Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 15–22

INVARIANT FACTORS OF PRODUCTS OVER ELEMENTARY DIVISOR DOMAINS

CRISTINA CALDEIRA AND JOÃO FILIPE QUEIRÓ

ABSTRACT: For matrices A and B, what can we say about the invariant factors of AB in terms of those of A and B? For matrices over principal ideal domains, the complete answer is known. In the present paper we consider the same problem for matrices over the larger class of elementary divisor domains.

KEYWORDS: Invariant factors. Principal ideal domains. Elementary divisor domains.
AMS SUBJECT CLASSIFICATION (2010): 15B33, 15B36.

1. Introduction

In this paper we are interested in describing the invariant factors of the product of two matrices over the most general class of integral domains for which the question makes sense. The problem has been completely solved for matrices over principal ideal domains (PIDs) and we begin in that setting. There is no loss of generality in restricting our study to square nonsingular matrices [14].

Let R be a PID and A an $n \times n$ nonsingular matrix over R. It is well known that A is equivalent to its *Smith normal form*, that is, there exist U and V unimodular (*i.e.* invertible over R) such that

$$UAV = \begin{bmatrix} a_n & 0 & \cdots & 0 \\ 0 & a_{n-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_1 \end{bmatrix},$$

where $a_n \mid a_{n-1} \mid \cdots \mid a_1$ are the *invariant factors* of A.

Received May 26, 2015.
This work was partially supported by the Centre for Mathematics of the University of Coimbra – UID/MAT/00324/2013, funded by the Portuguese Government through FCT/MEC and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020.

The invariant factors are uniquely determined by A, as follows from the characterization

$$a_{n-k+1} = \frac{d_k(A)}{d_{k-1}(A)}\ ,\ \ k = 1, \ldots, n,$$

where, for each k, $d_k(A)$, the so-called kth *determinantal divisor* of A, is the gcd of all $k \times k$ minors of A, $d_0 \equiv 1$. (This definition can of course be presented also for non-square matrices.) By the Cauchy-Binet theorem for determinants, the d_k are invariant under equivalence. That $d_{k-1}(A)$ divides $d_k(A)$ follows from Laplace's theorem.

The problem we are interested in is the following: What are the possible invariant factors $c_n \mid \cdots \mid c_1$ of a product AB, if A and B are $n \times n$ non-singular matrices over R with invariant factors $a_n \mid \cdots \mid a_1$ and $b_n \mid \cdots \mid b_1$, respectively?

For matrices over a PID, this problem has been solved with a variety of approaches, starting with its p-module version in [10], where p is a prime in R. Indeed, all approaches start by localizing the problem at an arbitrary prime p, working in that context, and then recovering the global solution.

To describe the solution in [10] we need some notation. For each fixed prime $p \in R$, we restrict our attention to matrices over the local ring R_p, that is, we just work with powers of p: $a_i \to p^{\alpha_i}$, $b_i \to p^{\beta_i}$, $c_i \to p^{\gamma_i}$, where $\alpha_1 \geq \cdots \geq \alpha_n, \beta_1 \geq \cdots \geq \beta_n, \gamma_1 \geq \cdots \geq \gamma_n$ are nonnegative integers.

Denote by $IF(\alpha, \beta)$ the set of possible γ in the invariant factor product problem. Introduce the notation $\Lambda_n = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n : \alpha_1 \geq \cdots \geq \alpha_n \geq 0\}$. What was proved in [10] was that $IF(\alpha, \beta) = LR(\alpha, \beta)$, where the latter is the set of $\gamma \in \Lambda_n$ which can be obtained from α and β using the combinatorial Littlewood-Richardson rule (for the description of the rule see *e.g.* [5]). Thus the invariant factor product problem, in its local "primary" version, has a complete and interesting solution, although not a clearly explicit one, via the Littlewood-Richardson rule. In particular, this solution is *not* given as a family of divisibility relations.

For each natural number r between 1 and n, denote by $Q_{r,n}$ the set of strictly increasing sequences with r elements taken from $\{1, 2, \ldots, n\}$. For many years, R. C. Thompson, who was aware of Klein's work, believed there should be a solution to the invariant factor product problem given by a family of divisibility relations of the type

$$c_{k_1} c_{k_2} \cdots c_{k_r} \mid a_{i_1} a_{i_2} \cdots a_{i_r} b_{j_1} b_{j_2} \cdots b_{j_r} \tag{1}$$

where $I = (i_1, \ldots, i_r), J = (j_1, \ldots, j_r), K = (k_1, \ldots, k_r) \in Q_{r,n}$. His main work on the subject, going a long way in that purpose, is the paper [16].

At the end of the 1990s, as a by-product to the solution of another well-known matrix problem – the description of the relations between the eigenvalues of two Hermitian matrices and those of their sum – a complete solution to the invariant factor problem in terms of divisibility relations of the type (1) was found. The reader interested in the details and in the connection between the two problems may consult the excellent survey [6], and also [11], [15]. Using the notation in [6], for $I = (i_1, \ldots, i_r) \in Q_{r,n}$ we define a decreasing r-sequence $\lambda(I)$ by

$$\lambda(I) = (i_r - r, i_{r-1} - (r-1), \ldots, i_2 - 2, i_1 - 1).$$

Then elements $c_n \mid \cdots \mid c_1$ occur as the invariant factors of a product AB where A and B have invariant factors $a_n \mid \cdots \mid a_1$ and $b_n \mid \cdots \mid b_1$, respectively, if and only if $c_1 \cdots c_n = a_1 \cdots a_n b_1 \cdots b_n$ and the relations (1) hold whenever $\lambda(K) \in LR(\lambda(I), \lambda(J))$ for all $r < n$. The proof is dependent on the localization argument mentioned above. The result also means that the valid divisibility relations are exactly those whose indices appear in the inequalities solving the Hermitian sum problem, the so-called Horn inequalities [6].

