850

STATE UNIVERSITY OF NEW YORK AT STONY BROOK

COLLEGE OF ENGINEERING

(NASA-CR-127767) SORTING AND ORDERING SPARSE LINEAR SYSTEMS R.P. Tewarson (State Univ. of New York) Jan. 1970 20 p CSCL 09B

N72-29158

Unclas
G3/08 15766

Report No. 151

SORTING AND ORDERING SPARSE LINEAR SYSTEMS

by

R. P. Tewarson

JANUARY 1970

Sorting and Ordering Sparse Linear Systems+

R. P. Tewarson*

1. Introduction.

Let us consider the solution of the system of simultaneous linear equations

$$Ax = b, \tag{1.1}$$

where A is a non-singular sparse matrix of order n, x and b are n element column vectors. It is well known that the Gaussian Elimination method for the solution of (1.1) is not only simple to implement on the computer but also gives fairly good results for the amount of computational work (Wilkinson, 1965, pp. 244-246). During the forward course of the Gaussian Elimination, generally new non-zero elements are created. But the back substitution part does not lead to any new non-zero elements. We would like to minimize the total number of such non-zero elements created dur- the entire forward course of the Gaussian Elimination. This leads not only to less roundoff errors (since computations involving zeroes are exact in most computers) but also saves the computer storage, because usually the storage released by column being eliminated at a particular stage of the elimination is not sufficient to store the additional non-zero elements created in the remaining columns. Furthermore, minimizing the number of such non-zero elements decreases the round-off errors not only in the forward course but also in the back substitution part of the

+ Invited paper. Conference on "Large Sparse Sets of Linear Equations", April 5-8, 1970 at Oxford University, England. This research was supported in part by the National Aeronautics and Space Administration, Washington, D.C., Grant No. NGR-33-015-013.

* State University of New York at Stony Brook, Stony Brook, New York, 11790, U.S.A.

Gaussian Elimination method, since whenever there is a zero element in the column under consideration no operations are performed on the corresponding element on the right hand side.

In view of the above facts, we would like to transform A by means of row column permutations to a form which leads to the creation of a minimum number of new non-zero elements during the forward course of the Gaussian Elimination. This is equivalent to the "a priori" determination of permutation matrices R and Q, such that

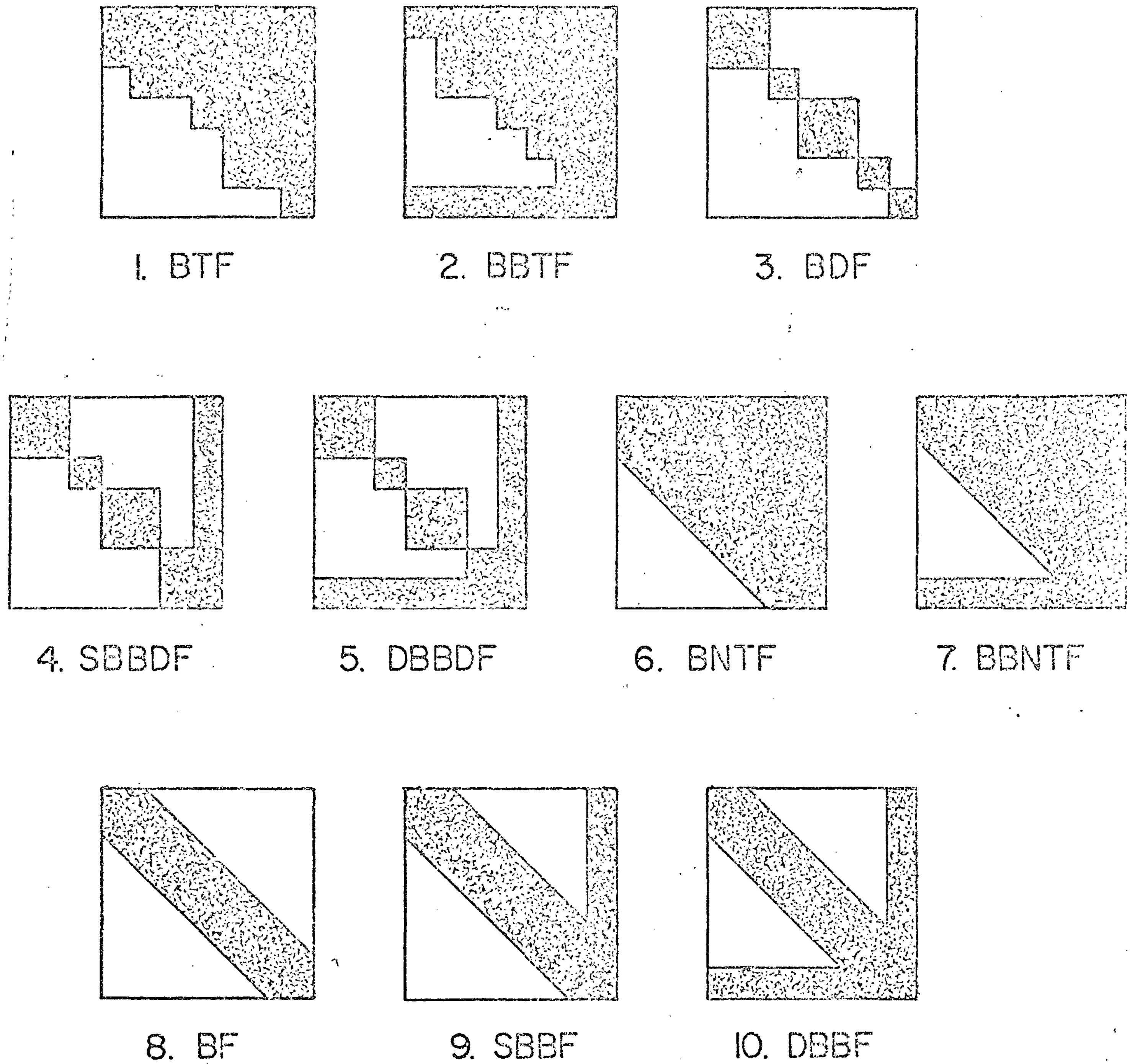
$$RAQ = G, \tag{1.2}$$

and if, $d = Rb$ and $Q'x = y$, then from (1.1) it follows that

$$Gy = d. \tag{1.3}$$

In Fig. 1, some of the forms that G could have, which are desirable for Gaussian elimination, are given, viz., (1) block triangular form(BTF), (2) bordered block triangular form(BBTF), (3) block diagonal form(BDF), (4) singly bordered block diagonal form(SBBDF), (5) doubly bordered block diagonal form(DBBDF), (6) band triangular form (BNTF), (7) bordered band triangular form(BBNTF), (8) band form(BF), (9) singly bordered band form(SBBF), and (10) doubly bordered band form(DBBF). The non-zero elements in each case lie only in the shaded areas. If in each case, the diagonal elements are chosen as pivots, then the new non-zero elements can only be created in the shaded areas during the elimination. If shaded areas contain no non-zero elements, then it is clear that during the elimination process no non-zero elements will be created.

