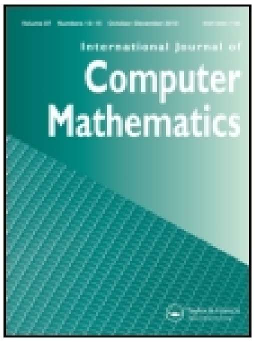

International Journal of Computer Mathematics

ISSN: 0020-7160 (Print) 1029-0265 (Online) Journal homepage: http://www.tandfonline.com/loi/gcom20

A comparison of gaussian and gauss-jordan elimination in regular algebra

R.C. Backhouse & B. A. Carre

To cite this article: R.C. Backhouse & B. A. Carre (1982) A comparison of gaussian and gauss-jordan elimination in regular algebra, International Journal of Computer Mathematics, 10:3-4, 311-325, DOI: 10.1080/00207168208803290

To link to this article: http://dx.doi.org/10.1080/00207168208803290

Intern. J. Computer Math. 1982. Vol. 10, pp. 311-325
0020-7160/82/1004-0311 $06.50/0

Printed in Great Britain

A Comparison of Gaussian and Gauss-Jordan Elimination in Regular Algebra

R. C. BACKHOUSE

Department of Computer Science, Heriot-Watt University, 79 Grassmarket, Edinburgh, EH1 2HJ.

and

B. A. CARRÉ

Department of Electronics, University of Southampton, Southampton, S09 5NH.

(Received February 1981)

A comparison is presented in regular algebra of the Gaussian and Gauss-Jordon elimination techniques for solving sparse systems of simultaneous equations. Specifically, the elimination form and product form of the star $\mathbf{A}^*$ of a matrix $\mathbf{A}$ are defined and it is then shown that the product form is never more sparse than the elimination form. This result generalises an earlier one due to Brayton, Gustavson and Willoughby in which it is shown that the product form of the inverse $\mathbf{A}^{-1}$ of a matrix $\mathbf{A}$ is never more sparse than the elimination form of the inverse. Our result applies both in linear algebra and, more generally, to path-finding problems.

KEY WORDS: Gaussian elimination, Gauss-Jordan elimination, regular algebra, linear algebra, path-finding, sparsity.

C.R. CATEGORIES: 5.14, 5.23, 5.25, 5.32.

1. INTRODUCTION

In the last two decades there has been considerable interest in methods for the solution of sparse simultaneous linear equations. In particular, the well-known Gaussian and Gauss-Jordan elimination methods for solving the matrix equation $\mathbf{Ax}=\mathbf{b}$ have been studied extensively by numerical analysts (Duff, [1]). Recently it has been observed that many path-finding algorithms are in fact variants of such elimination techniques (Carré, [2]; Backhouse and Carré, [3]). Examples are Floyd's shortest path algorithm (Floyd, [4]) and Warshall's transitive closure algorithm (Warshall, [5]), both of which are variants of Gauss-Jordan elimination. Most recently, interest in sparse matrix techniques has been aroused in a wide cross-section of computer scientists since the realisation that "new" algorithms developed to solve global data flow analysis problems can be regarded as applications of Gaussian elimination (Tarjan, [6]), or the equally well-known Gauss-Seidel iterative technique (Tarjan, [7]).

The framework for this unification is the algebra of regular languages. In Backhouse and Carré [3] the Gauss-Jordan and Gaussian elimination techniques were re-expressed using the $*$, $+$ and $.$ operators of regular algebra rather than the traditional inverse and $+$ and $.$ operators of linear algebra. The advantage of so doing is the greater generality of regular algebra: by reinterpreting a^* as $(1-a)^{-1}$ one immediately recovers the algorithms described in texts of numerical analysis, but also by interpreting the $*$, $+$ and $.$ operations in other ways it is possible to solve a wide variety of path-finding problems. A concrete example is finding shortest paths through a graph. Here by interpreting $a+b$ as the minimum of a and b, $a.b$ as the sum of a and b, and a^* as zero, Gauss-Jordan elimination reduces to Floyd's shortest path algorithm. Note, however, that the numerical analysts' formulation of Gauss-Jordan elimination cannot be applied because there is no meaningful interpretation of a^{-1}.

In this paper we employ regular algebra to give a novel presentation of Brayton *et al.*'s comparison of Gaussian and Gauss-Jordan elimination (Brayton, Gustavson and Willoughby, [8]). Our comparison adds insight to their result as well as being relevant to many path-finding problems.

The paper contains five sections. Section 2 reviews the properties of a regular algebra we require and Section 3 summarises the two algorithms. The formal comparison is presented in Section 4 whilst Section 5 discusses the meaning and implications of the comparison.

2. REGULAR ALGEBRA

A regular algebra consists of a set S which is closed under two binary operations $+$ and $.$ and one unary operation $*$. The following properties will be used without mention in the sequel.

