TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 190, 1974

A COMBINATORIAL APPROACH TO THE DIAGONAL N-REPRESENTABILITY PROBLEM(1)

BY

MARK LAURANCE YOSELOFF(2)

ABSTRACT. The problem considered is that of the diagonal N-representability of a pth-order reduced density matrix, $p \geq 2$, for a system of N identical fermions or bosons. A finite number M of allowable single particle states is assumed. The problem is divided into three cases, namely: Case I. $M = N + p$; Case II. $M < N + p$; Case III. $M > N + p$. Using the theory of polyhedral convex cones, a complete set of necessary and sufficient conditions is first found for Case I. This solution is then employed to find such conditions for Case II. For Case III, two algorithms are developed to generate solutions for the problem, and examples of the usage of these algorithms are given.

I. PRELIMINARIES

1. Physical problem.

a. *Wave functions and density matrices.* In the study of systems of identical symmetric (bosons) and antisymmetric (fermions) particles, the wavefunction provides a characterization of such assemblies. Because of the extremely complicated nature of the wavefunction, simpler objects may be introduced in order to facilitate an evaluation of various quantities associated with the system. These objects are reduced density matrices and are at the heart of the problem to which this paper is addressed.(3)

In the discussion which follows we will restrict our attention to fermions. However, by replacing the antisymmetry conditions by symmetry conditions, and Slater determinants by their permanents, all of the arguments work equally well for bosons.

As was mentioned above, given a system of N identical antisymmetric particles with coordinates $x_1, x_2, x_3, \cdots, x_N$, where each x_i is of the form

Received by the editors November 15, 1972.

AMS (MOS) subject classifications (1970). Primary 81A57, 15A57; Secondary 52A25.

(1) This work was supported by the National Science Foundation. The preparation of the final manuscript was supported by the Office of Naval Research under contract number N00014–67-A-0151–0010.

(2) I would like to express my sincerest thanks to Professor Harold Kuhn for his having presented this problem to me and for the many important suggestions which he made during the course of my work on it. I would also like to thank him for all his help and guidance during my graduate studies.

(3) For a more detailed discussion of the physical background, see [2], [5], [8], [9], [11], and [14].

(r_i, s_i) with r_i a space coordinate chosen from R^3, and s_i a spin coordinate chosen from Z_2, the system may be characterized by an N particle wavefunction

$$\Psi(x_1, x_2, x_3, \cdots, x_N).$$

This function is antisymmetric and normalizable.

In terms of such wavefunctions we may now define their *pth-order reduced density matrices*, or more simply their p matrices as

$$\begin{aligned}(1)\quad &\Gamma^{(p)}(x_1'x_2'x_3' \cdots x_p' \mid x_1x_2x_3 \cdots x_p) \\ &= \int \Psi^*(x_1', x_2', x_3', \cdots, x_p', x_{p+1}, \cdots, x_N)\Psi(x_1, \cdots, x_N)\, dx_{p+1} \cdots dx_N.\end{aligned}$$

Depending upon the context in which p matrices are being considered, other normalizations may be more appropriate. For example, Löwdin [8] multiplies the right-hand side of the above defining equation by $\binom{N}{p}$, and McWeeney [9] uses the factor $N!/p!$. Coleman [2], however, points out the general utility of the definition above as it stands. As will be pointed out below, the diagonal elements of a p matrix have a probabilistic interpretation, and this is most easily facilitated by employing the normalization of equation (1).

If Ω_{op} is a Hermitian operator representing a physical quantity associated with the system, it may be expanded as

$$(2)\qquad \Omega_{\text{op}} = \Omega_{(0)} + \sum_i \Omega_i + \frac{1}{2!} \sum_{ij}{}' \Omega_{ij} + \frac{1}{3!} \sum_{ijk}{}' \Omega_{ijk} + \cdots,$$

where the nth term is an $(n-1)$-particle operator and the prime on the summation indicates that only terms in which all indices are distinct are summed. We can evaluate the average value of this quantity $\langle \Omega_{\text{op}} \rangle_{\text{av}}$, in the situation given by a normalized wavefunction Ψ by using p matrices. In such a case, we have

$$\begin{aligned}(3)\quad \langle \Omega_{\text{op}} \rangle_{\text{av}} &= \Omega_{(0)} + \int \Omega_1 \Gamma^{(1)}(x_1' \mid x_1)\, dx_1 \\ &\quad + \int \Omega_{12} \Gamma^{(2)}(x_1'x_2' \mid x_1x_2)\, dx_1\, dx_2 + \cdots.\end{aligned}$$

By expanding the operator in this way, we can greatly simplify a computation of an approximation of its value. Löwdin [8] gives several examples of the usefulness of p matrices in this context.

In general, density matrices are bounded linear operators, of trace class, which satisfy the following conditions:

(i) they are Hermitian;

(ii) they are anitsymmetric;

(iii) the p and $(p-1)$ matrices associated with the same wavefunction are related by the equation

$$\Gamma^{(p-1)}(x'_1x'_2x'_3 \cdots x'_{p-1} \mid x_1x_2x_3 \cdots x_{p-1}) = p^{-1} \int \Gamma^{(p)}(x'_1x'_2x'_3 \cdots x'_p \mid x_1x_2x_3 \cdots x_p)\, dx_p. \tag{4}$$

However, in order for the expansion given by equation (3) to have physical significance, the matrices used must satisfy the condition of N-representability,(4) which is discussed in the next section.

b. *N-representability.* In this section we will define the properties, for a p matrix, of being pure state and ensemble N-representable. We will then specialize to the problem of diagonal N-representability and present a theorem connecting the pure state and ensemble problems in the diagonal case.

A p matrix $\Gamma^{(p)}(x'_1x'_2x'_3 \cdots x'_p \mid x_1x_2x_3 \cdots x_p)$ is said to be *pure state N-representable* if there exists some single normalized antisymmetric wavefunction Ψ of N particles such that $\Gamma^{(p)}$ and Ψ are related by equation (1). A p matrix is said to be *ensemble N-representable* if it can be written as a convex combination of pure state N-representable p matrices. That is, if there exists a set of pure state N-representable p matrices $\{\Gamma_i^{(p)}\}$ and weights w_i such that

$$\Gamma^{(p)} = \sum_i w_i \Gamma_i^{(p)}, \qquad \sum_i w_i = 1, \tag{5}$$

and, for all i, $w_i \geq 0$.

If we now restrict our attention to the diagonal elements(5) of a p matrix $\Gamma^{(p)}$, then it will be called *pure state diagonal N-representable* if there exists some single normalized antisymmetric wavefunction Ψ of N particles such that the elements $\Gamma^{(p)}(x_1x_2x_3 \cdots x_p \mid x_1x_2x_3 \cdots x_p)$ of $\Gamma^{(p)}$ and Ψ satisfy equation (1). As above, the diagonal of a p matrix will be called *ensemble diagonal N-representable* if it can be written as a convex combination of pure state diagonal N-representable p matrices.

We will now consider pure state diagonal N-representability in greater detail, and demonstrate its equivalence to ensemble diagonal N-representability.

As several authors have described, [2], [8], [13] and [14] in order to facilitate an investigation of density matrices, a countable set of single particle functions $f_i(x)$, $i = 1, 2, 3, \cdots$, may be selected. This set is complete, orthonormal, and is such that any normalizable single-particle function $f(x)$ may be expanded as

(4) See [2], [3], [5], [8], [9], [11], and [14].

(5) A p matrix is, in fact, not a matrix, but rather the kernel of a linear operator. Thus, the term diagonal really refers to the elements $\Gamma(x_1x_2 \cdots x_p \mid x_1x_2 \cdots x_p)$. In the finite-dimensional cases which we shall consider, the p matrix may be written as an actual matrix. In these cases, the elements $\Gamma(x_1x_2 \cdots x_p \mid x_1x_2 \cdots x_p)$ correspond to the diagonal entries of this matrix.

$$\Psi = \sum_K C_K D_K, \tag{6}$$

where K runs over all possible sets of N indices chosen from the natural numbers and, when $K - \{k_1, k_2, k_3, \cdots, k_N\}$,

$$\begin{aligned} C_K = \int \Psi(x_1, x_2, x_3, \cdots, x_N) \\ \times f^*_{k_1}(x_1) f^*_{k_2}(x_2) f^*_{k_3}(x_3) \cdots f^*_{k_N}(x_N)\, dx_1\, dx_2\, dx_3 \cdots dx_N, \end{aligned} \tag{7}$$

and D_K is the Slater determinant

$$(N!)^{-1/2} \begin{vmatrix} f_{k_1}(x_1) & f_{k_1}(x_2) & \cdots & f_{k_1}(x_N) \\ f_{k_2}(x_1) & f_{k_2}(x_2) & \cdots & f_{k_2}(x_N) \\ \vdots & & & \vdots \\ f_{k_N}(x_1) & f_{k_N}(x_2) & \cdots & f_{k_N}(x_N) \end{vmatrix}. \tag{8}$$

In addition, the normalization condition

$$\sum_K |C_K|^2 = \int |\Psi|^2\, dx_1\, dx_2\, dx_3 \cdots dx_N$$

is satisfied.

For purposes of actually carrying out a computation as in equation (3), it is usually necessary to select a finite set of M spin orbitals. Löwdin [8] discusses the problem of choosing the set of M spin orbitals which make the expression given by equation (6) best approximate the full expansion of the wavefunction, where the sets K are now restricted to the indices of the functions selected.

If f_i, $i = 1, 2, 3, \cdots, M$, is the set of spin orbitals chosen, then the N-representability problem can be asked for p matrices and wavefunctions expanded only on a subspace of L_2 of the configuration space spanned by all possible Slater determinants of these finitely many functions. From the defining properties of density matrices (4), they may actually be written as matrices on such finite-dimensional subspaces. For a p matrix, $\Gamma^{(p)}$, the diagonal elements are given by

$$\Gamma_{i_1 i_2 \cdots i_p} = \int g^*_{i_1 \cdots i_p} \Gamma^{(p)} g_{i_1 \cdots i_p}\, dx_1 \cdots dx_p\, dx'_1 \cdots dx'_p, \tag{9}$$

where $g_{i_1 \cdots i_p}$ is the $p \times p$ Slater determinant of the spin orbitals $f_{i_1}, \cdots, f_{i_p}$

and the variables $x_1, \cdots, x_p$. It can be shown that, if Ψ is an N-particle wavefunction, expanded as in equation (6), satisfying (7) and (8), Γ is the p matrix associated with Ψ by equation (2), and $L_{i_1 \cdots i_p} = \binom{N}{p}\Gamma_{i_1 \cdots i_p}$, then

$$L_{i_1 \cdots i_p} = \sum_{K(i_1, \cdots, i_p \in K)} |C_K|^2. \tag{10}$$

Conversely, given a p matrix Γ, if one could choose a set of C_K such that for all $i_1, \cdots, i_p$, condition (10) is satisfied and, in addition, the normalization condition

$$\sum_K |C_K|^2 = 1 \tag{11}$$

holds, then the diagonal elements of Γ and the wavefunction Ψ constructed from these C_K using equation (6) should satisfy equation (2). Thus the N-representability problem for diagonal elements is to find a set of C_K satisfying conditions (10) and (11).

We can now prove the following theorem, a slightly different version of which was proved by Ruskai [11].

Theorem 1.1. *Pure state diagonal N-representability is equivalent to ensemble diagonal N-representability.*

Proof. If the diagonal of a density matrix is pure state diagonal N-representable, then it is clearly ensemble diagonal N-representable, with the ensemble consisting of one p matrix. Conversely, if it is ensemble diagonal N-representable, then the elements $L_{i_1 \cdots i_p}$ may be written as $L_{i_1 \cdots i_p} = \Sigma_s w_s L^s_{i_1 \cdots i_p}$, where by (10), each $L^s_{i_1 \cdots i_p}$ satisfies

$$L^s_{i_1 \cdots i_p} = \sum_{K(i_1, \cdots, i_p \in K)} |C^s_K|^2.$$

Now, by letting $|C_K|^2 = \Sigma_s w_s |C^s_K|^2$,

$$L_{i_1 \cdots i_p} = \sum_{K(i_1, \cdots, i_p \in K)} \left(\sum_s w_s |C^s_K|^2 \right) = \sum_{K(i_1, \cdots, i_p \in K)} |C_K|^2.$$

Thus, it is pure state diagonal N-representable.

In the following considerations we will restrict our attention to pure state diagonal N-representability. Since by Theorem 1.1, pure state diagonal N-representability is equivalent to ensemble N-representability, our further results will hold for both. Therefore, we will use the phrase *diagonal N-representable* to mean either of these two equivalent properties.

2. Mathematical formulations of diagonal N-representability. In this section we will present three equivalent formulations of the diagonal N-representability problem [7]. In the following, let K be an arbitrary set of N distinct indices $\{i_1, i_2, \cdots, i_N\}$ chosen from $\{1, 2, \cdots, M\}$.