If A is symmetric and positive definite, then in (1.2) it is generally advantageous to have G also symmetric so that only the non-zero elements on and above the diagonal of G need to be stored, and the diagonal elements of A and G are same (though in different positions). A large number of sparse matrices occuring in various application areas



1. BTF
2. BBTF
3. BDF
4. SBBDF
5. DBBDF
6. BNTF
7. BBNTF
8. BF
9. SBBF
10. DBBF

Fig. 1

are symmetric and positive definite. In such cases, in place of (1.2), we have

$$Q'AQ = G, \tag{1.4}$$

and cases (3), (5), (8) and (10) in Fig. 1 are some of the desirable forms for G. In this paper, we shall be primarily concerned with the determination of Q such that G is either in the DBBDF, BF or DBDF (cases (5), (8) and (10) in Fig. 1). The case when G in in BDF has already been investigated (e.g. Harary, 1962 and Tewarson, 1967).

If A is not symmetric, then several methods are available for transforming it by row-column permutations to one of forms given in Fig. 1. A survey of such methods (as well as general computational methods) for spare matrices is given in (Tewarson, 1969).

In section 2 of this paper we will derive some results for matrices in BF, DBBF and DBBDF, and make use of these results in Section 3 for constructing algorithms to transform an arbitrary symmetric positive definite sparse matrix to BF, DBBF or DBBDF.

2. Matrices in band form, doubly bordered band form, and doubly bordered block diagonal form.

In this section we will derive some useful propoerties of matrices in BF, DBBF and DBBDF, which will be used in the next section for transforming symmetric sparse matrices to one of these forms. Let us assume that G is in band form such that

$$g_{ij} = 0 \text{ for } |i-j| > \lambda \text{ and } P(g_{ij} \neq 0 \text{ for } |i-j| \leq \lambda \text{ with } i \neq j) = p; \tag{2.1}$$

where g_{ij} is the i^{th} row and the j^{th} column element of G, λ is called the bandwidth of G and p is the probability that a non-diagonal element within the band is non-zero(P(...) = p denotes that the probability of '...' is p). The diagonal elements of G are all non-zero, since in view of (1.4) G is positive definite, for A is positive definite. If $p = 1$, then G is said to be a 'full' band matrix. We assume that n is large, $\lambda << n$ and p has a large value, say $0 < \frac{\sqrt{5}-1}{2} \le p \le 1$.

We will make use of a matrix B, which is obtained by replacing each non-zero element of G by unity. B is called the incidence matrix that correspondes to G. Let V be the n dimensional column vector of all ones and e_i the i^{th} column of the identity matrix I of order n. Evidently, $V = \sum_{i=1}^{n} e_i$. Let β_e denote the expected number of the non-zero elements of G (which is defined according to (2.1)). Then

$$\begin{aligned}\beta_e &= V'GV \\ &= n + 2\,[(n-1) + (n-2) + \ldots + (n-\lambda)]p \\ &= n + (2n-1)p\lambda - p\lambda^2.\end{aligned}$$

Solving for λ, we have

$$\lambda = n\left[1-\{1-n^{-1} + (2n)^{-2} - \beta_e\, p^{-1} n^{-2} + (pn)^{-1}\}^{\frac{1}{2}}\right] - \frac{1}{2}\,, \tag{2.2}$$

but $\lambda << n$ and $p \ge \frac{\sqrt{5}-1}{2}$ implies that β_e is of order n (and not n^2), therefore neglecting the terms of the order n^{-2} in (2.2), we have

$$\lambda \approx \frac{\beta_e - n}{2pn}\;. \tag{2.3}$$

If $p = 1$, then the number of non-zero elements in G, viz., β is given by

$$\beta = V'GV = n + (2n-1)\,\lambda - \lambda^2,$$

and

$$\begin{aligned}\lambda &= n\left[1-\{1-\beta n^{-2} + (2n)^{-2}\}^{\frac{1}{2}}\right] - \frac{1}{2}\,, \\ &\approx \frac{\beta - n}{2n}\;. \end{aligned}\tag{2.4}$$

We will need the Boolean powers of the incidence matrix B, which are defined as follows,

$$B^{(h+1)} = B^{(h)} * B,\; h = 1, 2, 3, \cdots, \tag{2.5}$$

where $*$ denotes that when computing the inner product of vectors (in the matrix multiplications), in place of usual addition, Boolean additon is used, viz., $1 + 1 = 1$, and $B^{(1)} \equiv B$. We now have

Theorem 2.1 If $p = 1$ and k is an integer $\leq \frac{n-1}{2\lambda}$, then $B^{(k)}$ is a full band matrix having a bandwidth of $k\lambda$.

In order to prove this theorem we need the following definitions for vectors whose entries consist of only zeroes and ones. Let u and v two such n dimensional column vectors. If $u'v \neq 0$ (or equivalently $u' * v = 1$), then u and v are said to 'intersect' and $u'v$ is the 'length of the intersection' between them (Tewarson 1968). Evidently $u'v = 0$ (or $u' * v = 0$) implies that u and v do not 'intersect'. The 'length' of u is defined as $u'u$ (or $u'V$). Throughout this paper we shall use the term 'length' in the above sense rather than the usual Euclidean length.