A1 $(P+Q)+R=P+(Q+R)$ A4 $P.(Q+R)=(P.Q)+(P.R)$

A2 $(P.Q).R=P.(Q.R)$ A5 $(P+Q).R=(P.R)+(Q.R)$

A3 $P+Q=Q+P$ A6 $P+P=P$

where $P, Q, R \in S$.
The set S contains a null element ϕ such that

A7 $P+\phi=P$ A8 $\phi . P=\phi=P . \phi$

and a unit element e such that

A9 $e . P=P=P . e.$

Finally, the *star* or *closure* operator obeys:

A10 $P^*=e+P . P^*$,

and

A11 $(P+Q)^*=(P^*Q)^*P^*$.

It is useful for us to introduce an additional unary operator, namely

$$P^+=PP^*.$$

P^+ is called the *weak closure* of P. From A10, we have

A12 $P^*=e+P^+$.

Note that addition is assumed to be idempotent (property A6). Thus we may define a partial ordering $\leqq$ on the set S by

$$P \leqq Q \Leftrightarrow P+Q=Q$$

Moreover, by A7,

$$\phi \leqq P \quad \text{for all } P \in S,$$

and by A10 and A12,

$$P \leqq P^+ \leqq P^* \quad \text{for all } P \in S.$$

Given any regular algebra R we can form a new regular algebra $M_n(R)$ consisting of all $n \times n$ matrices whose elements belong to R. In the algebra

$M_n(R)$ the operators $+$ and . and the order relation $\leqq$ are defined as follows: Let $\mathbf{A}=[a_{ij}]$ and $\mathbf{B}=[b_{ij}]$ be any $n\times n$ matrices with elements in R; then

$$\mathbf{A}+\mathbf{B}=[a_{ij}+b_{ij}],\ \mathbf{A}\,.\,\mathbf{B}=\left[\sum_{k=1}^{n} a_{ik}\,.\,b_{kj}\right],$$

and

$$\mathbf{A}\leqq\mathbf{B} \text{ if and only if } a_{ij}\leqq b_{ij} \quad \text{for all } i,j.$$

The *unit matrix* $\mathbf{E}=[e_{ij}]$ is defined as that $n\times n$ matrix with $e_{ij}=e$ if $i=j$ and $e_{ij}=\phi$ if $i\neq j$. The *null matrix*, denoted $\mathbf{\Phi}$, is that matrix all of whose elements are ϕ. The ith row and jth column of a matrix $\mathbf{A}$ will be denoted by $\mathbf{a}_{io}$ and $\mathbf{a}_{oj}$, respectively.

3. THE ALGORITHMS

In linear algebra the elimination and product forms of the inverse are economical representations of the inverse $\mathbf{A}^{-1}$ of a matrix $\mathbf{A}$, constructed by Gaussian elimination and Gauss-Jordan elimination respectively. Analogously, in regular algebra Gaussian elimination is used to construct the *elimination form of a star* (abbreviated EFS) and Gauss-Jordan elimination is used to construct the *product form of the star* (abbreviated PFS), both of which represent $\mathbf{A}^*$ for a given matrix $\mathbf{A}$.

Certain elementary matrices, which differ from the null matrix in only one column or one row, are the primary tool in both algorithms. An *elementary column matrix* is any matrix of the form

$$\mathbf{S}^{(k)}=\begin{bmatrix} & s_{1k} & \\ & \vdots & \\ \mathbf{\Phi} & s_{kk} & \mathbf{\Phi} \\ & \vdots & \\ & s_{nk} & \end{bmatrix}$$

and an *elementary row matrix* is any matrix of the form

$$\mathbf{T}^{(k)}=\begin{bmatrix} & & \mathbf{\Phi} & & \\ t_{k1} & \cdots & t_{kk} & \cdots & t_{kn} \\ & & \mathbf{\Phi} & & \end{bmatrix}$$

It is easy to compute the stars of such matrices. Specifically,

$$\mathbf{S}^{(k)*}=\begin{bmatrix} e & & & s_{1k}s_{kk}^{*} & & & \\ & \ddots & & \vdots & & & \\ & & e & s_{k-1,k}s_{kk}^{*} & & & \\ & & & s_{kk}^{*} & & & \\ & & & s_{k+1,k}s_{kk}^{*} & e & & \\ & & & \vdots & & \ddots & \\ & & & s_{nk}s_{kk}^{*} & & & e \end{bmatrix}$$

and

$$\mathbf{T}^{(k)*}=\begin{bmatrix} e & & & & & & \\ & \ddots & & & & & \\ & & e & & & & \\ t_{kk}^{*}t_{k1} & \cdots & t_{kk}^{*}t_{k,k-1} & t_{kk}^{*} & t_{kk}^{*}t_{k,k+1} & \cdots & t_{kk}^{*}t_{kn} \\ & & & & e & & \\ & & & & & \ddots & \\ & & & & & & e \end{bmatrix}$$

To justify the algorithms for obtaining the PFS and EFS, and to allow us to compare these forms, we shall in each case first give a concise derivation of it, originally presented in Backhouse and Carré [3]. Our notation here follows closely that in Carré [9], where the "path algebra" used is an example of a regular algebra.

