

APPENDIX 1

Discussion of the transcendental equation

Consider

$$\ln(p) + up + v = 0 \quad (A1.1)$$

which can be written in reduced form as

$$\bar{p} = \exp\{-\bar{u}\bar{p}\}, \quad (A1.2)$$

where

$$\bar{p} = p \exp\{v\}; \quad \text{and} \quad \bar{u} = u \exp\{-v\}. \quad (A1.3)$$

Using simple algebra it can readily be shown (see Fig. 1) that (A1.3) has:

- (i) one solution if $u \geq 0$.
- (ii) two solutions if $-\exp\{-1\} < u \exp\{-v\} < 0$.
- (iii) no solutions if $u < -\exp\{v-1\}$.

The coordinates of the critical point (\bar{u}_c, \bar{p}_c) are $(-e^{-1}, e)$.

In practice, (A1.1) can be solved for p by numerical iteration for p using Newton's method, *viz.*

$$p^{(1)} = [1 + up^{(0)}] \exp\{-up^{(0)} - v\} \\ \times [1 + u \exp\{-up^{(0)} - v\}]^{-1}. \quad (A1.4)$$

APPENDIX 2
Determination of λ_0

In practice, we have found it adequate to determine λ_0 in the SPE by successive iteration using Newton's

method. If $\rho^{(0,0)}$ is the solution (not necessarily a probability distribution) to (A1.1) obtained when $\lambda_0 = \lambda_0^{(0)}$, then the improved estimate for λ_0 is given by

$$\lambda_0^{(1)} = \lambda_0^{(0)} + \left[-1 + \sum_j \exp\{-U_j^{(0)} \rho_j^{(0,0)} - V_j^{(0)}\} \right] \\ \times \left[\sum_j \exp\{-U_j^{(0)} \rho_j^{(0,0)} - V_j^{(0)}\} / [1 + U_j^{(0)} \rho_j^{(0,0)}] \right]^{-1}, \quad (A2.1)$$

where $U_j^{(0)}$ and $V_j^{(0)}$ are evaluated in the trial probability distribution $p^{(0)}$. Notably, λ_0 may be refined without re-evaluation of $U_j^{(0)}$ and $V_j^{(0)}$, and hence does not involve heavy computation.

References

- COLLINS, D. M. (1982). *Nature (London)*, **298**, 49–51.
 GULL, S. F. & DANIELL, G. J. (1978). *Nature (London)*, **272**, 686–690.
 JAYNES, E. T. (1968). *Trans. IEEE Syst. Sci. Cybern.* **4**, 227–241.
 VARGHESE, J. N. & WILKINS, S. W. (1983). In *Crystallographic Computing*, edited by S. R. HALL & T. ASHIDA. Oxford Univ. Press. In the press.
 WERNECKE, S. J. (1977). *Radio Sci.* **12**, 831–845.
 WILKINS, S. W. (1983). *Acta Cryst.* **A39**, 896–898.
 WILKINS, S. W., VARGHESE, J. N. & LEHMANN, M. S. (1983). *Acta Cryst.* **A39**, 47–60.
 WILLINGALE, R. (1981). *Mon. Not. R. Astron. Soc.* **194**, 359–364.

Acta Cryst. (1983). **A39**, 896–898

Statistical Geometry. III. Accelerated Convergence using Contrast Amplification

BY STEPHEN W. WILKINS

CSIRO, Division of Chemical Physics, PO BOX 160, Clayton, Victoria, Australia 3168

(Received 12 September 1982; accepted 1 July 1983)

Abstract

A method for greatly improving the efficiency of numerical procedures for solving the fundamental equations of the statistical geometric method [Wilkins, Varghese & Lehmann (1983). *Acta Cryst.* **A39**, 47–60] is presented. The method involves optimizing the step length in a one-dimensional search based on two trial solutions. For constraint functions, f_r , which have derivatives $f_{r,j}^1 = \partial f_r / \partial p_j$, which are linear in \mathbf{p} , it is shown that the one-dimensional search does *not* involve any additional Fourier transforms (*i.e.* lengthy computations).