2. Elementary divisor domains

Invariant factors may be defined for matrices over more general rings. The more natural rings in this context are the *elementary divisor domains* (EDDs) introduced by Kaplansky in [9]. These are precisely the integral domains R where every matrix over R is equivalent to a Smith normal form exactly as above (Kaplansky allows zero divisors). One example of an elementary divisor domain which is not a principal ideal domain is the non Noetherian ring $H(\Omega)$ of all complex functions holomorphic in an open connected set $\Omega \subseteq \mathbb{C}$ [8]. Another example, relevant to Control Theory, can be found in [7]. So EDDs are a strictly larger class of rings than PIDs. Arguments using reduction to the primary case do not work here, as EDDs are not in general unique factorization domains.

As before, the determinantal divisors (and hence also the invariant factors) are invariant under equivalence, so two matrices over an EDD are equivalent if and only if they have the same invariant factors.

Kaplansky makes the interesting observation that for R to be an EDD it is enough to require that 2×2 matrices are equivalent to a diagonal. This allows him to give a characterization of EDDs with a simple algebraic condition: they are the domains where all finitely generated ideals are principal and whenever $\gcd(a,b,c)=1$ there exist p and q such that $\gcd(pa, pb+qc)=1$.

The question naturally arises: what can we say about properties of invariant factors of matrices over EDDs? Of course, results established using only the Smith normal form, without reduction to the primary case, immediately carry over to EDDs. Some examples can be found in [12]. But what about the huge family of divisibility relations, mentioned in the previous section, valid for invariant factors of products of matrices over PIDs? Extending those results to EDDs presents an interesting challenge, necessitating a change in the proofs.

We shall prove in Section 5 that all divisibility relations valid for invariant factors of products of matrices over PIDs (which give the complete answer to the product problem in that setting) remain valid for matrices over an EDD R. Our strategy — inspired by the Hermitian sum spectral problem [6] and the corresponding one for singular values of products [17]— is to establish extremal characterizations (for the divisibility order) for scattered products $a_{i_1}a_{i_2}\cdots a_{i_r}$ of invariant factors. The extremes will be taken over analogues of Schubert varieties of submodules of R^n. We do this in Section 4. For it to work over EDDs, we must restrict ourselves to the class of pure submodules. We dedicate the next section to the properties of these submodules that we shall need.

3. Pure submodules

Let R be an elementary divisor domain.

Definition. *Let M be an R-module and W a submodule of M. We say that W is a pure submodule of M if, for all $a\in R$, we have $W\cap aM = aW$.*

Remarks.
1. For modules over an integral domain the definition of pure submodule is usually presented in another form. For modules over an EDD the two definitions are equivalent [4, 13].
2. Every direct summand of a module M is a pure submodule of M.
3. Over an EDD, if both the module and the submodule are free with finite rank then W is pure in M if and only if W is a direct summand of M.

We begin this section with the generalization of the last remark to submodules, not necessarily free, of R^n.

Denote by $\mathbb{K}$ the quotient field of R. If F is a submodule of R^n then $\mathbb{K}F$ is a subspace of $\mathbb{K}^n$. Write $\operatorname{rank}(F) := \dim_{\mathbb{K}}(\mathbb{K}F)$. (If F is a free submodule of R^n then $\operatorname{rank}(F) = \dim_R(F)$, the usual rank of F).

The intersection of any non-empty family of pure submodules of R^n is a pure submodule of R^n. If F is a submodule of R^n, we denote by $\overline{F}$ the pure closure of F, that is, the intersection of all pure submodules of R^n containing F.

In the next Lemma we collect some straightforward results on these notions. For modules over a PID, results 2 to 5 can be found in [3].

Lemma 3.1. *Let F and G be submodules of R^n. We have*
1. $\mathbb{K}(F \cap G) = (\mathbb{K}F) \cap (\mathbb{K}G)$*;*
2. If $F \subseteq G$, F is pure in R^n and $\operatorname{rank}(F) = \operatorname{rank}(G)$*, then $F = G$;*
3. $\overline{F} = \{v \in R^n \colon \exists a \in R \setminus \{0\} \text{ s.t. } av \in F\}$*;*
4. $\overline{F \cap G} = \overline{F} \cap \overline{G}$*;*
5. $\operatorname{rank}(F) = \operatorname{rank}(\overline{F})$*.*

Next we prove that every pure submodule of R^n is the pure closure of a free submodule with finite basis and we use that result to generalize, for pure submodules of R^n, the basis extension theorem for finite dimension vector spaces.

Theorem 3.2. *Let L be a pure submodule of R^n with* $\operatorname{rank}(L) = r$*. Then there exist $x_1, \ldots, x_r \in L$, linearly independent, such that* $L = \overline{\operatorname{span}_R\{x_1, \ldots, x_r\}}$*.*

Proof. Let $\{v_1, v_2, \ldots, v_r\}$ be a basis of $\mathbb{K}L$. For $j = 1, \ldots, r$, $v_j = \alpha_j x_j$ with $\alpha_j \in \mathbb{K} \setminus \{0\}$ and $x_j \in L$. Clearly $x_1, \ldots, x_r$ are linearly independent. Put $F = \operatorname{span}_R\{x_1, \ldots, x_r\}$. Since L is pure and $F \subseteq L$ we have $\overline{F} \subseteq L$. On the other hand, $\operatorname{rank}(\overline{F}) = \operatorname{rank}(F) = r = \operatorname{rank}(L)$, whence $\overline{F} = L$. ■

