Acta Cryst. (1996). A**52**, 287–290

The Influence of the Completeness of the Data Set on the Charge Density Obtained with the Maximum-Entropy Method. A Re-examination of the Electron-Density Distribution in Si

MASAKI TAKATA AND MAKOTO SAKATA

Department of Applied Physics, Nagoya University, Nagoya 464-01, Japan

(Received 17 July 1995; accepted 30 October 1995)

Abstract

The charge densities derived with the maximum-entropy method (MEM) may be influenced to some extent by the completeness of the data set. In order to examine the effects of the incompleteness, structure-factor data of Si measured by the *Pendellösung* method [Saka & Kato (1986). *Acta Cryst.* A**42**, 469–478] were re-analysed by the MEM. This data set is incomplete: it contains all space-group-allowed reflections with $\sin\theta/\lambda = 0.86\,\text{Å}^{-1}$, and in addition 844 and 880 with $\sin\theta/\lambda = 1.04\,\text{Å}^{-1}$. Results of a MEM analysis of the complete subset of data are compared with those from the full but incomplete set published previously [Sakata & Sato (1990). *Acta Cryst.* A**46**, 263–270]. The smaller but complete set was found to give a smooth charge-density distribution that is consistent with previous theoretical work. It is found that the sharp peak maximum at the bond midpoint reported previously is exaggerated owing to the highest-order reflection 880. The completeness of the data set appears to be one of the key factors for obtaining reliable charge densities with MEM. The incompleteness of the data set may cause non-physical fine features of the MEM density distribution.

1. Introduction

The maximum-entropy method (MEM) is a statistical deduction (Bricogne, 1988) that can yield a high-resolution density distribution from a limited number of diffraction data without using a structural model (Collins, 1982; Sakata & Sato, 1990). It has been suggested that the MEM would be a suitable method for examining electron densities in the interatomic region, *e.g.* bonding densities. The first application of the MEM to a charge-density study has been done for Si and demonstrated the great power of the MEM in reducing truncation effects (Sakata & Sato, 1990). Since then, charge-density studies using X-ray diffraction data have been carried out for several substances and reveal some new aspects of the nature of bonding (Sakata, Uno, Takata & Mori, 1992; Takata, Kubota & Sakata, 1993; Kubota, Takata & Sakata, 1993).

The MEM allows a model-free reconstruction of the charge densities from measured X-ray diffraction data. Therefore, the reliability of the data is very important and will directly influence the results. Significant systematic errors in the observed data are expected to artificially deform the resultant MEM density. They will always tend to deteriorate the MEM map. Recently, Jauch (1994) has pointed out that MEM maps are susceptible to exhibiting similar artifacts to those inherent in Fourier inversion depending on data completeness, error accumulation at special positions *etc.* Consequently, it is very important to reduce all kinds of systematic errors as much as possible in order to construct an 'accurate' electron density from observed data with the MEM. Recently, it has been claimed that non-physical fine features found in the MEM density are probably due to the non-uniform residual distributions of χ^2 (Jauch & Palmer, 1993). From the simulation using the model charge density of a hypothetical crystal, De Vries, Briels & Feil (1994) have proposed a weighting scheme that leads to more uniform residual distributions and gives a smoother density. It is plausible that many factors may influence the MEM charge densities. It is necessary to study all these factors as much as possible before reaching a full understanding of the meaning of the MEM charge densities. In this paper, the influence of the completeness of the data set on the result of the MEM analysis is investigated by using the data set of Si (Saka & Kato, 1986) as part of the efforts towards a better understanding of MEM charge densities.