Solvability problem. Let the constants $y_{i_1 \cdots i_p} \geq 0$, for all $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, satisfying

$$\sum_{1 \leq i_1 < \cdots < i_p \leq M} y_{i_1 \cdots i_p} = \binom{N}{p} \tag{12}$$

be given. Under what conditions on these constants does the system of equations $\Sigma_{K(i_1, \cdots, i_p \in K)} t_K = y_{i_1 \cdots i_p}$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, have a nonnegative solution for the t_K?

By making the replacements $L_{i_1 \cdots i_p} = y_{i_1 \cdots i_p}$ and $|C_K|^2 = t_K$ it is clear that this problem is equivalent to the problem of the N-representability of diagonal elements as presented at the end of the last section.

Polyhedron problem. Let $X^K = (x^K_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, be the point in $\binom{M}{p}$-dimensional affine space defined by

$$x^K_{i_1 \cdots i_p} = \begin{cases} 1, & \text{if } \{i_1, \cdots, i_p\} \subset K, \\ 0, & \text{otherwise.} \end{cases}$$

Let C denote the convex hull of the points X^K. Characterize C as the intersection of halfspaces.

By letting $Y = (y_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, with the $y_{i_1 \cdots i_p}$ as in the solvability problem, it will be shown that Y satisfies the conditions of that problem if and only if it lies in C, as defined in the polyhedron problem. This equivalence will be made more precise by the introduction of polyhedral convex cones in the next section.

A combination of these two formulations will be considered in our solution. However, we will generally adopt the terminology of the solvability problem in stating results, in order to preserve the consistency of the presentation.

The final formulation which we present will not be considered explicitly in the following chapters. It is included here to give a more pictorial view of the problem.

Simplex problem. Let p, N and M, $2 \leq p < N < M$, be given, and consider the $(M-1)$-simplex S_{M-1}. For each $(p-1)$-face $(i_1, \cdots, i_p)$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, let weights $w_{i_1 \cdots i_p} \geq 0$ be given, with the property that

$$\sum_{1 \leq i_1 < \cdots < i_p \leq M} w_{i_1 \cdots i_p} = \binom{N}{p}.$$

Under what conditions on these weights is it possible to assign nonnegative weights t_K to the $(N-1)$-faces $(i_1, \cdots, i_N)$, in such a way that the weight of a $(p-1)$-face is equal to the sum of the weights of the $(N-1)$-faces which contain it?

This formulation is clearly equivalent to the solvability problem, with $y_{i_1 \cdots i_p}$ being replaced by $w_{i_1 \cdots i_p}$.

Figure 1 displays the simplex problem for the simplest possible case, namely $M = 4$, $N = 3$, and $p = 2$. Here, $K_1 = \{1, 2, 3\}$, $K_2 = \{1, 2, 4\}$, $K_3 = \{1, 3, 4\}$, and $K_4 = \{2, 3, 4\}$.

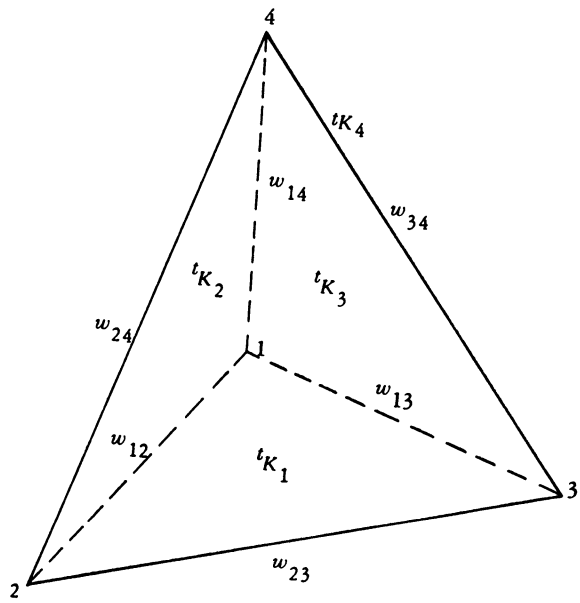


Figure 1. Simplex problem, with $M = 4$, $N = 3$, $p = 2$.

3. Combinatorial and cone theoretic background. The mathematical objects lying at the heart of the development in the following chapters are polyhedral convex cones. In this section we will outline some important properties of such cones.

Let R^n be real n-dimensional space. A *convex cone* C in R^n is a set of vectors satisfying the following:

(i) if $X \in C$, and $k \geq 0$, then $kX \in C$; and

(ii) if $X_1, X_2 \in C$, then $X_1 + X_2 \in C$.

A convex cone C is called polyhedral if it satisfies the additional condition

(iii) C is the intersection of finitely many closed halfspaces.

In general a halfspace H can be characterized as the set of vectors Y

satisfying $A \cdot Y \leq 0$ for some vector A. In particular, A is the exterior normal to H at the origin. Therefore, for a polyhedral convex cone C, condition (iii) may be restated as

(iii′) There exists a finite set of vectors $A_1, A_2, \cdots, A_m$, such that C may be written as $C = \{X \mid A_j \cdot X \leq 0, j = 1, 2, \cdots, m\}$.

In this way it can be seen that C may be characterized as the set of solutions of the system of homogeneous linear inequalities $\sum_{i=0}^{n} a_{ji} x_i \leq 0$, $j = 1, 2, \cdots, m$, where each vector $A_j = (a_{j1}, a_{j2}, \cdots, a_{jn})$.

At times it will be more convenient to write this system in matrix form. In that case, if A is the matrix (a_{ij}), $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, m$, then, for a vector X, $X \in C$ if and only if it satisfies $A \cdot X \leqq 0$,(6) where the X are now considered as column vectors.

For a convex cone C, we will say that a subset $B \subset C$ *spans* C if every vector in C can be written as a finite linear combination of vectors in B with nonnegative coefficients. In this case, we will say that C is the *convex-cone hull* of B. If B is a finite set, say $B = \{B_1, B_2, \cdots, B_r\}$, then

$$C = \left\{ X \mid X = \sum_{i=1}^{r} w_i B_i,\ w_i \geq 0,\ i = 1, 2, \cdots, r \right\}. \tag{13}$$

If we consider a set S of vectors, then the *polar* of S, denoted by S^* is defined by

$$S^* = \{X \mid X \cdot V \leq 0, \text{ for all } V \in S\}. \tag{14}$$

Clearly, if $X_1, X_2 \in S^*$, then for all $k \geq 0$, $kX_1 \in S^*$, and $X_1 + X_2 \in S^*$. Therefore, S^* is a convex cone.

In the case of a convex cone spanned by a finite set B, it can easily be seen that a vector X satisfies $X \cdot V \leq 0$, for all $V \in C$, if and only if it satisfies $X \cdot V \leq 0$, for all $V \in B$. Therefore, $C^* = B^*$. Now, since B is finite, say $B = \{B_1, B_2, \cdots, B_r\}$, B^* may be written as $B^* = \{X \mid B_j \cdot X \leq 0, j = 1, 2, \cdots, r\}$, and we therefore get

Theorem 1.2. *If C is the convex-cone hull of a finite subset, then C^* is a polyhedral convex cone.*

We will now state the following well-known results:

Theorem 1.3 (J. Farkas). *If C is the convex-cone hull of a finite subset B, then $C = C^{**} = B^{**}$.*

(6) In inequalities involving vectors or matrices, the symbols $\leq$ and $\leqq$ have different meanings. The former denotes that all entries on the left are less than or equal to the corresponding entries on the right and that at least one inequality is strict. In the case of $\leqq$, all of the corresponding entries may be equal.

For a proof of this, one may consult Goldman and Tucker [6], Gale [4], or Thrall and Tornheim [12]. In the last of these references this is called the double-description theorem.

Theorem 1.4 (H. Minkowski, J. Farkas). *If C is a polyhedral convex cone, then C is the convex-cone hull of finitely many vectors.*

For a proof of this, one may consult Goldman and Tucker [6] or Gale [4].

A convex cone C is said to be *pointed* if it does not contain any subspace. A nonzero vector $V \in C$ is called an *extreme vector* if $V = X_1 + X_2$, with $X_1, X_2 \in C$, implies that there exist $k_1, k_2 > 0$ such that $k_1X_1 = k_2X_2 = V$. An extreme vector is uniquely determined up to a positive multiple. For an extreme vector V, the halfline $L = \{kV \mid k \geq 0\}$ is called an *extreme ray*. In terms of this we may state the convex-cone analogue of the Krein-Milman Theorem for convex sets.

Theorem 1.5. *A pointed polyhedral convex cone is the convex-cone hull of its extreme rays.*

A proof of this result may be found in Goldman and Tucker [6].

If S is the set of all of the extreme rays of C, and if for each $L_i \in S$ one chooses a nonzero vector $V_i \in L_i$, then such a set of V_i will be called a *complete set of extreme* vectors. Theorem 1.5 may then be restated as

Theorem 1.5′. *A pointed polyhedral convex cone C is the convex-cone hull of any complete set of extreme vectors of C.*

We are now in a position to consider the following system of linear equations, given in matrix notation by

$$AX = B, \tag{15}$$

where A is a fixed $m \times n$ matrix. We now pose the following question: For which B does this system have a nonnegative solution for X?

In order to answer this, let $X^i = (x_j^i)$, $i = 1, 2, \cdots, n$, be the column vector of length n defined by

$$x_j^i = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

Now, let

$$B^i = AX^i, \qquad i = 1, 2, \cdots, n. \tag{17}$$

If for a given vector $\overline{B}$ the system given by equation (15) has a nonnegative solution $\overline{X}$, then clearly $\overline{X} = \Sigma_{i=1}^n k_i X^i$, with $k_i \geq 0$ for all $i = 1, 2, \cdots, n$. But in that case,

$$\bar{B} = A\bar{X} = A\left(\sum_{i=1}^{n} k_i X^i\right) = \sum_{i=1}^{n} k_i (AX^i) = \sum_{i=1}^{n} k_i B^i.$$

Conversely, if a vector $\bar{B}$ has the form $\bar{B} = \Sigma_{i=1}^{n} k_i B^i$, with $k_i \geq 0$ for all $i = 1, 2, \cdots, n$, then by reversing these steps we get

$$\bar{B} = \sum_{i=1}^{n} k_i B^i = \sum_{i=1}^{n} k_i (AX^i) = A\left(\sum_{i=1}^{n} k_i X^i\right).$$

Since $\Sigma_{i=1}^{n} k_i X^i$ is a nonnegative vector, such a $\bar{B}$ yields a nonnegative solution for the system given by equation (15). Thus we have

Theorem 1.6. *If A is a given $m \times n$ matrix, and vectors B^i are defined as in equation* (17), *then the system of linear equations $AX = B$ has a nonnegative solution for X if and only if B can be written as $B = \Sigma_{i=1}^{n} k_i B^i$, with $k_i \geq 0$ for all $i = 1, 2, \cdots, n$.*

Applying Theorems 1.2 through 1.4 to this, we get

Theorem 1.7. *If A is a given $m \times n$ matrix, then the set of vectors B for which the system of linear equations $AX = B$ has a nonnegative solution for X is the polyhedral convex cone C defined by*

$$C = \left\{ B \mid B = \sum_{i=1}^{n} k_i B^i,\ k_i \geq 0,\ i = 1, 2, \cdots, n \right\}, \tag{18}$$

where B^i is defined as in equation (17). *It is also clear from the way in which the B^i are defined that if A is of rank n, then the B^i are a complete set of extreme vectors of C.*

According to Theorem 1.3, the vectors contained in this cone C may be characterized by the following: $B \in C$ if and only if $B \in C^{**}$, and from the definition of polar cones this is if and only if $B \cdot W \leq 0$, for all $W \in C^*$. Going one step further, if C^* is pointed, and if we can find some set S which is a complete set of extreme vectors of C^*, then $B \in C$ if and only if $V \cdot W \leq 0$, for all $W \in S$.

We will now present a result which characterizes the extreme rays of C^* in the case in which A is an $n \times n$ nonsingular matrix. This result will be sufficient to handle the situation in which A is $m \times n$, with $m \geq n$ and $\operatorname{rank}(A) = n$, for the particular system of equations $AX = B$ which we will consider. In Chapter III we will examine the case in which A is $m \times n$, with $m < n$.