Theorem 3.3. *Let $W \subseteq M$ be pure submodules of R^n with* $\operatorname{rank}(W) = r$*,* $\operatorname{rank}(M) = k$ *and* $W = \overline{\operatorname{span}_R\{x_1, \ldots, x_r\}}$*, where $x_1, \ldots, x_r$ are linearly independent. Then there exist $x_{r+1}, \ldots, x_k \in M$ such that $x_1, \ldots, x_k$ are linearly independent and* $M = \overline{\operatorname{span}_R\{x_1, \ldots, x_k\}}$*.*

Proof. Let $y_1,\dots,y_k \in M$ be linearly independent and such that $M = \overline{\operatorname{span}_R\{y_1,\dots,y_k\}}$. Since $x_1,\dots,x_r \in M$ are linearly independent, there exist $b_1,\dots,b_r \in R\setminus\{0\}$, such that $b_1x_1,\dots,b_rx_r$ are linearly independent in $\operatorname{span}_R\{y_1,\dots,y_k\}$. Then there exist unimodular matrices U and V of orders k and r, respectively, and nonzero $a_1 \mid a_2 \mid \cdots \mid a_r$ such that

$$[b_1x_1\cdots b_rx_r] = [y_1\cdots y_k]U \begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_r \\ \hline & 0 & \end{bmatrix} V\,,$$

where we use the notation $[y_1\cdots y_k]$ for the matrix with columns $y_1,\dots,y_k$.

Let $Z = [z_1\cdots z_k] = [y_1\cdots y_k]U$. Since U is unimodular, $z_1,\dots,z_k$ are linearly independent and $\operatorname{span}_R\{z_1,\dots,z_k\} = \operatorname{span}_R\{y_1,\dots,y_k\}$. Therefore $M = \overline{\operatorname{span}_R\{y_1,\dots,y_k\}} = \overline{\operatorname{span}_R\{z_1,\dots,z_k\}}$.

On the other hand, from $[a_1z_1\cdots a_rz_r] = [b_1x_1\cdots b_rx_r]V^{-1}$ we get that $a_jz_j \in W\cap a_jM = a_jW$, for $j=1,\dots,r$. So $z_1,\dots,z_r \in W$ and are linearly independent. Therefore, as W is pure, $\overline{\operatorname{span}_R\{z_1,\dots,z_r\}} \subseteq W$. Equality holds as the two submodules are pure and have the same rank.

We claim that $M = \overline{\operatorname{span}_R\{x_1,\dots,x_r,z_{r+1},\dots,z_k\}}$.

Let $v \in M = \overline{\operatorname{span}_R\{z_1,\dots,z_k\}}$. There exist $a,c_1,\dots,c_k \in R$, with $a\neq 0$, such that $av = \sum_{i=1}^k c_iz_i$. Since $y = \sum_{i=1}^r c_iz_i \in W = \overline{\operatorname{span}_R\{x_1,\dots,x_r\}}$, there exist $b,d_1,\dots,d_r \in R$, with $b\neq 0$, such that $by = \sum_{i=1}^r d_ix_i$ and

$$abv = \sum_{i=1}^r d_ix_i + \sum_{i=r+1}^k bc_iz_i\,.$$

As $ab \in R\setminus\{0\}$, we get that $v \in \overline{\operatorname{span}_R\{x_1,\dots,x_r,z_{r+1},\dots,z_k\}}$. Therefore, $M \subseteq \overline{\operatorname{span}_R\{x_1,\dots,x_r,z_{r+1},\dots,z_k\}}$. The other inclusion follows from the fact that M is pure and contains $x_1,\dots,x_r,z_{r+1},\dots,z_k$.

That $x_1,\dots,x_r,z_{r+1},\dots,z_k$ are linearly independent follows promptly from $W = \overline{\operatorname{span}_R\{z_1,\dots,z_r\}}$, and the fact that both $z_1,\dots,z_k$ and $x_1,\dots,x_r$ are linearly independent. ∎

We now define the analogue of the usual Schubert varieties.

Definition. *Let $P=(P_1,P_2,\dots,P_n)$, with $P_1\subset P_2\subset\cdots\subset P_n$, be a chain of pure submodules of R^n such that* $\operatorname{rank}(P_i)=i$, *$i=1,2,\dots,n$. For*

$I = (i_1, \dots, i_r) \in Q_{r,n}$, we denote by $\Omega_I(P)$ the set of pure submodules L of R^n with rank r and such that, for $j = 1, \dots, r$, $\operatorname{rank}(L \cap P_{i_j}) \geq j$.

Theorem 3.4. *Let $P = (P_1, P_2, \dots, P_n)$ with $P_1 \subset P_2 \subset \cdots \subset P_n$ be a chain of pure submodules of R^n such that $\operatorname{rank}(P_i) = i$, $i = 1, 2, \dots, n$, and let $I = (i_1, \dots, i_r) \in Q_{r,n}$. A submodule L of R^n belongs to $\Omega_I(P)$ if and only if there exist linearly independent $x_1, \dots, x_r \in L$ such that $L = \overline{\operatorname{span}_R\{x_1, \dots, x_r\}}$ and $x_j \in P_{i_j}$ for $j = 1, \dots, r$.*

Proof. Let $L \in \Omega_I(P)$. Since $\operatorname{rank}(L \cap P_{i_1}) \geq 1$ there exists $x_1 \in L \cap P_{i_1}$ with $x_1 \neq 0$. $W = \overline{\operatorname{span}_R\{x_1\}}$ is a pure submodule of R^n and is contained in the pure submodule $L \cap P_{i_2}$. There exist $u_2, \dots, u_t \in L \cap P_{i_2}$ (with $t = \operatorname{rank}(L \cap P_{i_2}) \geq 2$) such that $x_1, u_2, \dots, u_t$ are linearly independent and $L \cap P_{i_2} = \overline{\operatorname{span}_R\{x_1, u_2, \dots, u_t\}}$. Take $x_2 = u_2$. Repeating the process we get that there exist linearly independent $x_1, \dots, x_r \in L$ such that $x_j \in P_{i_j}$ for $j = 1, \dots, r$ and $\overline{\operatorname{span}_R\{x_1, \dots, x_r\}} = L \cap P_{i_r} \subseteq L$. As $\overline{\operatorname{span}_R\{x_1, \dots, x_r\}}$ and L are both pure and have the same rank, we get that $\overline{\operatorname{span}_R\{x_1, \dots, x_r\}} = L$. ■