Proof of Theorem 2.1. The i^{th} row of $B^{(2)}$, (where $1 \leq i \leq \lambda + 1$) is given by $e_i' B^{(2)} = e_i' B * B$. But the i^{th} row of B (which is identical with its i^{th} column) intersects the first through the $(i + 2\lambda)^{th}$ columns of B. Therefore, the i^{th} row of $B^{(2)}$ has the first $2\lambda + i$ elements non-zero, in contrast with $\lambda + i$ such elemtns in the i^{th} row of B. Similarly, it can be seen that for $\lambda + 1 < i \leq n - \lambda$, 2λ elements on either side of the i^{th} diagonal element are non-zero and for $n - \lambda < i \leq n$, the last $2\lambda + 1 + n - i$ elements are non-zero. Therefore, $B^{(2)}$ is a band matrix of width 2λ. Proceeding in the above manner it can be easily shown that if $B^{(h)}$ is a band matrix of width $h\lambda$ then $B^{(h+1)}$, in (2.5), is also a band matrix of width $(h+1)\lambda$, provided that $2(h+1)\lambda + 1 \leq n$. Therefore, by induction on h, $B^{(k)}$ is a band matrix of band width $k\lambda$ for all k with $2k\lambda + 1 \leq n$ or $k \leq \frac{n-1}{2\lambda}$. This completes the proof of Theorem 2.1.

In order to make use of Theorem 2.1, when $0 < p < 1$, we will need the following.

Theorem 2.2. If the i^{th} elements of u and v are denoted by u_i and v_i, and it is known that either u_i or v_i, or both are equal to zero for a total of $n-\nu$ distinct values of i, and $P(u_i \neq 0) = P(v_i \neq 0) = p$ for ν values of i, then

$$P(u' * v \neq 0) = 1-(1-p^2)^\nu, \tag{2.6}$$

and the expected value of $u'v$ is given by

$$E(u'v) = \nu p^2 \tag{2.7}$$

Proof. Evidently the $n-\nu$ values of i for which u_i or v_i, or both are zero can be safely ignored and for the remaining ν distinct values of i, $P(u_i v_i \neq 0) = P(u_i \neq 0)\,P(v_i \neq 0) = p^2$ and $P(u_i v_i = 0) = 1-p^2$. Therefore $E(u'v) = E(\Sigma\, u_i v_i) = \nu p^2$, and since $u_i * v_i \equiv u_i v_i$, we have $P(u' * v = 0) = P(\Sigma\, u_i v_i = 0) = (1-p^2)^\nu$, which implies (2.6).

Corollary 2.2 If in Theorem 2.2, $P(u_i \neq 0) = P(v_i \neq 0) = p$, for only $\nu-2$ values of i; and for some i_1 and i_2 $(i_1 \neq i_2)$, $u_{i_1} = 1$, $P(v_{i_1} \neq 0) = p$, $v_{i_2} = 1$, $P(u_{i_2} \neq 0) = p$, then

$$P(u' * v \neq 0) = 1-(1-p^2)^{\nu-2}(1-p)^2, \tag{2.8}$$

and

$$E(u'v) = \nu p^2 + 2p(1-p). \tag{2.9}$$

Proof: Since $P(u_{i_1} v_{i_1} \neq 0) = P(u_{i_2} v_{i_2} \neq 0) = p$, or $P(u_{i_1} v_{i_1} = 0) = P(u_{i_2} v_{i_2} = 0) = 1-p$, therefore, similar to the proof of Theorem 2.2, it can be easily shown that $E(u'v) = (\nu-2)p^2+2p = \nu p^2+2p(1-p)$, and $P(u' * v = 0) = (1-p^2)^{\nu-2}(1-p)^2$, from which (2.8) directly follows.

We can now make use of Theroem 2.1 to prove

Theorem 2.3. If the i^{th} row and the j^{th} column element of $B^{(2)}$ is denoted by $b_{ij}^{(2)}$, and $P(b_{ij} \neq 0,\ |i-j| \leq \lambda,\ i \neq j) = p$, and $b_{ii} = 1$, then for $1 \leq i < j \leq n$

$$P(b_{ij}^{(2)} \neq 0) = p_{ij}^{(2)} = 1 - (1-p^2)^{\nu_{ij}}(1-p)^{\beta_{ij}}, \text{ for } |i-j| \leq 2\lambda,$$
$$= 0, \text{ otherwise} \tag{2.10}$$

where

(a) $\nu_{ij} = i + \lambda - 2$, $\beta_{ij} = 2$, for $1 \leq i < j \leq \lambda + 1$,

(b) $\nu_{ij} = i - j + 2\lambda - 1$, $\beta_{ij} = 2$, for $1 \leq i \leq \lambda + 1$ and $\lambda + 1 < j \leq i + \lambda$, or $\lambda + 1 < i < n - \lambda$ and $i < j \leq i + \lambda$,

(c) $\nu_{ij} = i - j + 2\lambda + 1$, $\beta_{ij} = 0$, for $1 \leq i \leq \lambda + 1$ and $i + \lambda < j \leq i + 2\lambda$, or $\lambda + 1 < i < n-\lambda$ and $i + \lambda < j \leq i + 2\lambda$,

(d) $\nu_{ij} = n - j + \lambda - 1$, $\beta_{ij} = 2$ for $n - \lambda < i < j \leq n$.

Proof. In view of Theorem 2.1 and the fact that $\lambda << n$, it is evident that $p_{ij}^{(2)} = 0$, for $|i-j| > 2\lambda$. For $|i-j| \leq 2\lambda$, we have $b_{ij}^{(2)} = e_i' B * Be_j = (Be_i)' * Be_j$. Thus $b_{ij}^{(2)} \neq 0$, if the i^{th} and the j^{th} columns of B have a non-zero intersection. If $1 \leq i < j \leq \lambda + 1$, then in view of Corollary 2.2, and the facts that $b_{ii} = 1$, $P(b_{ij} \neq 0) = p$, $b_{jj} = 1$, $P(b_{ji} \neq 0) = p$, and for only $i + \lambda - 2$ elements $P(b_{ti} \neq 0) = P(b_{tj} \neq 0) = p$; it follows that $P(b_{ij}^{(2)} \neq 0) = P\left[(Be_i)' * (Be_j) \neq 0\right] = 1 - (1-p^2)^{\nu-2}(1-p)^2$, where $\nu = i + \lambda$, and (2.10) follows since $\nu_{ij} = i + \lambda - 2 = \nu - 2$ and $\beta_{ij} = 2$ (case (a)). The proof for the other three cases follows exactly the same routine arguments and is omitted. It should be noted that in case (c), corresponding to the diagonal element of one column there is a zero in the other column, therefore we use Theorem 2.2 instead of Corollary 2.2. This accounts for the fact that $\beta_{ij} = 0$ in case (c).