In order to obtain the desired result in the case in which A is a nonsingular $n \times n$ matrix, we first define sets S_i, $i = 1, 2, \cdots, n$, where $S_i = \{1, 2, \cdots, n\} - \{i\}$. Now, let H_i be the hyperplane generated by the set of vectors T_i,

(19) $$T_i = \{B^j \mid j \in S_i\},$$

with B^j defined as in equation (17). If W^i is a vector such that

(20a) $$W^i \cdot V = 0, \text{ for all } V \in T_i,$$

and

(20b) $$W^i \cdot B^i < 0,$$

then clearly W^i is an extreme vector of the cone C^*, which is the polar cone of C, as defined in equation (18). If one takes the set $\{W^i\}$, $i = 1, 2, \cdots, n$, this is clearly a complete set of extreme vectors of the cone C^*. Thus we have

Theorem 1.8. *Let A be a nonsingular $n \times n$ matrix, $\{B^i\}$ be the set of vectors generated from the system of linear equations $AX = B$ as in equation* (17), *and C the cone defined from these B^i as in equation* (18). *If $\{W^i\}$, $i = 1, 2, \cdots, n$, is a set of vectors such that, for each i, W^i satisfies conditions* (20), *with T_i defined as in* (19), *then this set of W^i is a complete set of extreme vectors for C^*, the polar cone of the cone C.*

We may also observe that in the case in which A is a nonsingular $n \times n$ matrix the cone C is of full dimension, that is, it has a nonempty interior. This clearly implies that C^* is a pointed cone, so that we finally get

Theorem 1.9. *Let A be a nonsingular $n \times n$ matrix and let B be a given vector. The system of linear equations $AX = B$ will have a nonnegative solution for X if and only if $B \cdot W^i \leq 0$, for all $i = 1, 2, \cdots, n$, with W^i as in the statement of Theorem* 1.8.

With the background presented in this chapter we are now in a position to present a solution for the diagonal N-representability problem in the cases in which $M \leq N + p$.

Before closing this chapter, however, we will make the following note about the subscripting scheme employed in considering the N-representability problem. We start with the set of integers $\{1, 2, \cdots, M\}$. The vectors $Y = (y_{i_1 \cdots i_p})$ whose diagonal N-representability is to be tested are vectors whose entries are indexed by all possible subsets of p distinct indices $\{j_1, j_2, \cdots, j_p\} \subset \{1, 2, \cdots, M\}$. For convenience in referring to such a set we use the ordered p-tuple $i_1 i_2 \cdots i_p$, where $\{i_1, i_2, \cdots, i_p\} = \{j_1, j_2, \cdots, j_p\}$, and $1 \leq i_1 < i_2 < \cdots < i_p \leq M$. Similarly, the variables t_K are indexed by all possible subsets K of N distinct indices $\{k_1, k_2, \cdots, k_N\} \subset \{1, 2, \cdots, M\}$.

To give an example, let $M = 4$, $N = 3$, and $p = 2$. A vector Y is of the form $Y = (y_{12}, y_{13}, y_{14}, y_{23}, y_{24}, y_{34})$. The variables are t_{K_j}, $j = 1, 2, 3, 4$, where

$K_1 = \{1, 2, 3\}$, $K_2 = \{1, 2, 4\}$, $K_3 = \{1, 3, 4\}$, and $K_4 = \{2, 3, 4\}$. In general, Y will lie in an $\binom{M}{p}$-dimensional space, and there will be $\binom{M}{N}$ variables.

II. DIAGONAL N-REPRESENTABILITY WHEN $M \leq N + p$

1. Solution when $M = N + p$. In this section we will present an explicit solution to the diagonal N-representability problem when $M = N + p$. Using the terminology of the solvability problem presented in §I.2, we are given a vector $Y = (y_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, satisfying

$$\sum_{1 \leq i_1 < \cdots < i_p \leq M} y_{i_1 \cdots i_p} = \binom{N}{p},$$

and must determine whether the system of linear equations

$$\sum_{K(i_1, \cdots, i_p \in K)} t_K = Y_{i_1 \cdots i_p}$$

has a nonnegative solution for the t_K.

We begin by writing this system of equations in matrix form at $AT = Y$. In this form A is an $\binom{M}{p} \times \binom{M}{N}$ matrix, whose rows are indexed by all p-tuples of distinct indices $i_1 \cdots i_p$, $i \leq i_1 < i_2 < \cdots < i_p \leq M$, and whose columns are indexed by all sets K of N distinct indices. T is the vector (t_K), whose entries are indexed by all such sets K, and Y is the vector $(y_{i_1 \cdots i_p})$.

Since in this case $M = N + p$, we have that $\binom{M}{p} = \binom{M}{N}$. It is easily verified that A is a nonsingular square matrix. Because A takes this form, we may apply Theorem 1.9. In order to employ this theorem we must first determine the set of vectors $\{B^K\}$, where in this case the B's are indexed by all possible sets K of N distinct indices. According to equations (16) and (17), for a given set K, $B^K = AT^K$, where $T^K = (t^K_{K'})$, with

$$t^K_{K'} = \begin{cases} 1, & \text{if } K' = K, \\ 0, & \text{otherwise.} \end{cases}$$

From this we find $B^K = (b^K_{i_1 \cdots i_p})$, with

$$b^K_{i_1 \cdots i_p} = \begin{cases} 1, & \text{if } \{i_1, \cdots, i_p\} \subset K, \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

Having generated the B^K, we must now find a set $\{W^K\}$ satisfying conditions (20). To do this, for an arbitary but fixed N let the constants

$$w_{p-j} = (-1)^{j+1}(p-j)!(N-p+j-1)!/(N-p-1)!, \tag{22}$$

$j = 0, 1, \cdots, p$, be given. From these we define the vectors W^K, for each K, as $W^K = (w^K_{i_1 \cdots i_p})$, with

$$w^K_{i_1 \cdots i_p} = w_{p-j}, \tag{23}$$

where $|K \cap \{i_1, \cdots, i_p\}| = p - j$, and w_{p-j} is defined by equation (22). Here $|S|$ denotes the order of the set S. With this definition we prove the following, which is a more general result than is needed for this section:

Theorem 2.1. *Let p, N, and M be given, with $2 \leq p < N$, and $M = N + t$, $0 < t \leq p$. If $\{B^K\}$ is defined according to equation* (21), *and $\{W^K\}$ is defined according to equation* (23), *then the two sets satisfy conditions* (20).

Proof. In order to demonstrate condition (20a), let b and q be such that $1 \leq b \leq q < t$, then we will show that

$$\begin{aligned} &\frac{R(-1)^{b-1}(N-p+b-2)!}{(N-p-q+b-2)!}\binom{q-1}{b-2} + \sum_{i=b-1}^{q}\binom{N-q}{p-i}w_{p-i} \\ &\quad = \frac{R(-1)^{b}(N-p+b-1)!}{(N-p-q+b-1)!}\binom{q-1}{b-1} + \sum_{i=b}^{q}\binom{N-q}{p-i}\binom{q}{i}w_{p-i}, \end{aligned} \tag{24}$$

where $R = (N - q)!/(N - p)!$. To demonstrate the validity of equation (24), it is sufficient to prove that

$$\begin{aligned} &\frac{R(-1)^{b-1}(N-p+b-2)!}{(N-p-q+b-2)!}\binom{q-1}{b-2} + \binom{N-q}{p-b+1}\binom{q}{b-1}w_{p-b+1} \\ &\quad = \frac{R(-1)^{b}(N-p+b-1)!}{(N-p-q+b-1)!}\binom{q-1}{b-1}. \end{aligned} \tag{25}$$

But, the left-hand side of equation (25) is equal to

$$\begin{aligned} &\frac{R(-1)^{b-1}(N-p+b-2)!(q-1)!}{(N-p-q+b-2)!(b-2)!(q-b+1)!} \\ &\quad + \frac{(N-q)!q!}{(p-b+1)!(N-p-q+b-1)!(b-1)!(q-b+1)!} \\ &\qquad \times \frac{(-1)^{b}(p-b+1)!(N-p+b-2)!}{(N-p-1)!} \\ &= \frac{R(-1)^{b}(N-p+b-2)!(q-1)!}{(N-p-q+b-1)!(b-1)!(q-b+1)!} \times (-(N-p-q+b-1)(b-1) + q(N-p)) \\ &= \frac{R(-1)^{b}(N-p+b-1)!}{(N-p-q+b-1)!}\binom{q-1}{b-1}. \end{aligned}$$

Thus, equation (25) holds, which implies the validity of equation (24).

Now, to verify that $\{B^K\}$ and $\{W^K\}$ satisfy condition (20a), let K be any set of N distinct indices, and let K' be any other such set, where $K' \neq K$. Then, $|K' \cap K| = N - q$, where $0 < q \leq t$. For such a K and K'

$$W^K \cdot B^{K'} = \sum_{i=0}^{q} \binom{N-q}{p-i}\binom{q}{i} w_{p-i}. \tag{26}$$

Applying equation (24) q times in succession to equation (26), we have

$$\begin{aligned} W^K \cdot B^{K'} &= \frac{R(-1)^q(N-p+q-1)!}{(N-p-1)!}\binom{N-q}{p-q} w_{p-q} \\ &= \frac{R(-1)^q(N-p+q-1)!}{(N-p-1)!} + \frac{(N-q)!}{(p-q)!(N-p)!} \times \frac{(-1)^{q+1}(p-q)!(N-p+q-1)!}{(N-p-1)!} \\ &= \frac{R(-1)^q(N-p+q-1)!}{(N-p-1)!} + \frac{R(-1)^{q+1}(N-p+q-1)!}{(N-p-1)!} \\ &= 0. \end{aligned}$$

Thus $\{B^K\}$ and $\{W^K\}$ satisfy condition (20a).

To verify condition (20b), if K is any set of N distinct indices, then

$$B^K \cdot W^K = \sum_{i_1,\ldots,i_p \in K} w^K_{i_1\cdots i_p} = \binom{N}{p} w_p = -\binom{N}{p} p! < 0,$$

so that condition (20b) holds, and the theorem is proved.

We may now apply Theorem 1.9 to the result obtained in Theorem 2.1 to get

Theorem 2.2. *Let p and N be given such that $2 \leq p < N$, and let $M = N + p$. Also let the vector $Y = (y_{i_1\cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, be given such that $Y \geqq 0$, and*

$$\sum_{1 \leq i_1 < \cdots < i_p \leq M} y_{i_1\cdots i_p} = \binom{N}{p}.$$

Then the system of linear equations

$$\sum_{K(i_1,\ldots,i_p \in K)} t_K = y_{i_1\cdots i_p},$$

$1 \leq i_1 < i_2 < \cdots < i_p \leq M$, has a nonnegative solution for the t_K if and only if for all sets K of order N contained in $\{1, 2, \ldots, M\}$, $Y \cdot W^K \leq 0$, with W^K defined as in equation (23).

To rephrase this in terms of the diagonal N-representability problem we have

Theorem 2.3. *Let p and N be given such that $2 \leq p < N$, and let $M = N + p$. Let $L = (L_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq M$, be a given vector such that $L \geqq 0$. Then L is diagonal N-representable if and only if*

$$\text{(27a)} \qquad \text{(i)} \sum_{1 \leq i_1 < \cdots < i_p \leq M} L_{i_1 \cdots i_p} = \binom{N}{p},$$

and

(ii) *for all sets K of order N contained in $\{1, 2, \cdots, M\}$, if W^K is defined as in equation* (23), *then*

$$\text{(27b)} \qquad W^K \cdot L \leq 0.$$

2. **Solution when $M < N + p$.**

a. *Inequality conditions.* In this section we will present a set of inequalities as the solution to the diagonal N-representability problem in the cases in which $M = N + t$, with $1 \leq t < p$. The method used to obtain these results will be based upon "embedding" such problems as special cases of diagonal N-representability when $M = N + p$.

To this end, let $L = (L_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N + t$, with $1 \leq t < p$, be a vector whose diagonal N-representability is to be tested in the system with $M = N + t$. From this L we define $\bar{L} = (\bar{L}_{i_1 \cdots i_p})$, with $1 \leq i_1 < i_2 < \cdots < i_p \leq N + p$, where

$$\text{(28)} \qquad \bar{L}_{i_1 \cdots i_p} = \begin{cases} L_{i_1 \cdots i_p}, & \text{if } \{i_1, \cdots, i_p\} \subset \{1, 2, \cdots, N + t\}, \\ 0, & \text{otherwise.} \end{cases}$$

We then have

Theorem 2.4. *Let t, p, and N be given such that $1 \leq t < p < N$. Let $L = (L_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N + t$, be given. If $\bar{L} = (\bar{L}_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N + p$, is defined from L by equation* (28), *then L is diagonal N-representable in the system with $M = N + t$ if and only if $\bar{L}$ is diagonal N-representable in the system with $M = N + p$.*

Proof. If L is diagonal N-representable in the system with $M = N + t$, then

$$\text{(29a)} \qquad \sum_{1 \leq i_1 < \cdots < i_p \leq N + t} L_{i_1 \cdots i_p} = \binom{N}{p},$$

and for each $K \subset \{1, 2, \dots, N+t\}$, $|K| = N$, there exists a $t_K \geq 0$ such that

$$\sum_{K(i_1,\dots,i_p \in K)} t_K = L_{i_1 \cdots i_p}, \tag{29b}$$

for all $1 \leq i_1 < i_2 < \dots < i_p \leq N+t$.