Theorem 3.5. *Let $r \in \mathbb{N}_0$. The mapping that to each pure submodule L of R^n with rank r assigns $\mathbb{K}L$ is a bijection between the set of pure submodules of R^n with rank r and the set of r-dimensional subspaces of $\mathbb{K}^n$.*

Proof. Given a subspace E of $\mathbb{K}^n$ with dimension r, the set $L = \{x \in R^n \colon \exists\, \alpha \in \mathbb{K} \setminus \{0\} \text{ s.t. } \alpha x \in E\}$ is a pure submodule of R^n with $\mathbb{K}L = E$.

Let now W_1, W_2 be pure submodules of R^n such that $\mathbb{K}W_1 = \mathbb{K}W_2 =: E$. Consider $L = \{x \in R^n \colon \exists \alpha \in \mathbb{K} \setminus \{0\} \text{ s.t. } \alpha x \in E\}$. Clearly $W_1 \subseteq L$. Let $x \in L$. There exist $a, b \in R \setminus \{0\}$ such that $\frac{a}{b}x \in E = \mathbb{K}W_1$. Then $\frac{a}{b}x = \frac{c}{d}y$, with $c, d \in R \setminus \{0\}$ and $y \in W_1$. We get $adx = bcy \in W_1$, whence $x \in \overline{W_1} = W_1$. Therefore $L \subseteq W_1$. Similarly, $L = W_2$. ■

Theorem 3.6. *Let L be a pure submodule of R^n. Then, for every $I \in Q_{r,n}$,*

$$L \in \Omega_I(P) \iff \mathbb{K}L \in \Omega_I(\mathbb{K}P)\,,$$

where $\mathbb{K}P = (\mathbb{K}P_1, \mathbb{K}P_2, \dots, \mathbb{K}P_n)$.

Proof. This follows immediately from the fact that, for a submodule L of R^n, one has $\dim_{\mathbb{K}}(\mathbb{K}L \cap \mathbb{K}P_{i_j}) = \dim_{\mathbb{K}} \mathbb{K}(L \cap P_{i_j}) = \operatorname{rank}(L \cap P_{i_j})\,,\ j = 1, \dots, r\,.$ ■

4. Extremal characterizations

Our main inspiration in this section is [17].

Recall that $d_r(A)$ is the gcd of all $r\times r$ minors of A. If M is an $m\times n$ matrix and ω and η are strictly increasing sequences of elements of $\{1,\dots,m\}$ and $\{1,\dots,n\}$, respectively, $M[\omega|\eta]$ denotes the submatrix of M built with the rows and columns indexed by ω and η, respectively.

Theorem 4.1. *Let $A\in R^{n\times n}$, L a pure submodule of R^n and $x_1,\dots,x_r\in L$ linearly independent such that $L=\overline{\mathrm{span}_R\{x_1,\dots,x_r\}}$. Put $X=[x_1\cdots x_r]$. Then*
1. $d_r(X)\mid d_r(AX)$;
2. If $y_1,\dots,y_r\in L$ are linearly independent such that $L=\overline{\mathrm{span}_R\{y_1,\dots,y_r\}}$, and $Y=[y_1\cdots y_r]$, then

$$\frac{d_r(AX)}{d_r(X)}=\frac{d_r(AY)}{d_r(Y)}\,.$$

Proof. 1.

$$\begin{aligned} d_r(AX) &= \gcd_{\omega\in Q_{r,n}} \det\left(AX[\omega|1,\dots,r]\right)\\ &= \gcd_{\omega\in Q_{r,n}} \sum_{\gamma\in Q_{r,n}} \det(A[\omega|\gamma])\det(X[\gamma|1,\dots,r])\,. \end{aligned}$$

2. Since $x_1,\dots,x_r$ belong to $L=\overline{\mathrm{span}_R\{y_1,\dots,y_r\}}$ and are linearly independent, there exist $a_1,\dots,a_r\in R\setminus\{0\}$ and $S\in R^{r\times r}$ nonsingular such that $X\mathrm{diag}(a_1,\dots,a_r)=YS$. We then have

$$\begin{aligned} \frac{d_r(AX)}{d_r(X)} &= \frac{d_r(AX\mathrm{diag}(a_1,\dots,a_r))}{d_r(X\mathrm{diag}(a_1,\dots,a_r))}=\frac{d_r(AYS)}{d_r(YS)}\\ &= \frac{d_r(AY)\det(S)}{d_r(Y)\det(S)}=\frac{d_r(AY)}{d_r(Y)}\,. \blacksquare \end{aligned}$$

It follows from item 2 in the theorem that $\frac{d_r(AX)}{d_r(X)}$ does not depend on the choice of X, but only on the submodule L. We use this to present the definition of a kind of "Rayleigh functional" for A and L.

Definition. *Let L be a pure submodule of R^n, with* $\operatorname{rank}(L) = r$*. For A $n \times n$, we write*

$$\psi(A|_L) = \frac{d_r(AX)}{d_r(X)}\,,$$

where $x_1, \ldots, x_r \in L$ are linearly independent such that $L = \overline{\operatorname{span}_R\{x_1, \ldots, x_r\}}$ and X is the $n \times r$ matrix $[x_1 \cdots x_r]$.