0108-7673/83/060896-03\$01.50

1. Introduction

In order to make the statistical geometric (SG) method outlined in I (Wilkins, Varghese & Lehmann, 1983) a practical tool for the structure determination and refinement of biological macromolecules, it is essential that highly efficient methods for solving the fundamental equations (I.14) (see also Gull & Daniell, 1978) be developed. Simple iteration of these equations (see Gull & Daniell, 1978) is found to converge only weakly (in Borel sum) and slowly. In the second paper of this series (Wilkins, 1983, hereafter termed II), we outlined some improved methods of solution of the SG problem

© 1983 International Union of Crystallography

for the maximum-entropy structure (MES), $\mathbf{p}(\lambda)$, based on self-consistent solution of a single pixel equation (SPE). In the present paper we show how the convergence to the MES may be greatly accelerated in practice by the use of, say, two trial structures and a one-dimensional variational procedure. For cases where the constraint function $f_r(\mathbf{p})$ has a derivative with respect to p_j , *i.e.* $f_{r,j}^1(\mathbf{p})$ (defined in I), which is linear in \mathbf{p} , no additional derivative evaluations (*i.e.* Fourier transforms) are required in order to optimize the one-dimensional search (§ 4). In addition, the present contrast amplification procedure guarantees that the solution never gets worse in, say, a least-squares sense.

The notation and definitions of quantities used here are those introduced in I and II.

2. Contrast amplification

Given two arbitrary probability distributions [usually successive iterates for the self-consistent solution of (I.14) or (II.2) for given λ], say $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(1)}$, one may seek to obtain the *best estimate* for the next trial value of \mathbf{p} , say \mathbf{p}' . One approach to this problem (see also Wernecke & Addario, 1977) is as follows. Consider

$$\mathbf{p}'(\beta) = \mathbf{p}^{(1)} + \beta[\mathbf{p}^{(2)} - \mathbf{p}^{(1)}], \quad (1)$$

where β is a scalar, then clearly $\mathbf{p}'(\beta)$ is also a probability distribution provided

$$0 \leq p'_j \leq 1 \quad \text{for all } j,$$

i.e. provided β satisfies the following conditions, namely

$$-\min\{p_j^{(1)}/[p_j^{(2)} - p_j^{(1)}]\} \leq \beta \leq \min\{[p_j^{\max} - p_j^{(1)}] \div [p_j^{(2)} - p_j^{(1)}]\}, \quad \text{for } \forall p_j^{(2)} \geq p_j^{(1)} \quad (2)$$

and

$$-\min\{[p_j^{\max} - p_j^{(1)}]/[p_j^{(1)} - p_j^{(2)}]\} \leq \beta \leq \min\{p_j^{(1)} \div [p_j^{(1)} - p_j^{(2)}]\}, \quad \text{for } \forall p_j^{(1)} > p_j^{(2)},$$

where strictly $p_j^{\max} = 1$, although one may in practice choose a smaller value from other considerations (such as the maximum expected peak height based on known types of atoms present in the structure). Physically, the $[p_j^{(2)} - p_j^{(1)}]$ term in (1) corresponds to the contrast between successive structure iterates and thus β may be viewed as a *contrast-amplification factor*.

3. Accelerated convergence

Since \mathbf{p}' is a probability distribution it is a possible candidate for solving (I.14a) and one may seek to choose the β in (1) subject to (2), which gives the most nearly self-consistent solution to (I.14) in, say, a least-squares sense. Since we already know $\mathbf{p}'(\beta)$ for