Clearly, since $\bar{L}_{i_1 \cdots i_p} = 0$ for $\{i_1, \dots, i_p\} \not\subset \{1, 2, \dots, N+t\}$,

$$\sum_{1 \leq i_1 < \cdots < i_p \leq N+p} L_{i_1 \cdots i_p} = \binom{N}{p}. \tag{30a}$$

Now, for $K \subset \{1, 2, \dots, N+p\}$, $|K| = N$, let

$$\bar{t}_K = \begin{cases} t_K, & \text{if } K \cap \{N+t+1, \dots, N+p\} = \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

Then for $\{i_1, \dots, i_p\} \subset \{1, 2, \dots, N+t\}$, by (29b) we have

$$\bar{L}_{i_1 \cdots i_p} = L_{i_1 \cdots i_p} = \sum_{K(i_1,\dots,i_p \in K)} t_K = \sum_{K(i_1,\dots,i_p \in K)} \bar{t}_K. \tag{30b}$$

For $\{i_1, \dots, i_p\} \subset \{1, 2, \dots, N+p\}$ such that $\{i_1, \dots, i_p\} \cap \{N+t+1, \dots, N+p\} \neq \emptyset$, $\{i_1, \dots, i_p\} \subset K$ implies that $\bar{t}_K = 0$, so

$$\bar{L}_{i_1 \cdots i_p} = 0 = \sum_{K(i_1,\dots,i_p \in K)} \bar{t}_K. \tag{30c}$$

Hence $\bar{L}$ is diagonal N-representable in the system with $M = N+p$.

Conversely, if conditions (30) hold, then from (30a) and the definition of $\bar{L}$, (29a) holds. To demonstrate (29b), for $K \subset \{1, 2, \dots, N+t\}$, $|K| = N$, let $t_K = \bar{t}_K$, where $\bar{t}_K$ are those satisfying (30b). From (30c), if $K \cap \{N+t+1, \dots, N+p\} \neq \emptyset$, then $\bar{t}_K = 0$. Therefore, from (30b), if $\{i_1, \dots, i_p\} \subset \{1, 2, \dots, N+t\}$,

$$L_{i_1 \cdots i_p} = \bar{L}_{i_1 \cdots i_p} = \sum_{K(i_1,\dots,i_p \in K)} \bar{t}_K.$$

This sum may be broken into two parts to give

$$L_{i_1 \cdots i_p} = \sum_{\substack{K(i_1,\dots,i_p \in K) \\ K \cap \{N+t+1,\dots,N+p\} = \emptyset}} \bar{t}_K + \sum_{\substack{K(i_1,\dots,i_p \in K) \\ K \cap \{N+t+1,\dots,N+p\} \neq \emptyset}} \bar{t}_K$$

$$= \sum_{K(i_1,\dots,i_p \in K)} t_K + 0 = \sum_{K(i_1,\dots,i_p \in K)} t_K,$$

and (29b) holds.

Applying Theorem (2.3) to this yields

Theorem 2.5. *Let t, p, and N be given such that $1 \leq t < p < N$, and let $L = (L_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N + t$, be a given vector. If $\bar{L}$ is defined according to equation* (28), *then L is diagonal N-representable in the system with $M = N + t$ if and only if*

(31a) (i) $$\sum_{1 \leq i_1 < \cdots < i_p \leq N+t} L_{i_1 \cdots i_p} = \binom{N}{p},$$

and

(ii) *for all sets K of order N contained in $\{1, 2, \cdots, N + p\}$, if W^K is defined as in equation* (23), *then*

(31b) $$W^K \cdot \bar{L} \leq 0.$$

Theorem 2.5 thus gives a solution to the diagonal N-representability problem in the cases in which $M < N + p$. This solution is given in terms of the normalization condition (31a), and a system of inequality conditions (31b).

b. *Equality conditions.* In this section we will sharpen the results just obtained by showing that certain of the inequalities given by (31b) are in fact equalities. Here again we have t, p, and N given such that $1 \leq t < p < N$, and let $M = N + t$. Given a vector $L = (L_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N + t$, it is a trivial matter to test whether the normalization condition

$$\sum_{1 \leq i_1 < \cdots < i_p \leq N+t} L_{i_1 \cdots i_p} = \binom{N}{p}$$

holds, and we may consider this condition separately. Therefore, for the present we will consider the set of vectors L for which the system of linear equations

(32) $$\sum_{K(i_1, \cdots, i_p \in K)} t_K = L_{i_1 \cdots i_p},$$

$1 \leq i_1 < i_2 < \cdots < i_p \leq N + t$, has a nonnegative solution for the t_K, without regard for the sum of the entries of L. According to Theorem 1.7, this is the polyhedral convex cone C given by

(33) $$C = \left\{ L \mid L = \sum_K k_K B^K,\ k_K \geq 0 \text{ for all } K \right\},$$

where B^K is defined according to equation (21).

Now, define $\bar{B}^K = (\bar{b}^K_{i_1 \cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N + p$, by

$$\bar{b}^K_{i_1\cdots i_p} = \begin{cases} b^K_{i_1\cdots i_p}, & \text{if } \{i_1,\cdots, i_p\} \subset \{1, 2,\cdots, N+t\}, \\ 0, & \text{otherwise,} \end{cases} \tag{34}$$

and let $\bar{C}$ be the polyhedral convex cone given by

$$\bar{C} = \left\{ \bar{L} \mid \bar{L} = \sum_K k_K \bar{B}^K,\ k_K \geq 0 \text{ for all } K \right\}. \tag{35}$$

It is clear that if $L = (L_{i_1\cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N+t$, is a given vector, and $\bar{L} = (\bar{L}_{i_1\cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N+p$, is defined according to equation (28), then $L \in C$ if and only if $\bar{L} \in \bar{C}$.

We now turn our attention to the system of linear equations (32), with $1 \leq i_1 < i_2 < \cdots < i_p \leq N+p$. For this system let $\underline{B}^K = (\underline{b}^K_{i_1\cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N+p$, denote the vectors defined by equation (21). From equations (21) and (34), if $K \subset \{1, 2, \cdots, N+t\}$, then $\bar{B}^K = \underline{B}^K$. Now, let K' be any set of order N such that $K' \subset \{1, 2, \cdots, N+p\}$, and $K' \cap \{N+t+1, \cdots, N+p\} \neq \emptyset$, and K any set of order N such that $K \subset \{1, 2, \cdots, N+t\}$. Since $\bar{B}^K = \underline{B}^K$, we may apply Theorem 2.1 to get $\bar{B}^K \cdot W^{K'} = 0$, where $W^{K'}$ is defined by equation (23). From this we see that, for all $V \in \bar{C}$, $V \cdot W^{K'} = 0$, where $\bar{C}$ is defined by equation (35). Now, for the cone C defined by equation (33), since $L \in C$ if and only if $\bar{L} \in \bar{C}$, with $\bar{L}$ defined by equation (28), we get

Theorem 2.6. *Let t, p and N be given such that $1 \leq t < p < N$, and let $L = (L_{i_1\cdots i_p})$, $1 \leq i_1 < i_2 < \cdots < i_p \leq N+t$, be a given vector. If $\bar{L}$ is defined according to equation* (28), *then L is diagonal N-representable in the system with $M = N+t$ if and only if*

(36a) (i) $$\sum_{1 \leq i_1 < \cdots < i_p \leq N+t} L_{i_1\cdots i_p} = \binom{N}{p},$$

and

(ii) *for all sets K of order N contained in $\{1, 2, \cdots, N+p\}$, if W^K is defined as in equation* (23), *then*

$$W^K \cdot \bar{L} \begin{cases} \leq 0, & \text{(36b)} \quad \text{if } K \cap \{N+t+1, \cdots, N+p\} = \emptyset \\ = 0, & \text{(36c)} \quad \text{if } K \cap \{N+t+1, \cdots, N+p\} \neq \emptyset. \end{cases}$$

Theorem 2.6 thus gives a sharpened version of Theorem 2.5, with certain of the inequalities being replaced by equalities.

We may view conditions (36b) and (36c) as partitioning the W^K into two sets. The fact that certain of the W^K actually satisfy the equality condition for any $\bar{L}$ associated with an N-representable L can be interpreted from the point of

view of the geometry of C^*, the polar cone of the cone C defined by equation (33). In this interpretation the two sets into which the W^K are partitioned play different roles geometrically. This point is discussed in Chapter IV, where we consider the geometry of the polar cone for all values of M.

III. DIAGONAL N-REPRESENTABILITY WHEN $M > N + p$

1. Additional cone theoretic background. In this section we will present some additional cone theoretic results which will be useful in examining the diagonal N-representability problem when $M > N + p$. In this case, the problem is to identify those vectors B for which the system $AX = B$ has a nonnegative solution for X, where A is $m \times n$, with $m < n$. Recalling the discussion following Theorem 1.7, let C be the cone generated by the columns of A, C^* be the polar cone of the cone C, and S be a complete set of extreme vectors of C^*. Then such vectors B are characterized by the condition that $B \cdot W \leq 0$, for all $W \in S$.

The approach here will be to generate all of the extreme vectors of C^* by algorithmic means. Before proceeding to examine the methods employed it may be useful, for the arguments which follow, to give a brief geometric description of the situation. As we have said, A is $m \times n$, with $m < n$. The cone C^* lies in m-dimensional space. Each of the n columns of A is the outward pointing normal at the origin to a bounding hyperplane of C^*. Thus, C^* is the intersection of the set of halfspaces H^i, $i = 1, 2, \cdots, n$, where $H^i = \{V \mid V \cdot A^i \leq 0\}$, A^i being the ith column of A. Since $m < n$, the number of hyperplanes bounding C^* is greater than the dimension of the space in which C^* lies, and this leads to a rather complicated set of extreme vectors.

In constructing the algorithm used, the following theorem will be needed:

Theorem 3.1. *Let C be a polyhedral convex cone in real m-dimensional space given by $C = \bigcap_{i=1}^{n} H^i$, with H^i being the halfspace defined by $H^i = \{V \mid V \cdot A^i \leq 0\}$, where A^i is some given vector. Let W be an arbitrary vector, and let S_W be the set defined by $S_W = \{A^i \mid A^i \cdot W = 0\}$. Let the sets S_W, for all $W \in C$, be partially ordered by set inclusion. Then W is an extreme vector of C if and only if S_W is maximal with respect to this partial ordering.*

Proof. We will prove the contrapositive statement; namely, W is not an extreme vector of C if and only if S_W is not maximal with respect to this partial ordering. Assume W is not extreme. Then there exist $W_1, W_2 \in C$, not multiples of W, such that $W = W_1 + W_2$. For all $A^i \in S_W$, $A^i \cdot W = 0$ implies that $A^i \cdot W_1 + A^i \cdot W_2 = 0$. But, since $W_1, W_2 \in C$, $A^i \cdot W_1 \leq 0$ and $A^i \cdot W_2 \leq 0$. Therefore, $A^i \cdot W_1 = 0$ and $A^i \cdot W_2 = 0$, so $A^i \in S_{W_1}$ and $A^i \in S_{W_2}$. Thus, $S_W \subset S_{W_1}$ and $S_W \subset S_{W_2}$. If either of these inclusions is proper, then S_W is not maximal and the implication is proved. Therefore, we assume that $S_W = S_{W_1} = S_{W_2}$.

Now define the set J by $J = \{j \mid A^j \notin S_W\}$. For all $j \notin J$, let $k_j = W \cdot A^j$ and $c_j = W_1 \cdot A^j$. Since $j \in J$, $k_j < 0$ and $c_j < 0$. Setting $r_j = k_j/c_j$, let t be such that

$$r_t = \min_{j \in J} \{r_j\}. \tag{37}$$

Now, define a new vector W_3 by $W_3 = k_t W_1 - c_t W$. For all $j \notin J$,

$$A^j \cdot W_3 = k_t(W_1 \cdot A^j) - c_t(W \cdot A^j) = 0. \tag{38}$$

For all $j \in J$,

$$A^j \cdot W_3 = k_t(W_1 \cdot A^j) - c_t(W \cdot A^j) = k_t c_j - c_t k_j. \tag{39}$$

But, from equation (37), $r_t \leq r_j$, which implies that $k_t c_j \leq c_t k_j$. Applying this to equation (39) gives

$$A^j \cdot W_3 \leq 0. \tag{40}$$

Now (38) and (40) together imply that $W_3 \in C$. From equation (38) we get $S_W \subset S_{W_3}$. Furthermore, $A^t \notin S_W$, but $A^t \cdot W_3 = k_t c_t - c_t k_t = 0$. Thus the inclusion is proper and S_W is not maximal, proving the implication.

For the opposite implication, assume S_W is not maximal. Then there exists a vector $W_1 \in C$, such that $S_W \subset S_{W_1}$ and $S_{W_1} \neq S_W$. Define the set J by $J = \{j \mid A^j \notin S_{W_1}\}$, and let

$$k = \min_{j \in J} \left\{ \frac{A^j \cdot W}{A^j \cdot W_1} \right\}. \tag{41}$$

Now, define a new vector W_2 by $W_2 = W - kW_1$. For $j \notin J$, $A^j \cdot W_2 = A^j \cdot W \leq 0$, since $W \in C$. For $j \in J$,

$$A^j \cdot W_2 = A^j \cdot W - k(A^j \cdot W_1). \tag{42}$$

But, by equation (41), $k \leq (A^j \cdot W)/(A^j \cdot W_1)$. Since $A^j \cdot W_1 < 0$, this yields

$$A^j \cdot W \leq k(A^j \cdot W_1). \tag{43}$$

Applying (43) to (42) gives $A^j \cdot W_2 \leq 0$. Thus $W_2 \in C$. Furthermore, since $k > 0$, $kW_1 \in C$. Since W_1 is not a multiple of W, W_2 is not a multiple of W, and $W = kW_1 + W_2$, W is not an extreme vector and the theorem is proved.