Corollary 2.3. In Theorem 2.3, if either $\beta_{ij} = 2$, or $\beta_{ij} = 0$ but $\nu_{ij} \geq 2$ and $p \geq \frac{\sqrt{5}-1}{2}$, then $p_{ij}^{(2)} \geq p$.

Proof.

$$p_{ij}^{(2)} \geq p \Longleftrightarrow 1-(1-p^2)^{\nu_{ij}}(1-p)^{\beta_{ij}} \geq p \Longleftrightarrow (1-p^2)^{\nu_{ij}}(1-p)^{\beta_{ij}-1} \leq 1.$$

Therefore, for $\beta_{ij} = 2$,

$p_{ij}^{(2)} \geq p \Longleftrightarrow (1-p^2)^{\nu_{ij}}(1-p) \leq 1$, which holds for all $\nu_{ij} \geq 0$, since $p \leq 1$. On the other hand, for $\beta_{ij} = 0$,

$p_{ij}^{(2)} \geq p \Longleftrightarrow (1-p^2)^{\nu_{ij}} \leq (1-p) \Longleftrightarrow (1-p^2)^2 \leq (1-p) \Longleftrightarrow p^2 + p - 1 \geq 0$, the last inequality is true since $p \geq \frac{\sqrt{5}-1}{2}$.

From the above Corollary, it follows that, for all elements of $B^{(2)}$ within the band, $p_{ij}^{(2)} \geq p$, except those for which $\nu_{ij} = 1$ and $\beta_{ij} = 0$; and in the case of such elements $p_{ij}^{(2)} = p^2 < p$. But $\beta_{ij} = 0$ and $\nu_{ij} = 1$ for only $|i-j| = 2\lambda$; and if in B, the outermost elements in the band are non-zero viz., $b_{qt} = 1$ for $|q-t| = \lambda$, then for $|i-j| = 2\lambda$, $b_{ij}^{(2)} = (Be_i)' * Be_j = b_{i+\lambda,i} * b_{j-\lambda,j} = 1$, since $|i-j| = 2\lambda \Longrightarrow i + \lambda = j - \lambda$. In view of the above results and Corollary 2.3, we have.

Corollary 2.4. If in B, the outermost elements in the band are non-zero and p is the probability of the non-diagonal elements within the band being non-zero, then $p_{ij}^{(2)} \geq p$, $|i-j| \leq 2\lambda$.

We will now give a theorem for α_i, which is defined as the expected value of the sum of the 'intersections' of the i^{th} column of B with all the other columns. In other words,

$$\begin{aligned} \alpha_i &= E\left[\sum_{j\neq i}(Be_i)'(Be_j)\right] = E\left[\sum_j e_i'B^2e_j\right] \\ &= E\left[e_i'B^2(\sum_{j\neq i} e_j)\right] = E\left[e_i'B^2(V-e_i)\right], \end{aligned} \tag{2.11}$$

where B^2 is obtained by usual(not Boolean)matrix multiplication.

Theorem 2.4. If B is a band matrix and α_i is defined by (2.11), then

$$\alpha_i = p\left[p(\tfrac{3}{2}\lambda^2 - \tfrac{5}{2}\lambda + 2) + i(2\lambda p-2p+2) + 2(\lambda-1)\right], \ 1 \leq i \leq \lambda + 1, \tag{2.12}$$

$$= p\left[p(2\lambda^2-5\lambda-1) + 4\lambda + i\, p(2\lambda - \tfrac{i}{2} + \tfrac{3}{2})\right], \ \lambda + 2 \leq i \leq 2\lambda. \tag{2.13}$$

$$= 2p\lambda \left[p(2\lambda-1) + 2\right], \quad 2\lambda + 1 < i \le n - 2\lambda. \tag{2.14}$$

Proof. If $1 \le i \le \lambda + 1$, then at most $i + 2\lambda$ columns have a non-zero intersection with the i^{th} column. Out of these columns the diagonal elements have to be considered in the first $i + \lambda$ columns. If in Theorem 2.2 and Corollary 2.2, we let $u = Be_i$ and $v = Be_j$, then from (2.9) and (2.7) it follows that

$$E\left[(Be_i)'(Be_j)\right] = E(e_i'B^2e_j) = \nu_{ij}p^2 + 2p(1-p), \; 1 \le j \le i + \lambda, \; j \ne i,$$

$$= \nu_{ij}p^2, \; i + \lambda < j \le i + 2\lambda,$$

where $\nu_{ij} = j + \lambda, \; 1 \le j < i$

$$= i + \lambda, \; i < j \le \lambda + 1$$

$$= i - j + 2\lambda + 1, \; \lambda + 1 < j \le i + 2\lambda.$$

Therefore, in view of the above facts and (2.11) we have

$$\alpha_i = \sum_{j \ne i} E(e_i'B^2e_j), \quad 1 \le j \le i + 2\lambda,$$

$$= \sum_{j \le i+\lambda} \left[\nu_{ij}p^2 + 2p(1-p)\right] + \sum_{j > i+\lambda} \nu_{ij}p^2, \; 1 \le j \ne i \le i + 2\lambda$$

$$= p^2 \sum_{j \ne i} \nu_{ij} + 2(i+\lambda-1)\, p(1-p), \; 1 \le j \le i + 2\lambda$$

$$= 2p(1-p)(i+\lambda-1) + p^2\left[\sum_{j < i} (j+\lambda) + \sum_{i < j \le \lambda+1} (i+\lambda) + \sum_{\lambda+1<j} (i-j+2\lambda+1)\right]$$

which on simplification gives (2.12). Similar computations can be used to prove (2.13) and (2.14).