2. The MEM charge density of Si obtained previously

The first MEM charge density of Si (Sakata & Sato, 1990) was obtained from the highly accurate *Pendellösung* data measured by Saka & Kato (1986) and is reproduced in Fig. 1. De Vries *et al.* (1994) also applied their weighting scheme to the same data, but they did not arrive at a significant improvement and obtained almost the same density. However, the density in Fig. 1 also has 'non-physical fine features' such as artificial contamination in the MEM maps as pointed out in previous studies (Jauch & Palmer, 1993; Jauch, 1994; De Vries, Briels & Feil, 1994). Furthermore, the sharp maximum at the bond midpoint has been argued to be inconsistent with previous theoretical (*e.g.* Wang & Klein, 1981; Yin & Cohen, 1983) and experimental (Spackman, 1986)

©1996 International Union of Crystallography
Printed in Great Britain – all rights reserved

valence densities. In contrast, the MEM density of Be based on single-crystal data (Larsen & Hansen, 1984) is quite smooth and such fine features do not appear (Takata, Sakata, Kumazawa, Larsen & Iversen, 1994; Iversen, Larsen, Souhassou & Takata, 1995). In order to gain an insight into the cause of the difference in the quality of the MEM densities, the observed structure factors, F_{obs}, are listed in Table 1 together with the calculated ones, F_{MEM}, corresponding to Fig. 1. Evidently, the original data set is incomplete: reflection 931 is missing and ten reflections from 771 to 775 are also missing. On the other hand, the Be data set is complete ($\sin\theta/\lambda < 1.21$ Å^{-1}). Therefore, it may be suspected that the non-physical features in the Si MEM density are caused by the incompleteness of the data set in spite of the fact that each observed data point is of superb quality.

3. The influence of the incomplete data set

In order to see the influence of the completeness in the data set on the MEM density, we reanalysed the complete set up to 664 as listed in Table 1 with the computer program *MEED* (Kumazawa, Kubota, Takata, Sakata & Ishibashi, 1993). The number of pixels used is $120 \times 120 \times 120$. The newly obtained charge density is shown in Fig. 2. This density is very smooth and the fine features in Fig. 1 disappear. In order to investigate the origin of the difference between Fig. 1 and Fig. 2, the calculated scattering factors per atom, $F(hkl)/C(hkl)$ [$C(hkl)$ is the trigonometric factor], are plotted *versus* $\sin\theta/\lambda$ in Fig. 3 for both the MEM densities and the observations. For the map from the full but incomplete data set, the calculated value of 880 shows very good agreement with the observed one. However, it is found that the calculated value of $F(880)$ is relatively large compared to the general decrease of the scattering factor. For the complete set, $F(880)$ agrees with the trend although its calculated value is 6.5% smaller than the observed one. The discrepancy between the $F_{MEM}(880)$ of the complete set and that of the incomplete set is about 0.8. This is by far larger than the discrepancy of 0.1 between $F_{obs}(111)$ and $F_{MEM}(111)$, which has been claimed to be the cause of the non-physical features in the resultant MEM density (Jauch & Palmer, 1993). In this case, it can be conjectured that the information up to 664 is dominant in the process of the MEM analysis and the lack of information of the ten missing reflections from 771 to 775 is too important for inferring an appropriate value for $F(880)$. In the MEM charge density of Fig. 1, the fine features appear to be wavy and periodic. There are eight periods along [110]. This coincides with the spacing of the (880) plane.

In Fig. 4, the one-dimensional charge density between the Si atoms along the bonding direction is shown for both the complete and the incomplete data sets on a

Table 1. *Observed, F_{obs}, and calculated, F_{MEM}, structure factors for the MEM results based on the incomplete and complete data sets*

