SHORT COMMUNICATIONS

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 *words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible.*

Acta Cryst. (1993). A**49**, 590

Kitajgorodskij's categories. Corrigenda.

By A. J. C. WILSON, *Crystallographic Data Centre, University Chemical Laboratory, Cambridge CB2 1EW, England*

(*Received* 25 *January* 1993)

Abstract

Two errors in the paper by Wilson [*Acta Cryst.* (1993), A**49**, 210–212] should be corrected. In the heading of the final column of Table 1, s/mP is a misprint for $2/mP$, and in line 24 of the second column on page 212, *Fdd* is a misprint for *Fddd*.

All relevant information is given in the *Abstract*.

Acta Cryst. (1993). A**49**, 590–591

The maximum-entropy method in charge-density studies: aspects of reliability.

By W. JAUCH and A. PALMER, *Hahn-Meitner-Institut, Glienicker Strasse* 100, *D*-1000 *Berlin* 39, *Germany*

(*Received* 19 *May* 1992; *accepted* 13 *October* 1992)

Abstract

The maximum-entropy method (MEM) was applied to accurate γ-ray diffraction data from MnF_2 and NiF_2 to explore details of the charge-density distribution. For a fair judgement of the results, Si *Pendellösung* data [Saka & Kato (1986). *Acta Cryst.* A**42**, 469–478] were also treated. It is shown that conclusions drawn from MEM maps must be accepted with some reserve, particularly in the regions of interest in charge-density studies.

1. Introduction

In recent years, high expectations have been raised by the application of the maximum-entropy method (MEM) to the crystallographic inversion problem. Here, MEM is considered to be promising for the following reasons: the results are independent from structural models; the reconstructed maps must be positive; and Fourier components may be extrapolated from incomplete data sets. In particular, it has been suggested that MEM would be well suited for an accurate determination of the electron density distribution (Wei, 1985; Gull, Livesey & Sivia, 1987). Up to now, however, there have been only a few demonstrations of its use (Sakata & Sato, 1990). Therefore, the present note will concentrate on some practical aspects of the method and on the reliability of its results.

The conditional entropy of a map $\{p_i\}$ with respect to the map $\{m_i\}$ is given by

$$H(\{p_i\};\{m_i\}) = -\sum p_i \ln (p_i/m_i),$$

where $p_i = \rho_i/\sum \rho_i$ is the proportion of the electron density ρ in pixel i. The m_i represent the initial density and may be based on prior knowledge but are usually all set equal. The constraint function is given by

$$\chi^2 = \sum_{k=1}^{N} (|F_o| - |F_c|)^2/\sigma_k^2,$$

where $|F_c|$ is the structure-factor magnitude calculated from $\{\rho_i\}$, σ_k^2 is the variance of the kth observed structure factor $|F_o|$ and N is the number of observations. $\{\rho_i\}$ is to be determined so as to maximize H subject to the constraints $\chi^2 = N$ and a fixed number of electrons per unit cell. In addition, the signs of the structure factors are assumed to be known. The density obtained from MEM is an exponential.

The method was applied to extended γ-ray diffraction data sets from MnF_2 ($N = 324$; Jauch, Schultz & Schneider, 1988) and NiF_2 ($N = 298$; Palmer & Jauch, 1993). To investigate any potential influence of limited data quality on the results, MEM was also applied to highly accurate *Pendellösung* data from Si ($N = 30$; Saka & Kato, 1986), which was the subject of another study (Sakata & Sato, 1990). In the present work, the program *MEED* (Sakata, Mori, Kumazawa, Takata & Toraya, 1990) was used. Normally, the unit cell was divided into 64^3 pixels and convergence was reached after about 10^3 iterations.

2. Non-random scatter of residuals

In all cases studied, a highly non-random distribution in the contributions of individual reflections to χ^2 was found after convergence. The value of χ^2 is dominated by a small number of reflections, whereas the other observations are reproduced almost perfectly. As a consequence, the R factors are much smaller than their expected values based on the standard deviations of the data.

0108-7673/93/030590-02$06.00

For MnF_2, just one low-order reflection contributes 45% to the constraint function. This fact persists if the 'outlier' is omitted. In the remaining subset, a different observation, which fitted excellently before, now plays the role of the 'outlier'. With the Si data a contribution of 87% to χ^2 arises solely from the 111 and 220 structure factors.

It has been proposed that $\chi^2 = N$ is not the correct stopping criterion; values smaller than N are more appropriate (Gull, 1989). However, we did not observe a substantial adjustment of the residuals by tightening the convergence criterion. With the Si data, for example, the combined contribution from 111 and 220 still amounts to 64 and 49% if χ^2 is reduced to $N/10$ and $N/100$, respectively.

