

## CRYSTAL STRUCTURE ANALYSIS WITH THE MAXIMUM ENTROPY METHOD

RAMESH NARAYAN AND RAJARAM NITYANANDA

Raman Research Institute, Bangalore 560 080, India

### ABSTRACT

The application of the Maximum Entropy Method (MEM) to crystal structure analysis is investigated. The phase problem leads to new higher order phase relations. Simple one-dimensional simulations are presented where the method solves the structure from initial random phases. It is argued that the peak-seeking nature of the MEM is ideally suited to crystallographic applications. It is suggested that there is an intimate relation between the MEM and the maximum determinant method.

**I**N many fields of experimental physics, one measures the Fourier transform of the function of interest. It also invariably happens that only a part of the Fourier information is available. Therefore, an important problem is to extract the maximum information possible about the function of interest from partial Fourier data. Among the many schemes that have been developed for this, the so-called Maximum Entropy Method (MEM) has evoked much interest in the fields of geophysics<sup>1-3</sup> and radio astronomy<sup>4,5</sup>. In this paper we suggest that the MEM could be gainfully applied to the phase problem of crystal structure analysis. We argue that, in many respects, the MEM is ideally suited to crystallographic applications. We present some simple, one-dimensional simulations to bear this out.

We first review briefly relevant features of the MEM. Let  $\rho(r)$  be a real scalar function of a many-dimensional vector  $r$ . Let  $\rho(r)$  be either (a) non-zero over only a finite closed volume or (b) periodically repeated over a space lattice. In either case, the Fourier transform of  $\rho(r)$  is completely specified by giving its values  $F_{h_j}$  at a discrete lattice of points  $h_j$  in reciprocal space. By definition

$$F_{h_j} = \int \rho(r) \exp(2\pi i h_j \cdot r) dV_r, \quad (1)$$

which leads to the inverse relation

$$\rho(r) = \frac{1}{V_r} \sum_{h_j} F_{h_j} \exp(-2\pi i h_j \cdot r); \quad h_j \text{ on a lattice} \quad (2)$$

where  $V_r$  is the volume of the "unit cell" of  $\rho(r)$ .

Let us suppose, as is the case in many applications in geophysics and radio astronomy, that  $F_{h_j}$  is available both in amplitude and phase at some of the  $h_j$  but is unknown at the rest of the points. A simple-minded approach would set the unknown  $F_{h_j}$  to zero and compute  $\rho(r)$  through equation (2). The resultant function would generally suffer from loss of resolution and would also have termination ripple. The MEM instead

prescribes that the unknown  $F_{h_j}$  should be chosen so as to maximise the "entropy"  $S$  defined below.

$$S = \int F[\rho(r)] dV_r. \quad (3)$$

$F$  is a suitable real function of  $\rho(r)$  whose properties are discussed later. Substituting (2) in (3) and differentiating with respect to the unknown  $F_{h_j}$ , one obtains that

$$\int F'[\rho(r)] \exp(-2\pi i h_j \cdot r) dV_r = 0 \quad \text{for unknown } F_{h_j} \quad (4)$$

where the prime denotes differentiation with respect to the scalar argument. If  $G_{h_j}$  are the Fourier coefficients of  $F'[\rho(r)]$ , eq. (4) implies that

$$G_{h_j} = 0 \quad \text{for unknown } F_{h_j}. \quad (5)$$

We note that eqs. (5), which are well known<sup>2</sup>, generate as many conditions to be satisfied as there are unknowns. For a suitable choice of  $F$ , the solution is unique<sup>3,6</sup>. Iterative numerical algorithms can be developed<sup>6</sup> which seek to achieve the conditions in (5). In general, the MEM solution generates non-zero values for the unknown  $F_{h_j}$ . Thus, one reconstructs a  $\rho(r)$  with improved resolution as well as reduced ripple.

The results of the MEM depend on the choice of the function  $F$ . The following properties of  $F$  have been identified as crucial<sup>6</sup>:

- (a)  $F''$  should always be negative.
- (b)  $F'''$  should always be positive.

Traditionally, the following two forms of entropy have been studied, both of which satisfy the above requirements:

$$(a)^{1-5} \quad S_1 = \int \ln[\rho(r)] dV_r, \quad (6)$$

By (5) this leads to the result that  $[\rho(r)]^{-1}$  is "band-limited", i.e.,  $[\rho(r)]^{-1}$  has only a finite number of non-zero Fourier coefficients (they are in fact located at the  $h_j$  corresponding to the measured  $F_{h_j}$ ).

$$(b)^{7,8} \quad S_2 = - \int \rho(r) \ln[\rho(r)] dV_r. \quad (7)$$

This requires that  $\ln[\rho(r)]$  be band-limited.

