LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 315 (2000) 125–138
www.elsevier.com/locate/laa

The invariant polynomials degrees of the Kronecker sum of two linear operators and additive theory

Cristina Caldeira [a,1], J.A. Dias da Silva [b,*,2]

[a] *Departamento de Matemática, Universidade de Coimbra, 3000 Coimbra, Portugal*
[b] *Departamento de Matemática, da Universidade de Lisboa, Rua Ernesto de Vasconcelos, 1749-016 Lisboa, Portugal*

Received 2 April 1999; accepted 27 March 2000

Submitted by T.J. Laffey

Abstract

Let G be an abelian group. Let A and B be finite non-empty subsets of G. By $A + B$ we denote the set of all elements $a + b$ with $a \in A$ and $b \in B$. For $c \in A + B$, $\nu_c(A, B)$ is the cardinality of the set of pairs (a, b) such that $a + b = c$. We call $\nu_c(A, B)$ *the multiplicity of* c *(in* $A + B$).

Let i be a positive integer. We denote by $\mu_i(A, B)$ or briefly by μ_i the cardinality of the set of the elements of $A + B$ that have multiplicity greater than or equal to i.

Let $\mathbb{F}$ be a field. Let p be the characteristic of $\mathbb{F}$ in case of finite characteristic and ∞ if $\mathbb{F}$ has characteristic 0. Let A and B be finite non-empty subsets of $\mathbb{F}$.

We will prove that for every $\ell = 1, \ldots, \min\{|A|, |B|\}$ one has

$$\mu_1 + \cdots + \mu_\ell \geqslant \ell \min\{p, |A| + |B| - \ell\}. \tag{a}$$

This statement on the multiplicities of the elements of $A + B$ generalizes Cauchy–Davenport Theorem. In fact Cauchy–Davenport is exactly inequality (a) for $\ell = 1$. When $\mathbb{F} = \mathbb{Z}_p$ inequality (a) was proved in J.M. Pollard (J. London Math. Soc. 8 (1974) 460–462); see also M.B. Nathanson (Additive number theory: Inverse problems and the geometry of sumsets, Springer, New York, 1996).

* Corresponding author. Tel.: +351-1-790-4828; fax: +351-1-790-4700.
E-mail address: perdigao@hermite.cii.fc.ul.pt (J.A. Dias da Silva).

1 This research was done within the activities of “Centro de Matemática da Universidade de Coimbra” and partially supported by PRAXIS project “Álgebra e Matemáticas Discretas”.

2 This research was done within the activities of “Centro de Álgebra da Universidade de Lisboa” and partially supported by PRAXIS project “Álgebra e Matemáticas Discretas”.

PII: S 0 0 2 4 - 3 7 9 5 (0 0) 0 0 1 2 5 - 7

Keywords: Additive number theory; Derivations; Invariant polynomials

1. Introduction

Let G be an abelian group. Let A and B be finite non-empty subsets of G. By $A + B$ we denote the set of all elements $a + b$ with $a \in A$ and $b \in B$. For $c \in A + B$, $\nu_c(A, B)$ is the cardinality of the set of pairs (a, b) such that $a + b = c$. We call $\nu_c(A, B)$ *the multiplicity of c (in A + B)*.

Let i be a positive integer. We denote by $\mu_i(A, B)$ or briefly by μ_i the cardinality of the set of the elements of $A + B$ that have multiplicity greater than or equal to i.

Let X be a set. We denote by $|X|$ the cardinality of X. If $|X| = k$, we say that X is a k-set.

Let p be a prime number. If $G = \mathbb{Z}_p$, the Cauchy–Davenport Theorem [1–3] states that

$$|A + B| \geqslant \min\{p, |A| + |B| - 1\}.$$

In [4] the degree of the minimal polynomial of the Kronecker sum of two linear operators is studied and an alternative proof of Cauchy–Davenport Theorem is derived from this study.

Let $\mathbb{F}$ be a field. Let p be the characteristic of $\mathbb{F}$ in case of finite characteristic and ∞ if $\mathbb{F}$ has characteristic 0. Let A and B be finite non-empty subsets of $\mathbb{F}$. The main purpose of this article is to state lower bounds for the sum of the degrees of the initial segments of the (divisibility non-decreasing) chain of the invariant polynomials of the Kronecker sum of two linear operators and to get, from this study, new results on the multiplicities of the elements of $A + B$. In fact we will prove that for every $\ell = 1, \ldots, \min\{|A|, |B|\}$ we have

$$\mu_1 + \cdots + \mu_\ell \geqslant \ell \min\{p, |A| + |B| - \ell\}. \tag{1}$$

This statement on the multiplicities of the elements of $A + B$ generalizes Cauchy–Davenport Theorem. In fact Cauchy–Davenport is exactly inequality (1) for $\ell = 1$. When $\mathbb{F} = \mathbb{Z}_p$, inequality (1) was proved in [6] (see also [5]).

We can see (check the remark at the end of Section 3) that these lower bounds are tight and the equality, in the inequalities (1), is attained when A and B are arithmetic progressions of the same rate.