2. A double description algorithm. We are now in a position to derive an algorithm for generating all of the extreme rays of the cone C, in the cases in which $M > N + p$. Several authors have considered the problem of generating all of the extreme rays of a polyhedral convex cone which is given in terms of its

bounding hyperplanes [1] and [10]. The algorithm which we will present is a modification of the one given by Motzkin, Raiffa, Thompson, and Thrall [10]. The main difference in the two algorithms is the method of selecting those vectors which are to remain at the completion of each iteration.

In general terms, the method is as follows:

The process starts with some subset of the complete set of bounding hyperplanes for the cone in question. Also given along with this list of hyperplanes is a complete set of extreme vectors for the cone bounded by the subset of hyperplanes.

The remaining hyperplanes are considered one at a time. As each is introduced, a new cone, bounded by the previously considered hyperplanes and the one new hyperplane, is formed. A complete set of extreme vectors for the newly-formed cone is then generated from the complete set of extreme vectors for the previously considered cone, and the process repeats with another hyperplane. Once all of the hyperplanes have been used, what remains is a complete set of extreme vectors for the cone bounded by all of the original set of hyperplanes. This completes the algorithmic process.

We will postpone a discussion of the methods of obtaining a "start" for the algorithm until later in this chapter. The remainder of this section will detail the iterative step. Let H^i, $i = 1, 2, \cdots, n$, be the complete set of halfspaces generating the m-dimensional cone C, and let A^i, $i = 1, 2, \cdots, n$, be the corresponding normal vectors at the origin. Assume that the cone given by H^i, $i = 1, 2, \cdots, k-1$, has been considered, and that the vectors V_j, $j = 1, 2, \cdots, r$, are a complete set of extreme vectors for this cone.

The halfspace H^k is now introduced. Let $I_j = V_j \cdot A^k$, for all $j = 1, 2, \cdots, r$. Three cases must be considered:

Case I: $I_j \leq 0$, for all $j = 1, 2, \cdots, r$;

Case II: $I_j > 0$, for all $j = 1, 2, \cdots, r$;

Case III: There exists j_1 and j_2 such that $I_{j_1} \leq 0$ and $I_{j_2} > 0$.

In the event that Case I occurs, the cone generated by H^i, $i = 1, 2, \cdots, k-1$, lies entirely in the halfspace H^k. Therefore, intersecting this preceding cone with H^k leaves it unaltered, and the complete set of extreme vectors is unchanged. Thus, the process simply proceeds to the next step by introducing H^{k+1}.

If Case II occurs, then the cone generated by H^i, $i = 1, 2, \cdots, k-1$, lies entirely outside of the halfspace H^k. This means that the cone generated by the complete set H^i, $i = 1, 2, \cdots, n$, degenerates to a point, namely, the origin. In this case, no further iterations are necessary.

For the third case, a new set of extreme vectors must be generated. In order to do this, the vectors V_j are first partitioned into three sets, namely,

$$S_1 = \{V_j \mid V_j \cdot A^k < 0\}, \quad S_2 = \{V_j \mid V_j \cdot A^k = 0\}, \quad S_3 = \{V_j \mid V_j \cdot A^k > 0\}.$$

From S_1 and S_3 a new set, S_4, is constructed in the following manner:

For each vector $V_i \in S_1$ and each vector $V_j \in S_3$, let $V_{i,j} = I_i V_j - I_j V_i$. Then $S_4 = \{V_{i,j} \mid V_i \in S_1,\ V_j \in S_3\}$.

Denote the cone generated by H^i, $i = 1, 2, \cdots, s$, by C_s. The elements of S_1 are those extreme vectors of C_{k-1} which lie in the interior of the halfspace H^k. The elements of S_2 are those extreme vectors of C_{k-1} which lie on the bounding hyperplane of H^k. S_3 consists of the extreme vectors of C_{k-1} which lie outside of H^k; therefore, they are not in C_k. Finally, S_4 consists of a complete set of vectors representing all of the rays in which the hyperplane bounding H^k intersects the set of $(m-1)$-dimensional faces of C_{k-1}.

Now, to determine a complete set of extreme vectors of C_k once S_4 has been formed, it is no longer necessary to consider S_3, since all of its elements lie outside of C_k. Furthermore, it is clear that the extreme vectors of C_{k-1}, lying in the interior of H_k, are also extreme vectors of C_k, and, in fact, they are the only extreme vectors of C_k interior to H^k. So, all of the elements of S_1 are extreme vectors of C_k. Thus, it is only necessary to find those extreme vectors of C_k lying on the bounding hyperplane of H^k. This amounts to determining which vectors in $S_2 \cup S_4$ are extreme.

In order to do this we apply Theorem 3.1. Letting S_W be defined as in the statement of the theorem, we define $T = \{S_W \mid W \in S_2 \cup S_4\}$. From this we define $\bar{S}_2 = \{W \mid S_W$ is maximal with respect to the partial ordering on $T\}$. From Theorem 3.1 and the discussion above, we find that $S_1 \cup \bar{S}_2$ is a complete set of extreme vectors for C_k. Once this set is found, the process continues, with the introduction of a new halfspace.

3. Implementation of the double description method. In order to implement the double description algorithm explained above, a computerized version has been prepared. The results obtained by using this program will be presented in Chapter IV. In this section we will describe a method for generating a starting cone for the process.

To find such a starting cone, let H^i, $i = 1, 2, \cdots, n$, be the complete set of generating halfspaces for the m-dimensional cone C, whose extreme rays are to be determined. Let A^i, $i = 1, 2, \cdots, n$, be the corresponding normal vectors at the origin. Now, reindex the halfspaces so that the set $S = \{A^1, A^2, \cdots, A^m\}$ consists of m linearly independent vectors. From these, form a matrix A whose ith column is the vector A^i. By the way in which the set S was chosen, A is an $m \times m$ nonsingular matrix. In order to find a complete set of extreme vectors for the cone generated as the intersection of H^i, $i = 1, 2, \cdots, m$, we may now

apply Theorem 1.8. Since A is a square nonsingular matrix, we can find $-A^{-1}$. Theorem 1.8 asserts that the rows of $-A^{-1}$ are a complete set of extreme vectors for this cone. Thus, the columns of A and the rows of $-A^{-1}$ provide a start for the double description algorithm.

In applying the double description algorithm to the diagonal N-representability problem, only certain cases in which $p = 2$ and $N = 3$ have been computed. This is primarily because of storage limitations in the computer system which was used. For these cases the starting cones, that is, the matrices A and $-A^{-1}$, were generated recursively.

Before presenting the recursive scheme, we will describe the method employed in indexing the rows of A. Recalling the discussion at the end of Chapter I, the rows of A are indexed by all subsets of two elements chosen from $\{1, 2, \cdots, M\}$. Therefore, we need only consider the pairs of indices $i_1 i_2$ where $1 \leq i_1 < i_2 \leq M$. In indexing the rows of A, these pairs are enumerated by letting i_2 vary from 2 to M, and for each i_2, letting i_1 vary from 1 to $i_2 - 1$. Thus, the rows of A are indexed by $12, 13, 23, 14, 24, 34, \cdots, 1M, 2M, 3M, \cdots, (M-1)M$.

With this indexing in mind, the recursion is as follows:

Assume the A and $-A^{-1}$ have been found for $p = 2$, $N = 3$, and $M = g - 1$. We are looking for A and $-A^{-1}$ for $p = 2$, $N = 3$, and $M = g$. Let A_{g-1} and A_g denote the A matrices for $M = g - 1$ for $M = g$, respectively. In order to generate A_g, let S_{g-1} denote the set of columns of A_{g-1}, and define S_g by

$$S_g = S_{g-1} \cup \{A^{T_1}, A^{T_2}, \ldots, A^{T_{g-1}}\},$$

where T_1 is the triple $12g$ and T_i, $i = 2, 3, \cdots, g-1$, is the triple $(i-1)(g-1)g$. Letting the elements of S_g be the columns of A_g, then A_g has the form

$$A_g = \begin{pmatrix} A_{g-1} & B \\ 0 & C \end{pmatrix}.$$

In order to show that A_g is nonsingular, it is only necessary to show that C is nonsingular. But, C is of the form

$$C = \begin{pmatrix} 1 & 1 & & & & \\ 1 & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \\ 1 & 1 & 1 & 1 & \cdots & 1 \end{pmatrix},$$

where the super-diagonal and the last row consist of ones and the blank entries are zeros. In order to demonstrate the nonsingularity of C, we simply give its inverse, which is

$$C^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \cdots & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \cdots & -\frac{1}{2} & \frac{1}{2} \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix},$$

where the blank entries are zeros.

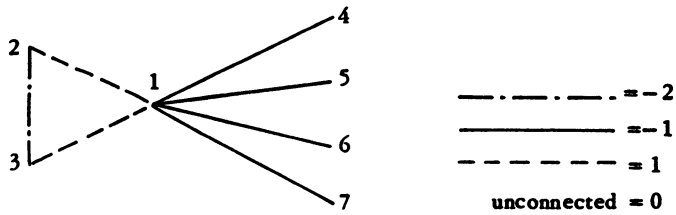
To complete the generation of the starting cone for $M = g$, we present A_g^{-1}, which takes the form

$$A_g^{-1} = \begin{pmatrix} A_{g-1}^{-1} & A_{g-1}^{-1} BC^{-1} \\ 0 & C^{-1} \end{pmatrix}.$$

4. **Collapsings and liftings.** In this section we will describe a technique which is useful in studying the structure of the cone C^*, the polar cone of the cone C, where C is the set of all diagonal N-representable vectors for some given p, N, and M. We fix our attention on the cases in which $p = 2$. For some fixed M, we will use weighted graphs to display vectors in $\binom{M}{2}$-dimensional space. The vertices of these graphs are the points $1, 2, \ldots, M$. The edge connecting i to j, $1 \leq i < j \leq M$, is the value of the ijth entry in the vector. For example, if $M = 7$, then the vector

$$V = (1, 1, -2, -1, 0, 0, -1, 0, 0, 0, -1, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0),$$

where the indexing of $\binom{7}{2}$-dimensional space follows the description in the last section, would be displayed in graph form as



The technique to be employed involves two operations, collapsing and lifting. If p and N are fixed, then a collapsing takes a vector from the space for $M = g$ into the space for $M = g - 1$, and a lifting takes a vector from the space for $M = g - 1$ into the space for $M = g$. Let $S = \{1, 2, \cdots, g\}$, and let $S_k = S - \{k\}$, $k = 1, 2, \cdots, g$. If $V = (v_{i_1 \cdots i_p})$ is a vector in $\binom{g}{p}$-dimensional space, then choosing some k, $1 \leq k \leq g$, we can form a new vector $V^k = (v^k_{i_1 \cdots i_p})$ in $\binom{g-1}{p}$-dimensional space, where for each p-tuple $i_1 \cdots i_p$, with $i_1, \cdots, i_p \in S_k$, we set $v^k_{i_1 \cdots i_p} = v_{i_1 \cdots i_p}$. This may be restated by saying that V^k is formed from V by simply deleting those entries of V which are indexed by p-tuples containing k. Such a vector V^k is called a *collapsing* of the vector V. Since this procedure may be carried out for all k, $1 \leq k \leq g$, there are g different collapsings for the vector V.

Now, if $V = (v_{i_1 \cdots i_p})$ is a vector in $\binom{g-1}{p}$-dimensional sapce, then we can form a new vector $\overline{V} = (\overline{v}_{i_1 \cdots i_p})$ in $\binom{g}{p}$-dimensional space by setting

$$\overline{v}_{i_1 \cdots i_p} = \begin{cases} v_{i_1 \cdots i_p}, & \text{if } 1 \leq i_1 < \cdots < i_p \leq g - 1, \\ \text{arbitrary}, & \text{otherwise.} \end{cases}$$

Such a vector $\overline{V}$ is called a *lifting* of V. We note that the collapsing $\overline{V}^g$ of $\overline{V}$ is the original vector V. In this sense, these two operations may be viewed as inverses of one another. In terms of these definitions we prove the following result:

Theorem 3.2. *If p and N are fixed, and if $V \in C^*_g$, then all of the collapsings V^k, $k = 1, 2, \cdots, g$, of V, are in the cone C^*_{g-1}.*

Proof. The assumption that $V \in C^*_g$ implies that for all $K \subset \{1, 2, \cdots, g\}$, $|K| = N$, $\Sigma_{i_1, \cdots, i_p \in K} v_{i_1 \cdots i_p} \leq 0$. This, in turn, implies that for all $K \subset S_k$, $|K| = N$, this same inequality holds. But, since $K \subset S_k$, $i_1, \cdots, i_p \in K$ implies that $i_1, \cdots, i_p \in S_k$, so

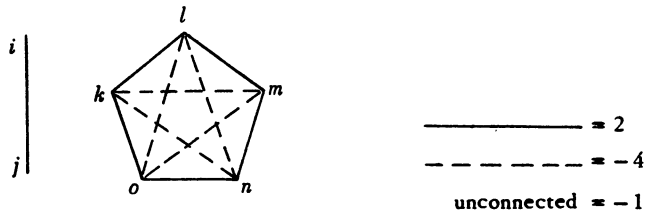
$$\sum_{i_1, \cdots, i_p \in K} v^k_{i_1 \cdots i_p} = \sum_{i_1, \cdots, i_p \in K} v_{i_1 \cdots i_p} \leq 0.$$

Therefore, $V^k \in C^*_{g-1}$.