Gauss-Jordan elimination

To obtain the PFS of an $n\times n$ matrix $\mathbf{A}$, let $\mathbf{A}^{(0)}=\mathbf{A}$ and let $\mathbf{A}^{(k)}$, $\mathbf{Q}^{(k)}$ and $\mathbf{R}^{(k)}$, $(k=1,2,\ldots,n)$ be the sequences of matrices defined as follows. We express $\mathbf{A}^{(k-1)}$ in the partitioned form

$$\mathbf{A}^{(k-1)}=\begin{bmatrix} \mathbf{A}_{11}^{(k-1)} & \mathbf{A}_{12}^{(k-1)} & \mathbf{A}_{13}^{(k-1)} \\ \mathbf{A}_{21}^{(k-1)} & \mathbf{A}_{22}^{(k-1)} & \mathbf{A}_{23}^{(k-1)} \\ \mathbf{A}_{31}^{(k-1)} & \mathbf{A}_{32}^{(k-1)} & \mathbf{A}_{33}^{(k-1)} \end{bmatrix} \tag{1}$$

where the diagonal submatrices $\mathbf{A}_{11}^{(k-1)}$, $\mathbf{A}_{22}^{(k-1)}$ and $\mathbf{A}_{33}^{(k-1)}$ are square, of order $k-1$, 1 and $n-k$ respectively. (Note that in $\mathbf{A}^{(0)}$, the first row and column of this partition do not exist.) Using the same partitioning,

$$\mathbf{Q}^{(k)} = \begin{bmatrix} \Phi & \mathbf{A}_{12}^{(k-1)} & \Phi \\ \Phi & \mathbf{A}_{22}^{(k-1)} & \Phi \\ \Phi & \mathbf{A}_{32}^{(k-1)} & \Phi \end{bmatrix}, \quad \mathbf{R}^{(k)} = \begin{bmatrix} \mathbf{A}_{11}^{(k-1)} & \Phi & \mathbf{A}_{13}^{(k-1)} \\ \mathbf{A}_{21}^{(k-1)} & \Phi & \mathbf{A}_{23}^{(k-1)} \\ \mathbf{A}_{31}^{(k-1)} & \Phi & \mathbf{A}_{33}^{(k-1)} \end{bmatrix} \tag{2}$$

and

$$\mathbf{A}^{(k)} = \mathbf{Q}^{(k)*}\mathbf{R}^{(k)}, \qquad (k = 1, 2, \ldots, n). \tag{3}$$

Here $\mathbf{Q}^{(k)}$ is an elementary column matrix, whose star is easily calculated.

Now since

$$\mathbf{A}^{(k-1)} = \mathbf{Q}^{(k)} + \mathbf{R}^{(k)}, \tag{4}$$

it follows from A11 that $\mathbf{A}^{(k-1)*} = (\mathbf{Q}^{(k)*}\mathbf{R}^{(k)})^*\mathbf{Q}^{(k)*}$, hence

$$\mathbf{A}^{(k-1)*} = \mathbf{A}^{(k)*}\mathbf{Q}^{(k)*}, \quad (k = 1, 2, \ldots, n), \tag{5}$$

and consequently

$$\mathbf{A}^* = \mathbf{A}^{(0)*} = \mathbf{A}^{(n)*}\mathbf{Q}^{(n)*}\mathbf{Q}^{(n-1)*}\ldots\mathbf{Q}^{(1)*}. \tag{6}$$

Also, it is easily proved, by induction on k, that the first k columns of $\mathbf{A}^{(k)}$ are null (Backhouse and Carré, [3]). Hence $\mathbf{A}^{(n)*} = \mathbf{E}$, and (6) gives

$$\mathbf{A}^* = \mathbf{Q}^{(n)*}\mathbf{Q}^{(n-1)*}\ldots\mathbf{Q}^{(1)*}, \tag{7}$$

which is called the *product form of the star* (PFS) of $\mathbf{A}$.

The fact that the first k columns of $\mathbf{A}^{(k)}$ are null suggests a simple and compact method of forming and storing the $\mathbf{Q}^{(k)}$-factors: we set $\mathbf{M}^{(0)} = \mathbf{A}$ initially, and then repeatedly modify this matrix, obtaining at each stage

$$\mathbf{M}^{(k)} = \sum_{j=1}^{k} \mathbf{Q}^{(j)} + \mathbf{Q}^{(k)*}\mathbf{R}^{(k)}, \quad (k = 1, 2, \ldots, n). \tag{8}$$

On termination, the columns of $\mathbf{M}^{(n)}$ define all the $\mathbf{Q}^{(k)}$-factors, as indicated in Figure 1. With the same partitioning as in (1) and (2), it follows from (8) that

$$\left[\begin{array}{c|c|c|c} Q^{(1)} & Q^{(2)} & \quad \cdot \quad \cdot \quad \cdot \quad & Q^{(n)} \end{array}\right]$$

