Combinatorial Verification of the Elementary Divisors of Tensor Products*

Richard A. Brualdi
Department of Mathematics
University of Wisconsin
Madison, Wisconsin 53706

Dedicated with admiration to Helmut Wielandt on his 75th birthday

Submitted by Emreic Deutsch

ABSTRACT

Using combinatorial arguments we determine the determinantal divisors for the tensor product of two matrices, thereby obtaining its elementary divisors and Jordan canonical form.

1. INTRODUCTION

The elementary divisor structure of the tensor product of two square matrices was determined independently and nearly simultaneously in the 1930s by Aitken [1] and Roth [8]. (Aitken's paper [1] contains an added note making reference to Roth's paper [8], while the latter carries a footnote which refers to a forthcoming paper by Aitken.) In their paper [6], Marcus and Robinson give a precise statement of the Aitken-Roth theorem and a complete, self-contained proof of it. The proof of Marcus and Robinson to some extent completes the argument given by Roth. Indeed, a statement made by Roth [8, p. 464] without comment is the difficult case of the proof of Marcus and Robinson, and its verification constitutes the most substantial part of their proof. The proof given by Aitken is also not beyond reproach. His method of proof (used also by Littlewood [5]) is combinatorial in nature and uses his "method of chains," which unfortunately is not correct as given. The invalidity of Aitken's "method of chains" has already been pointed out by Stanley

*Research partially supported by National Science Foundation Grant No. DMS-8320189.
In connection with his work on ordered sets. In [3, 9] a corrected version of Aitken's method of chains as it applies to generic matrices associated with acyclic digraphs is given. In this paper we shall correct Aitken's derivation of the elementary divisors of a tensor product. The proof proceeds by giving a combinatorial derivation of the determinantal divisors, from which the elementary divisors are then obtained in the classical way.

2. TENSOR PRODUCTS

Let $A$ and $B$ be $m \times m$ and $n \times n$ matrices, respectively, over a field $F$ that contains all the eigenvalues of $A$ and $B$. The tensor product of $A$ with $B$ is the $mn \times mn$ matrix $A \otimes B$ which is partitioned into $m^2$ blocks, the $(i, j)$ block being the $n \times n$ matrix $a_{i,j}B$ $(1 \leq i, j \leq m)$. Let

$$e_i(\lambda) = (\lambda - \lambda_i)^{p_i}, \quad i \leq i \leq s,$$

be the elementary divisors of $A$ (more precisely, the elementary divisors of the polynomial matrix $\lambda I - A$), and let

$$f_j(\lambda) = (\lambda - \mu_j)^{q_j}, \quad 1 \leq j \leq t,$$

be the elementary divisors of $B$. For $a \in F$ and $p$ a positive integer, define the $p \times p$ Jordan matrix by

$$J_p(a) = \begin{bmatrix} a & 1 & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & 1 \\ \vdots & \ddots & 0 \\ a & 1 & 0 \end{bmatrix},$$

where $J_p(a) = [a]$. Then there exist invertible matrices $P$ and $Q$ over $F$ such that

$$P^{-1}AP = J_{p_1}(\lambda_1) \oplus \cdots \oplus J_{p_s}(\lambda_s),$$

the Jordan canonical form of $A$, and

$$Q^{-1}BQ = J_{q_1}(\mu_1) \oplus \cdots \oplus J_{q_t}(\mu_t),$$

the Jordan canonical form of $B$. From properties of the tensor product [6] we
obtain

\[(P \otimes Q) \Lambda (A \otimes B)(P \otimes Q) = \bigoplus_{j=1}^{t} \bigoplus_{i=1}^{s} J_{p_i}(\lambda_i) \otimes J_{q_j}(\mu_j).\]

It follows that the Jordan canonical form of \(A \otimes B\) is the direct sum of the Jordan canonical forms of the \(J_{p_i}(\lambda_i) \otimes J_{q_j}(\mu_j)\); equivalently, the elementary divisors of \(A \otimes B\) are the union (with repetition) of the elementary divisors of the \(J_{p_i}(\lambda_i) \otimes J_{q_j}(\mu_j)\). Hence to determine the elementary divisors of \(A \otimes B\), it suffices to determine the elementary divisors of a tensor product of two Jordan matrices

\[J_{p_i}(a) \otimes J_{q_j}(b).\]  

(2.1)