Let A be an $n \times n$ matrix over R. There exist unimodular U, V such that $UAV = \operatorname{diag}(a_1, \ldots, a_n)$, with $a_n \mid a_{n-1} \mid \cdots \mid a_1$. Denote by $v_1, \ldots, v_n$ the columns of V, which form a basis of R^n. Consider the pure submodules of R^n defined by $V_i = \operatorname{span}_R\{v_1, \ldots, v_i\}$, $i = 1, \ldots, n$, and write $V = (V_1, \ldots, V_n)$.

For $I = (i_1, \ldots, i_r) \in Q_{r,n}$, we have that $\Omega_I(V)$ is nonempty, since $\operatorname{span}_R\{v_{i_1}, v_{i_2}, \ldots, v_{i_r}\} \in \Omega_I(V)$.

Our first extremal characterization is the following.

Theorem 4.2. *For every $I = (i_1, \ldots, i_r) \in Q_{r,n}$ we have*

$$a_{i_1} a_{i_2} \cdots a_{i_r} = \gcd_{L \in \Omega_I(V)} \psi\left(A_{|L}\right)\,.$$

Proof. Let $L \in \Omega_I(V)$, and let $x_1, ..., x_r \in L$ be linearly independent such that $L = \overline{\operatorname{span}_R\{x_1, \ldots, x_r\}}$ and $x_j \in V_{i_j}$ for all j.

Let B be the $n \times r$ matrix such that $X = [x_1 \cdots x_r] = VB$ and $D = \operatorname{diag}(a_1, \ldots, a_n)$. Then

$$d_r(AX) = d_r(UAVB) = d_r(DB) = \gcd_{\omega \in Q_{r,n}} \det(DB[\omega|1, \ldots, r])\,.$$

Since $x_j \in V_{i_j} = \operatorname{span}_R\{v_1, \ldots, v_{i_j}\}$, in column j of DB the entries below row i_j are zero, for $j = 1, \ldots, r$. Therefore, if $\omega(j) > i_j$ for some j, then $\det(DB[\omega|1, \ldots, r]) = 0$. Hence

$$d_r(AX) = \gcd_{\substack{\omega \in Q_{r,n} \\ \omega(j) \leq i_j, \forall j}} \det(B[\omega|1, \ldots, r]) \prod_{j=1}^{r} a_{\omega(j)}\,.$$

Similarly,

$$d_r(X) = d_r(VB) = d_r(B) = \gcd_{\substack{\omega \in Q_{r,n} \\ \omega(j) \leq i_j, \forall j}} \det(B[\omega|1, \ldots, r])\,.$$

For $\omega \in Q_{r,n}$ such that $\omega(j) \leq i_j$ for all j, let $c_\omega \in R$ be such that $\det(B[\omega|1,\ldots,r]) = c_\omega d_r(X)$.

Then we have

$$\psi(A|_L) = \frac{d_r(AX)}{d_r(X)} = \operatorname*{gcd}_{\substack{\omega\in Q_{r,n}\\ \omega(j)\leq i_j,\forall j}} \; c_\omega \prod_{j=1}^{r} a_{\omega(j)}$$

and, since $\omega(j) \leq i_j \Rightarrow a_{i_j} \mid a_{\omega(j)}$, we get that $a_{i_1}a_{i_2}\cdots a_{i_r} \mid \psi(A|_L)$.

Therefore, $a_{i_1}a_{i_2}\cdots a_{i_r}$ is a common divisor of the elements of the set $\{\psi(A|_L) : L \in \Omega_I(V)\}$.

On the other hand, $a_{i_1}a_{i_2}\cdots a_{i_r}$ belongs to that set because $F = \operatorname{span}_R\{v_{i_1}, v_{i_2}, \ldots, v_{i_r}\} \in \Omega_I(V)$, $v_{i_j} \in V_{i_j}$ for all j, and

$$\psi(A|_F) = \frac{d_r(A[v_{i_1}\cdots v_{i_r}])}{d_r([v_{i_1}\cdots v_{i_r}])} = \frac{d_r(D[e_{i_1}\cdots e_{i_r}])}{1} = a_{i_1}a_{i_2}\cdots a_{i_r}\,,$$

where $\{e_1,\ldots,e_n\}$ is the canonical basis of R^n. ■

The argument for the second extremal characterization is similar. We introduce a new notation. For $i = 1,\ldots,n$ write $V'_i = \operatorname{span}_R\{v_{n-i+1},\ldots,v_n\}$, and $V' = (V'_1,\ldots,V'_n)$. If $I' = (n-i_r+1,\ldots,n-i_1+1)$, then $\operatorname{span}_R\{v_{i_1},\ldots,v_{i_r}\} \in \Omega_{I'}(V')$.

Theorem 4.3. *For every $I = (i_1,\ldots,i_r) \in Q_{r,n}$ we have*

$$a_{i_1}a_{i_2}\cdots a_{i_r} = \operatorname*{lcm}_{L\in\Omega_{I'}(V')} \psi\left(A_{|L}\right)\,.$$

Proof. Let $L \in \Omega_{I'}(V')$, and let $x_1,\ldots,x_r \in L$ be linearly independent such that $L = \operatorname{span}_R\{x_1,\ldots,x_r\}$ and $x_j \in V'_{n-i_{r-j+1}+1}$ for all j.

Let B be the $n\times r$ matrix such that $X = [x_1\cdots x_r] = VB$ and $D = \operatorname{diag}(a_1,\ldots,a_n)$. We have

$$d_r(AX) = \operatorname*{gcd}_{\substack{\omega\in Q_{r,n}\\ \omega(j)\geq i_j,\forall j}} \det(B[\omega|1,\ldots,r]) \prod_{j=1}^{r} a_{\omega(j)}$$

and similarly

$$d_r(X) = \operatorname*{gcd}_{\substack{\omega\in Q_{r,n}\\ \omega(j)\geq i_j,\forall j}} \det(B[\omega|1,\ldots,r])\,.$$

For $\omega \in Q_{r,n}$ such that $\omega(j) \geq i_j$ for all j, let $c_\omega \in R$ be such that $\det(B[\omega|1,\ldots,r]) = c_\omega d_r(X)$.