FIGURE 1 Product Form of the Star

$$\mathbf{M}^{(k)} = \begin{bmatrix} \mathbf{M}_{11}^{(k-1)} & \mathbf{M}_{12}^{(k-1)} & \mathbf{M}_{13}^{(k-1)} + \mathbf{M}_{12}^{(k-1)} \mathbf{M}_{22}^{(k-1)*} \mathbf{M}_{23}^{(k-1)} \\ \mathbf{M}_{21}^{(k-1)} & \mathbf{M}_{22}^{(k-1)} & \mathbf{M}_{22}^{(k-1)*} \mathbf{M}_{23}^{(k-1)} \\ \mathbf{M}_{31}^{(k-1)} & \mathbf{M}_{32}^{(k-1)} & \mathbf{M}_{33}^{(k-1)} + \mathbf{M}_{32}^{(k-1)} \mathbf{M}_{22}^{(k-1)*} \mathbf{M}_{23}^{(k-1)} \end{bmatrix}. \tag{9}$$

Thus, the matrix $\mathbf{M}^{(n)}$ can be computed by the following simple algorithms:

for $k := 1$ **to** $n-1$ **do**

  for $i := 1$ **to** n **do**

    for $j := k+1$ **to** n **do**

      $m_{ij} := m_{ij} + m_{ik}\, m_{kk}^*\, m_{kj};$

When the PFS has been obtained in this way, it is possible to compute the least solution of a set of equations $\mathbf{y} = \mathbf{A}\mathbf{y} + \mathbf{b}$, viz

$$\mathbf{y} = \mathbf{A}^*\mathbf{b} = \mathbf{Q}^{(n)*}\, \mathbf{Q}^{(n-1)*} \ldots \mathbf{Q}^{(1)*}\, \mathbf{b}$$

by the following algorithm:

for $i := 1$ **to** n **do** $y_i := b_i;$

for $k := 1$ **to** n **do**

  for $i := 1$ **to** n **do**

    $y_i := y_i + m_{ik}\, m_{kk}^*\, y_k;$

Equation (7) expresses $\mathbf{A}^*$ as a product of elementary *column* matrices, but as is shown in Backhouse and Carré [3], $\mathbf{A}^*$ can also be obtained as a product of elementary *row* matrices. It is also demonstrated there that for triangular matrices it is particularly easy to obtain a PFS: specifically in the sequel we will use the fact that for a strictly upper triangular matrix $\mathbf{U}$,

$$\mathbf{U}^* = \mathbf{U}_1^* \mathbf{U}_2^* \ldots \mathbf{U}_{n-1}^* \tag{10}$$

where

$$\mathbf{U}_i = \mathbf{e}_{oi}\, \mathbf{u}_{io}, \qquad (i = 1, 2, \ldots, n). \tag{11}$$

Gauss elimination

To obtain the EFS of $\mathbf{A}$ we proceed as in the Jordan method, except that we replace the matrices $\mathbf{Q}^{(k)}$ and $\mathbf{R}^{(k)}$ of (2) by

$$\mathbf{Q}^{(k)} = \begin{bmatrix} \Phi & \Phi & \Phi \\ \Phi & \mathbf{A}_{22}^{(k-1)} & \Phi \\ \Phi & \mathbf{A}_{32}^{(k-1)} & \Phi \end{bmatrix}, \quad \mathbf{R}^{(k)} = \begin{bmatrix} \mathbf{A}_{11}^{(k-1)} & \mathbf{A}_{12}^{(k-1)} & \mathbf{A}_{13}^{(k-1)} \\ \mathbf{A}_{21}^{(k-1)} & \Phi & \mathbf{A}_{23}^{(k-1)} \\ \mathbf{A}_{31}^{(k-1)} & \Phi & \mathbf{A}_{33}^{(k-1)} \end{bmatrix}. \tag{12}$$

Since (4) still holds, (6) holds also. Furthermore, it is easily proved by induction on k that in each matrix $\mathbf{A}^{(k)}$, all entries on and below the principal diagonal, in the first k columns, are null. Thus $\mathbf{A}^{(n)}$ is strictly upper triangular, and if we denote this matrix by $\mathbf{U}$, Eq. (6) gives

$$\mathbf{A}^* = \mathbf{U}^* \mathbf{Q}^{(n)*} \mathbf{Q}^{(n-1)*} \ldots \mathbf{Q}^{(1)*},$$

or, in terms of the elementary row matrices $\mathbf{U}_i$ of (11),

$$\mathbf{A}^* = \mathbf{U}_1^* \mathbf{U}_2^* \ldots \mathbf{U}_{n-1}^* \mathbf{Q}^{(n)*} \mathbf{Q}^{(n-1)*} \ldots \mathbf{Q}^{(1)*} \tag{13}$$

which is called the *elimination form of the star* (EFS) of $\mathbf{A}$.