Let \(C\) be an \(n \times n\) matrix over \(F\) whose eigenvalues lie in \(F\), and let \(\lambda I - C\) be the characteristic matrix of \(C\). Then the determinantal divisors \(d_0(\lambda), d_1(\lambda), \ldots, d_n(\lambda)\) are defined by \(d_0(\lambda) = 1, \ d_1(\lambda) = \gcd \text{ of the determinants of the } i \times i \text{ submatrices of } \lambda I - C\) \((1 \leq i \leq n)\). The invariant factors \(h_1(\lambda), \ldots, h_n(\lambda)\) of \(C\) are then determined by

\[h_i(\lambda) = \frac{d_i(\lambda)}{d_{i-1}(\lambda)} \quad (1 \leq i \leq n),\]  

(2.2)

while the elementary divisors of \(C\) are the powers \((\lambda - c)^k\) of linear polynomials which occur in the prime factorizations of \(h_1(\lambda), \ldots, h_n(\lambda)\). Suppose \(C\) is a triangular matrix with constant main diagonal entries equal to \(c\). Then for \(i = 1, \ldots, n\), \(\lambda I - C\) has an \(i \times i\) submatrix with determinant equal to \((\lambda - c)^i\). It follows that \(d_i(\lambda) = (\lambda - c)^{k_i}\) for some nonnegative integer \(k_i\) \((i \geq 1)\), and hence the elementary divisors of \(C\) are the nonconstant invariant factors (2.2) of \(C\). Since the matrix (2.1) is triangular with constant main diagonal entries equal to \(ab\), it will suffice to determine the determinantal divisors of (2.1).

3. DIGRAPHS

Let \(C = [c_{ij}]\) be an \(n \times n\) matrix over \(F\). The digraph \(D = D(C)\) of \(C\) has vertices \(1, 2, \ldots, n\) with an arc \(ij\) from \(i\) to \(j\) whenever \(i \neq j\) and \(c_{ij} \neq 0\). A path \(\gamma\) of length \(t \geq 0\) in \(D\) is a sequence \(i_1, i_2, \ldots, i_{t+1}\) of \(t + 1\) distinct
vertices such that \( i_1 i_2, \ldots, i_i i_2 \) are all arcs. Following Gansner [3], we define a \( k \)-path in \( D \) to be a set \( \gamma \) of vertices which can be partitioned into sets \( W_1, \ldots, W_s \) for some \( s \leq k \) such that the vertices in each \( W_i \) can be ordered to form a path \( \gamma_i \). The size of the \( k \)-path \( \gamma \) is the number \(|\gamma|\) of vertices in \( \gamma \). Note that the set of vertices of a path of length \( t \) is a 1-path of size \( t + 1 \), and we sometimes do not distinguish between a path and its set of vertices. We define \( p_k(D) \) to be the largest size of a \( k \)-path in \( D \) \((k \geq 1)\) and set \( p_0(D) = 0 \).

Let \( D_1 \) and \( D_2 \) be two digraphs with vertex sets \( V_1 = \{1, \ldots, m\} \) and \( V_2 = \{1, \ldots, n\} \), respectively. The Cartesian product [4] \( D_1 \times D_2 \) of \( D_1 \) with \( D_2 \) is the digraph whose vertex set is the Cartesian product \( V_1 \times V_2 \) of the sets \( V_1 \) and \( V_2 \), with an arc from \((i, j)\) to \((k, l)\) if and only if \( i = k \) and there is an arc \( jl \) in \( D_2 \), or \( j = l \) and there is an arc \( ik \) in \( D_1 \). The Cartesian conjunction [4] \( D_1 \wedge D_2 \) of \( D_1 \) with \( D_2 \) also has vertex set \( V_1 \times V_2 \), and there is an arc from \((i, j)\) to \((k, l)\) if and only if there is an arc \( ik \) in \( D_1 \) and an arc \( jl \) in \( D_2 \). The tensor product \( D_1 \otimes D_2 \) of \( D_1 \) with \( D_2 \) has vertex set \( V_1 \times V_2 \) and arc set equal to the union of the arc sets of \( D_1 \times D_2 \) and \( D_1 \wedge D_2 \). Berge [2, p. 314] calls this the normal product. We denote by \( \Gamma_p \) the digraph with vertices 1,2,\( \cdots \), \( p \) and arcs 12,\( \cdots \),\((p-1)p\). Thus \( \Gamma_p \) is a path with \( p \) vertices.

**Lemma 3.1.** Let \( p \) and \( q \) be positive integers. Then for \( k = 1, \ldots, \min\{p, q\} \)

\[
p_k(\Gamma_p \otimes \Gamma_q) = p_k(\Gamma_p \times \Gamma_q) = \sum_{t=1}^{k} [p + q - (2t - 1)].
\]

**Proof.** First note then when \( k = \min\{p, q\} \), the sum in the equation above equals \( pq \). Hence \( p_k(\Gamma_p \otimes \Gamma_q) = p_k(\Gamma_p \times \Gamma_q) = pq \) for \( k \geq \min\{p, q\} \).