2. Generalized cyclic subspaces

Let $\mathbb{F}$ be an arbitrary field and denote by $\overline{\mathbb{F}}$ the algebraic closure of $\mathbb{F}$. Let $V \neq \{0\}$ be an n-dimensional vector space over $\mathbb{F}$. Let $\mathscr{B}$ be a basis of V. By I_V we denote the identity operator on V. Let g be a linear operator on V. We denote by P_g the minimal polynomial of g. For every $x \in V$ we denote by $\mathscr{C}_g(x)$ the g-cyclic space of x, i.e.

$$\mathscr{C}_g(x) = \langle g^i(x) \colon\ i \in \mathbb{N} \cup \{0\} \rangle,$$

where $\langle X \rangle$ means the linear closure of X. We use $\sigma(g)$ to denote the spectrum of g, i.e. $\sigma(g)$ is the family of the n characteristic roots of g in $\overline{\mathbb{F}}$, and $\alpha_{g,1}, \ldots, \alpha_{g,n}$, $(\alpha_{g,1} | \cdots | \alpha_{g,n})$ to denote the invariant polynomials of g. The following result is well-known.

Theorem 2.1 (Max–min). *The maximum dimension of the g-cyclic spaces*, $\mathscr{C}_g(x)$, *when x runs over* V, *is equal to the degree of* $P_g = \alpha_{g,n}$.

The purpose of this section is the generalization of this theorem.

Definition 2.2. Let $x_1, \ldots, x_\ell$ be linearly independent vectors of V and g a linear operator on V. We call *generalized g-cyclic subspace associated to* $x_1, \ldots, x_\ell$ the subspace

$$\mathscr{C}_g(x_1, \ldots, x_\ell) = \langle g^i(x_j) \colon\ i \in \mathbb{N} \cup \{0\},\ j = 1, \ldots, \ell \rangle.$$

The subspace $\mathscr{C}_g(x_1, \ldots, x_\ell)$ is the smallest g-invariant subspace containing $x_1, \ldots, x_\ell$.

We say that the pair $((x_1, \ldots, x_\ell), g)$ or the generalized g-cyclic subspace $\mathscr{C}_g(x_1, \ldots, x_\ell)$ are *completely controllable* if

$$\langle x_1, x_2, \ldots, x_\ell, g(x_1), \ldots, g(x_\ell), g^2(x_1), \ldots, g^2(x_\ell), \ldots \rangle = V. \tag{2}$$

Definition 2.3. Let g be a linear operator on V and $x_1, \ldots, x_\ell$ linearly independent vectors of V. A basis, $\mathscr{B}$, of $\mathscr{C}_g(x_1, \ldots, x_\ell)$ selected from the vectors of the sequence

$$x_1, x_2, \ldots, x_\ell, g(x_1), \ldots, g(x_\ell), g^2(x_1), \ldots, g^2(x_\ell), \ldots$$

is *nice* if, for $0 \leqslant i \leqslant k-1$, $g^i(x_j) \in \mathscr{B}$ provided that $g^k(x_j) \in \mathscr{B}$.

Let

$$\begin{aligned}\mathscr{B} = \{ & x_1, g(x_1), \ldots, g^{r_1-1}(x_1), x_2, g(x_2), \ldots, \\ & g^{r_2-1}(x_2), \ldots, x_\ell, g(x_\ell), \ldots, g^{r_\ell-1}(x_\ell)\}\end{aligned}$$

be a nice basis of $\mathscr{C}_g(x_1, \ldots, x_\ell)$. The non-negative integers $r_i,\ i = 1, \ldots, \ell$, are called *indices of* $\mathscr{B}$.

Let $\{x_1, \ldots, x_\ell\}$ be a linearly independent ℓ-set of vectors of V. If

$$\mathscr{I} = \bigcup_{i=1}^{\ell} \{x_i, g(x_i), g^2(x_i), \ldots, g^{s_i-1}(x_i)\}$$

is a linearly independent $(s_1 + \cdots + s_\ell)$-set, we say that $\mathscr{I}$ is a $((x_1, \ldots, x_\ell), g)$-*nice independent set* and we call the non-negative integers $s_1, \ldots, s_\ell$ *indices of* $\mathscr{I}$.

Definition 2.4. Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ be sequences of non-negative integers. Denote by $(\overline{a}_1, \ldots, \overline{a}_n)$ and $(\overline{b}_1, \ldots, \overline{b}_n)$ the reordering, in a non-increasing way, of a and b, respectively. We say that *a weakly-dominates b* and we write

$$a \sqsupseteq b$$

if

$$\sum_{i=1}^{k} \overline{a}_i \geqslant \sum_{i=1}^{k} \overline{b}_i, \quad k = 1, \ldots, n.$$

If also $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$, we say that *a dominates b* and we write $a \succeq b$.

In [7], the following result is proved.

Propositon 2.5. *Let $n = t + q$. Let $\alpha_1|\alpha_2|\cdots|\alpha_t$ be the invariant polynomials of the $t \times t$ matrix A. Let $\gamma_1, \ldots, \gamma_n$ be monic polynomials such that $\deg(\gamma_1 \cdots \gamma_n) = n$ and $\gamma_1|\cdots|\gamma_n$. Then there exist $C \in \mathbb{F}^{q\times t}$ and $D \in \mathbb{F}^{q\times q}$ such that the $n \times n$ matrix*

$$\begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$$

has invariant polynomials $\gamma_1, \ldots, \gamma_n$, if and only if

$$\gamma_i|\alpha_i|\gamma_{i+q}, \quad i = 1, \ldots, t.$$

The following result is proved in [9, Corollary 2.2] and states, for a fixed linear operator g on V and linearly independent vectors $x_1, \ldots x_\ell$ such that $\mathscr{C}_g(x_1, \ldots, x_\ell)$ is completely controllable, a necessary and sufficient condition for the existence of a nice basis of $\mathscr{C}_g(x_1, \ldots, x_\ell)$ with prescribed indices.