Construction of the matrix $\mathbf{M}^{(n)}$, by repeated use of (8), again conveniently gives the factors of the EFS, this time in the form shown in Figure 2. It is easily verified that with $\mathbf{Q}^{(k)}$ and $\mathbf{R}^{(k)}$ defined by (12),

$$\mathbf{M}^{(k)} = \begin{bmatrix} \mathbf{M}_{11}^{(k-1)} & \mathbf{M}_{12}^{(k-1)} & \mathbf{M}_{13}^{(k-1)} \\ \mathbf{M}_{21}^{(k-1)} & \mathbf{M}_{22}^{(k-1)} & \mathbf{M}_{22}^{(k-1)*} \mathbf{M}_{23}^{(k-1)} \\ \mathbf{M}_{31}^{(k-1)} & \mathbf{M}_{32}^{(k-1)} & \mathbf{M}_{33}^{(k-1)} + \mathbf{M}_{32}^{(k-1)} \mathbf{M}_{22}^{(k-1)*} \mathbf{M}_{23}^{(k-1)} \end{bmatrix}. \tag{14}$$

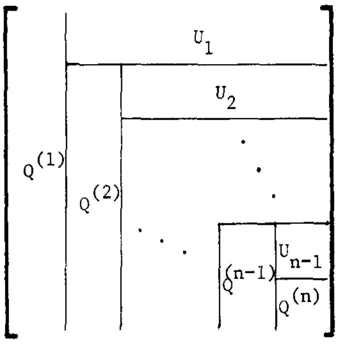

FIGURE 2 Elimination Form of the Star

Thus $\mathbf{M}^{(n)}$ can be computed by the following algorithm:

for $k := 1$ **to** n **do**

  for $i := k$ **to** n **do**

    for $j := k+1$ **to** n **do**

      $m_{ij} := m_{ij} + m_{ik}\, m_{kk}^{*}\, m_{kj}$;

When the EFS has been obtained in this way, the least solution $\mathbf{y} = \mathbf{A}^{*}\mathbf{b}$ of a system $\mathbf{y} = \mathbf{A}\mathbf{y} + \mathbf{b}$ can be obtained using

$$\mathbf{y} = (\mathbf{U}_1^{*}\,(\mathbf{U}_2^{*} \ldots (\mathbf{U}_{n-1}^{*}(\mathbf{Q}^{(n)*}(\mathbf{Q}^{(n-1)*} \ldots (\mathbf{Q}^{(1)*}\,\mathbf{b}) \ldots))) \ldots))$$

The algorithm is as follows:

for $i := 1$ **to** n **do** $y_i := \mathbf{b}_i$;

forward-substitution:

  for $k := 1$ **to** n **do**

    for $i := 1$ **to** n **do**

      $y_i := y_i + m_{ik}\, m_{kk}^{*}\, y_k$;

back-substitution:

for $k:=n-1$ **downto** 1 **do**

for $j:=k+1$ **to** n **do**

$y_k:=y_k+m_{kj}\,y_j;$

4. THE COMPARISON

In practice the matrix **A** is often very large, but sparse. To make elimination methods feasible in such circumstances, it is important to exploit sparsity, by storing and manipulating only non-null matrix elements at each stage of the computation. The effectiveness of this technique depends on the extent to which sparsity is "preserved" in constructing the EFS or PFS, and therefore we are interested in the relative sparsity of these two forms of the star.

In numerical linear algebra, it is well-known that the *elimination form of the inverse* (EFI)—which corresponds to our EFS—has no more non-null entries, and often considerably less, than the *product form of the inverse* (PFI)—which corresponds to our PFS. This was rigorously established by Brayton *et al.*, [8], using an algebraic relationship between the EFI and PFI.

Our purpose in this section is to present a relationship between the EFS and PFS, analogous to that which exists between the EFI and PFI; this algebraic relationship enables us immediately to compare the sparsities of the two forms of the star.

To distinguish between the $\mathbf{M}^{(k)}$-matrices produced in the Gauss and Gauss-Jordan methods, we shall henceforth denote these by $\mathbf{M}^{G(k)}$ and $\mathbf{M}^{J(k)}$ respectively; similar notations will be used for the $\mathbf{Q}^{(k)}$ and $\mathbf{R}^{(k)}$ matrices.

Figure 3 summarises the argument which follows. Essentially, we regard $\mathbf{M}^{G(k)}$ as composed of two matrices **L** (a lower triangular matrix) and **U** (a strictly upper triangular matrix). The matrix **U** is further decomposed into those rows from row k to row n (inclusive) and the remainder. Now, we argue that the corresponding shaded portions of Figures 3(a) and (b) are equal and the unshaded portion of Figure 3(b) is the weak closure of the unshaded portion of Figure 3(a). Consequently, we can conclude that the upper triangle of the PFS is the weak closure of the upper triangle of the EFS.

