

The Variational Calculation of Reduced Density Matrices

M. ROSINA

*Faculty for Natural Sciences and Technology and J. Stefan Institute,
University of Ljubljana, Ljubljana, Yugoslavia*

AND

C. GARROD

University of California, Davis, California

Received November 5, 1974

We discuss the computational problem encountered in making direct variational calculations of the reduced density matrices of many-particle systems. The problem is one of minimizing a linear function within a convex domain defined by a finite set of nonlinear constraints. Two different algorithms are presented for which working programs have been written.

1. INTRODUCTION

In recent years a method has been developed for calculating the one-particle and two-particle density matrices of a many-particle system directly, without first constructing the system wavefunction [1–6]. We shall call that method the density matrix method. In this paper we shall describe two algorithms which have been successfully employed to solve certain difficult nonlinear minimization problems peculiar to the density matrix method. The computational techniques may have applications in other areas involving nonlinear programming problems.

2. THE DENSITY MATRIX METHOD

The ground state energy of a many-particle system with one- and two-body interactions may be written as a linear function of the one-body and two-body density matrices of the ground state as follows.

$$E = \sum T_{ij} \gamma_{ji} + \frac{1}{2} \sum V_{ijkl} \Gamma_{kl ij} = \text{Tr}(T\gamma) + \frac{1}{2} \text{Tr}(V\Gamma), \quad (1)$$

where

$$\gamma_{ij} = \langle g | a_j^\dagger a_i | g \rangle, \quad (2)$$

$$\Gamma_{ijkl} = \langle g | a_i^\dagger a_k^\dagger a_l a_j | g \rangle, \quad (3)$$

and $|g\rangle$ is the many-particle ground state. Given a pair of matrices γ and Γ , any many-particle state $|g\rangle$ which satisfies Eqs. (2) and (3) is termed a *representation* of the given density matrices. A pair of density matrices (γ, Γ) for which at least one representation exists is said to be representable. In the density matrix method the linear function E is minimized with respect to the elements of the matrices γ and Γ subject to certain constraints called representability conditions which guarantee that the density matrices allowed in the minimization procedure are at least close to representable ones. (Two density matrices are said to be close to one another if they give similar expectation values for all physically important one- and two-body operators.)

The representability constraints that were utilized in the calculations to be described were of two forms; (a) a finite set of linear equalities, and (b) a finite set of nonlinear inequalities. The linear equalities are consequences of the known symmetry properties of the ground state of the system. If a pair of density matrices (γ, Γ) is representable by a wavefunction with certain known symmetry properties such as

$$\hat{N} |g\rangle = N |g\rangle \quad (4)$$

and

$$\hat{L}^2 |g\rangle = L(L+1) |g\rangle, \quad (5)$$

then corresponding linear relations involving the density matrix and the matrix elements of the symmetry operators exist. Thus Eq. (4) implies that

$$\sum_k \Gamma_{ikjk} = (N-1) \gamma_{ij} \quad (6)$$

and

$$\text{Tr}(\gamma) = N. \quad (7)$$

Similar consequences of angular momentum and spin symmetry can be easily derived. For more details the reader is advised to consult Refs. [1, 2, 6]. The other set of representability conditions is a consequence of the fact that the ground state wavefunction is antisymmetric in all variables. Those which we have utilized are of the following form.

(a) The two-body density matrix Γ must be a nonnegative matrix. That is, all its eigenvalues must be nonnegative numbers.

(b) The particle-hole matrix G_{ijkl} must be a nonnegative matrix. G is related to the density matrices γ and Γ by

$$G_{ijkl} \equiv \langle g | a_j^+ a_i a_k^+ a_l | g \rangle = \Gamma_{ijik} + \delta_{ik} \gamma_{ij} \quad (8)$$

(c) The hole-hole matrix Q_{ijkl} must be a nonnegative matrix, where

$$Q_{ijkl} \equiv \langle g | a_j a_i a_k^+ a_l^+ | g \rangle = \Gamma_{ijkl} + \delta_{ik} \bar{\gamma}_{jl} - \delta_{jk} \bar{\gamma}_{il} + \delta_{il} \gamma_{jk} - \delta_{jl} \gamma_{ik} \quad (9)$$

and $\bar{\gamma}_{ij} = \delta_{ij} - \gamma_{ji}$ is the one-hole matrix.