Similar to α_i, another useful quantity is $\gamma_{\mu j}$ which is given by

Theorem 2.5. If $\gamma_{\mu j}$ is the expected value of the sum of the lengths of intersections of the j^{th} column with the first μ columns of a band matrix B, then

$$\gamma_{\mu j} = \mu\, p \left[p(\lambda + \frac{\mu}{2} - \frac{3}{2}) + 2\right], \; 1 \le \mu < j \le \lambda + 1, \tag{2.15}$$

$$= p\left[p\{\mu(2\lambda - \frac{1}{2} + \frac{\mu}{2} - j) + 2(j-\lambda-1)\} + 2(\mu-j+\lambda+1)\right], \; 1 \le \mu \le \lambda + 1$$

and $\lambda + 2 \le j \le 2\lambda$, (2.16)

$$= \frac{p^2}{2}(2\lambda + \mu - j)(2\lambda + 2 + \mu - j), \quad 1 \le \mu \le \lambda \text{ and } 2\lambda + 1 \le j \le 3\lambda, \text{ or}$$

$$\lambda + 1 \le \mu \le 2\lambda \text{ and } 3\lambda + 1 \le j \le 4\lambda, \text{ or}$$

$$2\lambda + 1 \le \mu \text{ and } \mu + \lambda < j \le \mu + 2\lambda, \quad (2.17)$$

$$= p\left[p\{\mu(2\lambda - \frac{1}{2} + \frac{1}{2}\mu - j) - 2(\lambda+1-j)\} + 2(\lambda+1+\mu-j)\right], \quad \lambda + 1 \le \mu \le 2\lambda,$$

$$\lambda + 2 \le j \le 2\lambda + 1, \quad (2.18)$$

$$= p(\mu-j)\left[\frac{p}{2}(\mu-j) + p(2\lambda - \frac{1}{2}) + 2\right] + p(\lambda+1)(2\lambda p-p+2),$$

$$\lambda + 1 \le \mu \le 2\lambda \text{ and } 2\lambda + 1 \le j \le 3\lambda, \text{ or } 2\lambda + 1 < \mu < j \le \mu + \lambda. \quad (2.19)$$

Proof. Let $1 \le \mu < j \le \lambda + 1$, then

$$\gamma_{\mu j} = E\left[\sum_{i=1}^{\mu} (Be_i)'(Be_j)\right] = \sum_{i=1}^{\mu} E\left[(Be_i)'(Be_j)\right]$$

$$= \sum_{i=1}^{\mu}\left[\nu_{ij}p^2 + 2p(1-p)\right], \text{ using (2.9).}$$

$$= \sum_{i=1}^{\mu}\left[(i+\lambda)p^2 + 2p(1-p)\right], \text{ since } \nu_{ij} = i + \lambda.$$

$$= \frac{\mu(\mu+1)}{2}p^2 + \mu\left[\lambda p^2 + 2p(1-p)\right]$$

$$= \mu p\left[p(\lambda + \frac{\mu}{2} - \frac{3}{2}) + 2\right].$$

This proves (2.15). In similar manner (2.16)-(2.19) can be proved.

In case B is of doubly bordered band form (case 10, in Fig. 1) and σ is the width of the border, then we have

Theorem 2.6. If B is DBBF and for $i \ne j$,

$$P(b_{ij} \ne 0) = p, \text{ for } 1 \le i,\ j \le n - \sigma \text{ and } |i-j| \le \lambda,$$

$$= \hat{p}, \text{ for either } i \text{ or } j \text{ or both in } [n-\sigma+1, n],$$

and α_i is defined according to (2.11), then

$$\alpha_i = p\left[\lambda\, p\left(\frac{3}{2}\lambda + 2i - \frac{5}{2}\right) - 2p(i-1) + 2(\lambda + i - 1)\right] + \sigma\hat{p}\left[(\lambda+i-1)p+(n-1)\hat{p}+1\right],\ 1 \le i \le \lambda + 1, \qquad (2.20)$$

$$= p\left[p(2\lambda^2 - 5\lambda - 1) + 4\lambda + ip\left(2\lambda - \frac{i}{2} + \frac{3}{2}\right)\right] + \sigma\hat{p}\left[(\lambda + i - 1)p + (n-1)\hat{p} + 1\right],\ \lambda + 2 \le i \le 2\lambda \qquad (2.21)$$

$$= 2p\lambda\left[(2\lambda-1)p+2\right] + \sigma\hat{p}\left[2(\lambda p+1)+\hat{p}(n-2)\right],\ 2\lambda + 1 \le i \le n-2\lambda-\sigma \qquad (2.22)$$

$$= \hat{p}\left[\{2(n-\sigma)-\lambda-1\}\lambda p + (2n-2)+(\sigma-1)(2n-\sigma-2)\hat{p}\right],\ n-\sigma < i \le n. \qquad (2.23)$$

Proof. The proof of this theorem follows the same routine arguments as those of Theorem 2.4 and is therefore omitted.

Theorem 2.7 If B is DBBF or DBBDF such that $P(b_{ij} \neq 0) = \hat{p} \le \frac{\sqrt{5}-1}{2}$ for $i \neq j$ and i or j or both in $[n-\sigma+1,n]$, and $\sigma > 1$, then

$$P(b_{ij}^{(2)} \neq 0) \ge \hat{p},\ \text{for all } 1 \le i,\ j \le n.$$

Proof. For all $1 \le i,\ j \le n$ we have, $b_{ij}^{(2)} = e_i'B^{(2)}e_j = (Be_i)' * Be_j$. Since for the last elements of both the i^{th} and the j^{th} columns of B, $P(b_{ti} \neq 0) = P(b_{tj} \neq 0) \ge \hat{p}$(in fact, the inequality holds for only the diagonal elements), therefore in view of (2.6), Corollary 2.3 and the fact that $\sigma > 1$, we have

$$P(b_{ij}^{(2)} \neq 0) \ge 1 - (1-\hat{p}^2)^\sigma \ge \hat{p}\ .$$

3. Permuting matrices to BF, DBBF and DBBDF.

In the preceding section, we gave some results for matrices in BF, DBBF and DBBDF. In this section, we will show how these results can be used to transform an arbitrary symmetric positive definite matrix A to one of these forms. Let S be the matrix obtained from A by replacing each non-zero element of A by one. In view of (1.4), and the definitions of B it is evident that,

$$Q'SQ = B. \qquad (3.1)$$

In the above equation, S is known and we would like to find Q such that B is in BF, DBBF or DBBDF. We assume that S is sparse viz., $V'SV=\beta=o(n)$ and not $o(n^2)$. In order to describe an algorithm for the determination of Q and B we will need a few simple theorems which follow easily from the results given in Section 2.