In Figure 1 we have drawn \( \Gamma_p \otimes \Gamma_q \). The horizontal and vertical arcs are the arcs of the Cartesian product \( \Gamma_3 \times \Gamma_4 \) while the 45° arcs are those of the Cartesian conjunction \( \Gamma_3 \wedge \Gamma_4 \). In referring to \( \Gamma_p \otimes \Gamma_q \) or any subdigraph, we assume it has been drawn in this way. Since every arc of \( \Gamma_p \otimes \Gamma_q \) either goes to the right, or goes up, or both, it is clear that

\[
p_1(\Gamma_p \otimes \Gamma_q) = p_1(\Gamma_p \times \Gamma_q) = p + q - 1.
\]

We first prove that \( p_k(\Gamma_p \otimes \Gamma_q) = p_k(\Gamma_p \times \Gamma_q) \). Consider \( k \)-paths \( \gamma \) of \( \Gamma_p \otimes \Gamma_q \) having size equal to \( p_k(\Gamma_p \otimes \Gamma_q) \). Such \( \gamma \) can be partitioned into paths \( \gamma_1, \ldots, \gamma_k \) (the number of paths may be smaller than \( k \), but for notational
Fig. 1. The digraph $\Gamma_3 \otimes \Gamma_4$ ($\Gamma_3 \times \Gamma_4$ when the $45^\circ$ arcs are removed; $\Gamma_3 \wedge \Gamma_4$ when the horizontal and vertical arcs are removed).

purposes we assume there are $k$ by allowing some to be empty. We assume $y$ and $y_1, \ldots, y_k$ have been chosen to satisfy in order:

(3.1) the number of $45^\circ$ arcs in the paths $y_1, \ldots, y_k$ is as small as possible.

and, when there is a $45^\circ$ arc,

(3.2) there is a $45^\circ$ arc as far to the right as possible.

If the number of $45^\circ$ arcs is 0, then $\gamma$ is a $k$-path of $\Gamma_p \times \Gamma_q$. Now assume that there is at least one $45^\circ$ arc. Consider the arc $xy$ which is chosen to satisfy (3.2) and which is furthest towards the bottom in the drawing of $\Gamma_p \otimes \Gamma_q$. Let it be an arc of $\gamma_1$. Refer now to Figure 2.

The vertex $z$ must be in $\gamma$, for otherwise we could replace $x, y$ in $\gamma_1$ by $x, z, y$, contradicting the fact that the size of $\gamma$ is $p_k(\Gamma_p)$. Let $z$ be a vertex of $\gamma_2$. First suppose that $z$ is on the right edge of the drawing of $\Gamma_p \otimes \Gamma_q$. Then $z$ is the last vertex of $\gamma_2$. Removing $z$ from $\gamma_2$ and replacing $x, y$ in $\gamma_1$ by $x, z, y$, we contradict (3.1). If $z$ is on the bottom edge of the drawing of $\Gamma_p \otimes \Gamma_q$, then $z$ is the first vertex of $\gamma_2$ and we contradict (3.1) again. Now suppose that $z$ is on neither the bottom edge nor the right edge of $\Gamma_p \otimes \Gamma_q$. Using the above argument, we may also suppose that $z$ is neither the first or
last vertex of $\gamma_2$. Then by choice of $xy$ we have the situation shown in Figure 3. We then replace $x, y$ in $\gamma_1$ with $x, z, y,$ and $u, z, v$ in $\gamma_2$ with $u, v.$ But now $\gamma_2$ contains a $45^\circ$ arc further to the right than $xy,$ contradicting the choice of $xy$ to satisfy (3.2). It follows that there is no $45^\circ$ arc and $\gamma$ is a $k$-path of $\Gamma_p \times \Gamma_q.$ Hence $p_k(\Gamma_p \otimes \Gamma_q) = p_k(\Gamma_p \times \Gamma_q)$.