Theorem 2.6. *Let g be a linear operator on V. Let $r_1, \ldots, r_\ell$ be positive integers. Then there exist linearly independent vectors $x_1, \ldots, x_\ell$ and a nice basis $\mathscr{B}$, of $\mathscr{C}_g(x_1, \ldots, x_\ell)$, with indices $r_1, \ldots, r_\ell$ such that $\mathscr{C}_g(x_1, \ldots, x_\ell)$ is completely controllable if and only if the following conditions hold*:

$$\alpha_{g,i} = 1, \quad i = 1, \ldots, n - \ell,$$

and

$$(r_1, \ldots, r_\ell) \preceq (\deg(\alpha_{g,n}), \ldots, \deg(\alpha_{g,n-\ell+1})).$$

The next theorem states a necessary condition for the existence of nice bases with prescribed indices, where the constraint of complete controllability is skipped.

Theorem 2.7. *Let g be a linear operator on V. Let $r_1, \ldots, r_\ell$ be positive integers. If there exist linearly independent vectors $x_1, \ldots, x_\ell$ and a nice basis $\mathscr{B}$, of $\mathscr{C}_g(x_1, \ldots, x_\ell)$, with indices $r_1, \ldots, r_\ell$, then the following condition holds*:

$$(r_1, \ldots, r_\ell) \sqsubseteq (\deg(\alpha_{g,n}), \ldots, \deg(\alpha_{g,n-\ell+1})).$$

Proof. Let $U = \mathscr{C}_g(x_1, \ldots, x_\ell)$. By definition $\mathscr{C}_{g|U}(x_1, \ldots, x_\ell)$ is completely controllable. Assume that $\dim(\mathscr{C}_g(x_1, ..., x_\ell)) = \dim(U) = t$ and that $q = n - t$. Then Theorem 2.6 guarantees that

$$(r_1, \ldots, r_\ell) \preceq (\deg(\alpha_{g|U,t}), \ldots, \deg(\alpha_{g|U,t-\ell+1})). \tag{3}$$

By the transposed version of Proposition 2.5, we know that

$$\alpha_{g,i} | \alpha_{g|U,i} | \alpha_{g,i+q}, \quad i = 1, \ldots, t.$$

Therefore,

$$\alpha_{g|U,t}\alpha_{g|U,t-1} \cdots \alpha_{g|U,t-j} | \alpha_{g,n}\alpha_{g,n-1} \cdots \alpha_{g,n-j}, \quad j = 0, \ldots, t-1. \tag{4}$$

Taking degrees in (4) and bearing in mind (3) we get

$$(r_1, \ldots, r_\ell) \sqsubseteq (\deg(\alpha_{g,n}), \ldots, \deg(\alpha_{g,n-\ell+1})). \qquad \square$$

Corollary 2.8. *Let g be a linear operator on V. Let $s_1, \ldots, s_\ell$ be positive integers. If there exist linearly independent vectors $v_1, \ldots, v_\ell$ such that*

$$\bigcup_{i=1}^{\ell} \{v_i, g(v_i), g^2(v_i), \ldots, g^{s_i-1}(v_i)\}$$

is a linearly independent $(s_1 + \cdots + s_\ell)$-set, then the following condition holds:

$$(s_1, \ldots, s_\ell) \sqsubseteq (\deg(\alpha_{g,n}), \ldots, \deg(\alpha_{g,n-\ell+1})).$$

Proof. Complete the set

$$\bigcup_{i=1}^{\ell} \{v_i, g(v_i), g^2(v_i), \ldots, g^{s_i-1}(v_i)\}$$

to a nice basis of $\mathscr{C}_g(v_1, \ldots, v_\ell)$. This completion is always possible as can be easily seen. In fact, for $q \in \{1, \ldots, \ell\}$, let t_q be the positive integer such that

$$\left(\bigcup_{j=1}^{q} \{v_j, g(v_j), \ldots, g^{t_j-1}(v_j)\}\right) \cup \left(\bigcup_{i=q+1}^{\ell} \{v_i, g(v_i), g^2(v_i), \ldots, g^{s_i-1}(v_i)\}\right)$$

is a linearly independent $(t_1 + \cdots + t_q + s_{q+1} + \cdots + s_\ell)$-set and

$$g^{t_q}(v_q) \in \left\langle \left(\bigcup_{j=1}^{q} \{v_j, g(v_j), \ldots, g^{t_j-1}(v_j)\}\right) \right.$$
$$\left. \cup \left(\bigcup_{i=q+1}^{\ell} \{v_i, g(v_i), g^2(v_i), \ldots, g^{s_i-1}(v_i)\}\right) \right\rangle.$$

It is obvious, from the definitions, that

$$g\left(\left\langle\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}\right\rangle\right)\subseteq\left\langle\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}\right\rangle. \tag{5}$$