Theorem 3.1. If $S^{(2)} = S * S$, and there exists a permutation matrix Q such that $Q'SQ = B$, where B is either DBBF or DBBDF, then for all i

$$E(e_i' S^{(2)}V) = o(n). \tag{3.2}$$

Proof. Since Q has only one non-zero element in each row and column, therefore $Q'*Q = Q'Q = I$, $Q'V = V$, and $B = Q'SQ = Q'* S * Q$. Thus, for $1 \le i \le n$,

$$e_i' S^{(2)}V = e_i' S * S\ V = e_i' Q\ B\ Q'* Q\ B\ Q'V = e_i' Q(B * B)V$$

$$= e_j' B^{(2)}V, \text{ for } 1 \le j \le n.$$

But from Theorem 2.7, $P(b_{ij}^{(2)} \neq 0) \ge \hat{p}$, which implies that

$$E\left[e_i' S^{(2)}V\right] = E\left[e_j' B^{(2)}V\right] \ge \hat{p}\ n = o(n).$$

Corollary 3.1. If in Theorem 3.1, B is a band matrix, then

$$E(e_i' S^{(2)}V) = o(\lambda) << o(n). \tag{3.3}$$

Proof. From Corollary 2.3, $P(b_{ij}^{(2)} \neq 0) \ge p$(except for the outermost element in the band,) therefore, $E(e_i' S^{(2)}V) = E\left[e_j' B^{(2)}V\right] \ge p(2\lambda) = o(\lambda) << o(n)$, since $\lambda << n$.

In making use of the above Corollary, λ can be estimated by using (2.4), where $\beta = V'S\ V$. It should be noted that, in view of (2.3) and the fact that $0 < p \le 1$, the value of λ so obtained is generally an underestimate.

In order to find the rows and columns of S, that correspond to the last σ rows and columns of B(when B is DBBF or DBBDF), we will use the following Theorem. Let Γ denote the set of indices of those rows and columns of S which after permutation according to (3.1), become the last σ rows and columns of B, then we have

Theorem 3.2. If in (3.1), B is in DBBF with $p = \hat{p}$, $\lambda = \sigma$, and $Q' e_i = e_j$, then

$$E\left[e_i' \, S^2(V-e_i)\right] \approx 2np\left[(2\lambda - 1)p + 1\right], \; i \subset \Gamma \tag{3.4}$$

and $$\max_i E\left[e_i' \, S^2(V - e_i)\right] \approx \lambda p^2 n, \; i \not\subset \Gamma. \tag{3.5}$$

Proof. From (2.11), (3.1) and the facts that $QV = V$, $e_i = Qe_j$, we have

$$\alpha_j = E\left[e_j' \, B^2(V - e_j)\right] = E\left[e_j' \, Q' \, S^2 \, Q(V - e_j)\right] = E\left[e_i' \, S^2(V - e_i)\right].$$

But from (2.23) and the fact that $\lambda << n$, we have

$$\begin{aligned}\alpha_j &= p\left[\lambda p(2n - 3\lambda - 1) + (2n - 2) + (\lambda - 1)(2n - \lambda - 2)p\right] \\ &= 2p\left[\lambda p(2n - 2\lambda - 1) + (n - 1)(1 - p)\right] \\ &\approx 2np\left[(2\lambda - 1)p + 1\right], \text{ which proves (3.4).}\end{aligned}$$

On the other hand, for $i \not\subset \Gamma$, $E\left[e_i \, S^2(V - e_i)\right]$ will be maximum for $2\lambda + 1 \le i \le n - \sigma - 2\lambda$, and from (2.22) it follows that

$$\begin{aligned}\alpha_j &= \lambda p\left[p(6\lambda + n - 4) + 6\right] \\ &\approx \lambda p^2 n, \text{ which proves (3.5)}\end{aligned}$$

From the above theorem it follows that

$$E\left[e_i' \, S^2(V - e_i)\right] \approx \theta \, E\left[e_j' \, S^2(V - e_j)\right], \tag{3.6}$$

where $i \subset \Gamma$ and $j \not\subset \Gamma$, and $\theta = 4 + \frac{2}{\lambda}(\frac{1}{p} - 1) \ge 4$, since $0 < p \le 1$. It can be shown that (3.6) also holds for DBBDF, if we assume that the diagonal blocks are of average size λ. However, in this case $\theta \ge 3$. Therefore, we can generally make use of S^2 to determine the rows and columns of S which belong to Γ. If such rows and columns are removed from S, then we need to determine whether the remaining matrix can be transformed to the BDF or BF. To this end we will need the following.

Theorem 3.3. If B is in band form and $S^{(h)} = S^{(h)} * S$, $h = 1, 2, \ldots$, and $k \geq \frac{n}{\lambda}$, then

$$E(e_i' S^{(k)} V) = o(n). \tag{3.7}$$

Proof. In the proof of Theorem 2.1 we have seen that the bandwidth of $B^{(k+1)}$ is λ more than that for $B^{(k)}$ if $p = 1$. Therefore it follows that $B^{(k)} = V'V$ for $k \geq \frac{n}{\lambda}$. In case $0 < p < 1$, then from Corollary 2.3 it follows that for nearly all elements $b_{ij}^{(k)}$ of $B^{(k)}$, $P(b_{ij}^{(k)} \neq 0) \geq p$. Therefore $E(e_i' B^{(k)} V) = o(n)$, since $1 \geq p \geq \frac{\sqrt{5}-1}{2}$.