We now prove by induction on $k$ that $p_k(\Gamma_p \otimes \Gamma_q) = \Sigma_{t=1}^{k} [m + n - (2t - 1)]$ for $k = 1, \ldots, \min \{p, q\}.$ This has already been noted for $k = 1,$ and we now assume $k > 1.$ Let $\gamma$ be a $k$-path of $\Gamma_p \times \Gamma_q$ with size $p_k(\Gamma_p \times \Gamma_q)$ which is partitioned into paths $\gamma_1, \ldots, \gamma_k.$ We first show that $\gamma_1$ can be chosen so that it joins the vertex in the lower left corner to the vertex in the upper right corner. Suppose that $\gamma_1$ joins vertex $x$ to vertex $y,$ and let there be a vertex $u$ below $x$ (see Figure 4). Thus $u$ is a vertex of one of $\gamma_2, \ldots, \gamma_k,$ for otherwise we may begin $\gamma_1$ at $u,$ contradicting the maximum size of $\gamma.$ Let $\gamma_2 = u_1, \ldots, u_t = u, u_{t+1}, \ldots, u_s.$ Then replacing $\gamma_1 = x, \ldots, y$ by $\gamma'_1 = u_1, \ldots, u_t, x, \ldots, y$ and $\gamma_2$ by $\gamma'_2 = u_{t+1}, \ldots, u_s,$ we obtain another partition of $\gamma$ into $k$ paths with $\gamma'_1$ having longer length than $\gamma_1.$ A similar argument applies when there is a vertex to the left of $x,$ of a vertex to the right or above
A finite number of applications of this argument leads to a partitioning of \( y \) into \( k \) paths in which one of the paths joins the vertex \((1, 1)\) in the lower left corner to the vertex \((p, q)\) in the upper right corner. We may take this path to be \( \gamma_1 \).

Since the paths \( \gamma_2, \ldots, \gamma_k \) share no vertex with \( \gamma_1 \), each lies entirely above \( \gamma_1 \) or entirely below \( \gamma_1 \) (see Figure 5). Hence the paths \( \gamma_2, \ldots, \gamma_k \) correspond, respectively, to pairwise vertex disjoint paths \( \gamma_2', \ldots, \gamma_k' \) of \( \Gamma_{p-1} \times \Gamma_{q-1} \) of the same length (see Figure 5). Hence \( \gamma' \), which is the union of the vertices in \( \gamma_2', \ldots, \gamma_k' \), is a \((k - 1)\)-path of \( \Gamma_{p-1} \times \Gamma_{q-1} \). Thus by the inductive hypothesis

\[
|\gamma'| < p_{k-1}(\Gamma_{p-1} \times \Gamma_{q-1}) \leq \sum_{t=1}^{k-1} \left( (p - 1) + (q - 1) - (2t - 1) \right)
= \sum_{t=2}^{k} \left[ p + q - (2t - 1) \right].
\]
The canonical $k$-path of $\Gamma_p \times \Gamma_q$ for $1 \leq k \leq \min \{p, q\}$ now shows that

$$p_k(\Gamma_p \times \Gamma_q) = \sum_{t=1}^{k} [(p + q) - (2t - 1)].$$

The canonical $k$-path of $\Gamma_p \times \Gamma_q$ defined by Figure 6 now shows that

$$p_k(\Gamma_p \times \Gamma_q) = \sum_{t=1}^{k} [(p + q) - (2t - 1)].$$

Again let $D_1$ and $D_2$ be two digraphs with vertex sets $V_1 = \{1, \ldots, m\}$ and $V_2 = \{1, \ldots, n\}$, respectively. Let $\overline{K}_m$ and $\overline{K}_n$ be respectively the digraphs with vertex sets $V_1$ and $V_2$ and no arcs. Then the Cartesian product $D_1 \times \overline{K}_n$ is obtained from $D_1 \times D_2$ by removing the vertical arcs, while $\overline{K}_m \times D_2$ is obtained from $D_1 \times D_2$ by removing the horizontal arcs (thus $D_1 \times D_2$ is the "arc disjoint union" of $D_1 \times \overline{K}_n$ and $\overline{K}_m \times D_2$). We define the left tensor product $D_1 \otimes_1 D_2$ to be the digraph with vertex set $V_1 \times V_2$ and
with arc set equal to the union of the arc sets of $D_1 \land D_2$ and $D_1 \times \bar{K}_n$. The right tensor product $D_1 \otimes D_2$ is defined in a similar way using $\bar{K}_m \times D_2$ in place of $D_1 \times \bar{K}_n$.

