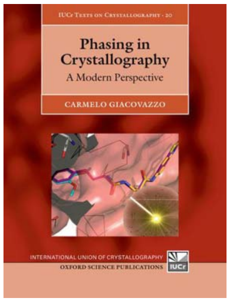
Acta Crystallographica Section A
Foundations and Advances
ISSN 2053-2733

book reviews

Works intended for this column should be sent direct to the Book-Review Editor, whose address appears on the journal web site. All reviews are available from **Crystallography Journals Online**, supplemented where possible with direct links to the publisher's information.



Phasing in Crystallography: A Modern Perspective. By Carmelo Giacovazzo. IUCr Texts on Crystallography, No. 20. International Union of Crystallography/ Oxford University Press, 2013. Pp. 432. Price GBP 65.00 (hardback). ISBN 978-0-19-968699-5.

Carmelo Giacovazzo's newly published book, *Phasing in Crystallography: A Modern Perspective*, is a capstone on his impressive bibliography of publications: his 1992 textbook *Fundamentals of Crystallography*, authored along with a team of special-topics co-authors; his comprehensive 1998 monograph *Direct Phasing in Crystallography: Fundamentals and Applications*; his authoritative 2008 chapter entitled *Direct Methods* in *International Tables for Crystallography*, Vol. B, *Reciprocal Space*; and his and his co-workers almost 300 specialized papers published in *Acta Crystallographica* since 1973.

A chapter-by-chapter description of the new book's contents is not needed here, because in the book's *Preface* Professor Giacovazzo gives succinct, clear summaries of the point and purpose of each of the 15 chapters and 35 (mainly mathematical) chapter appendices. The careful organization of the book is outlined clearly in a detailed *Table of Contents*, and the work is thoroughly documented in an 18-page list of some 750 references. One could, however, wish that the book's terse, three-page index, containing only some 300 brief entries, were more detailed. The book notes the rich crystallographic legacy that continues to accumulate in the archives of, at present, ~750 000 small-unit-cell molecular and materials crystal structures in the Cambridge Structural Database (CSD) and the Inorganic Crystal Structure Database (ICSD), and ~100 000 biomacromolecular crystal structures in the Protein Data Bank (PDB).

Early on, the book states that the 'basic postulate of structural crystallography' is a biunique correspondence between a crystal structure and a set of experimental diffraction amplitudes: 'Only one chemically sound crystal structure exists that is compatible with the experimental diffraction data.' In symbols,

$$\text{crystal structure} \Leftrightarrow \{\mathbf{r}_A - \mathbf{r}_B\} \Leftrightarrow \{|F_{hkl}|\},$$

where $\mathbf{r}_A$ and $\mathbf{r}_B$ are atomic position vectors and the set of $\mathbf{r}_A \rightarrow \mathbf{r}_B$ interatomic distance vectors defines the crystal structure, which must meet the requirement of sensible chemical geometry and must obey the physical constraints of the non-negativity and atomicity of the distribution of scattering matter in the unit cell.

An equally or perhaps more basic postulate is what has been called the 'fundamental theorem' of structural crystallography: The unit-cell scattering density distribution and the crystal structure factors are related by Fourier transformation:

$$F_{hkl} = |F_{hkl}| \exp(i\varphi_{hkl}) \underset{\mathrm{FT}^{-1}}{\overset{\mathrm{FT}}{\rightleftharpoons}} \rho(x, y, z),$$

$$\begin{cases} \mathrm{FT}[F_{hkl}] = \rho(x, y, z) & \text{Fourier synthesis,} \\ \mathrm{FT}^{-1}[\rho(x, y, z)] = F_{hkl} & \text{Fourier analysis.} \end{cases}$$

Of course, this fundamental relationship is implicit throughout the book, but an explicit expression would have been worthwhile, because so many modern phasing procedures are based on iterative recycling back-and-forth between phase adjustments in *hkl* reciprocal space and density adjustments in *xyz* crystal space. Examples include 'shake-and-bake', 'half-bake', the (Bari) SIR program suites, density modification, solvent flattening, non-crystallographic symmetry averaging, charge flipping, the VLD (*vive la difference* $\rho_{\rm total} - \rho_{\rm model}$) method and others.