We are going to show that

$$\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}$$

is a maximal linear independent set contained in $\langle g^j(v_i)\,|\,i=1,\ldots,\ell,\ \ j\in\mathbb{N}\cup\{0\}\rangle$. Assume, in order to get a contradiction, that for some $i\in\{1,\ldots,\ell\}$ and some $r\in\mathbb{N}$, $g^r(v_i)\notin\langle\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}\rangle$. Wlog we can suppose that r is the smallest integer with this property. Then

$$g^{r-1}(v_i)\in\left\langle\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}\right\rangle.$$

Therefore,

$$g^r(v_i)\in g\left(\left\langle\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}\right\rangle\right).$$

Using (5) we get

$$g^r(v_i)\in\left\langle\bigcup_{i=1}^{\ell}\{v_i,\ldots,g^{t_i-1}(v_i)\}\right\rangle.$$

Contradiction.

By Theorem 2.7 we can conclude that

$$(t_1,\ldots,t_\ell)\sqsubseteq(\deg(\alpha_{g,n}),\ldots,\deg(\alpha_{g,n-\ell+1})).$$

But, since by construction, we have $s_i\leqslant t_i,\ \ i=1,\ldots,\ell$, we get from the former inequalities

$$(s_1,\ldots,s_\ell)\sqsubseteq(\deg(\alpha_{g,n}),\ldots,\deg(\alpha_{g,n-\ell+1})). \qquad \square$$

3. Main results

Notation. Let A and B be subsets of the field $\mathbb{F}$. Recall that, if i is a positive integer, $\mu_i(A,B)$ (or μ_i) is the cardinality of the set $\{x\in A+B: \nu_x(A,B)\geqslant i\}$.

Theorem 3.1. *Let V and W be non-zero finite-dimensional vector spaces over the field $\mathbb{F}$ with dimensions n and m, respectively. Let p be the characteristic of $\mathbb{F}$ in the case of finite characteristic and ∞ if $\mathbb{F}$ has characteristic 0. Assume that ℓ is a positive integer satisfying*

$$\ell \leqslant \min\{\deg(P_f), \deg(P_g)\}.$$

Then we have

$$\sum_{i=1}^{\ell} \deg(\alpha_{f\otimes I_W + I_V \otimes g, mn-i+1}) \geqslant \ell \min\{p, \deg(P_f) + \deg(P_g) - \ell\}.$$

Theorem 3.2. *Let A and B be finite non-empty subsets of* $\mathbb{F}$. *Then, for* $\ell = 1, 2, \ldots,$ $\min\{|A|, |B|\}$,

$$\sum_{i=1}^{\ell} \mu_i \geqslant \ell \min\{p, |A| + |B| - \ell\}.$$

4. Proofs

Let $V \neq \{0\}$ be an n-dimensional vector space over the field $\mathbb{F}$ and let h be a linear operator on V. Let i be a positive integer. Denote by $m_i(h)$ the cardinality of the elements of $\sigma(h)$ whose algebraic multiplicity is greater than or equal to i. The following proposition is an easy consequence of basic results on Linear Algebra.

Propositon 4.1. *Let h be a diagonalizable linear operator on the n-dimensional vector space V. Then, if* $j \leqslant n$, *we have*

$$m_1(h) + \cdots + m_j(h) = \sum_{i=1}^{j} \deg(\alpha_{h,n-i+1}).$$

Propositon 4.2. *Given non-empty finite subsets of* $\mathbb{F}$, *A and B, let V and W be vector spaces over* $\mathbb{F}$ *of dimensions* $|A|$ *and* $|B|$, *respectively. Let f be a linear operator on V with spectrum* $\sigma(f) = A$ *and g be a linear operator on W with spectrum* $\sigma(g) = B$. *Then*

$$m_i(f \otimes I_W + I_V \otimes g) = \mu_i(A, B), \quad i = 1, \ldots, \min\{|A|, |B|\}.$$

Proof. It could be easily derived from the definitions that the spectrum of $f \otimes I_W + I_V \otimes g$ is the family

$$(a + b)_{(a,b)\in A\times B}.$$

Then, for $1 \leqslant i \leqslant \min\{|A|, |B|\}$, we have

$$\begin{aligned} m_i(f \otimes I_W + I_V \otimes g) &= |\{x \in A + B : \ |\{(a, b) \in A \times B : \ a + b = x\}| \geqslant i\}| \\ &= \mu_i(A, B). \qquad \square \end{aligned}$$

Lemma 4.3. *Let $e_1, \ldots, e_n$ be linearly independent vectors of the vector space V. Let $v_1, \ldots, v_t \in \langle e_1, \ldots, e_n \rangle$. Let $k \in \{1, \ldots, n\}$ and $r \in \{1, \ldots, t\}$ and denote by π the projection of $\langle e_1, \ldots, e_n \rangle$ onto $\langle e_{k+1}, \ldots, e_n \rangle$ along $\langle e_1, \ldots, e_k \rangle$. If $v_1, \ldots, v_t$ satisfy the following conditions*:
(1) $v_1, \ldots, v_r \in \langle e_1, \ldots, e_k \rangle$,
(2) $\pi(v_{r+1}), \ldots, \pi(v_t)$ *are linearly independent*,
then

$$\langle v_1, \ldots, v_t \rangle = \langle v_1, \ldots, v_r \rangle \oplus \langle v_{r+1}, \ldots, v_t \rangle.$$