**Lemma 3.2.** Let $p$ and $q$ be positive integers. Then

$$p_k(I; p \otimes \Gamma_q) = p_k(\Gamma_p \times \bar{K}_q) = kp \quad \text{for } k = 1, \ldots, q$$

and

$$p_k(I; p \otimes \Gamma_q) = p_k(\bar{K}_p \times \Gamma_q) = kq \quad \text{for } k = 1, \ldots, p.$$  

**Proof.** Since $\Gamma_p \times K_q$ consists of $q$ "disjoint copies" of $\Gamma_p$, clearly $p_k(\Gamma_p \times \bar{K}_q) = kp$ for $k = 1, \ldots, q$. The proof that $p_k(I; p \otimes \Gamma_q) = \bar{p}_k(I; p \otimes \Gamma_q)$ is along the lines of the proof of $p_k(I; \Gamma_p \otimes \Gamma_q) = p_k(I; \bar{K}_p \times \Gamma_q)$ in Lemma 3.1. Indeed, choose a $k$-path $\gamma$ of $\Gamma_p \otimes \Gamma_q$ of size $p_k(\Gamma_p \otimes \Gamma_q)$ and a partition of $\gamma$ into paths $\gamma_1, \ldots, \gamma_k$ satisfying (3.1) and (3.2). Let $xy$ be an arc satisfying (3.2) which is furthest towards the bottom in the drawing of $\Gamma_p \otimes \Gamma_q$, and suppose $xy$ is an arc of $\gamma_1$. Refer to Figure 7. If the vertex $z$ were not a vertex in $\gamma$, then we could replace $x, y, u$ in $\gamma_1$ with $x, z, u$, contradicting (3.2) [or if $y$ is the terminal vertex of $\gamma_1$, replace $x, y$ in $\gamma_1$ with $x, z$, contradicting (3.1)]. Thus $z$ is in $\gamma$. Then by the choice of $xy$ and the fact that there are no vertical arcs, $z$ is the initial vertex of one of $\gamma_2, \ldots, \gamma_k$, say $\gamma_2$, and all the arcs of $\gamma_2$ are horizontal. We may then replace $\gamma_2$ by the portion of $\gamma_1$ beginning with $y, u, \ldots$ and replace $y, u, \ldots$ of $\gamma_1$ by $\gamma_2$. This gives a new decomposition of $\gamma$ into $k$ paths with one fewer $45^\circ$ arc, contradicting (3.1). It follows that there can be no $45^\circ$ arc and $\gamma$ is a $k$-path of $\Gamma_p \times \bar{K}_q$. Hence

---

**Fig. 7.**
\[ p_k(\Gamma_p \otimes \Gamma_q) = p_k(\Gamma_p \times \overline{\Gamma_q}). \] Since \( \Gamma_p \otimes \Gamma_q \) is isomorphic to \( \Gamma_q \otimes \Gamma_p \) and \( \overline{\Gamma_p} \times \Gamma_q \) is isomorphic to \( \Gamma_q \times \overline{\Gamma_p} \), the lemma now follows.

Finally we prove the following.

**Lemma 3.3.** Let \( p \) and \( q \) be positive integers. Let \( s = \min\{p, q\} \) and \( t = |p - q| + 1 \). Then

\[
p_1(\Gamma_p \wedge \Gamma_q) = ks \quad \text{for} \quad k = 1, \ldots, t,
\]

\[
p_{i+k}(\Gamma_p \wedge \Gamma_q) = ts + 2 \sum_{i=1}^{\lfloor k/2 \rfloor} (s - i) + \left( k - 2 \left\lfloor \frac{k}{2} \right\rfloor \right) \left( s - \left\lfloor \frac{k}{2} \right\rfloor - 1 \right)
\]

for \( k = 1, \ldots, 2(s - 1) \).

**Proof.** The digraph \( \Gamma_p \wedge \Gamma_q \) contains only the 45° arcs and is a disjoint union of paths of various lengths (see Figure 8). The maximum size of a path of \( \Gamma_p \wedge \Gamma_q \) is \( s = \min\{p, q\} \), and there are \( t = |p - q| + 1 \) such paths. The removal of the edges of these \( t \) paths of size \( s \) leaves \( 2(s - 1) \) paths of sizes \( s - 1, s - 1, s - 2, s - 2, \ldots, 1, 1 \). The lemma now follows.

In the next section we shall use Lemmas 3.1 to 3.3 to calculate the elementary divisors of the matrix tensor product (2.1).

### 4. Elementary Divisors of Tensor Products

We now compute the elementary divisors of the tensor product of two matrices. As noted in Section 2, it suffices to compute the elementary divisors...
of a tensor product $C = J_\alpha(a) \otimes J_\beta(b)$ of two Jordan matrices, and since $C$ is triangular with constant main diagonal, the elementary divisors of $C$ are its nonconstant invariant factors, and these are given by quotients of consecutive determinantal divisors. It is necessary to distinguish three cases according to whether both, one, or neither of $a$ and $b$ equals $0$. But first we prove the following two lemmas.