The MEM as developed so far cannot be directly applied to crystal structure analysis because of the phase problem. It is therefore necessary to generate new conditions, in addition to (5), corresponding to the unknown phases. Substituting (2) in (3) and differentiating with respect to the unknown phases of the measured amplitudes, one obtains

$$\int F'[\rho(r)] [F_{h_j} \exp(-2\pi i h_j \cdot r) - F_{-h_j} \exp(2\pi i h_j \cdot r)] dV_r = 0, \quad (8)$$

where we have employed the property that  $F_{h_j}$  and  $F_{-h_j}$  have opposite phases as a consequence of  $\rho(r)$  being real. Eq. (8) leads to the result that

$$G_{-h_j} F_{h_j} - G_{h_j} F_{-h_j} = 0. \quad (9)$$

Since  $\rho(r)$  and  $F'[\rho(r)]$  are real, inversion related Fourier coefficients are complex conjugates of each other. Hence

$$G_{-h_j} F_{h_j} = (G_{h_j} F_{-h_j})^*. \quad (10)$$

This implies that  $G_{-h_j} F_{h_j}$  is real and therefore

$$\text{ph}(F_{h_j}) - \text{ph}(G_{h_j}) = 0 \text{ or } \pi \quad (11)$$

where "ph" means "the phase of". Eqs. (11) generate one phase relation corresponding to every

unknown phase. It is possible to introduce eqs. (11) into the numerical algorithm proposed in Ref. 6. We can thus implement the MEM in problems of crystal structure analysis.

We have carried out simulations on simple one-dimensional "structures" using the "entropy"  $S_1$  defined in (6). Typical results on a centrosymmetric and a non-centrosymmetric structure are shown in Figs. 1 and 2. In both cases, the algorithm was initially given the correct amplitudes of the structure factors and *random* phases. As it happened, the initial random phase map of the non-centrosymmetric structure was nearly centrosymmetric and *vice versa* [Figs. 1(b) and 2(b)]. The converged solutions [Figs. 1(c) and 2(c)] are surprisingly close to the true structures. There is, of course, an origin shift in both solutions and the absolute configuration of Fig. 2(c) is opposite to that of 2(a). These arise because of the random nature of the initial phases.

It should be noted that all centrosymmetric functions automatically satisfy the phase relations (11). Therefore, to obtain non-centrosymmetric solutions, it is necessary to take suitable precautions. These details, as well as the algorithm employed, are being published elsewhere.

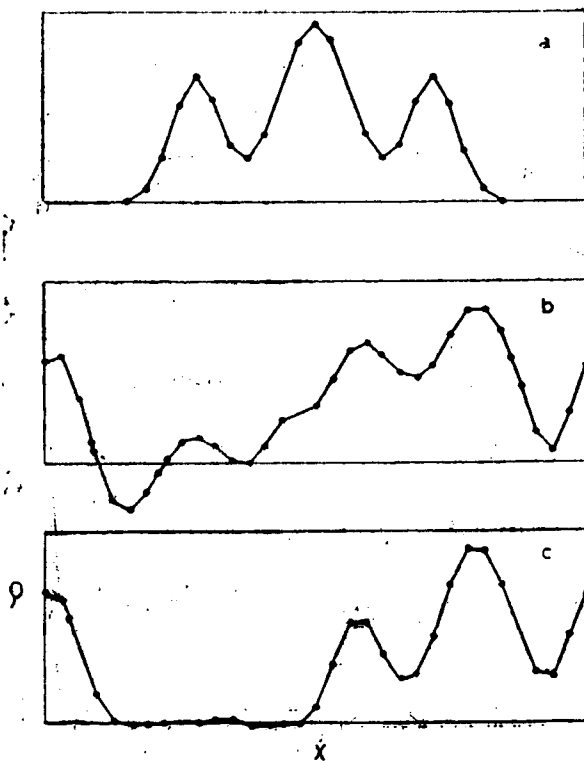


FIG. 1. (a) Model one-dimensional centrosymmetric structure. (b) Structure with the same Fourier amplitudes as in (a) but random phases. (c) Result of the MEM algorithm starting with (b).