Proof. Let $x \in \langle v_1, \ldots, v_r \rangle \cap \langle v_{r+1}, \ldots, v_t \rangle$. Then

$$x = \lambda_1 v_1 + \cdots + \lambda_k v_r = \gamma_{r+1} v_{r+1} + \cdots + \gamma_t v_t.$$

Then

$$0 = \pi(x) = \gamma_{r+1}\pi(v_{r+1}) + \cdots + \gamma_t \pi(v_t).$$

Therefore,

$$\gamma_{k+1} = \cdots = \gamma_t = 0,$$

and then $x = 0$. $\square$

Lemma 4.4. *Let p be the characteristic of $\mathbb{F}$ in the case of finite characteristic and ∞ if $\mathbb{F}$ has characteristic 0. Let u, v, t, q be positive integers satisfying*
(i) $v + q \leqslant u$,
(ii) $t \leqslant u$,
(iii) $t \leqslant q + 1$,
(iv) $u < p$.
Then the matrix over $\mathbb{F}$

$$B_{u,v,t,q} = \left[\binom{u-i+1}{v-i+j}\right]_{\substack{i=1,\ldots,t \\ j=1,\ldots,q+1}}$$

has rank t. We use the convention $\binom{u}{m} = 0$ if $m < 0$.

Proof. Let φ and ψ be maps from $\mathbb{Z}$ into $\mathbb{N} \cup \{0\}$ defined in the following way:

$$\varphi(t) = \begin{cases} 1 & \text{if } t \geqslant 0, \\ 0 & \text{if } t < 0, \end{cases} \qquad t \in \mathbb{Z},$$

$$\psi(t) = \begin{cases} t & \text{if } t > 0, \\ 1 & \text{if } t \leqslant 0, \end{cases} \qquad t \in \mathbb{Z}\ .$$

It is easy to check that $B_{u,v,t,q}$ is equivalent to

$$\left[\varphi(v-i+j)\frac{1}{(u-v-j+1)!\,\psi(v-i+j)!}\right]_{\substack{i=1,\ldots,t \\ j=1,\ldots,q+1}}. \tag{6}$$

Multiplying the column j of matrix (6) by $(v+j-1)!(u-v-j+1)!$ we can show that the former matrix (and then matrix $B_{u,v,t,q}$) is equivalent to

$$C_{v,t,q} = \left[\varphi(v-i+j)\frac{(v+j-1)!}{\psi(v-i+j)!}\right]_{\substack{i=1,\dots,t\\ j=1,\dots,q+1}}.$$

We are going to prove, by induction on t, that $C_{v,t,q}$ has rank t. If $t=1$, the result is obviously true. On the other hand, let J denote the $(q+1)\times(q+1)$-matrix, with the entries $(i,i+1)$ equal to 1, $i=1,\dots,q$, and the remaining entries equal to 0. We have

$$C_{v,t,q}(I_{q+1}-J) = \begin{bmatrix} 1 & 0 \\ * & A' \end{bmatrix},$$

where A' is equivalent to the matrix

$$C_{v,t-1,q-1} = \left[\varphi(v-i+j)\frac{(v+j-1)!}{\psi(v-i+j)!}\right]_{\substack{i=1,\dots,t-1\\ j=1,\dots,q}}.$$

Using, now, the induction hypothesis A' has rank equal to $t-1$. Then $C_{v,t,q}$ (which is equivalent to $B(u,v,t,q)$) has rank equal to t. $\square$

4.1. *Proof of main theorems*

Let $v \in V$ and $w \in W$. Let f be a linear operator on V and g a linear operator on W. Suppose that $\{v, f(v), \dots, f^{k-1}(v)\}$ is a basis of $\mathscr{C}_f(v)$ and $\{w, g(w), \dots, g^{r-1}(w)\}$ is a basis of $\mathscr{C}_g(w)$. Then, it is well known that

$$\mathscr{B} = \{f^i(v)\otimes g^j(w)\colon\ 0 \leqslant i \leqslant k-1,\ 0 \leqslant j \leqslant r-1\}$$

is a basis of $\mathscr{C}_f(v)\otimes\mathscr{C}_g(w)$. Let z be a vector of $\mathscr{C}_f(v)\otimes\mathscr{C}_g(w)$,

$$z = \sum_{i=0}^{k-1}\sum_{j=0}^{r-1}\gamma_{ij} f^i(v)\otimes g^j(w).$$

We say that $z \in \mathscr{C}_f(v)\otimes\mathscr{C}_g(w)$ has weight t if

$$t = \max\{i+j\colon\ 0 \leqslant i \leqslant k-1,\ 0 \leqslant j \leqslant r-1 \text{ and } \gamma_{ij} \neq 0\}.$$