Let $T$ be an $n \times n$ matrix. Let $1 \leq r \leq n$, and let $1 \leq i_1 < \cdots < i_r \leq n$. The $r \times r$ submatrix of $T$ formed by rows $i_1, \cdots, i_r$ and columns $j_1, \cdots, j_r$ is denoted by $T[i_1, \cdots, i_r | j_1, \cdots, j_r]$. The $(n-r) \times (n-r)$ submatrix formed by the rows complementary to those of $i_1, \cdots, i_r$ and the columns complementary to those of $j_1, \cdots, j_r$ is denoted by $T(i_1, \cdots, i_r | j_1, \cdots, j_r)$ and is called the complementary submatrix of $T[i_1, \cdots, i_r | j_1, \cdots, j_r]$. The term rank $p(T)$ of $T$ is the maximum number of nonzero entries of $T$ having the property that no two come from the same row or column.

**Lemma 4.1.** Let $T$ be an $n \times n$ triangular matrix with 0's on and below its main diagonal. Let $X = T[i_1, \cdots, i_r | j_1, \cdots, j_r]$ be an $r \times r$ submatrix of $T$ with $p(X) = r$. Then the complementary submatrix $Y = T[k_1, \cdots, k_{n-r} | l_1, \cdots, l_{n-r}]$ of $X$ has only 0's on (and below) its main diagonal.

**Proof.** Let $X = [x_{pq}]$ ($p, q = 1, \cdots, r$). Suppose that for some $s$ with $1 \leq s \leq r$, we have $j_s < i_s$. Then it follows from the definition of $T$ that $x_{pq} = 0$ for $p = j_s, \cdots, r$ and $q = 1, \cdots, s$. Hence $X$ has a $(r-s+1) \times s$ zero submatrix with $(r-s+1)+s = r+1$ and $X$ cannot satisfy $p(X) = r$. Thus $j_s > i_s$ for each $s = 1, \cdots, r$. But now it follows that $k_i > l_i$ for $i = 1, \cdots, n-r$ and $Y$ has only 0's on and below its main diagonal.

**Lemma 4.2.** Let $T$, $X$, and $Y$ be as in Lemma 4.1. Let $P$ be a set of $r$ nonzero entries of $X$, no two from the same row or column. Then the arcs of the digraph $D(T)$ corresponding to $P$ can be partitioned into at most $n-r$ paths joining vertices in $\{l_1, \cdots, l_{n-r}\}$ to vertices in $\{k_1, \cdots, k_{n-r}\}$, and thus the set of vertices on these paths is an $(n-r)$-path of $D(T)$.

**Proof.** By Lemma 4.1, $Y$ contains only 0's on its main diagonal. Let $T'$ and $Y'$ be the matrices obtained from $T$ and $Y$, respectively, by replacing the main diagonal entries of $Y$ with 1's. Let $P'$ be the union of $P$ and the main diagonal entries of $Y'$. Then $P'$ corresponds to a set of $n$ arcs of the digraph $D(T')$, one going into each vertex and simultaneously one leaving each vertex. Thus these $n$ arcs can be partitioned into cycles $\mu_1, \cdots, \mu_t$ for some $t > 1$. Since $T$ has 0's on and below its main diagonal, $D(T)$ is acyclic (that is,
\(D(T)\) has no cycles of any length \(1, 2, \cdots, n\). Hence removing from \(\mu_1, \cdots, \mu_n\) the arcs corresponding to the main diagonal entries of \(Y\), we are left with paths. Since the removal of each arc in turn increases the number of paths by at most 1, the total number of paths is at most \(n - r\), and these paths join a terminal vertex of an arc corresponding to a main diagonal entry of \(Y\) to an initial vertex of an arc corresponding to a main diagonal entry of \(Y\). The lemma now follows.

The following three lemmas suffice to determine the elementary divisors of a tensor product.

**Lemma 4.3.** If \(a\) and \(b\) are both different from zero, then the elementary divisors of \(C = J_p(a) \otimes J_q(b)\) are

\[
(\lambda - ab)^{p + q - (2k - 1)} \quad \text{for} \quad k = 1, 2, \cdots, \min\{p, q\}.
\]

**Proof.** For \(i = 0, 1, \cdots, pq\), let \(d_i(\lambda)\) be the \(i\)th determinantal divisor of \(\lambda I - C\). Then \(d_{pq} = \det(\lambda I - C) = (\lambda - ab)^{pq}\). The matrix \(C\) is the tensor product of a \(p \times p\) and a \(q \times q\) matrix, and we label its rows and columns using \((i, j)\) where \(1 \leq i \leq p\) and \(1 \leq j \leq q\). Since \(a \neq 0\) and \(b \neq 0\), the digraph \(D(C)\) is \(\Gamma_p \otimes \Gamma_q\) (see Figure 9).

