ARS MATHEMATICA CONTEMPORANEA

ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 16 (2019) 49–58
https://doi.org/10.26493/1855-3974.1547.454
(Also available at http://amc-journal.eu)

On the parameters of intertwining codes*

Stephen P. Glasby , Cheryl E. Praeger

Centre for Mathematics of Symmetry and Computation,
University of Western Australia, 35 Stirling Highway, Crawley 6009, Australia

Received 8 December 2017, accepted 6 March 2018, published online 13 September 2018

Abstract

Let F be a field and let $F^{r\times s}$ denote the space of $r\times s$ matrices over F. Given equinumerous subsets $\mathcal{A}=\{A_i \mid i\in I\}\subseteq F^{r\times r}$ and $\mathcal{B}=\{B_i \mid i\in I\}\subseteq F^{s\times s}$ we call the subspace $C(\mathcal{A},\mathcal{B}) := \{X\in F^{r\times s} \mid A_iX = XB_i \text{ for } i\in I\}$ an *intertwining code*. We show that if $C(\mathcal{A},\mathcal{B})\neq\{0\}$, then for each $i\in I$, the characteristic polynomials of A_i and B_i and share a nontrivial factor. We give an exact formula for $k=\dim(C(\mathcal{A},\mathcal{B}))$ and give upper and lower bounds. This generalizes previous work. Finally we construct intertwining codes with large minimum distance when the field is not 'too small'. We give examples of codes where $d=rs/k=1/R$ is large where the minimum distance, dimension, and rate of the linear code $C(\mathcal{A},\mathcal{B})$ are denoted by d, k, and $R=k/rs$, respectively.

Keywords: Linear code, dimension, distance.

Math. Subj. Class.: 94B65, 60C05

1 Introduction

Let F be a field and let $F^{r\times s}$ denote the space of $r\times s$ matrices over F. Given equinumerous subsets $\mathcal{A}=\{A_i \mid i\in I\}\subseteq F^{r\times r}$ and $\mathcal{B}=\{B_i \mid i\in I\}\subseteq F^{s\times s}$ we call the subspace $C(\mathcal{A},\mathcal{B}) := \{X\in F^{r\times s} \mid A_iX = XB_i \text{ for } i\in I\}$ an *intertwining code*. The parameters of this linear code are denoted $[n,k,d]$ where $n=rs$, $k:=\dim(C(\mathcal{A},B))$ and d is the *minimum distance* of $C(\mathcal{A},\mathcal{B})$. Given $u,v\in F^n$ the *Hamming distance* $d(u,v)=|\{i \mid u_i\neq v_i\}|$ is the number of different coordinate entries, and a subspace $C\leqslant F^n$ has minimal (Hamming) distance $d(C) := \min\{d(u,v) \mid u\neq v\}$ which equals $\min\{d(0,w) \mid w\in V \text{ where } w\neq 0\}$. If $|I|=1$ we write $C(A,B)$ instead

*The authors acknowledge the contribution of Robin Chapman who emailed us in August 2016 a proof of Theorem 2.8. His formula for $\dim(C(N_\lambda,N_\mu))$ involved a double sum which can be reduced to a single sum using Theorem 3.1. The authors gratefully acknowledge the support of the Australian Research Council Discovery Grant DP160102323.

E-mail addresses: Stephen.Glasby@uwa.edu.au (Stephen P. Glasby), Cheryl.Praeger@uwa.edu.au (Cheryl E. Praeger)

of $C(\{A\},\{B\})$. Centralizer codes [1] have the form $C(A,A)$ and twisted centralizer codes [2, 3] have the form $C(A,\alpha A)$ where $A \in F^{r\times s}$ and $\alpha \in F$. Intertwining codes $C(A,B)$ are more general still, so our dimension formula (Theorem 2.8) has particularly wide applicability. Furthermore, the greater abundance of intertwining codes turns out to help us construct intertwining codes with large minimum distance, cf. Theorem 4.3 and [3, Theorem 3.2]. Intertwining codes have the advantage of a short description, and fast matrix multiplication algorithms give rise to efficient syndrome computations which, in turn, may be used for decoding as described in [3, §3].

Given representations $g_i \mapsto A_i$ and $g_i \mapsto B_i$ a group algebra $F\langle g_i \mid i \in I\rangle$, elements of $C(\mathcal{A},\mathcal{B})$ are homomorphisms between the associated modules. Hence Lemma 2.2 generalizes the fact that irreducible representations with distinct characters are inequivalent.

An exact formula for $k := \dim(C(A,B))$ is given in Theorems 2.9 and 2.8 of Section 2. The formula for k is simplified by an identity involving partitions proved in Section 4. Simpler upper and lower bounds for k are given in Section 5. In Theorem 4.3 in Section 4, we give an algorithm to construct A, B for which the minimum distance is $d(C(A,B)) = \lfloor r/k\rfloor s$. These examples have $dR \leqslant 1$ where $R = \frac{k}{rs}$ is the rate of $C(A,B)$.