The essential fact for direct-methods phasing is that, although the phases of the X-ray waves scattered by a crystal are 'lost' in a diffraction experiment because only the intensities $I_{hkl} \propto F_{hkl}F^*_{hkl} = |F_{hkl}|^2$ or squared amplitudes of the waves can be measured, there is nevertheless accessible phase information implicit in relationships among the amplitudes for certain *structure invariant* linear combinations of phases, $\varphi_{\mathbf{h}1} + \varphi_{\mathbf{h}2} + \ldots + \varphi_{\mathbf{h}n}$ with $\mathbf{h}_1 + \mathbf{h}_2 + \ldots + \mathbf{h}_n = 0$, where the reciprocal-lattice vector $\mathbf{h}_1$ has the component indices $h_1k_1l_1$, and so forth.

The especially important triplet phase relationship $\varphi_{\mathbf{h}1} + \varphi_{\mathbf{h}2} + \varphi_{-(\mathbf{h}1+\mathbf{h}2)} \approx 0$ for a three-phase structure invariant that corresponds to large amplitude values $|F_{\mathbf{h}1}|$, $|F_{\mathbf{h}2}|$ and $|F_{-(\mathbf{h}1+\mathbf{h}2)}|$ follows directly from the insightful analysis by David Sayre in 1952 of the $\rho^2(x, y, z)$ 'squared structure' convolution product for a hypothetical structure of resolved equal atoms. Sayre's result, $F_{\mathbf{h}} = (1/\gamma V)\sum_{\mathbf{k}} F_{\mathbf{k}}F_{\mathbf{h}-\mathbf{k}}$, in which $\gamma = f^{\rm sq}_{\mathbf{h}}/f_{\mathbf{h}}$ is the scaling ratio for the 'squared atom' scattering factor, is now universally called the Sayre equation. This was a founding contribution to direct phasing methods, and leads straightaway to the triplet relationship and tangent formula, which have been historically hugely important. It is regrettable that a description of Sayre's analysis is not included in the new book.

The mathematical approach to direct phasing methods that Hauptman and Karle introduced in the 1950s is that the phase information implicit in the amplitudes data is accessible *via* joint, marginal and conditional probability theory,

$$\frac{p_{\rm J}\left(|F_{\mathbf{h}1}|, |F_{\mathbf{h}2}|, \ldots, |F_{\mathbf{h}n}|, \varphi_{\mathbf{h}1}, \varphi_{\mathbf{h}2}, \ldots, \varphi_{\mathbf{h}n}\right)}{p_{\rm M}\left(|F_{\mathbf{h}1}|, |F_{\mathbf{h}2}|, \ldots, |F_{\mathbf{h}n}|\right)} = p_{\rm C}\left(\varphi_{\mathbf{h}1}, \varphi_{\mathbf{h}2}, \ldots, \varphi_{\mathbf{h}n}\Big||F_{\mathbf{h}1}|, |F_{\mathbf{h}2}|, \ldots, |F_{\mathbf{h}n}|\right).$$

Professor Giacovazzo's book relates in thorough and rigorous mathematical detail how, over the last 50 or so years, this theory has been and continues to be developed and applied by various research groups (collaborating and competing) internationally. No small part of that work has been done by the school of crystallographic researchers that Professor Giacovazzo leads in Bari, Italy.

The new book introduces the principles of probability distributions and statistical expectation values by reviewing the derivations by Arthur Wilson in 1942 of intensity and structure-factor statistics for hypothetical $P1$ and $P\bar{1}$ crystal structures with uniform random distributions of atomic positions. As an interesting aside: in a 2007 conversation with Herbert Hauptman, Eaton Lattman remarked that, when he had first studied the 1953 Hauptman and Karle monograph, *The Solution of the Phase Problem. 1. The Centrosymmetric Crystal*, he was surprised that the assumption of a uniform random distribution of non-interacting atoms actually worked. Hauptman replied, 'So were we.'