$\beta = 0$ and $\beta = 1$, one further trial evaluation of $\mathbf{p}(\lambda)$ suffices to allow determination of a local optimum for β , say β^{\min} , in a least-squares sense. Because approaches to the solution of (I.14) are often quite slow (see, *e.g.*, Gull & Daniell, 1978), such methods as this can lead to a very significant reduction in the number of iterations required to attain convergence of $\mathbf{p}^{(n)}(\lambda)$, for given λ . The method is not restricted to the simple iterative method (see, *e.g.*, Gull & Daniell, 1978) or the SPA and may readily be applied to all structure or image processing problems based on a probability distribution. Essentially, the method replaces an N -dimensional search in \mathbf{p} space by a one-dimensional search in β space. In practice, one may evaluate the sum of squared deviations (s.s.d.), Φ , as a function of β for, say, one additional value of β and then determine β^{\min} by quadratic fitting [a more efficient procedure involving no extra explicit evaluations of (I.14) is possible under certain conditions, and this is discussed in § 4]. A cautionary note regarding the use of (1) is that although mathematically (1) guarantees normalization of $\mathbf{p}'(\beta)$ for arbitrary \mathbf{p} , in practice it has been found prudent to normalize $\mathbf{p}'(\beta)$, especially when $[p_j^{(2)} - p_j^{(1)}]$ is small and β may be large. In such cases there is the danger that 'noise' as well as 'signal' is being amplified and some form of adaptive smoothing may prove beneficial, although we have not tried this. It should also be pointed out that the method of contrast amplification never leads to a worse solution in, say, a least-squares sense, and so helps to avoid problems of 'overshoot' or oscillation.

4. Scaling properties

For constraints such that $f_{r,j}^1$ is *linear* [or has been linearized, see, *e.g.* (II.1)] in p_j , some special scaling properties apply so that $\Phi(\beta)$ and hence β^{\min} may be determined without any further explicit evaluations of $f_{r,j}^1$ (*i.e.* Fourier transform operations). To illustrate this, consider constraint (1) in Table 1 of I, for which

$$f_{1,j}^1 = \frac{1}{N_1} \sum_{k \in D_1} \frac{(P_k - E_k)}{\sigma_{k,1}^2} \exp\{-2\pi ijk/N\} \quad (3)$$

and is clearly linear in p_j . The fundamental equation (I.14) may be written as

$$p_j = \exp\{-\lambda_1 f_{1,j}^1(\mathbf{p})\} / \bar{z}(\lambda), \quad (4)$$

where, for simplicity, we are working with the \mathbf{f}_j^1 formulation rather than \mathbf{f}_j^j (*i.e.* we have chosen $C_1^1 = 0$ and \bar{z} is the appropriate normalizing factor).

The aim is then to find the self-consistent solution to (4) [*e.g.* minimum of, say, Φ giving the s.s.d.'s for (4)]. To this end, we note that substitution of (1) into (4) leads to the result that:

$$\text{l.h.s. of (4)} \equiv p_j(\beta) \equiv p_j^{(1)} + \beta[p_j^{(2)} - p_j^{(1)}] \quad (5a)$$

$$\text{r.h.s. of (4)} \equiv R[\beta; p_j^{(1)'(\lambda_1), p_j^{(2)'(\lambda_1), \tilde{z}_1(\lambda_1), \tilde{z}_2(\lambda_1)] \\ \div \tilde{z}[\lambda_1; \beta], \quad (5b)$$

where

$$R(\beta) = [p_j^{(1)'}]^{1-\beta} [p_j^{(2)'}]^\beta \tilde{z}_1(\lambda_1)^{1-\beta} \tilde{z}_2(\lambda_1)^\beta. \quad (6)$$

In (5) $p_j^{(1)'}$, e.g., here denotes the iterate of (4) produced when $p_j^{(1)}$ is substituted on the r.h.s. of (4) {and corresponds to $\mathbf{p}[\lambda_1; \mathbf{p}^{(1)}]$ in I}, while $\tilde{z}(\lambda_1; \beta)$ is the normalizing factor appropriate to the r.h.s. of (4) when $\mathbf{p}(\beta)$ is substituted there.