3. THE COMPUTATIONAL PROBLEM

As stated, the computational problem associated with the density matrix method has the following form. One chooses a set of variational parameters x_1, \dots, x_k which are taken to be elements of γ , $\bar{\gamma}$, and Γ . (γ and $\bar{\gamma}$ are included in order to make Eq. (11) valid. The relationships among γ , $\bar{\gamma}$, and Γ are taken into account by means of linear constraints.) Of course use is made of the fact that the matrices are symmetric and that Γ_{ijkl} is antisymmetric in the indices (i, j) and (k, l) to reduce the number of variational parameters needed. The linear equality constraints then impose a number L of conditions of the form

$$\sum_n C_n^\alpha x_n + B^\alpha = 0, \quad \alpha = 1, \dots, L. \quad (10)$$

The inequalities have the form of nonnegativity conditions on a set of matrices. If we consider the separate matrices Γ , G , and Q as the diagonal blocks of a block diagonal matrix M whose dimension is equal to the sum of the dimensions of the three separate matrices, then the three separate nonnegativity conditions are equivalent to the single condition that M be nonnegative. The elements of the matrix M are linear functions of the variational parameters x_1, \dots, x_k .

$$M_{ij} = \sum_n T_{ijn} x_n. \quad (11)$$

All eigenvalues of M are constrained to be nonnegative. Since the eigenvalues of a matrix are nonlinear functions of the matrix elements for which no explicit formula exists these constraints produce a distinctly challenging computational problem. The complexity of the problem is enhanced by the fact that any calculation of real value in atomic, molecular, or nuclear physics requires a single-particle basis large enough to produce from 100 to 300 variational parameters. Thus any method used must be capable of operating in very large-dimensional parameter spaces. Two computational algorithms which have proved useful in this problem are presented in the following two sections. One very favorable aspect of the computational

problem is that the region in parameter space defined by the constraints can easily be shown to be convex. Thus the minimization problem has a unique solution and no local constrained minima exist.

4. THE PENALTY FUNCTION ALGORITHM

In the penalty function method one replaces the linear function $E(x) = \sum H_n x_n$ by a nonlinear function

$$F(x) = \sum_n H_n x_n + P(x). \quad (12)$$

The penalty function $P(x)$ is a function which is zero within the region of parameter space satisfying the nonlinear inequalities but becomes smoothly and rapidly large and positive as x moves out of the allowed region. The particular form of $P(x)$ we have chosen is

$$P(x) = \frac{1}{2} W \sum_{n=1}^{\text{NEG}} (\lambda_n)^2, \quad (13)$$

where λ_n is the n th eigenvalue of M (the eigenvalues are ordered in a nondecreasing sequence) and NEG is the number of negative eigenvalues. W is a positive constant which should be chosen large enough to keep the result close to the allowed region yet small enough not to cause numerical difficulties by making $P(x)$ too rapidly varying.

With this penalty function it is possible to construct at any point an explicit second-order polynomial approximation to $P(x)$. This is true in spite of the fact that the negative eigenvalues are likely to be nearly or exactly degenerate close to the boundary of the allowed region. In the case of degeneracy no single eigenvalue can be expanded to second order in x ; however, the sum over all degenerate eigenvalues is an analytic function of the matrix elements. The sum of the squares of all degenerate eigenvalues is also an analytic function. The desired quadratic approximation for $P(x)$, valid in the neighborhood of any point x_1^0, \dots, x_n^0 to second order in the quantities $x_n - x_n^0$, is of the form

$$P(x) = \frac{1}{2} W \sum P_{mn} x_m x_n, \quad (14)$$

where the coefficients P_{mn} depend upon x_1^0, \dots, x_n^0 . The detailed formulas for the calculation of P_{mn} are given in a later section. The total function to be minimized is thus approximated by

$$F(x) = \sum H_n x_n + \frac{1}{2} W \sum P_{mn} x_m x_n. \quad (15)$$

The condition for a minimum of this quadratic approximation, subject to the L linear constraints

$$\sum_n C_n^\alpha x_n + B^\alpha = 0, \quad \alpha = 1, \dots, L, \quad (16)$$

is that the gradient of F be perpendicular to every vector in the $K - L$ dimensional subspace defined by the constraint equations. This gives the following $K - L$ linear equations.

$$W \sum_{m,n} V_m^\alpha P_{mn} x_n + \sum_m V_m^\alpha H_m = 0, \quad \alpha = L + 1, \dots, K, \quad (17)$$

where the vectors V^{L+1}, \dots, V^K span the orthogonal complement of the L -dimensional subspace spanned by the vectors C^1, \dots, C^L . The computational algorithm is as follows.