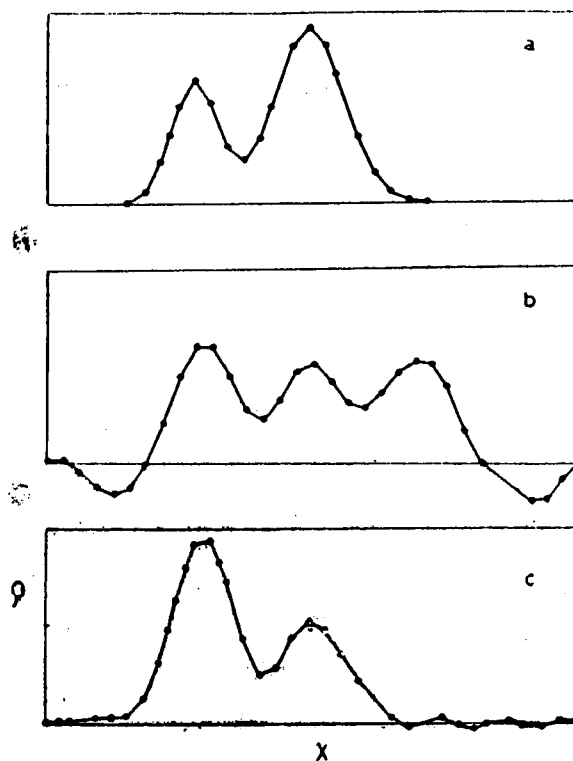


FIG. 2. (a) Model one-dimensional non-centrosymmetric structure. (b) Structure with the same Fourier amplitudes as in (a) but random phases. (c) Result of the MEM algorithm starting with (b).

The general experience of the authors<sup>5</sup> with the MEM is that it is ideally suited for certain classes of applications. Whenever  $\rho(r)$  is close to zero (or any constant value) over a large range of  $r$  and is positive and "peaky" over a relatively smaller range of  $r$ , the restorations obtained with the MEM are very good. In particular, the *positions* and (to a lesser extent) the *strengths* of peaks are reproduced very well. On the other hand, the *shapes* of the peaks are usually not reliable. Also, if  $\rho(r)$  has "plateaus" *i.e.*, fairly extensive regions of approximately constant "height", the MEM restorations have excessive ripple on the plateaus.

It is remarkable that the weaknesses of the MEM are quite immaterial for crystallographic applications. What is important in crystal structure analysis is the ability to pinpoint the positions of the peaks, and this is precisely what the MEM is most efficient at (once the peaks have been identified, chemical information and least squares usually take over to refine the structure). This coupled with our earlier experience<sup>5,6</sup> leads us to believe that the MEM is ideally suited for crystal structure analysis. Figs. 1 and 2 support this view. We should mention that reconstructions from radio astronomical data without phases have been presented by Gull and Daniell<sup>8</sup> using  $S_2$  and the method of least squares.

In the one-dimensional case, it has been noted<sup>3,9,5</sup> that maximisation of  $S_1$  is equivalent to maximising a

certain determinant built up from structure factors. This is strongly reminiscent of the maximum determinant rule<sup>10,11</sup> which has been studied in connection with the phase problem. We are currently investigating this relationship.

1. Burg, J. P., Paper presented at the Annual International Society of Exploration Geophysicists Meeting, Oklahoma City, Oklahoma, 1967.
2. —, Thesis *Maximum Entropy Spectral Analysis*, Department of Geophysics, Stanford University, 1975.
3. Smylie, D. E., Clarke, G. K. C. and Ulrych, T. J., In *Methods in Computational Physics*, Academic Press, New York, 1973, 13.
4. Ables, J. G., *Astron. Astrophys. Suppl.*, 1974, 15, 383.
5. Komesaroff, M. M., Narayan, R. and Nityananda, R., *Astron. and Astrophys.*, 1981 (In press).
6. Nityananda, R. and Narayan, R., to be published.
7. Frieden, B. R., *J. Opt. Soc. Am.*, 1972, 62, 511.
8. Gull, S. F. and Daniell, G. J., *Nature (London)*, 1978, 272, 686.
9. van den Boss, *IEEE Trans. Inform. Theory*, 1971, IT-17, 493.
10. Lajzerowicz, J. and Lajzerowicz, J., *Acta Cryst.*, 1966, B21, 8.
11. Tsoucaris, G., *Ibid.*, 1970, A26, 492.