From Theorem 3.2 we observe that every vector in C^*_g and, in particular, every extreme vector is a lifting of some vector in C^*_{g-1}. Therefore, if one could characterize the set of vectors in C^*_{g-1} which are collapsings of extreme vectors in C^*_g, then by choosing the appropriate liftings and applying them to this set, one could generate C^*_g from C^*_{g-1}. However, the problem of determining such a set still remains open.

Even without such a set it is possible to generate a partial list of extreme

vectors for C_g^* from C_{g-1}^*. This can be done by determining those liftings of the extreme vectors and zero vector of C_{g-1}^* which are extreme vectors of C_g^*. Once these have been found, the problem described above reduces to characterizing those *interior* vectors of C_{g-1}^* which are collapsings of extreme vectors of C_g^*. It should be pointed out that there do exist situations in which all of the collapsings of an extreme vector for a given C_g^* are, indeed, interior vectors of C_{g-1}^*. In particular, for $p = 2$, $N = 3$, and $M = 7$, the vector



where i, j, k, l, m, n, o is any permutation of $1, 2, \cdots, 7$ is extreme in C_7^*, but none of its collapsings are extreme vectors of C_6^*.

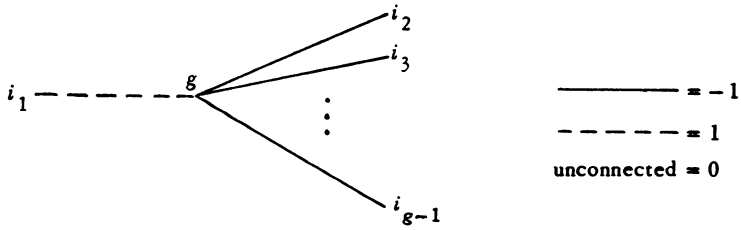
Fixing our attention on the cases in which $p = 2$ and $N = 3$, we prove the following result concerning the liftings of the zero vector in C_{g-1}^* which produce extreme vectors in C_g^*:

Theorem 3.3. *If $p = 2$, $N = 3$, and $g > 5$, then the only liftings of the zero vector in C_{g-1}^* which are extreme vectors in C_g^* are of two types. Up to a positive multiple these are:*

Type I:

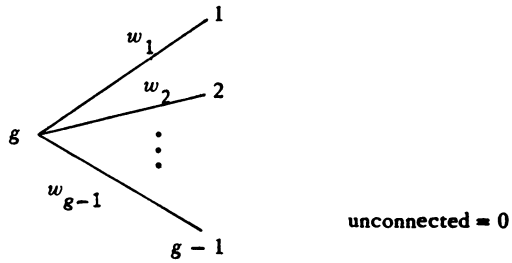
where $i = 1, 2, \cdots, g-1$, and

Type II:



where $i_1, i_2, \cdots, i_{g-1}$ is any permutation of $1, 2, \cdots, g-1$.

Proof. For any vector $V = (v_{ij})$, $1 \leq i < j \leq g$, which is a lifting of the zero vector in C_{g-1}^*, V has the form



where w_i is the vector entry v_{ig}. Furthermore, if two of these weights, w_i and w_j, are positive, then if $K = \{i, j, g\}$,

$$\sum_{m,n \in K} v_{mn} = w_i + w_j > 0,$$

which implies that $V \notin C_g$. Thus, there are two cases to consider:

Case I: $w_i \leq 0$ for all $i = 1, 2, \cdots, g-1$.

Case II: For some i, $w_i > 0$, for all $k \neq i$, $w_k \leq 0$.

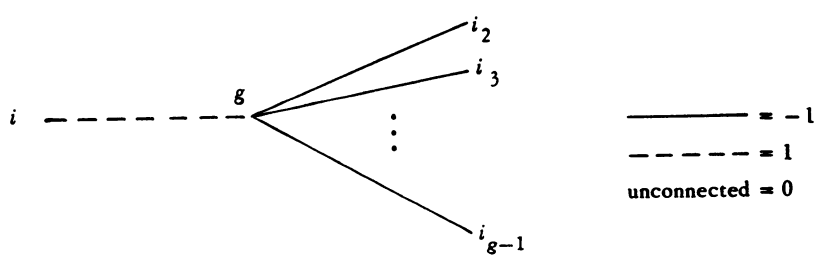
For Case I, since $V \neq 0$, there exists some i such that $w_i < 0$. By assumption $w_j \leq 0$ for all j; therefore, for all sets K of the form $K = \{i, j, g\}$, $j = 1, 2, \cdots, g-1$, $j \neq i$, $\Sigma_{m,n \in K} v_{mn} \neq 0$.

Now, if some other weight $w_j \neq 0$, then since $g > 5$, there exists some $k \neq i, j, g$ and, for the set $K = \{j, k, g\}$, $\Sigma_{m,n \in K} v_{mn} \neq 0$. However, if $W^i = (w^i_{jk})$ denotes the type I vector

then $\Sigma_{m,n \in K} w^i_{mn} \neq 0$ if and only if K is of the form $\{i, j, g\}$, $j \neq i, g$. This implies that for all V satisfying the conditions of Case I, if V is not of type I, then there exists some W^i which is of type I such that $S_V \subset S_{W^i}$ and $S_V \neq S_{W^i}$. Therefore, applying Theorem 3.1, no such V can be an extreme vector.

For Case II, if $w_i > 0$, then since $V \in C^*_g$ implies $\Sigma_{m,n \in K} v_{mn} \leq 0$ for all K, it must be the case that $w_k \leq -w_i$ for all $k \neq i$. But, this implies that for all $K = \{j, k, g\}$, $j, k \neq i, g$, $\Sigma_{m,n \in K} v_{mn} \neq 0$.

Now, if for some $j \neq i$, $w_j < -w_i$, then for $k = \{i, j, g\}$ it is also the case that $\Sigma_{m,n \in K} v_{mn} \neq 0$. However, if $U^i = (u^i_{jk})$ denotes the type II vector



where $\{i_2, i_3, \cdots, i_{g-1}\} = \{1, 2, \cdots, g-1\} - \{i\}$, then $\Sigma_{m,n \in K} u^i_{mn} \neq 0$ if and only if K is of the form $\{j, k, g\}$, $j, k \neq i, g$. Thus, again applying Theorem 3.1, if V satisfies the conditions of Case II and is not of type II, then V is not an extreme vector.

Letting $T = \{S_W | W \in C^*_g\}$, we have shown that the set $T' = \{S_W | W$ is of type I or type II$\}$ contains the only possible maximal elements of T such that $S \subset S_W$, where

$$S = \{A^K \mid K \subset \{1, 2, \cdots, g-1\}, |K| = 3\}.$$

That is to say, the vectors of type I and type II are the only liftings of the zero vector in C^*_{g-1} which can be extreme vectors in C^*_g.

To complete the proof of the theorem, it is only necessary to show that all of the vectors of type I and type II are extreme. For this it is sufficient to show that no element of T' contains any other element of T' as a proper subset. To see this, if W^i is the type I vector described above, then

$$S_{W^i} = S \cup \{A^K \mid K = \{i, j, g\}, j \neq i, g\}.$$

If U^i is the type II vector described above, then

$$S_{U^i} = S \cup \{A^K \mid K = \{j, k, g\}, j, k \neq i, g\}.$$

It is clear by inspection that, for all i, j, $S_{U^i} \not\subset S_{Uj}$, $S_{Wi} \not\subset S_{Wj}$, $S_{U^i} \not\subset S_{Wj}$, and $S_{Wj} \not\subset S_{U^i}$. Thus, all of the elements of T' are maximal elements of T, and the theorem is proved.

The remainder of this section will be concerned with generating extreme vectors of C^*_g by lifting extreme vectors of C^*_{g-1} in the cases in which $p = 2$ and $N = 3$. Recalling that C_g is the cone of those vectors B such that $AX = B$ has a nonnegative solution for X, we have that C^*_g is the cone of vectors such that $WA \leqq 0$. Using the indexing scheme described in the last section, if we let A_k denote the A matrix when $M = k$ (note that this usage of the notation A_k is different from that of the last section), then we have

$$A_g = \begin{pmatrix} A_{g-1} & I_{g-1} \\ 0 & B_{g-1} \end{pmatrix} \tag{44}$$

where I_{g-1} is the $\binom{g-1}{2} \times \binom{g-1}{2}$ identity matrix and B_{g-1} is $(g-1) \times \binom{g-1}{2}$.

If W^1 denotes an extreme vector of C^*_{g-1}, then a lifting of W^1 is of the form $W = (W^1, W^2)$, where W^2 is a vector in $(g-1)$-dimensional space. The following theorem will be useful in determining those W which are extreme vectors of C^*_g:

Theorem 3.4. *If W^1 is an extreme vector of C^*_{g-1}, and if the lifting $W = (W^1, W^2)$ is an extreme vector of C^*_g, then W^2 is an extreme point of $\{Y \mid YB_{g-1} \leqq -W^1\}$.*

Proof. From the form of A_g displayed in equation (44), we see that the condition $WA_g \leqq 0$ breaks into two parts as

$$W^1 A_{g-1} \leqq 0, \tag{45a}$$

$$W^1 + W^2 B_{g-1} \leqq 0. \tag{45b}$$

Assume W^2 is not an extreme point, then there exist Y^1, Y^2, k_1, k_2 such that $W^2 = k_1 Y^1 + k_2 Y^2$, $k_1 + k_2 = 1$, and $Y^i B_{g-1} \leqq -W^1$, $k_i > 0$, $i = 1, 2$. Define two new vectors V^1 and V^2 in $\binom{g}{2}$-dimensional space by $V^1 = (W^1, Y^1)$, $V^2 = (W^1, Y^2)$. Now, $W^1 A_{g-1} \leqq 0$ and $Y^i B_{g-1} \leqq -W^1$ for $i = 1, 2$. Therefore, V^1, $V^2 \in C_g$. But,

$$W = (W^1, W^2) = (W^1, k_1 Y^1 + k_2 Y^2) = k_1(W^1, Y^1) + k_2(W^1, Y^2) = k_1 V^1 + k_2 V^2,$$

which contradicts the hypothesis that W is an extreme vector of C^*_g. Thus, W^2 is an extreme point of $\{Y \mid YB_{g-1} \leqq -W^1\}$.

Theorem 3.4 implies that if $W = (W^1, W^2)$ is an extreme vector of C^*_g, then there exists some $(g-1) \times (g-1)$ nonsingular submatrix M of B_{g-1} such that $W^2 M = -W^1_M$, where W^1_M is the $(g-1)$-dimensional vector of components of the vector W^1 corresponding to the indices of the columns of M. Thus, in order to find all of the liftings of W^1 which are extreme vectors of C^*_g, we consider the elements of the set L_{W^1}, where

$$L_{W^1} = \{W \mid W = (W^1, W^2), \text{ where } W^2 M = -W^1_M, \text{ for some } M \in N\},$$

with

$$N = \{M \mid M \text{ is a } (g-1) \times (g-1) \text{ nonsingular submatrix of } B_{g-1}\}.$$

According to Theorem 3.4, any lifting of W^1 which is extreme in C^*_g is contained in L_{W^1}. Therefore, it is sufficient to determine which elements in L_{W^1} are extreme vectors of C^*_g. But, a vector $V \in C^*_g$ is extreme if and only if V makes a zero inner product with a set of $\binom{g}{2} - 1$ linearly independent columns of A_g. Since W^1 is extreme in C^*_{g-1}, W^1 makes a zero inner product with a set S_1 of $\binom{g-1}{2} - 1$ linearly independent columns of A_{g-1}. Furthermore, for $W \in L_{W^1}$, there exists a set S_2 of $g-1$ linearly independent columns of B_{g-1} such that, for each column $B^i_{g-1} \in S_2$, $W^2 B^i_{g-1} = -W^{1i}$, where W^{1i} is the corresponding component of W^1.