1. One chooses some initial value of W and an initial point x which satisfies all linear constraints. If the initial point x is not known to be close to the solution of the minimization problem then W should not be chosen too large. Too large a value of W will make the penalty function vary very rapidly and it will be necessary to take many small moves along the constraint surface to reach the minimum.

2. One constructs the quadratic approximation to $F(x)$ and determines the constrained minimum of the quadratic approximation. If the predicted minimum occurs at the point y one moves to a point z given by

$$z = x + \alpha(y - x),$$

where α is chosen by minimizing the function $F(z)$ with respect to α . If the quadratic approximation is valid in a region containing x and y then the minimizing value of α will be close to one. In the initial stages of the calculation when the number of negative eigenvalues of M is likely to be different at y than at x simply moving to y may cause severe overshooting of the minimum and slow the convergence.

3. With x replaced by z one repeats steps 1 and 2 until a minimum of the function $F = E + P$ is found for the given value of the constraint parameter W . The value of W is then increased and whole process repeated again. For any finite value of W the minimum of $F(x)$ is given by a point outside the allowed region. However, as W is increased the degree to which the constraints are violated at the minimum of $F(x)$ approaches zero. Thus the point converges to that point which minimizes E alone subject to the linear and nonlinear constraints.

The matrix U may be written in the form

$$U = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} 0 & v \\ -\tilde{v} & 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -v\tilde{v} & 0 \\ 0 & -\tilde{v}v \end{bmatrix}, \tag{23}$$

where v_{ar} is a $\text{NEG} \times (\text{NDIM-NEG})$ matrix. The required matrix v is easily seen to be a solution of the equation

$$Av - vB = \epsilon. \tag{24}$$

The matrix block C contains the lowest NEG eigenvalues and thus the sum of the squares of those eigenvalues is equal to the trace of the square of C .

$$\sum_a \lambda_a^2 = \sum_a (C^2)_{aa} = \sum_{a,b} C_{ab}^2. \tag{25}$$

By referring to Eqs. (21), (22), and (23) we see that

$$C = A + \frac{1}{2}(v\tilde{\epsilon} + \epsilon\tilde{v}). \tag{26}$$

It is clear from the above formula that we need calculate v only to first order in small quantities in order to obtain C to second order. We may therefore replace A and B in Eq. (24) by their zero-order approximations $A_{ab} = -\mu_a\delta_{ab}$ and $B_{rs} = \nu_r\delta_{rs}$. Equation (24) can then easily be solved to yield

$$v_{ar} = -\epsilon_{ar}/(\mu_a + \nu_r). \tag{27}$$

Since ν_r is positive by assumption the denominator can never vanish. However, if ν_r is small the region in which the quadratic approximation is valid may be limited. Using the above formula for v_{ar} in Eq. (26) and again neglecting terms of order higher than the second we obtain the following expression for the trace of C^2 .

$$\text{Tr}(C^2) = \sum A_{ab}^2 + 2 \sum \mu_a \frac{\epsilon_{ar}^2}{\mu_a + \nu_r}. \tag{28}$$

We must now transform this expression from the representation in which M^0 is diagonal back to the original representation in which $M = Tx$ is valid. This is done by means of the orthogonal matrix U^0 . That expression is then transformed, by means of the transformation T , into a quadratic expression in the components of x . The result is

$$\lambda_a^2 = \sum P_{mn}x_mx_n, \tag{29}$$

where

$$P_{mn} = \sum S_{abm}S_{abn} + \sum \frac{2\mu_a}{\mu_a + \nu_r} S_{arm}S_{arn} \tag{30}$$

and

$$S_{ijm} = \sum_{k,l} U_{ik}^0 U_{jl}^0 T_{klm}. \quad (i, j, k, l = 1, \dots, \text{NDIM}). \tag{31}$$

6. THE CUTTING PLANE ALGORITHM

We shall describe a procedure in which the nonlinear constraints are written as an infinite set of linear inequalities. This can be done because the nonlinear constraint defines a convex region in the parameter space. In the proposed procedure only K linear equalities and inequalities are ensured in each iteration, where K is the number of variational parameters x_n . In each iteration one new inequality is imposed and one of the old inequalities is discarded. The procedure may be utilized in general minimization problems which have the following characteristics.

- (i) The function to be minimized is linear.
- (ii) The allowed region in the parameter space is convex.
- (iii) For each point outside the allowed region it is possible to construct a violated condition in a straightforward way. The convergence of the procedure depends, however, on how efficiently conditions can be constructed for each point.