			Incomplete set		Complete set	
h	k	l	F_{obs}	F_{MEM}	F_{obs}	F_{MEM}
1	1	1	60.13 (5)	59.97	60.13 (5)	59.98
2	2	0	67.34 (5)	67.52	67.34 (5)	67.52
3	1	1	43.63 (3)	43.61	43.63 (3)	43.61
4	0	0	56.23 (4)	56.22	56.23 (4)	56.22
3	3	1	38.22 (3)	38.23	28.22 (3)	38.23
4	2	2	49.11 (3)	49.06	49.11 (3)	49.06
5	1	1	32.94 (2)	32.95	32.94 (2)	32.95
3	3	3	32.83 (2)	32.84	32.83 (2)	32.84
4	4	0	42.88 (3)	42.88	42.88 (3)	42.88
5	3	1	28.81 (2)	28.82	28.81 (2)	28.82
6	2	0	37.59 (6)	37.53	37.59 (6)	37.53
5	3	3	25.36 (4)	25.35	25.36 (4)	25.35
4	4	4	33.18 (5)	33.19	33.18 (5)	33.18
7	1	1	22.37 (3)	22.37	22.37 (3)	22.37
5	5	1	22.42 (3)	22.41	22.42 (3)	22.41
6	4	2	29.42 (4)	29.43	29.42 (4)	29.44
5	5	3	19.98 (3)	19.97	19.98 (3)	19.97
7	3	1	19.90 (3)	19.90	19.90 (3)	19.91
8	0	0	26.23 (4)	26.23	26.23 (4)	26.23
7	3	3	17.83 (3)	17.83	17.83 (3)	17.83
8	2	2	23.48 (4)	23.49	23.48 (4)	23.49
6	6	0	23.48 (4)	23.49	23.48 (4)	23.49
7	5	1	15.98 (2)	15.98	15.98 (2)	15.98
5	5	5	15.98 (2)	15.98	15.98 (2)	15.98
8	4	0	21.15 (3)	21.15	21.15 (3)	21.15
9	1	1	14.46 (2)	14.46	14.46 (2)	14.46
7	5	3	14.43 (2)	14.43	14.43 (2)	14.43
6	6	4	19.13 (3)	19.13	19.13 (3)	19.13
9	3	1		12.97		12.93
8	4	4	17.44 (3)	17.44		17.22
7	7	1		11.91		11.67
9	3	3		11.78		11.68
7	5	5		11.86		11.75
10	2	0		15.65		15.59
8	6	2		15.75		15.55
9	5	1		10.70		10.59
7	7	3		10.75		10.61
9	5	3		9.80		9.61
10	4	2		12.95		12.77
7	7	5		8.82		8.75
8	8	0	12.41 (2)	12.41		11.61

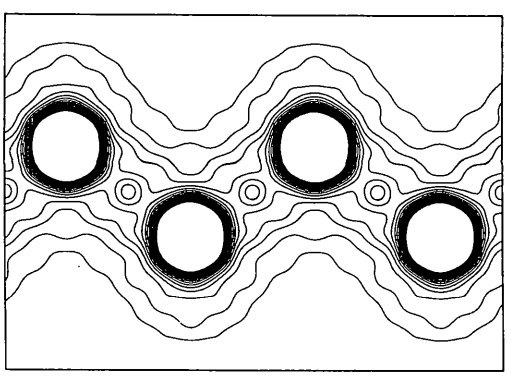

Fig. 1. The (110) MEM charge density of Si based on the full set of data measured by Saka & Kato. The contour lines are drawn from 0.0 to 2.0 e Å^{-3} with a step of 0.1 e Å^{-3}.

logarithmic scale. It is clearly seen that the new density based on the complete set shows much smaller bond midpoint maxima than the previous one. The period of the modulation in Fig. 2 coincides with the distance along [110] between the atomic site and the bond midpoint. From the above consideration, it is concluded that the fine features in the MEM map of Fig. 1 are non-physical artificial modulations caused by the failure of the MEM to extrapolate over the gap of missing data from 931 to 880 in the case of the incomplete set and that the maxima at the bond midpoint in the previous charge density are exaggerated artificially owing to the incompleteness of the data set used for the MEM.

4. The value of the forbidden $F(222)$

In the present study, the complete data set modifies the feature of the bonding charge as mentioned above. Therefore, it is expected that the value of $F(222)$ for the new MEM charge density should be different from the previous one since the forbidden $F(222)$ is well known to emanate from the antisymmetric distribution of the charge density owing to the bonding nature or the anharmonic thermal vibrations.