Theorem 3.4. If in (3.1), B is in BDF with $P(b_{ij} \neq 0) = p$ for i, j in any of the diagonal blocks and zero otherwise, and m is the size of the largest diagonal block, then

$$\operatorname*{Max}_{i,\, k} \; E(e_i' S^{(k)} V) \leq m. \tag{3.8}$$

Proof: Since only the columns belonging to the same diagonal blocks can have a non-zero intersection and the Boolean powers of B increase the probability(of being non zero)of those elements that lie in the diagonal blocks, therefore at most m elements can be non-zero in any row or column, and (3.8) follows. This completes the proof of the theorem.

If we know that S can be permuted to the form of a band matrix, then we need the following results for ordering the rows and columns of S (viz., to determine Q).

From the proof of Theorem 3.2, we have

$$E\left[e_i S^2 (V - e_i)\right] = E\left[e_j B^2 (V - e_i)\right] = \alpha_j, \text{ where } Q\, e_j = e_i; \tag{3.9}$$

and from (2.12) it follows that

$$\alpha_j - \alpha_i = p(2\lambda p - 2p + 2)(j - i), \; 1 \leq i < j \leq \lambda + 1$$

and $\min_{i,j} (\alpha_j - \alpha_i) = p(2\lambda p - 2p + 2)$, $1 \leq i < j \leq \lambda + 1$

$$> \frac{\lambda+1}{2}, \text{ since } p > \frac{1}{2}. \tag{3.10}$$

Let V_μ be the vector obtained from V by replacing its last $n - \mu$ elements by zero. Then $\gamma_{\mu j}$, which was defined in Theorem 2.5, can be expressed as

$$\gamma_{\mu j} = E(e_j' \; B^2 V_\mu), \; j > \mu. \tag{3.11}$$

If we let $QV_\mu = \Omega_\mu$ and $Qe_j = e_i$, then from (3.10) and (3.1) it follows that

$$\gamma_{\mu j} = E(e_j' \; B^2 V_\mu) = E(e_i' \; S^2 \; \Omega_\mu). \tag{3.12}$$

We are now finally in a position to describe an algorithm for finding a permutation matrix Q corresponding to a given sparse symmetric positive definite matrix A such that the matrix G defined according to (1.4) is in DBBF, DBBDF, or BF.

Algorithm 3.1.

1. Construct S, the incidence matrix corresponding to A and compute S^2. From S^2, construct the corresponding incidence matrix $S^{(2)}$. If for all i, $e_i' \; S^{(2)} V = o(n)$, then go to step 6 (In view of Theorem 3.1 and Corollary 3.1, B can be either DBBDF, or DBBF but not in BF or BDF).

2. Compute $\beta = V'SV$, $\lambda \approx \frac{\beta - n}{2n}$ and $S^{(k)}$, where $k \geq \frac{n}{\lambda}$. If $\underset{i}{\mathrm{Max}} \; e_i \; S^{(k)} V = o(n)$, then go to step 4(B is in band form-this follows from Theorems 3.3 and 3.4 and the fact that $m << n$, since A is sparse. It should be noted that $\lambda \geq \underset{i}{\mathrm{Max}}(e_i' \; S \; V - 1)$, since $2\lambda + 1$ is the maximum number of non-zero elements in any row of B; also in view of (2.3), the value of λ given by $\lambda \approx \frac{\beta - n}{2n}$ is generally an underestimate).

3. Compute $S^{(n)}$ and denote its i^{th} row and j^{th} column element by $s_{ij}^{(n)}$. Then $s_{ij}^{(n)} \neq 0$, for all columns(rows)of S which belong to the same diagonal block as the i^{th} column(row). Starting with the first column, assign each column(row)of S to a particular diagonal block. This determines Q such that $Q'SQ$ is in BDF(Harary 1962, Tewarson 1967). Stop.

4. Determine 2λ values of η for which

$\hat{\alpha}_\eta = e'_\eta S^2(V-e_\eta) \le e'_i S^2(V-e_i), i \ne \eta, 1 \le i \le n$. Separate these values of η into two sets as follows. If $s^{(2)}_{\eta_r \eta_k} = 0$ (or $e'_{\eta_r} S^2 e_{\eta_k} = 0$), $r \ne k$, then η_r and η_k belong to different sets. Within each set arrange the values of η's in the order of ascending values of $\hat{\alpha}_\eta$. Let $\eta_1, \eta_2, \ldots, \eta_\lambda$ and $\bar{\eta}_1, \bar{\eta}_2, \ldots, \bar{\eta}_\lambda$ be the resulting arrangements for the η's in the first and the second set respectively, then $e_{\eta_1}, e_{\eta_2}, \ldots, e_{\eta_\lambda}$ are the first λ columns and $e_{\bar{\eta}_\lambda}, \ldots, e_{\bar{\eta}_2}, e_{\bar{\eta}_1}$ are the last λ columns of Q. (Remarks: Note that λ was estimated in step 2 of this algorithm. Furthermore, from (3.9) and (3.10) it follows that for the η's in each set, the values of $\hat{\alpha}_\eta$'s are generally distinct. Ties can be broken by using $e'_\eta SV$.). Construct an n dimensional column vector Ω which has unity in positions $\eta_1, \eta_2, \ldots, \eta_\lambda$ and zeroes elsewhere.

5. Compute $\hat{\gamma}_\tau = \max_i e'_i S^2 \Omega$, $i \ne \alpha_\eta$, then e_τ is the next column of Q. (This follows from (3.12), (2.15) and (2.16). It can easily be shown that if τ has more than one value, then the corresponding columns of B are very close together. We can use $e'_\tau S^2(V-e_\tau)$ to break the ties in the beginning if any.). Make the τ^{th} element of Ω a one. Similarly the additional columns of Q from the right hand side are also determined by using $\bar{\Omega}$, which has unity in positions $\bar{\eta}_1, \bar{\eta}_2, \ldots, \bar{\eta}_\lambda$. Repeat the current step of the algorithm until all columns of S have been exhausted, viz., $\Omega + \bar{\Omega} = V$, and Q has been determined. Stop.