Now, each element in S_1 and S_2 is a portion of a column of A_g. Let S be

the set of columns of A_g which are these extensions of the elements of S_1 and S_2. Clearly, S is a set of linearly independent vectors, and since $|S_1| = \binom{g-1}{2} - 1$ and $|S_2| = g - 1$, we have $|S| = \binom{g}{2} - 1$. Furthermore, for any column $A_g^i \in S$, $WA_g^i = 0$. From this we conclude that for any vector $W \in L_{W^1}$, W is an extreme vector of C_g^* if and only if $W \in C_g^*$. This is equivalent to the condition that

$$WA_g^i \leqq 0 \tag{46}$$

for all $A_g^i \in S$. To carry this one step further, if A_g is partitioned into two submatrices:

$$C_g = \begin{pmatrix} A_{g-1} \\ 0 \end{pmatrix} \quad \text{and} \quad D_g = \begin{pmatrix} I_{g-1} \\ B_{g-1} \end{pmatrix},$$

then, since $W^1 \in C_g^*$, $WC_g \leqq 0$. Therefore, in order to test condition (46), it is sufficient to consider only the set of columns T, where $A_g^i \in T$ if and only if A_g^i is a column of D_g and $A_g^i \notin S$.

We will now organize the results of the preceding discussion into an algorithm for generating all of the vectors W which are extreme vectors of C_g^* from an extreme vector W^1 of C_{g-1}^*. The procedure is the following:

Let

$$N = \{M \mid M \text{ is a set of } g-1 \text{ linearly independent columns of } B_{g-1}\}.$$

For each $M_1 \in N$, let M_2 denote the complementary set of columns in B_{g-1}. This partitioning of the columns of B_{g-1} induces a partition on the entries of W^1 and the columns of A_{g-1}. In terms of this, the condition $WA_g \leqq 0$ becomes

$$W_1^1 A_{g-1}^1 + W_2^1 A_{g-1}^2 \leqq 0, \qquad W_1^1 + W^2 M_1 \leqq 0, \qquad W_2^1 + W^2 M_2 \leqq 0.$$

Let $E = M_1^{-1}$ and $F = EM_2$. Set $W^2 = -W_1^1 E$. The condition that $W \in C_g^*$ is that $W^2 M_2 \leqq -W_2^1$. This can be rewritten as

$$W_2^1 \leqq W_1^1 F. \tag{47}$$

Now, using this construction for each $M_1 \in N$, we have $W = (W^1, W^2)$ is an extreme vector of C_g^* if and only if condition (47) holds. A computerized version of this algorithm has been prepared. In this version of the algorithm the symmetric nature of the structure of both cones C_{g-1}^* and C_g^* has been exploited to produce a complete list of extreme vectors of C_g^* which are liftings of extreme vectors of C_{g-1}^*. The method allows the generation of this complete list without having to use all of the set N.

To understand this, we will say that two extreme vectors of some C_k^* are equivalent if one can be gotten from the other by exchanging its entries according to some permutation induced by a permutation of the indices $1, 2, \cdots, k$. In a

similar fashion, two sets $M, M' \in N$ are equivalent if the triples indexing the columns of one can be gotten from the triples indexing the columns of the other by a permutation induced by a permutation of the indices $1, 2, \cdots, g$, which leaves g fixed.

Now, we employ one representative from each equivalence class of the sets $M \in N$ and apply it to a complete set of extreme vectors for C^*_{g-1}. It is clear that at least one representative of each class of extreme vectors which are liftings of extreme vectors of C^*_{g-1} will be generated in this manner. This is the method used in the computerized algorithm. Results employing this program will be presented in the next chapter.

IV. OBSERVATIONS AND SUMMARY OF RESULTS

1. Some remarks on the geometry of C and C^*. In the preceding chapters the diagonal N-representability problem has been divided into three cases, depending upon the values of the parameters p, N, and M. These are

Case I: $M = N + p$;

Case II: $M < N + p$;

Case III: $M > N + p$.

In this section we will describe the geometry of the cones C and C^*, which will be useful in understanding why this is a natural approach to the problem.

Recalling the definition given in §I.3, a cone is said to be pointed if it does not contain any subspace. The term *blunted* will be used to describe a cone which does contain a subspace. It is clear that if C is a blunted cone in a vector space V, then V has a direct sum decomposition $V = V_1 \oplus V_2$ such that C decomposes as $C = C_1 \oplus V_2$, where C_1 is a pointed cone in V_1.

For a given p, N, and M, let $S = \{K | K \subset \{1, 2, \cdots, M\}, |K| = N\}$. For each $K \in S$, define the vector $E^K = (e^K_{i_1 \cdots i_p})$ by

$$e^K_{i_1 \cdots i_p} = \begin{cases} 1, & \text{if } i_1, \cdots, i_p \in K, \\ 0, & \text{otherwise.} \end{cases}$$

The cone C has as a complete set of extreme vectors the set

$$T = \{E^K \mid K \in S\}. \tag{48}$$

From this the cone C^* can be defined as

$$C^* = \bigcap_{K \in S} H^K, \tag{49}$$

where $H^K = \{V \mid V \cdot E^K \leq 0\}$. In general, the cones C and C^* lie in real $\binom{M}{p}$-dimensional space, and $|S| = \binom{M}{N}$.

In Case I, since $M = N + p$, $\binom{M}{p} - \binom{M}{N}$. Also, the elements of T are linearly independent. From this it follows that C is pointed and of full dimension. Employing Theorem 1.8, for each E^K there is a uniquely associated (up to a positive multiple) extreme vector W^K of C^*. Letting $\overline{T} = \{W^K | K \in S\}$, the elements of $\overline{T}$ are linearly independent, and C^* is pointed and of full dimension.

For each $K \in S$, let F^K denote the bounding hyperplane of H^K. That is,

$$F^K = \{V \mid V \cdot E^K = 0\}. \tag{50}$$

In a similar manner, for each $W^K \in \overline{T}$, we can define

$$\overline{H}^K = \{V \mid V \cdot W^K \leq 0\} \quad \text{and} \quad \overline{F}^K = \{V \mid V \cdot W^K = 0\}.$$

In terms of these definitions, we can describe the geometric connection between C and C^* in a particularly simple manner. To do this, let R^K and $\overline{R}^K$ be the extreme rays defined by

$$R^K = \{kW^K \mid k \geq 0\} \quad \text{and} \quad \overline{R}^K = \{kE^K \mid k \geq 0\}.$$

Then, from Theorem 1.8, each R^K is the intersection of $\binom{M}{p} - 1$ hyperplanes, namely $R^K = \bigcap_{K' \neq K} F^{K'}$. Similarly, $\overline{R}^K$ is also the intersection of $\binom{M}{p} - 1$ hyperplanes, namely $\overline{R}^K = \bigcap_{K' \neq K} \overline{F}^{K'}$. That is, the set of extreme rays of C^* is precisely the set of all possible intersections of $\binom{M}{p} - 1$ of the $\binom{M}{p}$ hyperplanes F^K, and the set of extreme rays of C is the set of all possible intersections of $\binom{M}{p} - 1$ of the $\binom{M}{p}$ hyperplanes $\overline{F}^K$.

For Case II, since $M < N + p$ and $N > p$, $\binom{M}{N} < \binom{M}{p}$. Let V be the vector subspace of $\binom{M}{p}$-dimensional space having the set T as a basis, where T is defined by equation (48). The cone C_V (C restricted to V) is pointed and of full dimension. The cone C^* is then the direct sum of the polar cone of C_V in V, denoted by C_V^* and the subspace $V^\perp$, which is the orthogonal complement of V in $\binom{M}{p}$-dimensional space.

With this description of C^*, we can now explain the geometric content of Theorem 2.6, which gives the solution to the diagonal N-representability problem when $M < N + p$. Let W^K be defined as in equation (23). For each such W^K, let Z^K be the vector in $\binom{M}{p}$-dimensional space defined by

$$z^K_{i_1 \cdots i_p} = w^K_{i_1 \cdots i_p}, \qquad 1 \leq i_1 < \cdots < i_p \leq M.$$

That is, Z^K is obtained from W^K by deleting all entries which are indexed by p-tuples involving any of the individual indices $M + 1, M + 2, \cdots, N + p$.

By Theorem 2.6, the W^K are partitioned into two sets. This induces a partition of the Z^K into two sets, namely

$$S_1 = \{Z^K \mid K \subset \{1, 2, \cdots, M\}\} \quad \text{and} \quad S_2 = \{Z^K \mid K \not\subset \{1, 2, \cdots, M\}\}.$$

From the result of Theorem 2.6, it is then clear that S_1 is a complete set of extreme vectors for C_V^*, and S_2 is a spanning set for $V^\perp$. This, then, clarifies the remark which follows Theorem 2.6.

For Case III, a complete set of extreme vectors for C^* can be described as follows: Let T be the set of vectors defined by equation (48). C^* is the cone in $\binom{M}{p}$-dimensional space defined by equation (49). Now a vector V in $\binom{M}{p}$-dimensional space is an extreme vector of C^* if and only if $V \in C^*$, and it is contained in the intersection of $\binom{M}{p} - 1$ faces F^K of C^*, whose outward pointing normals at the origin are linearly independent. That is, $V \in C^*$ is an extreme vector of C^* if and only if there exists some set $T' \subset T$, such that $|T'| = \binom{M}{p} - 1$, the elements of T' are linearly independent, and if $I = \{K \mid E^K \in T'\}$, then $V \in \bigcap_{K \in I} F^K$, where F^K is defined by equation (50). From this it is easily seen that the number of possible extreme rays of C^* is given by

$$\binom{\binom{M}{N}}{\binom{M}{p} - 1},$$

which grows quite rapidly as the parameters M, N, and p increase. For example, if $M = 8$, $N = 3$, and $p = 2$, then

$$\binom{\binom{8}{3}}{\binom{8}{2} - 1} = \binom{56}{27} \approx 7.387 \times 10^{15}.$$

2. Summary of results. In this section we will summarize the results obtained for the diagonal N-representability problem and display solutions to some special cases when $M > N + p$. If $M = N + p$, then a vector L in $\binom{M}{p}$-dimensional space is diagonal N-representable if and only if

(i) $$\sum_{1 \le i_1 < \cdots < i_p \le M} L_{i_1 \cdots i_p} = \binom{N}{p}$$

and

(ii) *for all sets K of order N contained in $\{1, 2, \ldots, M\}$, if W^K is defined as in equation* (23), *then* $W^K \cdot L \le 0$.

If $M < N + p$, and if L is a vector in $\binom{M}{p}$-dimensional space, then L is diagonal N-representable if and only if

(i) $$\sum_{1 \le i_1 < \cdots < i_p \le M} L_{i_1 \cdots i_p} = \binom{N}{p}$$

and

(ii) for all sets K of order N contained in $\{1, 2, \cdots, N+p\}$, if W^K is defined as in equation (23), and $\overline{L}$ is defined as in equation (28), then

$$W^K \cdot \overline{L} \begin{cases} \leq 0, & \text{if } K \cap \{N+t+1, \cdots, N+p\} = \varnothing, \\ = 0, & \text{if } K \cap \{N+t+1, \cdots, N+p\} \neq \varnothing. \end{cases}$$

For the case in which $M > N + p$, a vector L in $\binom{M}{p}$-dimensional space is diagonal N-representable if and only if

(i) $$\sum_{1 \leq i_1 < \cdots < i_p \leq M} L_{i_1 \cdots i_p} = \binom{N}{p}$$

and

(ii) if S is a complete set of extreme vectors for the cone C^* defined as in equation (49), then, for all $W \in S$, $W \cdot L \leq 0$.

We now display the set S of extreme vectors in several special cases.