Let \(1 \leq k \leq \min\{p, q\}\). As pointed out in Section 2, \(d_{pq-k}(\lambda)\) is a power of \(\lambda - ab\). Let \(M\) be a \((pq - k) \times (pq - k)\) matrix obtained from \(\lambda I - C\) by deleting \(k\) rows and \(k\) columns. Thus \(\det M\) is a polynomial in \(\lambda - ab\). Since \(\lambda I - C\) has 0's below its main diagonal and \(\lambda - ab\) on its main diagonal, it follows from Lemmas 4.1 and 4.2 that there is a term of the form \(a(\lambda - ab)^{pq-s}\), \(a\) a nonzero scalar, in the determinant expansion of \(M\) only if the digraph \(D(C)\) has a \(k\)-path of size \(s\). Hence the lowest degree of a nonzero term in the determinant expansion of \(M\) is at least \(pq - p_k(D(C))\). Since \(D(C) = \Gamma_p \otimes \Gamma_q\), it follows from Lemma 3.1 that the degree of \(d_{pq-k}(\lambda)\) is at least

\[
pq - p_k(D(C)) = pq - \sum_{t=1}^{k} [p + q - (2t - 1)] \quad \text{(4.1)}
\]

for \(k = 1, \cdots, \min\{p, q\}\).

Consider now the *canonical \(k\)-path* \(\gamma\) of \(\Gamma_p \otimes \Gamma_q\) of size \(p_k(\Gamma_p \otimes \Gamma_q)\) as defined in the proof of Lemma 3.1. Then \(\gamma\) can be partitioned into paths \(\gamma_1, \cdots, \gamma_k\) where \(\gamma_i\) joins \((1, i)\) to \((p - i + 1, q)\) for \(i = 1, \cdots, k\). Let \(M_k\) be the submatrix of \(C\) obtained by deleting the columns labeled \((1, 1), \cdots, (1, k)\) and
The matrix $C = J_3(a) \otimes J_3(b)$ with $c = ab \neq 0$ and its digraph $D(C)$.
the rows labeled \((p - k + 1, q), \ldots, (p, q)\). The path \(y_i\) has length \(p + q - (2i - 1)\) and corresponds in \(M_k\) to a permutation cycle of cardinality \(p + q - 2i(1 + \text{the number of arcs of } y_i)\). Since \(y_i\) has \(p - i\) horizontal arcs (of "value" \(b\)) and \(q - i\) vertical arcs (of "value" \(a\)), there is a nonzero term

in the determinant expansion of \(M_k\) equal to

\[
(1)^{\tau}a^{(a-1)+\cdots+(a-k)b^{(b-1)+\cdots+(b-k)}}(\lambda ab)^{pq - p_k(D(C))},
\]

where \(\tau = \sum_{i=1}^{k}(p + q - 2i - 1)\). We now show that each nonzero term in the determinant expansion of \(M_k\) with degree \(pq - p_k(D(C))\) has this value. But it follows from Lemma 4.2 that such a term corresponds to a \(k\)-path \(\gamma\) of size \(p_k(D(C))\) which can be partitioned into paths joining the vertices in \(\{(1,1), (1,2), \ldots, (1,k)\}\) to the vertices in \(\{(p, q), (p - 1, q), \ldots, (p - k + 1, q)\}\). Now a path from \((1, i)\) to \((p - i + 1, q)\) has length at most \(p + q - (2i - 1)\) [and so size \(p + q - (2i - 1)\)], with equality if and only if each arc of the path is either horizontal or vertical. Since the paths into which \(y'\) is partitioned cannot intersect, it follows that \(\gamma'\) is partitioned into paths \(\gamma'_1, \ldots, \gamma'_i\) where for each \(i\), \(\gamma'_i\) joins \((1, i)\) to \((p - i + 1, q)\), \(\gamma'_i\) has only horizontal and vertical arcs, and \(\gamma'_i\) has length \(p + q - (2i - 1)\). Hence each nonzero term in the determinant expansion of \(M_k\) with degree \(pq - p_k(D(C))\) has value (4.2), and it follows that

\[
\det M_k = a(\lambda - a)^{pq - p_k(D(C))} + \text{(higher degree terms)}
\]

for some \(\alpha \neq 0\). Hence the degree of \(d_{pq-k}(\lambda)\) is exactly (4.1), and

\[
d_{pq-k}(\lambda) = (\lambda - ab)^{pq - p_k(D(C))} \quad \text{for } k = 1, \ldots, \min\{p, q\}.
\]