It may be argued that these apparent 'outliers' are due to the assumption of a uniform initial map where $F_c = 0$ for any reflection. Hence, the discrepancies are large for strong reflections and thus they might be selected to be 'outliers' at the beginning of the iteration algorithm. It would seem preferable to start from a more realistic density distribution since the initial dynamic range of the χ^2 components is then largely reduced. To test this conjecture, we examined the influence of non-uniform initial distributions in the case of Si.

3. Influence of the initial distribution

The entropy is maximized with respect to the shape of the initial distribution and thus even small deviations from uniform initial maps can lead to large effects in the results. If, for example, in the initial map the density at the pixel of the Si site is increased from 0.7 to 1.0 e $Å^{-3}$, the final map shows atomic peaks each with a sharp spike at the center (see, for example, Papoular & Livesey, 1989).

An obvious initial distribution is a Gaussian, embedded in a non-zero background density. Then there are two tunable parameters, the r.m.s. width and the peak-to-background ratio. We explored various combinations but none of them resulted in more uniform residuals. The atomic peaks in the final maps are hardly affected by the choice of different initial parameters, whereas remarkable changes show up in the bonding region. Therefore, any inference from low-density regions of MEM maps has to be drawn with caution, unless the special role of maximum prior ignorance can be proven unambiguously. Furthermore, it should be emphasized that the 'outliers' generally are low-order reflections, which are most important for the valence density.

4. Extrapolation of structure factors

Extrapolation or prediction of unmeasured structure factors from a limited amount of data is considered as a particular strength of the MEM. Omission of the 111 reflection from the Si data set, for example, yields $F_c(111) = 46.9$ and $F_c(222) = 0.06$ as compared to the observed values of 60.1 and 1.5, respectively. Generally, the absence of strong reflections leads to reduced atomic peak heights and, in consequence, to rather poor extrapolations.

MEM maps of MnF_2 and NiF_2 calculated from a limited $\sin\theta/\lambda$ range result in extrapolated high-order structure factors that are substantially underestimated. Neglected reflections within the considered $\sin\theta/\lambda$ range, however, are reproduced more reliably. Acceptance of the most precisely measured reflections only, *e.g.* $\sigma/F \leq 0.01$ (about $\frac{1}{4}$ of the data), leads to the appearance of ghost peaks randomly distributed in the cell. These features have a height of a few e $Å^{-3}$ and become more pronounced as the σ/F threshold is lowered.

5. Concluding remarks

The MEM emphasizes any sharp features occurring in a map. The quite subtle effects of chemical bonding and crystal environment on the atomic charge density, however, generally show up as smooth features in the low-density region. From our experience, it is particularly in this region of interest that the MEM maps tend to be contaminated by artifacts and should therefore be interpreted with great caution. The potential shortcomings inherent in the generality of the MEM were anticipated by Collins & Mahar (1983), who stressed that non-physical features in a map will persist as long as they are not in disagreement with the data. Obviously, a MEM analysis does not provide a suitable alternative to least-squares refinements relying on elaborate physical models.

We wish to thank Drs Makoto Sakata and Masaki Takata for providing us with the *MEED* program and Dr Detmar Welz for stimulating discussions.

References

COLLINS, D. M. & MAHAR, M. C. (1983). *Acta Cryst.* A**39**, 252-256.

GULL, F. S. (1989). *Maximum Entropy and Bayesian Methods*, edited by J. SKILLING, pp. 53-71. Dordrecht: Kluwer Academic Publishers.

GULL, F. S., LIVESEY, A. K. & SIVIA, D. S. (1987). *Acta Cryst.* A**43**, 112-117.

JAUCH, W., SCHULTZ, A. J. & SCHNEIDER, J. R. (1988). *J. Appl. Cryst.* **21**, 975-979.

PALMER, A. & JAUCH, W. (1993). In preparation.

PAPOULAR, R. J. & LIVESEY, A. K. (1989). *Maximum Entropy and Bayesian Methods*, edited by J. SKILLING, pp. 163-173. Dordrecht: Kluwer Academic Publishers.

SAKA, T. & KATO, N. (1986). *Acta Cryst.* A**42**, 469-478.

SAKATA, M., MORI, R., KUMAZAWA, S., TAKATA, M. & TORAYA, H. (1990). *J. Appl. Cryst.* **23**, 526-534.

SAKATA, M. & SATO, M. (1990). *Acta Cryst.* A**46**, 263-270.

WEI, W. (1985). *J. Appl. Cryst.* **18**, 442-445.