The MEM enables us to predict the unobserved structure factors. This is one of the differences from the conventional Fourier method. The $F(222)$ value for Fig. 2, 1.550, is found to be not significantly different from that of Fig. 1, 1.547. At the bond midpoint, the present charge density peaks at $0.61\,e\,Å^{-3}$ and shows close agreement with previously reported experimental and theoretical maps (Wang & Klein, 1981; Yin & Cohen, 1983; Spackman, 1986), except for details such as the twin peak feature reported in the static valence density maps. Although the maximum of $\sim 0.77\,e\,Å^{-3}$ in the previous density represents a different value, the base part of the bond charge appears to be identical to that of the new density as seen in Fig. 4. $F(222)$ is entirely the Fourier transform of the antisymmetric component of the charge density. Therefore, it seems reasonable to suppose that the artificial modulation in Fig. 1 does not contribute seriously to the antisymmetric component of the charge density as a whole.

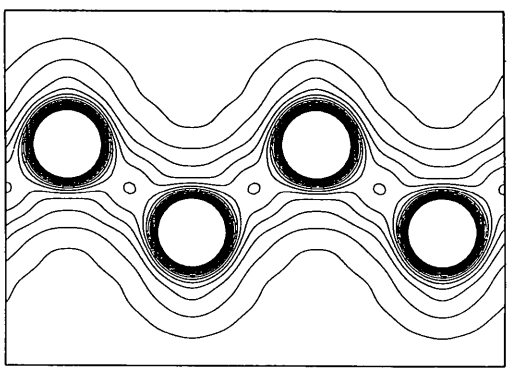

Fig. 2. The (110) MEM charge density of Si based on the data set complete out to 664. The contour lines are drawn from 0.0 to $2.0\,e\,Å^{-3}$ with a step of $0.1\,e\,Å^{-3}$.

5. Discussion

The observed value of $F(880)$ is likely to be accurate since the other experimental values (Aldred & Hart, 1973; Teworte & Bonse, 1984) are all well accounted for. However, the predicted value of $F(880)$ in the MEM case applied to the complete set is smaller than the observed one. In the MEM study, it is difficult to predict the appropriate value of $F(880)$ by extrapolating the structure factors from $F(931)$ to $F(880)$.

In terms of the reconstruction of the diffraction data, the structure factors of lower-order reflections, *i.e.* forward scattering, include considerable contributions from the weak charge density in the interatomic region. On the other hand, high-order reflections mainly

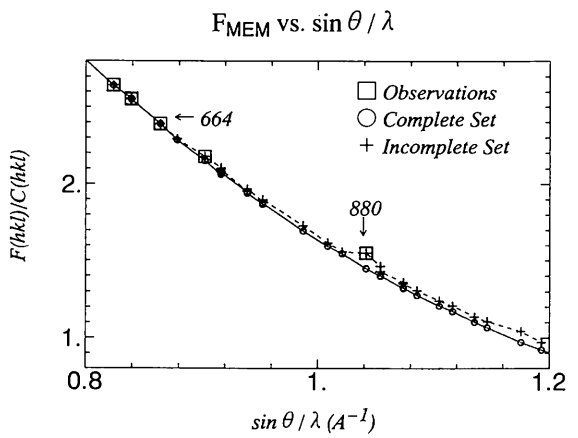

Fig. 3. The observed and calculated values of $F(hkl)/C(hkl)$ against $\sin\theta/\lambda$. $C(hkl)$ is the trigonometric factor [$C = 8$ (even order), $4 \times 2^{1/2}$ (odd order)].