Then we have

$$\psi(A|_L) = \frac{d_r(AX)}{d_r(X)} = \gcd_{\substack{\omega\in Q_{r,n} \\ \omega(j)\geq i_j,\forall j}} c_\omega \prod_{j=1}^{r} a_{\omega(j)}$$

and, since $\omega(j) \geq i_j \Rightarrow a_{\omega(j)} \mid a_{i_j}$, we get that $\psi(A|_L)$ divides

$$\gcd_{\substack{\omega\in Q_{r,n} \\ \omega(j)\geq i_j,\forall j}} c_\omega a_{i_1}a_{i_2}\cdots a_{i_r} = (a_{i_1}a_{i_2}\cdots a_{i_r}) \gcd_{\substack{\omega\in Q_{r,n} \\ \omega(j)\geq i_j,\forall j}} c_\omega = a_{i_1}a_{i_2}\cdots a_{i_r}\,.$$

Therefore, $a_{i_1}a_{i_2}\cdots a_{i_r}$ is a common multiple of the elements of the set $\{\psi(A|_L) : L \in \Omega_{I'}(V')\}$.

On the other hand, $a_{i_1}a_{i_2}\cdots a_{i_r}$ belongs to that set because $F = \operatorname{span}_R\{v_{i_1}, v_{i_2}, \ldots, v_{i_r}\} \in \Omega_{I'}(V')$, $v_{i_{r-j+1}} \in V'_{n-i_{r-j+1}+1}$ for all j, and

$$\psi(A|_F) = \frac{d_r(A[v_{i_1}\cdots v_{i_r}])}{d_r([v_{i_1}\cdots v_{i_r}])} = \frac{d_r(D[e_{i_1}\cdots e_{i_r}])}{1} = a_{i_1}a_{i_2}\cdots a_{i_r}\,. \blacksquare$$

5. Schubert intersections and divisibility relations

The basic result which allows us to prove our divisibility relations is the following.

Theorem 5.1. *Let $A, B \in R^{n\times n}$ and L, M be pure submodules of R^n such that* $\operatorname{rank}(M) = \operatorname{rank}(L)$ *and M contains $BL := \{Bv : v \in L\}$. Then*

$$\psi\left(AB_{|L}\right) = \psi\left(A_{|M}\right)\psi\left(B_{|L}\right)\,.$$

Proof. Let $r = \operatorname{rank}(L) = \operatorname{rank}(M)$ and consider $X = [x_1\cdots x_r]$, with $x_1, \ldots, x_r \in L$ linearly independent and such that $L = \operatorname{span}_R\{x_1, \ldots, x_r\}$. Put $Y = [y_1\cdots y_r]$, with $y_1, \ldots, y_r \in M$ linearly independent such that $M = \operatorname{span}_R\{y_1, \ldots, y_r\}$. Since $BL \subseteq M$, there exist $c_1, \ldots, c_r \in R\setminus\{0\}$ and

$Z \in R^{r\times r}$ such that $BX\operatorname{diag}(c_1,\dots,c_r) = YZ$. Therefore

$$\begin{aligned}
\psi(AB|_L) &= \frac{d_r(ABX)}{d_r(X)} = \frac{d_r(ABX\operatorname{diag}(c_1,\dots,c_r))}{d_r(X)c_1\cdots c_r} \\
&= \frac{d_r(AYZ)}{d_r(X)c_1\cdots c_r} = \frac{d_r(AY)\det(Z)}{d_r(X)c_1\cdots c_r} \\
&= \frac{d_r(AY)d_r(Y)\det(Z)}{d_r(Y)d_r(X)c_1\cdots c_r} = \frac{d_r(AY)}{d_r(Y)}\,\frac{d_r(YZ)}{d_r(X)c_1\cdots c_r} \\
&= \psi(A|_M)\,\frac{d_r(BX\operatorname{diag}(c_1,\dots,c_r))}{d_r(X)c_1\cdots c_r} \\
&= \psi(A|_M)\,\frac{d_r(BX)}{d_r(X)} = \psi(A|_M)\psi(B|_L)\,. \ \blacksquare
\end{aligned}$$

Lemma 5.2. *Let $B \in R^{n\times n}$ be nonsingular and L and S submodules of R^n, with S pure. Then $\overline{BL} \cap S = \overline{B(L \cap \overline{\operatorname{adj}(B)S})}$.*

Proof. Let $x \in BL \cap S$. Then $x = By$ with $y \in L$. Hence $\det(B)y = \operatorname{adj}(B)x$ and so $y \in \overline{\operatorname{adj}(B)S}$. Therefore $y \in L \cap \overline{\operatorname{adj}(B)S}$ and $x = By \in B(L \cap \overline{\operatorname{adj}(B)S})$. So we have proved that $BL \cap S \subseteq B(L \cap \overline{\operatorname{adj}(B)S})$. Thus $\overline{BL} \cap S = \overline{BL \cap S} \subseteq \overline{B(L \cap \overline{\operatorname{adj}(B)S})}$.