Hence for \(k = 1, \ldots, \min\{p, q\}\),

\[
\lambda^a - (p \text{ times}).
\]

**Lemma 4.4.** If \(a \neq 0\) and \(b = 0\), then the elementary divisors of \(C = I_p(a) \otimes I_q(b)\) are
Fig. 10. The matrix $C = I_3(a) \otimes I_3(b)$, where $a \neq 0$ and $b = 0$, and its digraph $D(C)$. 
If $a = 0$ and $b \neq 0$, then the elementary divisors of $C = I_p(a) \otimes I_q(b)$ are

$$\lambda^p \quad (q \text{ times}).$$

**Proof.** Since $I_p(a) \otimes I_q(b)$ is similar (by means of a permutation matrix) to $I_q(b) \otimes I_p(a)$, we need only prove the first statement of the lemma. Thus suppose $C = I_p(a) \otimes I_q(b)$ where $a \neq 0$ and $b = 0$. The digraph $D(C)$ is then the right-tensor product $\Gamma_p \otimes \Gamma_q$ (see Figure 10). By Lemma 3.2 $p_k(D(C)) = kq$ for $k = 1, \ldots, p$. It follows as in the proof of Lemma 4.3 that the degree of $d_{pq-k}(\lambda)$ is at least $pq - kq$ for $k = 1, \ldots, p$. Let $M_k$ be the submatrix of $C$ obtained by deleting columns $(1,1), \ldots, (k,1)$ and rows $(1,q), \ldots, (k,q)$. There is a unique collection of $k$ vertex disjoint paths which join the vertices in $\{(1,1), \ldots, (k,1)\}$ to the vertices in $\{(1,q), \ldots, (k,q)\}$, and these paths are "vertical" paths of length $q-1$ joining $(1,1)$ to $(1,q), \ldots, (k,1)$ to $(k,q)$. It follows as in the proof of Lemma 4.3 that

$$\det M_k = \alpha \lambda^{pq-kq} + \text{(higher degree terms)}$$

for some $\alpha \neq 0$ and hence that

$$d_{pq-k}(\lambda) = \lambda^{pq-kq} \quad \text{for} \quad k = 1, \ldots, p.$$

Hence for $k = 1, \ldots, p$, $e_k(\lambda) = \lambda^q$ and the lemma holds.

**Lemma 4.5.** If $a = b = 0$, then the elementary divisors of $C = I_p(a) \otimes I_q(b)$ are

$$\lambda^k \quad \text{(twice)} \quad \text{for} \quad k = 1, \ldots, \min\{p,q\} - 1$$

and

$$\lambda^{\min\{p,q\}} \quad (|p-q| + 1 \text{ times}).$$

**Proof.** In this case $D(C)$ is the conjunction $\Gamma_p \wedge \Gamma_q$, whose arcs are the $45^\circ$ arcs. The value of $p_k(\Gamma_p \wedge \Gamma_q)$ is given in Lemma 3.3 for $k = 1, \ldots, 2(\min\{p,q\} - 1)$. The rest of the proof follows as in the proofs of Lemma 4.3 and 4.4.

Combining Lemmas 4.3, 4.4, and 4.5 with the observations made in Section 2, we obtain the theorem of Aitken [1] and Roth [8] (cf. Theorem 1 of [7]).

**Theorem 4.6.** Let $A$ and $B$ be $m \times m$ and $n \times n$ matrices respectively over a field $F$ which contains all the eigenvalues of $A$ and $B$. Then the
complete list of elementary divisors of the tensor product $A \otimes B$ is obtained as follows. To each pair consisting of an elementary divisor $(\lambda - a)^p$ of $A$ and an elementary divisor $(\lambda - b)^q$ of $B$, there correspond the following elementary divisors of $A \otimes B$:

1. When $a \neq 0$ and $b \neq 0$,
   \[
   (\lambda - ab)^{p+q-(2k-1)} \quad \text{for} \quad k = 1, 2, \cdots, \min\{p, q\}.
   \]

2. When $a \neq 0$ and $b = 0$,
   \[
   \lambda^q \quad (p \text{ times}).
   \]

3. When $a = 0$ and $b \neq 0$,
   \[
   \lambda^p \quad (q \text{ times}).
   \]

4. When $a = 0$ and $b = 0$,
   \[
   \lambda^k \quad (\text{twice}) \quad \text{for} \quad k = 1, 2, \cdots, \min\{p, q\} - 1,
   \]
   \[
   \lambda^{\min\{p, q\}} \quad (|p - q| + 1 \text{ times}).
   \]

REFERENCES


Received 25 October 1984