LEMMA 1 *The first k rows of $\mathbf{M}^{G(k)}$ are identical to the corresponding rows of $\mathbf{M}^{G(n)}$, for $k=1,2,\ldots,n-1$.*

Proof By (14), the first k rows of $\mathbf{M}^{G(k)}$ are identical to the

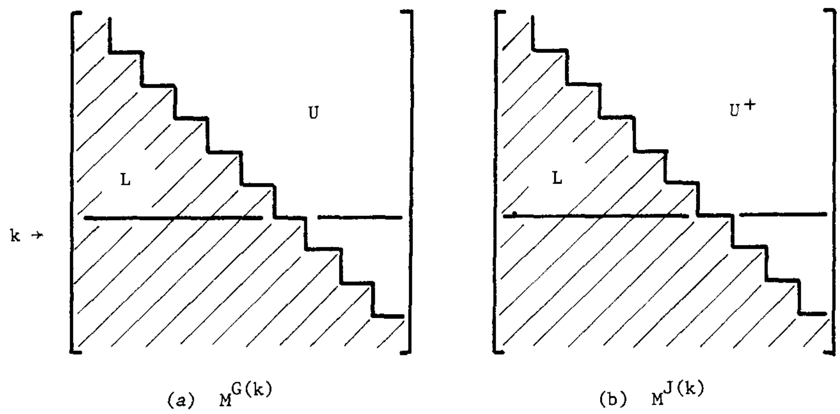

FIGURE 3 Relationship between $\mathbf{M}^{G(k)}$ and $\mathbf{M}^{J(k)}$

corresponding rows of $\mathbf{M}^{G(k+1)}$, for $k=1, 2,\ldots,n-1$, from which the lemma follows.

LEMMA 2 *The last $n-k+1$ rows of $\mathbf{M}^{J(k)}$ are identical to the corresponding rows of $\mathbf{M}^{G(k)}$, for $k=1, 2,\ldots,n$.*

Proof Since $\mathbf{M}^{G(0)}=\mathbf{M}^{J(0)}=\mathbf{A}$, and (by comparison of (2) and (12)) $\mathbf{Q}^{G(1)}=\mathbf{Q}^{J(1)}$ and $\mathbf{R}^{G(1)}=\mathbf{R}^{J(1)}$, the matrices $\mathbf{M}^{G(1)}$ and $\mathbf{M}^{J(1)}$ defined by (8) are identical. Now let us suppose that the lemma holds for $k-1$, where $k>0$. Then in the block forms (9) and (14) of $\mathbf{M}^{J(k)}$ and $\mathbf{M}^{G(k)}$, each block in the second and third row of $\mathbf{M}^{J(k)}$ is identical to the corresponding block of $\mathbf{M}^{G(k)}$, which proves the lemma.

LEMMA 3 *The* kth *row of $\mathbf{M}^{J(k)}$ is identical to the* kth *row of $\mathbf{M}^{G(n)}$, for $k=1,2,\ldots,n$.*

Proof This follows immediately from Lemma 1 and Lemma 2.

Now let us express $\mathbf{M}^{G(n)}$ as the sum of two matrices,

$$\mathbf{M}^{G(n)}=\mathbf{L}+\mathbf{U} \tag{15}$$

where $\mathbf{L}$ is lower triangular and $\mathbf{U}$ is strictly upper triangular. We shall use a notation similar to (11) to describe the elementary matrices associated with the rows of $\mathbf{L}$:

$$\mathbf{L}_i=\mathbf{e}_{oi}\,\mathbf{l}_{io}, \quad (i=1,2,\ldots,n). \tag{16}$$

LEMMA 4 *The matrix $\tilde{\mathbf{M}}^{J(k)}$, obtained from $\mathbf{M}^{J(k)}$ by nullifying its last*

$n-k$ rows, satisfies the equation

$$\tilde{\mathbf{M}}^{J(k)} = \left(\sum_{i=1}^{k} \mathbf{L}_i\right) + \left(\sum_{i=1}^{k} \mathbf{U}_i\right)\left(\prod_{i=1}^{k} \mathbf{U}_k^*\right), \quad \text{for } k=1,2,\ldots,n. \tag{17}$$

Proof From Lemma 3 it follows that

$$\tilde{\mathbf{M}}^{J(1)} = \mathbf{L}_1 + \mathbf{U}_1.$$

Since $\mathbf{U}_i\mathbf{U}_j = \mathbf{\Phi}$ if $i \geqq j$, it follows that

$$\tilde{\mathbf{M}}^{J(1)} = \mathbf{L}_1 + \mathbf{U}_1(\mathbf{E} + \mathbf{U}_1)$$

and since $\mathbf{U}_i^* = \mathbf{E} + \mathbf{U}_i$, for $i = 1, 2, \ldots, n$,

$$\tilde{\mathbf{M}}^{J(1)} = \mathbf{L}_1 + \mathbf{U}_1\mathbf{U}_1^*.$$

Consequently Eq. (17) holds for $k=1$. Now let us suppose that (17) holds for $k-1$, where $k-1 \geqq 0$. From (9).