Proof of Theorem 3.1. Let $v \in V$, $w \in W$ be such that $k = \dim\mathscr{C}_f(v) = \deg(P_f)$ and $r = \dim\mathscr{C}_g(w) = \deg(P_g)$. Let $s = k+r-1$. We are going to prove that we can extract a $(v\otimes w, f(v)\otimes w, \dots, f^{\ell-1}(v)\otimes w, f\otimes I_W + I_V\otimes g)$-nice independent set,

$$\mathscr{M} = \{(f\otimes I_W + I_V\otimes g)^b(f^m(v)\otimes w)\colon\ 0 \leqslant m \leqslant \ell-1,\ 0 \leqslant b \leqslant \min\{p-1, s-\ell\}\}$$

with all indices equal to $\min\{p, s-\ell+1\}$, from the family

$$((f \otimes I_W + I_V \otimes g)^b (f^m(v) \otimes w))_{\substack{b=0,\dots,s-1 \\ m=0,\dots,\ell-1}}. \tag{7}$$

Since for $0 \leqslant m \leqslant \ell - 1$ and $0 \leqslant b \leqslant \min\{p-1, s-\ell\}$ the tensor

$$z_{b,m} = (f \otimes I_W + I_V \otimes g)^b (f^m(v) \otimes w)$$

has weight $b + m$, the maximum weight of the tensors of $\mathscr{M}$ is

$$M_\ell = \min\{p + \ell - 2, s - 1\}.$$

For $u = 0, \dots, M_\ell$ denote by $\mathscr{S}_u$ the index set of the subset of the elements of $\mathscr{M}$ of weight u, i.e.

$$\begin{aligned}\mathscr{S}_u &= \{(b,m)\colon\ z_{b,m} \in \mathscr{M} \text{ and } b + m = u\} \\ &= \{(b,m) \in (\mathbb{N} \cup \{0\}) \times (\mathbb{N} \cup \{0\})\colon\ 0 \leqslant b \leqslant \min\{p-1, s-\ell\}, \\ &\qquad 0 \leqslant m \leqslant \ell - 1 \text{ and } b + m = u\}.\end{aligned}$$

Let

$$b_u = \max\{0, u - p + 1, u - s + \ell\} \quad \text{and} \quad d_u = \min\{u, \ell - 1\}.$$

Then we get from the former equalities,

$$\mathscr{S}_u = \{(u - m, m) \in (\mathbb{N} \cup \{0\}) \times (\mathbb{N} \cup \{0\})\colon\ b_u \leqslant m \leqslant d_u\}.$$

Let x_u be the cardinality of $\mathscr{S}_u$, i.e. $x_u = d_u - b_u + 1$.

It is easy to see that $\mathscr{M}$ is the disjoint union of the subsets indexed by the $\mathscr{S}_u$'s, i.e.

$$\mathscr{M} = \dot{\bigcup_{u=0}^{M_\ell}} \{z_{b,m}\colon\ (b,m) \in \mathscr{S}_u\}. \tag{8}$$

Claim 1. *The set* $\{z_{b,m}\colon\ (b,m) \in \mathscr{S}_u\}$ *is linearly independent.*

Let $\mathscr{B}_u$ be the set of tensors of weight u of the basis

$$\{f^i(v) \otimes g^j(w)\colon\ 0 \leqslant i \leqslant k-1,\ 0 \leqslant j \leqslant r-1\}.$$

Let π_u be the projection of $\mathscr{C}_f(v) \otimes \mathscr{C}_g(w)$ onto $\langle \mathscr{B}_u \rangle$ along $\bigoplus_{\gamma=0,\gamma\neq u}^{s-1} \langle \mathscr{B}_\gamma \rangle$. If we define

$$\zeta_u = \max\{0, u - r + 1\}$$

and

$$\tau_u = \min\{k - 1, u\},$$

then π_u is a projection onto the subspace spanned by

$$\begin{aligned}\mathscr{B}_u &= \{f^i(v) \otimes g^j(w)\colon\ 0 \leqslant i \leqslant k-1,\ 0 \leqslant j \leqslant r-1, i + j = u\} \\ &= \{f^i(v) \otimes g^{u-i}(w)\colon\ \zeta_u \leqslant i \leqslant \tau_u\}.\end{aligned}$$

By expanding $(f \otimes I_W + I_V \otimes g)^b$ we can easily see that for $u \in \{0, \ldots, M_\ell\}$, and $(b, m) \in \mathscr{S}_u$

$$z_{b,m} = \sum_{t=0}^{u-m} \binom{u-m}{t} f^{m+t}(v) \otimes g^{u-m-t}(w).$$

Then, since for $m + t \geqslant k$ or for $u - m - t \geqslant r$ the tensor

$$f^{m+t}(v) \otimes g^{u-m-t}(w)$$

has weight less than or equal to $u - 1$, we have

$$\pi_u(f^{m+t}(v) \otimes g^{u-m-t}(w)) = 0 \quad \text{if } t \geqslant k - m \ \text{ or } \ t \leqslant u - m - r.$$

Then

$$\pi_u(z_{b,m}) = \sum_{t=\max\{0,u-m-r+1\}}^{\min\{u-m,k-1-m\}} \binom{u-m}{t} f^{m+t}(v) \otimes g^{u-m-t}(w).$$

Let us order the projection onto $\langle \mathscr{B}_u \rangle$ of the elements indexed by $\mathscr{S}_u$ following the values of the second coordinate,

$$y_j = \pi_u(z_{u-j-b_u+1, j+b_u-1}), \quad j = 1, \ldots, x_u.$$

Claim 1 can be reformulated in the following way.