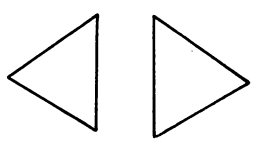
Case I: $p = 2$, $N = 3$, *and* $M = 6$. In this case S consists of 70 vectors. Employing the equivalence relation described in §III.4, these vectors fall into four classes, namely:

Class 1: (15 vectors)

——— = −1

unconnected = 0

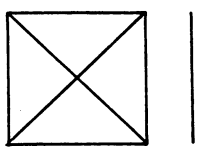
Class 2: (10 vectors)



——— = −2

unconnected = 1

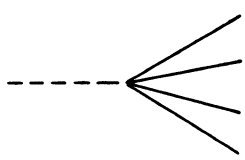
Class 3: (15 vectors)



——— = −2

unconnected = 1

Class 4: (30 vectors)



——— = −1

– – – – – – = 1

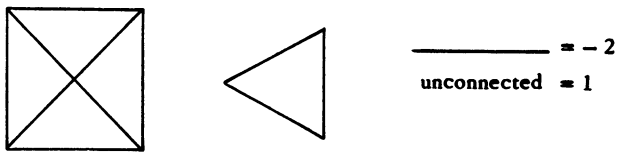
unconnected = 0

Case II: $p = 2$, $N = 3$, *and* $M = 7$. In this case S consists of 896 vectors falling into 7 classes, namely:

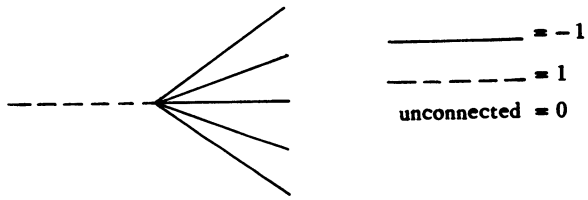
Class 1: (21 vectors)

Class 2: (21 vectors)

Class 3: (35 vectors)

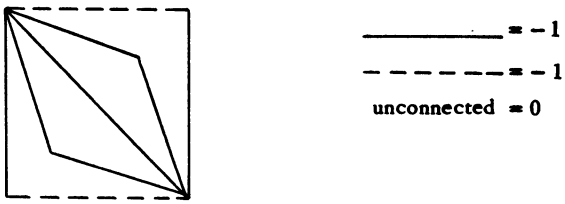


Class 4: (105 vectors)

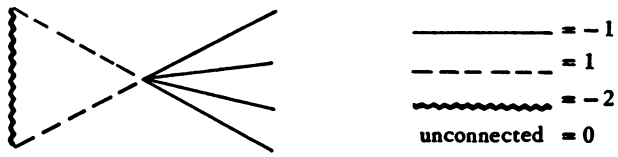


Class 5: (252 vectors)

Class 6: (420 vectors)



Class 7: (105 vectors)

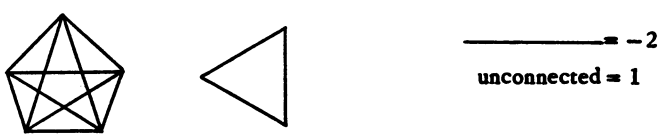


Both of the above solutions were generated by employing the double description algorithm presented in §§III.2 and III.3. For the case in which $p = 2$, $N = 3$, and $M = 8$, the problem becomes too large to generate the complete set S using this algorithm. For this case a portion of the set S was obtained. This was done by employing the lifting algoritm described in §III.4 and applying the double description algorithm to a face of C_8^*, rather than all of C_8^*. A complete set of those extreme vectors of C_8^* which are liftings of extreme vectors of C_7^* consists of 29,127 vectors falling into 15 classes. These are:

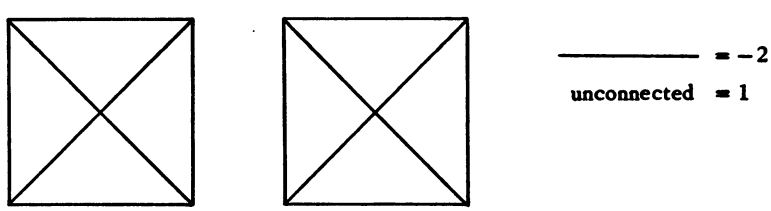
Class 1: (28 vectors)

Class 2: (28 vectors)

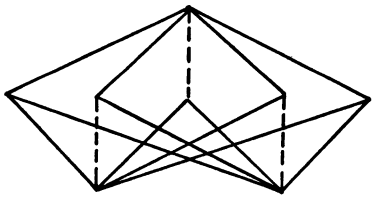
Class 3: (56 vectors)



Class 4: (35 vectors)

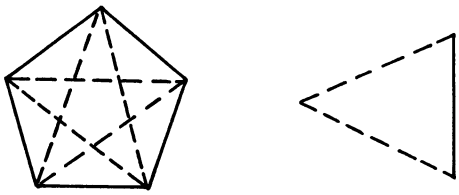


Class 5: (3,360 vectors)



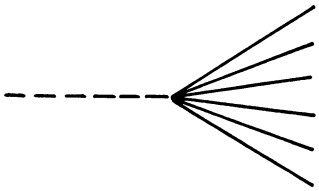
——— = −4
– – – – – = 2
unconnected = −1

Class 6: (2,016 vectors)



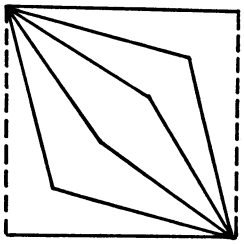
——— = −1
– – – – – = 1
unconnected = 0

Class 7: (56 vectors)



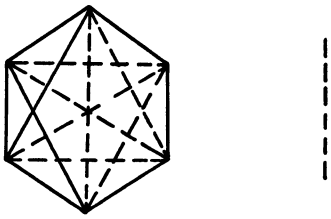
——— = −1
– – – – – = 1
unconnected = 0

Class 8: (840 vectors)



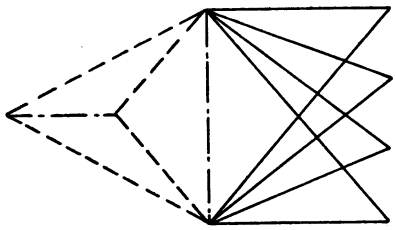
——— = −1
– – – – – = 1
unconnected = 0

Class 9: (5,040 vectors)



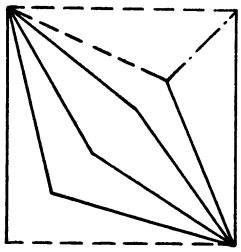
——— = −4
– – – – – = 2
unconnected = −1

Class 10: (420 vectors)



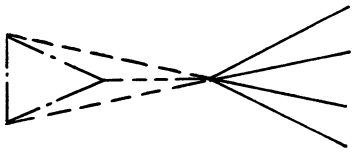
= −1
= 1
= −2
unconnected = 0

Class 11: (3,360 vectors)



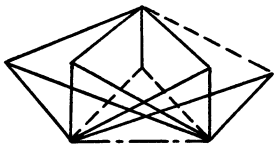
= −1
= 1
= −2
unconnected = 0

Class 12: (280 vectors)



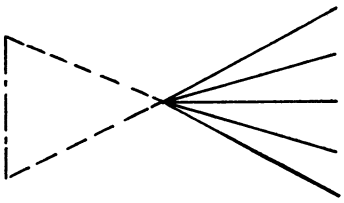
= −1
= 1
= −2
unconnected = 0

Class 13: (3,360 vectors)



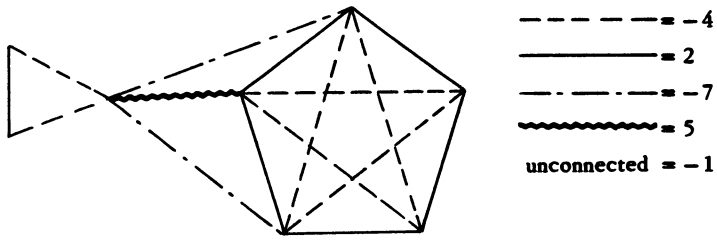
= −1
= 1
= −2
unconnected = 0

Class 14: (168 vectors)



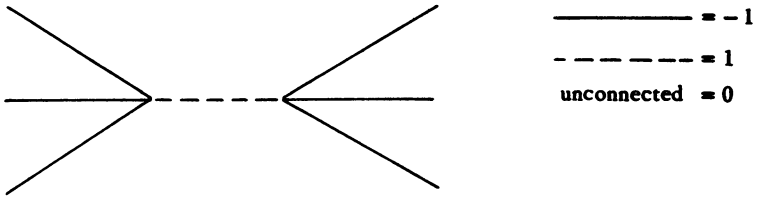
= −1
= 1
= −2
unconnected = 0

Class 15: (10,080 vectors)

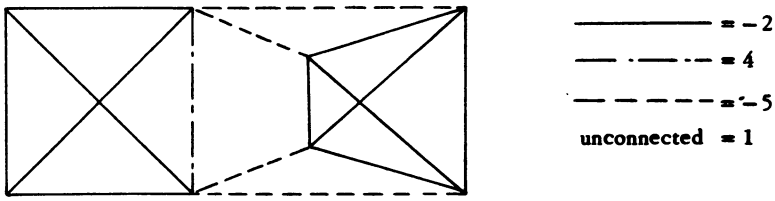


In addition to these, the double description algorithm applied to a face of C_8^* produced 23,240 vectors, none of whose collapsings are extreme in C_7^*. These vectors fall into four classes, namely:

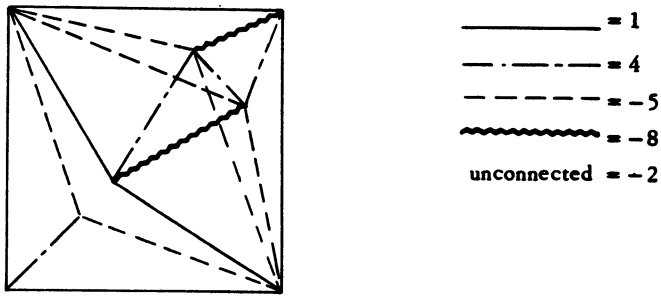
Class 1: (560 vectors)



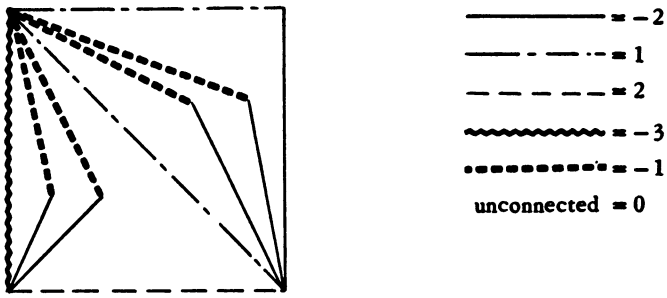
Class 2: (2,520 vectors)



Class 3: (10,080 vectors)



Class 4: (10,080 vectors)



3. **A probabilistic interpretation.** An interpretation of the vectors L can be given in terms of particles filling spin orbitals. From this point of view, for a p matrix, there are N particles and M positions, each of which can hold one particle. Then, $L_{i_1 \cdots i_p}$ is viewed as the product of the total number of p-tuples of particles and the probability that the spin orbitals $i_1, \cdots, i_p$ are all filled by a particular p-tuple of particles. More simply, $L_{i_1 \cdots i_p}$ is just the expected number of p-tuples of particles which lie in the p-tuple of spin orbitals $i_1, \cdots, i_p$. The Pauli principle states that this value lies between zero and one.

As an example of the form that the diagonal N-representability conditions take, in terms of this interpretation, we consider the 2-matrix case (i.e. $p = 2$), when $M = N + p$. Then, from (27a),

$$\sum_{1 \le i_1 < i_2 \le N+2} L_{i_1 i_2} = \binom{N}{2},$$

which simply states that the sum, over all pairs of spin orbitals, of the expected number of pairs of particles lying in a pair of spin orbitals is equal to the total number of pairs of particles.

By applying (27a) and (iii) of §I.1.a (equation (4)), condition (27b) may be written as $L_{i_1 i_2} \geq L_{i_1} + L_{i_2} - 1$, for all i_1, i_2. That is, the expected number of pairs of particles lying in a pair of spin orbitals is greater than or equal to the sum of the expected number of particles in each separate spin orbital minus one. This condition means that if both members of a pair have a large expectation of being filled by individual particles, then the pair of spin orbitals has a large expectation of being filled by a pair of particles. Similar interpretations in terms of expected occupancy can be derived for all p matrices.

BIBLIOGRAPHY

1. M. Balinski, Ph.D. Thesis, Princeton University, Princeton, N. J., 1959.

2. A. J. Coleman, *Structure of fermion density matrices*, Rev. Mod. Phys. 35 (1963), 668–689. MR 27 #5571.

3. E. R. Davidson, *Linear inequalities for density matrices*, J. Mathematical Phys. 10 (1969), 725–734. MR 39 #6592.

4. D. Gale, *The theory of linear economic models*, McGraw-Hill, New York, 1960. MR 22 #6599.

5. C. Garrod and J. Percus, *Reduction of the N-particle variational problem*, J. Mathematical Phys. 5 (1964), 1756–1776. MR 30 #896.

6. A. J. Goldman and A. W. Tucker, *Polyhedral convex cones. Linear inequalities and related systems*, Ann. of Math. Studies, no. 38, Princeton Univ. Press, Princeton, N. J., 1956, pp. 19–40. MR 19,446.

7. H. W. Kuhn, *Linear inequalities and the Pauli principal*, Proc. Sympos. Appl. Math., vol. 10, Amer. Math. Soc., Providence, R. I., 1960, pp. 141–147. MR 22 #13104.

8. P.-O. Lowdin, *Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spinorbitals, and convergence problems in the method of configurational interaction*, Phys. Rev. (2) 97 (1955), 1474–1489. MR 16, 983.

9. R. McWeeny, *Some recent advances in density matrix theory*, Rev. Mod. Phys. 32 (1960), 335–369. MR 22 #3527.

10. T. S. Motzkin, H. Raiffa, G. L. Thompson and R. M. Thrall, Ann. of Math. Studies, no. 28, Princeton Univ. Press, Princeton, N. J., 1963, p. 51.

11. M. B. Ruskai, Ph.D. Thesis, University of Wisconsin, Madison, Wis., 1969.

12. R. M. Thrall and L. Tornheim, *Vector spaces and matrices*, Wiley, New York, 1957. MR 19, 241.

13. E. B. Wilson and F. Weinhold, J. Chem. Phys. 46 (1967), 2752.

14. ———, J. Chem. Phys. 47 (1967), 2298.

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08540