Let now $x \in \overline{B(L \cap \overline{\operatorname{adj}(B)S})}$. There exist $a \in R\setminus\{0\}$ and $y \in L \cap \overline{\operatorname{adj}(B)S}$ such that $ax = By$. Also, there exist $b \in R \setminus \{0\}$ and $z \in S$ such that $by = \operatorname{adj}(B)\,z$. So $abx = Bby = \det(B)z$ and, therefore, $x \in \overline{S} = S$. On the other hand, $ax = By \in BL$, so $x \in \overline{BL}$. $\blacksquare$

Theorem 5.3. *Let $S = (S_1,\dots,S_n)$ where $S_1 \subset \cdots \subset S_n$ are pure submodules of R^n such that $\operatorname{rank}(S_k) = k$, for $k = 1,\dots,n$. Let $B \in R^{n\times n}$ be nonsingular and, for $k = 1,\dots,n$, put $T_k = \overline{\operatorname{adj}(B)S_k}$. Then $T_1 \subset \cdots \subset T_n$ are pure submodules of R^n such that $\operatorname{rank}(T_k) = k$, $k = 1,\dots,n$. Additionally, if $T = (T_1,\dots,T_n)$ and $I = (i_1,\dots,i_r) \in Q_{r,n}$, and L is a pure submodule of R^n, we have*

$$L \in \Omega_I(T) \iff \overline{BL} \in \Omega_I(S)\,.$$

Proof. For $k = 1, \ldots, n-1$, we have

$$S_k \subset S_{k+1} \Rightarrow \operatorname{adj}(B)S_k \subseteq \operatorname{adj}(B)S_{k+1} \Rightarrow T_k \subseteq T_{k+1}\,.$$

On the other hand, for all k,

$$\operatorname{rank}(T_k) = \operatorname{rank}(\operatorname{adj}(B)S_k) = \dim_{\mathbb{K}} \mathbb{K}(\operatorname{adj}(B)S_k) = \dim_{\mathbb{K}}(\mathbb{K}S_k) = \operatorname{rank}(S_k) = k\,.$$

Let L be a pure submodule of R^n. Since B is nonsingular, we have that $\operatorname{rank}(\overline{BL}) = \operatorname{rank}(BL) = \dim_{\mathbb{K}} \mathbb{K}(BL) = \dim_{\mathbb{K}} \mathbb{K}L = \operatorname{rank}(L)$. On the the other hand, for $j = 1, \ldots, r$,

$$\operatorname{rank}(\overline{BL} \cap S_{i_j}) = \operatorname{rank}(\overline{B(L \cap \overline{\operatorname{adj}(B)S_{i_j}})})$$

$$= \operatorname{rank}(\overline{B(L \cap T_{i_j})}) = \operatorname{rank}(B(L \cap T_{i_j})) = \operatorname{rank}(L \cap T_{i_j})\,,$$

and we get the result. ■

Let A, V, V' and $a_n \mid a_{n-1} \mid \cdots \mid a_1$ as before.

Let $B \in R^{n\times n}$, $C = AB$, with invariant factors $b_n \mid b_{n-1} \mid \cdots \mid b_1$ and $c_n \mid c_{n-1} \mid \cdots \mid c_1$, respectively. Let W, W' and P, P' be chains of pure submodules of R^n defined from the columns of unimodular matrices V_1, V_2 such that $U_1BV_1 = \operatorname{diag}(b_1, \ldots, b_n)$ and $U_2CV_2 = \operatorname{diag}(c_1, \ldots, c_n)$, respectively (with U_1, U_2 unimodular). Let $I = (i_1, \ldots, i_r)$, $J = (j_1, \ldots, j_r)$, $K = (k_1, \ldots, k_r) \in Q_{r,n}$.

When is a intersection of the type $\Omega_K(\mathbb{K}P) \cap \Omega_{I'}(\mathbb{K}V') \cap \Omega_{J'}(\mathbb{K}W')$ nonempty? In a recent paper [1] it is proved that this happens when $\lambda(K)$ can be obtained from $\lambda(I)$ and $\lambda(J)$ in only one way using the Littlewood-Richardson rule, or, in the language of Littlewood-Richardson coefficients [6], when $c^{\lambda(K)}_{\lambda(I)\lambda(J)} = 1$.

We then have:

Theorem 5.4. *If $\lambda(K)$ can be obtained from $\lambda(I)$ and $\lambda(J)$ in only one way using the Littlewood-Richardson rule then*

$$c_{k_1}c_{k_2}\cdots c_{k_r} \mid a_{i_1}a_{i_2}\cdots a_{i_r}b_{j_1}b_{j_2}\cdots b_{j_r}\,.$$

Proof. For $i = 1, \ldots, n$ write $T_i = \overline{\operatorname{adj}(B)V'_i}$ and consider $T = (T_1, \ldots, T_n)$. Under the hypothesis we have

$$\Omega_K(\mathbb{K}P) \cap \Omega_{I'}(\mathbb{K}T) \cap \Omega_{J'}(\mathbb{K}W') \neq \varnothing$$

whence

$$\Omega_K(P) \cap \Omega_{I'}(T) \cap \Omega_{J'}(W') \neq \varnothing\,.$$

Let $L \in \Omega_K(P) \cap \Omega_{I'}(T) \cap \Omega_{J'}(W')$. Then $\overline{BL} \in \Omega_{I'}(V')$ and

$$\begin{aligned} c_{k_1} c_{k_2} \cdots c_{k_r} &= \gcd_{S \in \Omega_K(P)} \psi(AB_{|S}) \mid \psi(AB_{|L}) = \psi(A_{|\overline{BL}}) \psi(B_{|L}) \\ &\mid \operatorname*{lcm}_{S \in \Omega_{I'}(V')} \psi(A_{|S}) \operatorname*{lcm}_{S \in \Omega_{J'}(W')} \psi(B_{|S}) = a_{i_1} a_{i_2} \cdots a_{i_r} b_{j_1} b_{j_2} \cdots b_{j_r} . \quad \blacksquare \end{aligned}$$

In the general case when $\lambda(K)$ can be obtained from $\lambda(I)$ and $\lambda(J)$ in one or more ways, or $c_{\lambda(I)\lambda(J)}^{\lambda(K)} \geq 1$, the intersection of the Schubert varieties may be empty but the corresponding divisibility relation follows from those in the theorem (see [6] and its references). So we get that all "Horn relations", *i.e.* those whose indices appear in the inequalities solving the Hermitian sum problem, remain valid in our setting.