Claim 1′. *The tensors $y_1, \ldots, y_{x_u}$ are linearly independent.*

Proof of Claim 1′. Let $\{\theta_i : \ \zeta_u \leqslant i \leqslant \tau_u\}$ be the dual basis of the basis, $\mathscr{B}_u$, of $\langle \mathscr{B}_u \rangle$, i.e. $\theta_i(f^j(v) \otimes g^{u-j}(w)) = \delta_{ij}, \ \zeta_u \leqslant i, j \leqslant \tau_u$, where δ_{ij} is the Kronecker symbol.

We are going to split the proof of Claim 1′ in two cases.

Case 1: $\zeta_u \leqslant b_u$. Let $X_i = \theta_{i+b_u-1}, \ i = 1, \ldots, x_u$. Observe now that the matrix $(X_i(y_j))_{i,j=1,\ldots,x_u}$ is a lower triangular matrix with principal elements equal to 1. In fact, we have

$$\begin{aligned} X_i(y_j) &= \theta_{i+b_u-1}(\pi_u(z_{u-j-b_u+1, j+b_u-1})) \\ &= \theta_{i+b_u-1}\Bigg(\sum_{t=0}^{\tau_u-j-b_u+1} \binom{u-j-b_u+1}{t} f^{t+j+b_u-1}(v) \\ &\qquad \otimes g^{u-t-j-b_u+1}(w) \Bigg) \\ &= \sum_{t=0}^{\tau_u-j-b_u+1} \binom{u-j-b_u+1}{t} \theta_{i+b_u-1}\left(f^{t+j+b_u-1}(v) \right. \\ &\qquad \left. \otimes g^{u-t-j-b_u+1}(w) \right). \end{aligned}$$

Denote η_u the upper bound of the value allowed for t in the previous sum, i.e.

$$\eta_u = \tau_u - j - b_u + 1.$$

Then

$$X_i(y_j) = \begin{cases} 0 & \text{if } i - j \notin \{0, \ldots, \eta_u\}, \\ \binom{u - j - b_u + 1}{i - j} & \text{if } i - j \in \{0, \ldots, \eta_u\}. \end{cases}$$

We know from the definitions that

$$\tau_u \geqslant d_u.$$

Therefore

$$i \leqslant x_u = d_u - b_u + 1 \leqslant \tau_u - b_u + 1, \quad i = 1, \ldots, x_u.$$

Subtracting j in each side of the inequalities of the former expression, we obtain

$$i - j \leqslant d_u - b_u - j + 1 \leqslant \tau_u - j - b_u + 1.$$

Then, for $i = 1, \ldots, x_u$, we have that $i - j \notin \{0, \ldots, \eta_u\}$ if and only if $i < j$. Therefore

$$X_i(y_j) = \begin{cases} 0 & \text{if } j > i, \\ \binom{u - j - b_u + 1}{i - j} & \text{if } i \geqslant j. \end{cases}$$

Since, for $i = 1, \ldots, x_u$

$$X_i(y_i) = \binom{u - i - b_u + 1}{0} = 1,$$

we have proved that $(X_i(y_j))_{i,j=1,\ldots,x_u}$ is lower triangular with principal elements equal to 1. Thus, $y_1, \ldots, y_{x_u}$ is a linearly independent family.

Case 2: $\zeta_u > b_u$. Let $X_i = \theta_{i+\zeta_u-1}$, $i = 1, 2, \ldots, \tau_u - \zeta_u + 1$. Arguing in a similar way, we have used in case $\zeta_u \leqslant b_u$, we can prove that the (i, j)-entry of the matrix $(X_i(y_j))_{\substack{i=1,\ldots,\tau_u-\zeta_u+1 \\ j=1,\ldots,x_u}}$ whose columns are the coordinate vectors of $y_1, \ldots, y_{x_u}$ is

$$X_i(y_j) = \begin{cases} 0 & \text{if } i - j + (\zeta_u - b_u) < 0, \\ \binom{u - j - b_u + 1}{i - j + \zeta_u - b_u} & \text{if } i - j + (\zeta_u - b_u) \geqslant 0. \end{cases}$$

It is now easy to see that

$$(X_i(y_j))_{\substack{i=1,\ldots,\tau_u-\zeta_u+1 \\ j=1,\ldots,x_u}} = (B_{u-b_u,\zeta_u-b_u,x_u,\tau_u-\zeta_u})^{\mathrm{T}}.$$

We can easily see that the conditions for application of Lemma 4.4 are fulfilled. Then, $y_1, \ldots, y_{x_u}$ is a linearly independent family. $\square$

Proof of Theorem 3.1 (*continued*). Now we see from (8) that

$$\langle \mathscr{M} \rangle = \sum_{u=0}^{M_\ell} \langle z_{b,m} \colon\ (b,m) \in \mathscr{S}_u \rangle.$$

Using now Lemma 4.3 and Claim 1′, we get from the former equality

$$\langle \mathscr{M} \rangle = \bigoplus_{u=0}^{M_\ell} \langle z_{b,m} \colon\ (b,m) \in \mathscr{S}_u \rangle.$$

Then $\mathscr{M}$ is linearly independent, therefore a $(v \otimes w, f(v) \otimes w, \ldots, f^{\ell-1}(v) \otimes w, f \otimes I_W + I_V \otimes g)$-nice independent set with all indices equal to $\min\{p, s-\ell+1\}$.