The constraint that the matrix $M_{ij} = \sum_n T_{ijn}x_n$ be nonnegative can be written in the form that the expectation value of the matrix M with respect to any vector Z be nonnegative.

$$\sum_{ij} z_i M_{ij} z_j \equiv \sum_{ijn} (z_i T_{ijn} z_j) x_n \geq 0. \tag{32}$$

This is an infinite set of linear constraints. The matrix nonnegativity constraint has the convenient feature that it is easy to construct for any point x outside the allowed

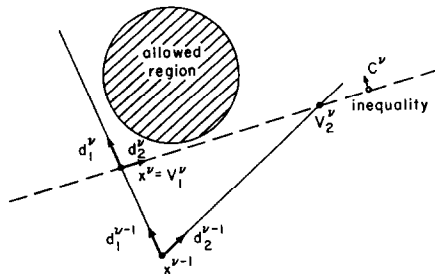


FIG. 1. The geometrical illustration of the numerical procedure.

region the most violated condition. The most violated condition is obtained by diagonalizing the matrix M and by taking the eigenvector z_i^{\min} corresponding to the most negative eigenvalue λ^{\min} . The expression $\sum_{ijn} z_i^{\min} T_{ijn} z_j^{\min} x_n \equiv \lambda^{\min}$ then violates the inequality $\sum_{ijn} (z_i^{\min} T_{ijn} z_j^{\min}) x_n \geq 0$ by the amount λ^{\min} which is the maximum possible amount for any choice of z_i . Our task is to minimize the function $E = \sum_{i=1}^n H_i x_i$, where H_i are given coefficients and x_i are variational parameters. The variational parameters must satisfy the following equalities and inequalities.

$$\sum_n C_n^\alpha x_n + B^\alpha = 0, \quad \alpha = 1, \dots, L, \tag{33}$$

$$\sum_n C_n^\beta x_n + B^\beta \geq 0, \quad \beta = L + 1, \dots, \infty. \tag{34}$$

The coefficients of equalities C_n^α and B^α are given in advance, while the coefficients of inequalities C_n^β and B^β are generated in consecutive iterations.

The algorithm is as follows.

(i) In the zeroth iteration, the function E is minimized with the constraints $l_n \leq x_n \leq u_n$. The lower and upper bounds l_i and u_i are given in advance. One can choose any bounds which reasonably confine the range of parameters. However, they should be implied by the constraints (34). These bounds serve only to find the initial point

$$x_n^0 = \begin{cases} l_n & \text{if } H_n \geq 0, \\ u_n & \text{if } H_n < 0, \end{cases} \tag{35}$$

and they are gradually discarded in consecutive iterations. The geometrical picture of these bounds is a rectangular box containing the allowed region. The point x^0 is the lowest vertex of this box, where we define the upward direction by the vector H . The vertex x^0 is the intersection of K hyperplanes which are faces of the box. There are K edges pointing upward from the vertex x^0 . Each edge is the intersection of $K - 1$ hyperplanes defining the vertex x^0 . They are easily constructed for the rectangular box

$$d_{mn}^0 = \begin{cases} \delta_{mn} & \text{if } H_n \geq 0, \\ -\delta_{mn} & \text{if } H_n < 0, \end{cases} \quad m = 1, \dots, K, \tag{36}$$

where d_{mn}^0 is the n th component of the m th edge. The signs are chosen so that all edges point upward ($\sum H_n d_{mn}^0 \geq 0$).

(ii) The equalities are imposed one in each iteration by intersecting the edges with the corresponding hyperplane and choosing the lowest intersection. The

consecutive iterations will be indexed by the integer ν . The distance from the point $x^{\nu-1}$ to the m th intersection is

$$D_m^\nu = - \left(B^\nu + \sum_n C_n^\nu x_n^{\nu-1} \right) / \sum_n C_n^\nu d_{mn}^{\nu-1}. \tag{37}$$

The m th intersection v_{mn}^ν is then

$$v_{mn}^\nu = x_n^{\nu-1} + D_m d_{mn}^{\nu-1}. \tag{38}$$

Let us write the index of the lowest intersection as \tilde{m}_ν .

$$\sum_n H_n v_{\tilde{m}n}^\nu \leq \sum_n H_n v_{mn}^\nu,$$

for any m such that $D_m > 0, D_{\tilde{m}} > 0$.