A commendable feature of the new book is that it pulls together under the rubric of phasing methods both small-unit-cell (mainly single-crystal but also powder diffraction) crystallography and macromolecular (mainly protein) crystallography. The book defines as 'direct methods' any procedures that yield phases directly from diffraction-amplitudes data without recourse to deconvolution of the Patterson function that maps the distribution of interatomic vector density in the unit cell. Discussion of traditional, mathematically based direct methods for small-unit-cell crystal structures occupies most of the first half of the book, while most of the second half presents discussions of the more physically based practical methods for macromolecular crystal structures.

The macromolecular methods discussed include:

SIR or MIR (<u>S</u>ingle or <u>M</u>ultiple derivative <u>I</u>somorphous <u>R</u>eplacement methods),

SAD or MAD (<u>S</u>ingle or <u>M</u>ultiple wavelength <u>A</u>nomalous <u>D</u>ispersion methods),

SIRAS (a hybrid <u>S</u>ingle derivative <u>I</u>somorphous <u>R</u>eplacement/<u>A</u>nomalous <u>S</u>cattering method),

MR (<u>M</u>olecular <u>R</u>eplacement and rotation-translation Patterson search methods), and

EDM (<u>E</u>lectron <u>D</u>ensity <u>M</u>odification methods).

All except the last of these macromolecular methods rely on analysis of the Patterson $|F|^2$ cosine Fourier synthesis. Indeed, the largest and growing majority of new biomacromolecular structure determinations entering the PDB results from applications of Patterson-based MR methods, which become stronger methods as the database of known biomolecular structures for use as Patterson search models grows. It is a fitting tribute to Lindo Patterson that most crystal structures are still being determined using the methods that Patterson himself always modestly called $|F|^2$ synthesis methods.

The macromolecular 'indirect' phasing methods are discussed at the same high level of mathematical and physical treatment as the small-unit-cell 'direct' methods proper, and the macromolecular methods are presented in language and notation consistent with the presentation of the small-unit-cell methods. One peeve is the use of the notation RES, which looks like a programmer's variable name for resolution, rather that the notation $d_{\min}$, which is standard usage among macromolecular crystallographers for the diffraction resolution limit $d_{\min} = \lambda/(2\sin\theta_{\max})$ of the amplitudes data.

The new book is laid out well; it has a wide left or right page margin on, respectively, even or odd numbered pages to accommodate tables, figures, readers notes and other marginalia; it is composed in clear, uncrowded typography for both its linguistic and mathematical text; and it is helpfully illustrated using well designed, well drafted, uncluttered, black-and-white line drawings.

One hopes that the book will find a sizeable readership. Direct-methods phasing is a mature field and its procedures have been cleverly encoded in sophisticated, user-friendly software packages. Most users of the software are primarily interested in contributing to the enormous growth of physical, chemical or biological knowledge resulting from crystallographic (and other) structural studies. Few users are inclined to delve into the mathematical theory that is the underpinning of the automated crystallographic phasing algorithms. Consequently, the field of direct-methods research is these days a rather small one with rather few researchers working on new developments. Indeed, a fair fraction of the number of workers presently doing direct-methods research are or have been working in Professor Giacovazzo's Instituto di Cristallografia in Bari.

References

Giacovazzo, C. (1998). *Direct Phasing in Crystallography: Fundamentals and Applications*. IUCr Monographs on Crystallography, No. 8. International Union of Crystallography/Oxford University Press. [Chs. 14 and 17 touch upon macromolecular crystallography.]

Giacovazzo, C. (2008). *Direct methods*. In *International Tables for Crystallography*, Vol. B, *Reciprocal Space*, 3rd ed., edited by U. Shmueli. Heidelberg: Springer.

Giacovazzo, C., Monaco, H. L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, G. & Catti, M. (1992). *Fundamentals of Crystallography*. IUCr Texts on Crystallography, No. 7. International Union of Crystallography/Oxford University Press.

Robert H. Blessing

Hauptman–Woodward Institute, 700 Ellicott Street, Buffalo, New York 14203, USA