We can now use Corollary 2.8 to get

$$\sum_{i=1}^{\ell} \deg(\alpha_{f \otimes I_W + I_V \otimes g, mn-i+1}) \geqslant \ell \min\{p, \deg(P_f) + \deg(P_g) - \ell\}. \qquad \square$$

We are now going to prove Theorem 3.2.

Proof of Theorem 3.2. Let $|A| = n$ and $|B| = m$. Let f be a diagonalizable linear operator whose spectrum is A and g be a diagonalizable linear operator whose spectrum is B. Then $f \otimes I + I \otimes g$ is diagonalizable with spectrum $A + B$. Using Propositions 4.1 and 4.2 we have

$$\mu_1 + \cdots + \mu_j = \sum_{i=1}^{j} \deg(\alpha_{f \otimes I + I \otimes g, mn-i+1}), \qquad j = 1, \ldots, \min\{|A|, |B|\}.$$

Then using Theorem 3.1 we can conclude that

$$\sum_{i=1}^{\ell} \mu_i \geqslant \ell \min\{p, |A| + |B| - \ell\}. \qquad \square$$

Remark. If x is an integer, denote by $\overline{x}$ the element of $\mathbb{F}$, $x1_{\mathbb{F}}$. Suppose A and B are arithmetic progressions of the same rate. Then $p \geqslant |A|$ and $p \geqslant |B|$. Assume that $|A| \geqslant |B|$. Let $s = |A| + |B| - 1$. Let $A' = \{\overline{0}, \overline{1}, \ldots, \overline{|A|-1}\}$ and $B' = \{\overline{0}, \overline{1}, \ldots, \overline{|B|-1}\}$. It is easy to see that

$$\mu_i(A, B) = \mu_i(A', B'), \quad i \in \mathbb{N}.$$

For $\overline{x} \in A' + B' = \{\overline{0}, \overline{1}, \ldots, \overline{\min\{p-1, s-1\}}\}$, we have:

- If $p \leqslant s-1$,

$$\nu_{\overline{x}}(A', B') = \begin{cases} s-p+1 & \text{if } \overline{x} \in \{\overline{0}, \ldots, \overline{s-p-1}\}, \\ x+1 & \text{if } \overline{x} \in \{\overline{s-p}, \ldots, \overline{|B|-1}\}, \\ |B| & \text{if } \overline{x} \in \{\overline{|B|}, \ldots, \overline{|A|-1}\}, \\ s-x & \text{if } \overline{x} \in \{\overline{|A|}, \ldots, \overline{p-1}\}. \end{cases}$$

- If $p > s-1$,

$$\nu_{\overline{x}}(A', B') = \begin{cases} x+1 & \text{if } \overline{x} \in \{\overline{0}, \dots, \overline{|B|-1}\}, \\ |B| & \text{if } \overline{x} \in \{\overline{|B|}, \dots, \overline{|A|-1}\}, \\ s-x & \text{if } \overline{x} \in \{\overline{|A|}, \dots, \overline{s-1}\}. \end{cases}$$

Then, for $i = 1, \dots, \min\{|A|, |B|\} = |B|$, we have

$$\begin{aligned} \mu_i(A, B) &= \mu_i(A', B') \\ &= |\{\overline{x} \in A' + B' : \ \nu_{\overline{x}}(A', B') \geqslant i\}| \\ &= \begin{cases} p & \text{if } 1 \leqslant i \leqslant s-p+1, \\ s-2i+2 & \text{if } \max\{1, s-p+2\} \leqslant i \leqslant |B|. \end{cases} \end{aligned}$$

It follows that, for $\ell = 1, 2, \dots, |B|$,

$$\sum_{i=1}^{\ell} \mu_i = \begin{cases} \ell p & \text{if } \ell \leqslant s-p+1, \\ \ell(s-\ell+1) & \text{if } \ell \geqslant s-p+2. \end{cases}$$

and equality holds in Theorem 3.2.

References

[1] A. Cauchy, Recherche sur les nombres, J. École Polytech. 9 (1813) 99–116.
[2] H. Davenport, On the addition of residue classes, J. London Math. Soc. 10 (1935) 30–32.
[3] H. Davenport, A historical note, J. London Math. Soc. 22 (1947) 100–101.
[4] J.A. Dias da Silva, Y.O. Hamidoune, A note on the minimal polynomial of the Kronecker sum of two linear operators, Linear Algebra Appl. 141 (1990) 283–287.
[5] M.B. Nathanson, Additive Number Theory: Inverse Problems and the Geometry of Sumsets, Springer, New York, 1996.
[6] J.M. Pollard, A generalization of a theorem of Cauchy and Davenport, J. London Math. Soc. 8 (1974) 460–462.
[7] I. Zaballa, Matrices with prescribed rows and invariant factors, Linear Algebra Appl. 87 (1987) 113–146.
[8] I. Zaballa, Interlacing inequalities and control theory, Linear Algebra Appl. 101 (1988) 9–31.
[9] I. Zaballa, Controllability and Hermite indices of matrix pairs, Int. J. Control 68 (1) (1997) 61–86.