6. Compute $\hat{\alpha}_j = e_j S^2(V-e_j)$, $j = 1, 2, \ldots, n$. Determine the set Γ, such that if $\rho \subset \Gamma$ and $k \not\subset \Gamma$ then $\hat{\alpha}_\rho$ is significantly greater than $\hat{\alpha}_k$. (For example, $\hat{\alpha}_\rho \approx \theta \hat{\alpha}_k$, where $\theta \approx 4$, this follows from (3.6)). Let $\rho_1, \rho_2, \ldots, \rho_\sigma \subset \Gamma$. Then $e_{\rho_1}, e_{\rho_2}, \ldots, e_{\rho_\sigma}$ are the last columns of Q. Now delete the rows and columns of S which belong to Γ and we have a matrix of order $n-\sigma$, which is either in BF or BDF. Go to step 2 with n replaced

by n-σ to determine the first n-σ columns of Q. This completes Algorithm 3.1.

We shall now make a few pertinent remarks about the above algorithm. Let φ be the undirected graph which corresponds to S such that it has n nodes and there is an edge between its i^{th} and j^{th} nodes if and only if $s_{ij} = s_{ji} = 1$, (Busacker and Saaty, 1965). Then the permutation of the rows and the columns of S (according to (3.1))is equivalent to the rearrangement of the nodes of φ to get an undirected graph ψ which corresponds to B (matrix B is in BF, DBBF or DBBDF). In view of these definitions of φ and ψ, it is evident that the equation $e_i' S^2 V = o(n)$ in the first step of Algorithm 3.1 implies that there is a path of length two or less between most of the nodes of φ (or ψ). Furthermore, in step 6, we determine and delete some nodes and the associated edges of φ, such that the remaining graph does not have most of its nodes connected by paths of length two or less (the associated matrix can be permuted to BF or BDF). In step 3, we make use of the connectivity matrix $S^{(n)}$ to determine the nodes belonging to each connected subgraph of φ (the diagonal blocks of B). The determination of Q in steps 4 and 5 generally does not lead to a matrix which has bandwidth close to the one estimated in step 2, mainly due to the non-uniqueness of the quantities $\hat{\alpha}_\eta$ and $\hat{\gamma}_\tau$, however the rows and columns which will minimize the bandwidth are in general fairly close together in $Q'S\,Q$ at the conclusion of these steps. Therefore, a few additional interchanges of rows and columns might at times be desirable.

The above Algorithm is based on the assumption that there exists a Q such that $Q'\,S\,Q = B$; where B is either in BF, DBBF, or BBBDF and the probability of its elements (within the shaded areas in cases 8, 10 or 5 in Fig. 1) being non-zero is $p \geq \frac{\sqrt{5}-1}{2}$, and m, σ, λ are of same order of magnitude, but much less than n. The closer p is to unity the more efficient the algorithm will be. For arbitrary symmetric matrix S with non-zeros on the diagonal, the efficiency of this Algorithm will have to be decided on the basis of a large number of computational experiments. In any case, the algorithm should certainly do better than the present methods in literature that the author is familiar, due to the following reasons. First, the rows and columns of S which would keep us from minimizing λ or m are put in the set Γ; and second, at each stage of the algorithm we have used more information from the rows and columns of both S and the desired form B than other methods seem to utilize.

We conclude this paper with a brief description of the methods for matrix bandwidth minimization presently available in literature. If we let $\pi_i = i - j$, $j \le i$ and zero otherwise, where a_{ij} is the left most non-zero element of A in the i^{th} row, then Akyuz and Utku (1968) give an iterative program for finding the quantity $\xi = \min_Q \frac{1}{n} \sum_{i=1}^{n} \pi_i$. Their method is based on interchanging two successive rows of A if bandwidth is decreased or a row with large number of zeroes goes away from the central row. The above problem can also be expressed as a Linear Programming problem (Tewarson, 1967). The related problem of finding $\bar{\xi} = \min_Q \max_i \pi_i$ is discussed by Alway and Martin (1965), Cuthill and McKee (1969) and Rosen (1968). Alway and Martin (1965) have constructed a program which by means of an educated search of possible permutations determines Q. Rosen's (1968) program is an iterative scheme which is based on interchanging a pair of diagonal elements of A, such that either $\max_i \pi_i$ is decreased or in certain cases remains the same. Cuthill and McKee (1969) base their scheme on renumbering the diagonal elements of A by looking at a few permutations suggested by the structure of φ (the associated graph).

The Algorithm given in this paper should be especially useful where many problems with similar pattern of non-zero elements but differing values have to be solved. It will perhaps be advantageous to use powers of S greater than two in steps 4 and 5 of the algorithm for greater expected seperation between the $\hat{\alpha}_j$'s and $\hat{\gamma}_\tau$'s. We hope that the probabilistic approach used in this paper will in the future lead to additional algorithms.

December 15, 1969
State University of New York, Stony Brook, New York

References

Wilkinson, J. H. 1965. The Algebraic Eigenvalue Problem, Oxford University Press, London.

Harary, F. 1962. A graph theoretic approach to matrix inversion by partioning, Numer. Math. 4, 128-135.

Tewarson, R. P. 1967. Row-column permutation of sparse matrices. Computer J. 10, 300-305.

Tewarson, R. P. 1970. Computations with sparse matrices, SIAM Rev. 12, (Invited paper; Oct 1, 1969).

Tewarson, R. P. 1968. On the orthonormalization of sparse vectors, Computing 3, 268-279.

Busacker, R. G. and Saaty, T. L. 1965. Finite Graphs and Networks, McGraw-Hill, New York.

Akyuz, F. A. and Utku, S. 1968. An automatic relabeling scheme for bandwidth minimization of stiffness matrices, AIAA Journal 6, 728-730.

Alway, G. G. and Martin, D. W. 1965. An algorithm for reducing the bandwidth of a matrix of symmetrical configuration, Comp. J. 8, 264-272.

Cuthill, E. and McKee, J. 1969. Reducing the bandwidth of sparse symmetric matrices, Applied Math. Lab., Naval Ship Research and Development Centre, Washington, D. C. Tech. Note. AML-40-69.

Rosen, R. 1968. Matrix bandwidth minimization, Proceedings of 23rd National Conference of ACM, Publication P-68, Brandon Systems Press, Princeton, N. J., 585-595.