$$\mathbf{M}^{J(k)} = \begin{bmatrix} \mathbf{M}_{11}^{J(k-1)} & \mathbf{M}_{12}^{J(k-1)} & \mathbf{M}_{13}^{J(k-1)} \\ \mathbf{M}_{21}^{J(k-1)} & \mathbf{M}_{22}^{J(k-1)} & \mathbf{M}_{23}^{J(k-1)} \\ \mathbf{M}_{31}^{J(k-1)} & \mathbf{M}_{32}^{J(k-1)} & \mathbf{M}_{33}^{J(k-1)} \end{bmatrix} \begin{bmatrix} \mathbf{E} & \mathbf{\Phi} & \mathbf{\Phi} \\ \mathbf{\Phi} & \mathbf{E} & \mathbf{M}_{22}^{J(k-1)*}\mathbf{M}_{23}^{J(k-1)} \\ \mathbf{\Phi} & \mathbf{\Phi} & \mathbf{E} \end{bmatrix}.$$

By Lemma 3, the second block matrix here is $(\mathbf{E} + \mathbf{U}_k) = \mathbf{U}_k^*$, and therefore

$$\mathbf{M}^{J(k)} = \mathbf{M}^{J(k-1)}\mathbf{U}_k^*.$$

It follows that the first $k-1$ rows of $\tilde{\mathbf{M}}^{J(k)}$ are given by

$$\mathbf{e}_{io}\tilde{\mathbf{M}}^{J(k)} = \mathbf{e}_{io}\tilde{\mathbf{M}}^{J(k-1)}\mathbf{U}_k^*, \quad \text{for } 1 \leqq i < k,$$

and by Lemma 3 again, the *kth* row of $\tilde{\mathbf{M}}^{J(k)}$ can be written as

$$\mathbf{e}_{ko}\tilde{\mathbf{M}}^{J(k)} = \mathbf{e}_{ko}(\mathbf{L}_k + \mathbf{U}_k).$$

Combining these two equations,

$$\tilde{\mathbf{M}}^{J(k)} = \tilde{\mathbf{M}}^{J(k-1)}\mathbf{U}_k^* + \mathbf{L}_k + \mathbf{U}_k,$$

and substituting from (17) we obtain

$$\tilde{\mathbf{M}}^{J(k)}=\left[\left(\sum_{i=1}^{k-1}\mathbf{L}_i\right)+\left(\sum_{i=1}^{k-1}\mathbf{U}_i\right)\left(\prod_{i=1}^{k-1}\mathbf{U}_i^*\right)\right]\mathbf{U}_k^*+\mathbf{L}_k+\mathbf{U}_k$$

Now if $i<k$, $\mathbf{L}_i\mathbf{U}_k=\mathbf{\Phi}$ and therefore $\mathbf{L}_i\mathbf{U}_k^*=\mathbf{L}_i$, so that

$$\tilde{\mathbf{M}}^{J(k)}=\left(\sum_{i=1}^{k-1}\mathbf{L}_i\right)+\left(\sum_{i=1}^{k-1}\mathbf{U}_i\right)\left(\prod_{i=1}^{k}\mathbf{U}_i^*\right)+\mathbf{L}_k+\mathbf{U}_k$$

$$=\left(\sum_{i=1}^{k}\mathbf{L}_i\right)+\left(\sum_{i=1}^{k-1}\mathbf{U}_i\right)\left(\prod_{i=1}^{k}\mathbf{U}_i^*\right)+\mathbf{U}_k.$$

In the same way, if $i\leqq k$ then $\mathbf{U}_k\mathbf{U}_i=\mathbf{\Phi}$ and therefore $\mathbf{U}_k\mathbf{U}_i^*=\mathbf{U}_k$, so that

$$\tilde{\mathbf{M}}^{J(k)}=\left(\sum_{i=1}^{k}\mathbf{L}_i\right)+\left(\sum_{i=1}^{k}\mathbf{U}_i\right)\left(\prod_{i=1}^{k}\mathbf{U}_i^*\right),$$

as required.

THEOREM $\mathbf{M}^{J(n)}=\mathbf{L}+\mathbf{U}^+$.

Proof Since $\mathbf{M}^{J(n)}=\tilde{\mathbf{M}}^{J(n)}$, it follows from Lemma 4 that

$$\mathbf{M}^{J(n)}=\left(\sum_{i=1}^{n}\mathbf{L}_i\right)+\left(\sum_{i=1}^{n}\mathbf{U}_i\right)\left(\prod_{i=1}^{n}\mathbf{U}_i^*\right)$$

$$=\mathbf{L}+\mathbf{U}\left(\prod_{i=1}^{n}\mathbf{U}_i^*\right).$$

It follows from (10) that

$$\mathbf{M}^{J(n)}=\mathbf{L}+\mathbf{U}\mathbf{U}^*$$

$$=\mathbf{L}+\mathbf{U}^+,$$

as required.