Then one can construct the new vertex and the new edges

$$x_n^\nu = v_{\tilde{m}n}^\nu; \tag{39}$$

$$\tilde{d}_{mn}^\nu = \pm (v_{mn}^\nu - x_n^\nu) \quad \text{if} \quad \begin{cases} D_m > 0, & m \neq \tilde{m}_\nu, \\ D_m < 0; \end{cases} \tag{40}$$

$$d_{mn}^\nu = \tilde{d}_{mn}^\nu / \left(\sum_n (\tilde{d}_{mn}^\nu)^2 \right)^{1/2}. \tag{41}$$

In each iteration, the index m runs over $m = 1, \dots, K, m \neq \tilde{m}_1, \tilde{m}_2, \dots, \tilde{m}_{\nu-1}$ so in each iteration one additional value of m is omitted.

(iii) After the introduction of all the equalities, in each iteration one inequality is imposed by the same algorithm as in step (ii) except that the index $m = \tilde{m}_\nu$ is no longer omitted. The index $m = 1, \dots, n; m \neq \tilde{m}_1, \dots, \tilde{m}_L$. In this way, the inequalities for $\beta = \nu - K + L, \dots, \nu$ are satisfied while some of the previous ones may get violated again. In the case of inequalities, the \tilde{m}_ν th edge is needed for the next iteration; it equals $d_{\tilde{m}_\nu}^\nu = d_{\tilde{m}_\nu}^{\nu-1}$.

(iv) The program is terminated when one of the following conditions is fulfilled.

- (a) $E^\nu \leq E^{\nu+\Delta\nu} + \epsilon$ (we take $\epsilon = 10^{-6}, \Delta\nu = 10$)
- (b) when no inequality violated by more than ϵ is found.
- (c) $\nu > \nu_{\max}$, where ν_{\max} is given in advance.

In the calculation, caution should be taken for divisions with denominators too small. If for any $\tilde{m}, |\sum_n C_n^\nu d_{\tilde{m}n}^\nu| < \epsilon$, one must avoid choosing $\tilde{m} = \tilde{m}$ and one puts $d_{\tilde{m}i}^\nu = d_{\tilde{m}i}^{\nu-1}$. If an equality is almost linearly dependent on the previous ones (if $|\sum_n C_n^\nu d_{mn}^{\nu-1}| < \epsilon$ for all m and $|B^\nu + \sum_n C_n^\nu x_n^{\nu-1}| < \epsilon$) it is ignored. If an

equality is incompatible with the previous ones (if $|\sum_j C_n^v d_{mn}^{v-1}| < \epsilon$ for all m and $|B^v + \sum_n C_j^v x_n^{v-1}| > \epsilon$) the calculation is terminated. If an equality happens to be satisfied by the point x under consideration (if $|B^v + \sum_n C_n^v x_n^{v-1}| < \epsilon$ but some $|\sum_n C_n^v d_{mn}^{v-1}| > \epsilon$), then $x_n^v = x_n^{v-1}$ and the new edges are constructed by the following trick. $x_n^{v-1} \rightarrow x_n^{v-1} - 10\epsilon C_n^v$ and then the normal procedure can be applied.

This algorithm has the feature that the function E rises monotonically.

However, whether it converges and the speed with which it does so depend on certain aspects of the sequence of inequalities which are generated.

This algorithm is an adaptation of the Revised Simplex Method (7). It is designed for problems (such as the one considered above) in which one can directly construct the most violated inequality from a continuum of inequalities rather than check through a finite list of inequalities given at the start of the program.

REFERENCES

1. C. GARROD AND J. K. PERCUS, *J. Math. Phys.* **5** (1964), 1756.
2. A. J. COLEMAN AND R. M. ERDAHL, Eds., "Reduced Density Matrices with Applications to Physical and Chemical Systems," Queen's Papers on Pure and Applied Mathematics No. 11, Queen's University, Kingston, Ontario, 1968.
3. M. V. MIHAILOVIĆ AND M. ROSINA, *Nucl. Phys.* **A130** (1969), 386.
4. E. R. DAVIDSON, *J. Math. Phys.* **10** (1969), 725.
5. L. KIJEWski, *Phys. Rev.* **6** (1972), 31.
6. M. BOUTEN, P. VAN LEUVEN, M. V. MIHAILOVIĆ, AND M. ROSINA, *Nucl. Phys.* **A202** (1973), 127.
7. D. G. LUENBERGER, "Introduction to Linear and Nonlinear Programming," p. 55, Addison-Wesley, Reading, Mass., 1973.