6. Extension to GCD domains

The above proof of divisibility relations for matrices over EDDs allows a further extension to a even larger class of rings. We briefly describe this technique, already used by Kaplansky in [9].

An integral domain is a *valuation domain* if, up to products by units, divisibility is a total order.

If R is an integral domain, we say R is *integrally closed* if it contains the roots of monic polynomials over R. A result by Krull states that such an R is equal to the intersection of all valuation domains that contain it. Therefore, a divisibility relation holds in R if and only if it holds in every valuation domain containing R. Trivially valuation domains are EDDs. Hence divisibility relations proved for arbitrary EDDs may be used to obtain statements valid for integrally closed domains.

We are interested in the class of GCD domains, defined by the condition that every finite set of elements has a gcd in the ring. This class contains EDDs (or, more generally, Bézout domains, *i.e.* domains in which every finitely generated ideal is principal) and also unique factorization domains. GCD domains are easily seen to be integrally closed.

For a matrix over a GCD domain we can define invariant factors as quotients of determinantal divisors as in the Introduction. The very fact that the invariant factors form a divisibility chain is an example of a divisibility relation that extends from EDDs to GCD domains using the argument above (we don't know a direct proof of that).

We can now present the desired extension.

Theorem 6.1. *Let A, B be nonsingular $n \times n$ matrices over a GCD domain and let $a_n \mid \cdots \mid a_1$, $b_n \mid \cdots \mid b_1$ and $c_n \mid \cdots \mid c_1$ be the invariant factors of A, B and AB. For any $r < n$ and $I, J, K \in Q_{r,n}$, if $\lambda(K) \in LR(\lambda(I), \lambda(J))$ then*

$$c_{k_1} c_{k_2} \cdots c_{k_r} \mid a_{i_1} a_{i_2} \cdots a_{i_r} b_{j_1} b_{j_2} \cdots b_{j_r} .$$

7. Final remarks

A literature search shows that ideas similar to those in this paper appear in two papers separated by 30 years. In [3], which deals with many other subjects, different extremal characterizations for products of invariant factors of matrices over a PID are presented but not used to obtain divisibility relations. In the very recent paper [2], the Horn relations for invariant factors are proved for modules over a PID using the intersection of Schubert varieties, a different technique from that presented in [6], where the main connection was via representation theory.

It is natural to ask if the relations presented in Theorem 5.4, together with the equality for $r = n$, constitute the complete answer for the invariant factor problem for EDDs (as they are for PIDs), *i.e.* if they are sufficient for the existence of matrices A and B such that the given elements are the invariant factors of A, B and AB.

References

[1] H. Bercovici, B. Collins, K. Dykema, W. S. Li and D. Timotin, Intersections of Schubert varieties and eigenvalue inequalities in an arbitrary finite factor, *Journal of Functional Analysis* 258 (2010), 1579-1627.

[2] H. Bercovici, K. Dykema and W. S. Li, The Horn inequalities for submodules, *Acta Sci. Math. (Szeged)* 79 (2013), 17-30.

[3] D. Carlson and E. Marques de Sá, Generalized Minimax and Interlacing Theorems, *Linear and Multilinear Algebra*, 15 (1984), 77-103.

[4] C. Faith, *Algebra II: Ring Theory*, Berlin, Springer, 1976.

[5] W. Fulton and J. Harris, *Representation Theory*, New York, Springer, 1991.

[6] W. Fulton, Eigenvalues, invariant factors, highest weights, and Schubert calculus, *Bulletin AMS* 37 (2000), 209-249.

[7] H. Glüsing-Lüerssen, A behavioral approach to delay-differential systems, *SIAM J. Control Optim.* 35 (1997), 480-499.

[8] O. Helmer, The elementary divisor theorem for certain rings without chain condition, *Bulletin AMS* 49 (1943), 225-236.

[9] I. Kaplansky, Elementary divisors and modules, *Transactions AMS* 66 (1949), 464-491.

[10] T. Klein, The multiplication of Schur functions and extensions of p-modules, *J. London Math. Soc.* 43 (1968), 280-284.
[11] A. Knutson, The symplectic and algebraic geometry of Horn's problem, *Linear Algebra Appl.*, 319 (2000), 61-81.
[12] J. F. Queiró, Invariant factors as approximation numbers, *Linear Algebra Appl.* 49 (1983), 131-136.
[13] J. J. Rotman, *An Introduction to Homological Algebra*, Acad. Press, 1979.
[14] E. Marques de Sá, *Interlacing Problems for Invariant Factors*, University of Coimbra, Textos de Matemática no. 2 (1998).
[15] A. P. Santana, J. F. Queiró and E. Marques de Sá, Group representations and matrix spectral problems, *Linear and Multilinear Algebra* 46 (1999), 1-23.
[16] R. C. Thompson, Smith invariants of a product of integral matrices, in *Linear Algebra and Its Role in Systems Theory*, Contemp. Math. 47, AMS (1985), 401-435.
[17] R. C. Thompson and S. Therianos, On the singular values of matrix products - I, *Scripta Mathematica* 29 (1973), 99-110.

Cristina Caldeira
CMUC, Department of Mathematics, University of Coimbra
P-3001 454 Coimbra, Portugal

E-mail address: caldeira@mat.uc.pt

João Filipe Queiró
CMUC, Department of Mathematics, University of Coimbra
P-3001 454 Coimbra, Portugal

E-mail address: jfqueiro@mat.uc.pt