Thus, the elements which lie on and below the diagonal of the PFS are identical to the corresponding elements of the EFS; whereas the elements which lie above the diagonal of the PFS are the elements of $\mathbf{U}^+$ (and $\mathbf{U}^*$), where $\mathbf{U}$ is the upper triangular matrix obtained at the end of the forward course of Gaussian elimination.

Since $\mathbf{U} \leqq \mathbf{U}^+$, we have

$$\mathbf{M}^{G(n)} \leqq \mathbf{M}^{J(n)},$$

hence the EFS has at least as many null elements as the PFS.

5. DISCUSSION

Brayton *et al.*'s [8] comparison of Gaussian and Gauss-Jordan elimination is probably well-known to numerical analysts, so what contribution is made by our comparison? Well, firstly, it clarifies their comparison. Gaussian elimination is often claimed to be superior to Gauss-Jordan elimination because "$\mathbf{U}^{-1}$ has more nonzeros than $\mathbf{U}$" (Tewarson, [10] p. 101). But the latter statement is paradoxical. How can $\mathbf{U}^{-1}$ have more nonzeros than $\mathbf{U}$ when $(\mathbf{U}^{-1})^{-1} = \mathbf{U}$? No such paradox occurs when comparing $\mathbf{U}^+$ to $\mathbf{U}$ since $(\mathbf{U}^+)^+ = \mathbf{U}^+$. The statement $\mathbf{U}^+ \geqq \mathbf{U}$ is an unequivocal description of the relative sparsity of $\mathbf{U}^+$ and $\mathbf{U}$.

Secondly, the comparison of $\mathbf{U}^+$ and $\mathbf{U}$ is a very meaningful one. If we interpret $\mathbf{U}$ as a graph having n nodes with an arc connecting node i to node j whenever $u_{ij} \neq \phi$ then $\mathbf{U}^+$ is the graph describing *paths* (i.e. sequences of arcs) through $\mathbf{U}$. That is, there is an arc from i to j in $\mathbf{U}^+$ if and only if there is a path from i to j in $\mathbf{U}$. In many applications the nature of the problem enables one to predict that $\mathbf{U}^+$ will always be full whilst $\mathbf{U}$ may be very sparse. For example, if one were using Gauss-Jordan elimination to find shortest distances between given towns in Britain it is inevitable that $\mathbf{U}^+$ will be full (because one can always drive from any one town to any other).

Finally, our comparison is more general than Brayton *et al.*'s. It is important to observe that our comparison makes no use of properties which are not valid in real arithmetic. By replacing P^* by $(1-P)^{-1}$ everywhere in Sections 3 and 4 the reader will obtain a valid comparison of Gaussian and Gauss-Jordan elimination applied to the computation of $(\mathbf{1}-\mathbf{A})^{-1}$. However, our comparison is equally valid in other algebras in which no inverse operator can be defined. Practical applications such as global data flow analysis have already been referenced in the introduction.

The result presented here should be viewed alongside Carré's comparison [9] of three iterative techniques (Jacobi, Gauss-Seidel and Yen's double-sweep method). The advantage of Gaussian over Gauss-Jordan elimination is very similar to the advantage of the double-sweep method over the Gauss-Seidel method. That is, Gaussian elimination avoids the explicit determination of $\mathbf{U}^*$ by using the rows of $\mathbf{U}$ in *reverse order* in the subsequent solution process rather than in forward order as in Gauss-Jordan elimination. Likewise, the double-sweep method processes the rows of the upper triangle of **A** in *reverse order* rather than forward order as in the Gauss-Seidel method. Both Carré's comparison and our present one attest to the usefulness of regular algebra in describing and comparing such algorithms.

References

[1] I. S. Duff, A survey of sparse matrix research. *Proceedings of the IEEE* **65** (1977), 500–535.
[2] B. A. Carré, An algebra for network routing problems. *J. Inst. Maths. Applics.* **7** (1971), 273–294.
[3] R. C. Backhouse and B. A. Carré, Regular algebra applied to path-finding problems. *J. Inst. Maths. Applics.* **15** (1975), 161–186.
[4] R. W. Floyd, Algorithm 97: shortest path. *Comm. ACM* **5** (1962), 345.
[5] S. Warshall, A theorem on Boolean matrices. *J. ACM* **9** (1962), 11–12.
[6] R. E. Tarjan, A unified approach to path problems. Technical Report STAN-CS-79-729, Computer Science Dept., Stanford University, 1979.
[7] R. E. Tarjan, Iterative algorithms for global flow analysis, in *Algorithms and Complexity: New Directions and Recent Results*, J. F. Traub (Ed.), Academic Press, New York, 1976, pp. 91–102.
[8] R. K. Brayton, F. G. Gustavson and R. A. Willoughby, Some results on sparse matrices, *Math. of Comp.* **24** (1970), 937–954.
[9] B. A. Carré, *Graphs and Networks*, Oxford Applied Mathematics and Computing Science Series, Oxford University Press, Oxford 1979.
[10] R. P. Tewarson, *Sparse Matrices*, Academic Press, New York, 1973.