e., the odd-order invariant multipoles give a contribution roughly orthogonal to all other contributions to the structure factor.
Figure 1. Scatter plot of the log of the ratio of amplitudes $F_{\text{InvOdd}}^h / F_{\text{Mono+InvEven+NonInv}}^h$ vs. the value of the angle between the lines of $F_{\text{InvOdd}}^h$ and $F_{\text{Mono+InvEven+NonInv}}^h$, for 2748 reflexions up to 0.43 Å measured from a Pna2$_1$ crystal of DL-alanine (DL-Ala) [26].

3. Discussion

The problems associated with the refinement of populations for odd-order multipoles invariant under the point group of the crystal have been discussed in the literature [9–12], and especially the authors of [11] have gone to some lengths in order to formally analyse the reasons for the high correlations encountered when refining those populations, although their analysis is restricted to the individual atomic scattering factors. In some cases, the problem has been tackled simply by omitting refinement of the ill-determined poles [12]; in other cases, extra-constraints on these population coefficients were chosen so as to stabilise the refinement. In the majority of cases though, no particular problems in the refinement of the odd-order invariant poles were reported—without an explanation being offered.

Here, we have used group theory and analysed multipolar contributions to the structure factor (and not the atomic scattering factor), distinguishing between invariant and non-invariant ones, in an attempt to make the present exposition of the problem clearer and more stringent. Based on the formulae derived in this manuscript, we can predict that the refinement of these poles will be troublesome when the phase of $F_{\text{NonInv}}^h$ is similar to the phase of $F_{\text{Mono+InvEven}}^h$ for a large number of structure factors, so that $F_{\text{InvOdd}}^h$ is orthogonal to all other contributions to the structure factor.

One of the most important consequences of such troublesome multipolar refinements would be artificially featureless residual densities: since the ill-refined multipoles do not alter the calculated amplitude to first-order, and the phase of the experimental structure factor is not measured, but is taken to be equal to the one of the calculated $F$, the amplitudes of many of the Fourier coefficients of the residual density are artificially close to 0, notwithstanding the fact that the model obtained from the ill-determined refinement may contain physically meaningless contributions from the odd-order invariant multipoles.

If orbital-based density functions, or the libraries of Extremely Localized Molecular Orbitals, recently described as a possible way of modelling asphericity in atoms in macromolecules [27,28], were to be used instead of multipoles with refined populations, the results presented in this paper would still be relevant to the contributions to the structure factors from the odd-order invariant molecular orbitals. Since the coefficients associated with those orbitals are not refined though,
the problems would be limited (no correlation would occur among orbital populations or between orbital populations and, say, coordinates).

The larger the number of atoms in the asymmetric unit, the less likely the occurrence of the troublesome conditions. Even if a certain subset of atoms (such as heavy atoms in metallo-enzymes [5]) were the main contributors to certain structure factors, and the values of the odd-order invariant poles were to be poorly determined by refinement against those structure factors, the larger the numbers of experimentally derived amplitudes, the higher the chance that the values of the odd-order invariant pole populations can be determined by other structure factors for which the problem does not arise. We conclude that ultra high-resolution X-ray diffraction data from crystals of biological macromolecules are unlikely to be plagued by ill-determined refinement of odd-order invariant multipoles. There is sometimes safety in numbers.

Acknowledgments: Pietro Roversi was the recipient of a Milano University PhD studentship for the years 1993–1996. Thanks are due to Gérard Bricogne and the MRC-LMB in Cambridge (England) where part of this work was done. We are grateful to Carlo Gatti, Fausto Cargnoni and Richard Morris for helpful suggestions, and to the reviewers of the first version of this manuscript for their constructive feedback.

Author Contributions: P.R. conceived and designed the work, and carried out the derivations. R.D. computed the multipolar structure factors for DL-Alanine. P.R. and R.D. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).