The key point of these manipulations is that both the l.h.s. and r.h.s. of (4) can be evaluated for arbitrary β without any need to re-evaluate $f_{1,j}^1$. Thus the s.s.d., Φ , is given by

$$\Phi[\beta; \mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \lambda_1] = \sum_{j=1}^N \{ p_j(\beta) \\ - R[\beta; p_j^{(1)'}, p_j^{(2)'}, \tilde{z}_1(\lambda_1), \tilde{z}_2(\lambda_1)] \\ \div \tilde{z}(\lambda_1; \beta) \}^2, \quad (7)$$

which may be evaluated for arbitrary β , and so β in (1) optimized, without much extra computation (i.e. no extra Fourier transform operations).*

5. Discussion and conclusions

The present contrast-amplification procedure, involving a one-dimensional search in β space, has been used in numerical methods for solving the fundamental equations of the statistical geometric method (I.14) via the single pixel approximation (see II) and found to improve greatly the efficiency of the solution procedure (Fig. 1). In addition, the procedure guarantees that the solution never gets worse in, say, a least-squares sense and can give information on nearby solutions in \mathbf{p} space.

Although presented for the simple case of two trial distributions, the contrast-amplification procedure may clearly be extended to t trial distributions leading to optimization of $\Phi(\beta)$ in a $(t-1)$ -dimensional β space. A special case which merits discussion is that of the 'flat-map' distribution $p_j = 1/N$ for $\forall j = 1, \dots, N$. This distribution may always be added to (1), usually without involving any extra Fourier transforms (since $P_k = 0$ for $\forall k \neq 0$), and so provides a means for optimizing the background level in the map at each stage.

* Note added in proof: An alternative approach is to evaluate the derivatives of $\Phi(\beta)$ with respect to β about $\beta = 0$.

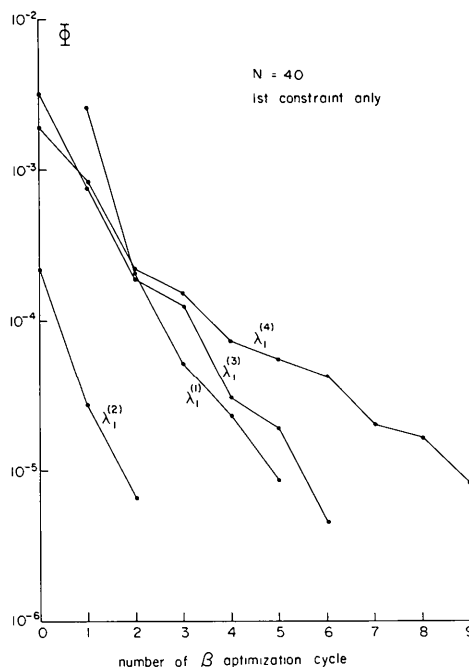


Fig. 1. Plot of s.s.d., Φ , for solution of eqs. (I.14) via the single pixel approximation, as a function of the number of β -refinement cycles. Convergence of the solution was taken to have occurred when $\Phi < 10^{-5}$, which corresponds to better than 2% in p_j relative to the 'flat map', i.e. $p_j = 1/N$ for all j . The simulated structure refinement is for the same model as in § 4.6 of I.

In the following paper in this series (Varghese & Wilkins, 1983), we shall give results for simulated structure determinations on a simple one-dimensional model structure, which were derived using the SPA and the contrast-amplification procedure in concert.

I am grateful to Drs J. N. Varghese, C. H. J. Johnson, and A. McL. Mathieson and J. Skilling for helpful discussions and to Miss R. Wong for assistance.

References

- GULL, S. F. & DANIELL, G. J. (1978). *Nature (London)*, **272**, 686–690.
 VARGHESE, J. N. & WILKINS, S. W. (1983). In *Crystallographic Computing*, edited by S. R. HALL & T. ASHIDA. Oxford Univ. Press. In the press.
 WERNECKE, S. J. & ADDARIO, L. R. (1977). *IEEE Trans Comput.* **C-26**, 351–364.
 WILKINS, S. W. (1983). *Acta Cryst.* **A39**, 892–896.
 WILKINS, S. W., VARGHESE, J. N. & LEHMANN, M. S. (1983). *Acta Cryst.* **A39**, 47–60.