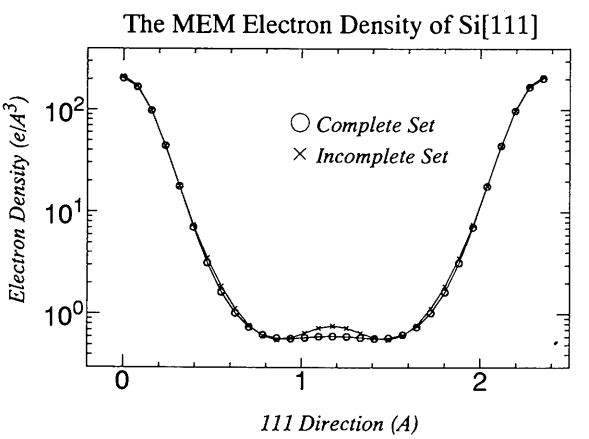

Fig. 4. The one-dimensional charge densities of Figs. 1 and 2 are plotted on a logarithmic scale along the bonding direction between the atoms.

contribute to the charge density of the core region of the atom. In Fig. 3, the MEM analysis using the complete set appears to underestimate the scattering factors for the higher-order reflections than those in the incomplete set. This effect is known to lower the peak height at the atomic position in spite of the high resolution achievable in principle by MEM. As has been pointed out by Sakata & Sato (1990), in order to obtain a good estimate of the true electron density at an atomic position, we have to measure the very high order reflections.

6. Concluding remarks

The completeness of the data set is one of the key factors in the MEM analysis to produce expected smooth features of the charge density even in the lower-density region. From the viewpoint of model-free reconstruction of diffraction data, it can be suggested that incompleteness in the data set would act as a kind of systematic error. It is, therefore, recommended to measure a complete set for accurate charge-density studies by the MEM.

The authors are grateful to Professor N. Kato for helpful comments. They also acknowledge the partial support of this work by the Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

References

Aldred, P. J. E. & Hart, M. (1973). *Proc. R. Soc. London Ser. A*, **332**, 223–254.
Bricogne, G. (1988). *Acta Cryst.* A**44**, 517–545.
Collins, D. M. (1982). *Nature (London)*, **298**, 49–51.
De Vries, R. Y., Briels, W. J. & Feil, D. (1994). *Acta Cryst.* A**50**, 383–391.
Iversen, B. B., Larsen, F. K., Souhassou, M. & Takata, M. (1995). *Acta Cryst.* B**51**, 580–591.
Jauch, W. (1994). *Acta Cryst.* A**50**, 650–652.
Jauch, W. & Palmer, A. (1993). *Acta Cryst.* A**49**, 590–591.
Kubota, Y., Takata, M. & Sakata, M. (1993). *J. Phys. Condens. Matter*, **5**, 8245–8254.
Kumazawa, S., Kubota, Y., Takata, M., Sakata, M. & Ishibashi, Y. (1993). *J. Appl. Cryst.* **26**, 453–457.
Larsen, F. K. & Hansen, N. K. (1984). *Acta Cryst.* A**40**, 169–179.
Saka, T. & Kato, N. (1986). *Acta Cryst.* A**42**, 469–478.
Sakata, M. & Sato, M. (1990). *Acta Cryst.* A**46**, 263–270.
Sakata, M., Uno, T., Takata, M. & Mori, R. (1992). *Acta Cryst.* B**48**, 591–598.
Spackman, M. (1986). *Acta Cryst.* A**42**, 271–281.
Takata, M., Kubota, Y. & Sakata, M. (1993). *Z. Naturforsch. Teil A*, **48**, 75–80.
Takata, M., Sakata, M., Kumazawa, S., Larsen, F. K. & Iversen, B. B. (1994). *Acta Cryst.* A**50**, 330–337.
Teworte, R. & Bonse, U. (1984). *Phys. Rev. B*, **29**, 2102–2108.
Wang, C. S. & Klein, B. M. (1981). *Phys. Rev. B*, **24**, 3393–3416.
Yin, M. T. & Cohen, M. L. (1983). *Phys. Rev. Lett.* **50**, 1172.