Corollary 4.4 to Theorem 4.3 shows that there exist matrices $A \in F^{r\times r}$ and $B \in F^{s\times s}$ such that the intertwining code $C(A,B)$ has dimension $\min\{r,s\}$ and minimum distance $\max\{r,s\}$. We wonder how much this result can be improved.

2 A formula for $\dim_F(C(\mathcal{A},\mathcal{B}))$

Throughout this section $\mathcal{A} = \{A_i \mid i \in I\} \subset F^{r\times r}$ and $\mathcal{B} = \{B_i \mid i \in I\} \subset F^{s\times s}$ for a field F. The idea underlying this section is to use the Jordan form over the algebraic closure $\overline{F}$ of F to compute $\dim_F(C(\mathcal{A},\mathcal{B}))$. To implement this idea we must simultaneously conjugate each $A_i \in \mathcal{A}$, and each $B_i \in \mathcal{B}$, into Jordan form. This is always possible when $|I| = 1$.

Let $\mathrm{GL}(r,F)$ denote the general linear group of $r \times r$ invertible matrices over F. An ordered pair $(R,S) \in \mathrm{GL}(r,F) \times \mathrm{GL}(s,F)$ acts on $F^{r\times s}$ via $X^{(R,S)} = R^{-1}XS$. Clearly

$$\begin{aligned}(X^{(R_1,S_1)})^{(R_2,S_2)} &= X^{(R_1R_2,S_1S_2)},\\ (XS_1)^{(R,S)} &= X^{(R,S)}S_1^S, \quad \text{and}\\ (R_1X)^{(R,S)} &= R_1^R X^{(R,S)}.\end{aligned}$$

Lemma 2.1. *If $(R,S) \in \mathrm{GL}(r,F) \times \mathrm{GL}(s,F)$, then*

$$C(\mathcal{A},\mathcal{B})^{(R,S)} = R^{-1}C(\mathcal{A},\mathcal{B})S = C(\mathcal{A}^R,\mathcal{B}^S)$$

where $\mathcal{A}^R := \{R^{-1}A_iR \mid i \in I\}$ and $\mathcal{B}^S := \{S^{-1}B_iS \mid i \in I\}$.

Proof. For each $i \in I$, the equation $A_iX = XB_i$ is equivalent to

$$A_i^R X^{(R,S)} = (A_iX)^{(R,S)} = (XB_i)^{(R,S)} = X^{(R,S)}B_i^S. \qquad \square$$

Let $c_A(t) = \det(tI - A)$ be the characteristic polynomial of A.

Lemma 2.2. *If $C(\mathcal{A},\mathcal{B}) \neq \{0\}$, then $\gcd(c_{A_i}(t), c_{B_i}(t)) \neq 1$ for all $i \in I$.*

Proof. Suppose that for some $i \in I$ we have $\gcd(c_{A_i}(t), c_{B_i}(t)) = 1$. Then there exist polynomials $f(t), g(t)$ such that $f(t)c_{A_i}(t) + g(t)c_{B_i}(t) = 1$. Evaluating this equation at $t = B_i$, and noting that $c_{B_i}(B_i) = 0$, shows $f(B_i)c_{A_i}(B_i) = I$. Hence $c_{A_i}(B_i)$ is invertible. For $X \in C(\mathcal{A}, \mathcal{B})$, we have $A_i X = X B_i$. Thus

$$\left(\sum_{k \geqslant 0} \alpha_k A_i^k \right) X = X \left(\sum_{k \geqslant 0} \alpha_k B_i^k \right),$$

for all $\alpha_k \in F$, and therefore $c_{A_i}(A_i)X = X c_{A_i}(B_i)$. Since $c_{A_i}(A_i) = 0$, post-multiplying by $c_{A_i}(B_i)^{-1}$ shows that $X = 0$, and hence $C(\mathcal{A}, \mathcal{B}) = \{0\}$. □

Henceforth when we wish to emphasize the field F, we write $C_F(\mathcal{A}, \mathcal{B})$. Lemma 3.1 of [2], in essence, says $C_{\overline{F}}(\mathcal{A}, \mathcal{B}) = C_F(\mathcal{A}, \mathcal{B}) \otimes \overline{F}$. This immediately yields Lemma 2.3.

Lemma 2.3. *If K is an extension field of F, then $\dim_F(C_F(\mathcal{A}, \mathcal{B})) = \dim_K(C_K(\mathcal{A}, \mathcal{B}))$. In particular, $\dim_F(C_F(\mathcal{A}, \mathcal{B})) = \dim_{\overline{F}}(C_{\overline{F}}(\mathcal{A}, \mathcal{B}))$ where $\overline{F}$ is the algebraic closure of F.*

Lemma 2.3 allows us to assume that F is algebraically closed, which we shall do for the rest of this section. Given $A \in F^{r \times r}$ and $B \in F^{s \times s}$ define $A \oplus B$ to be the block diagonal matrix $\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$, and define $\mathcal{A} \oplus \mathcal{B}$ to be $\{A_i \oplus B_i \mid i \in I\} \subseteq F^{(r+s)\times(r+s)}$.

Lemma 2.4. *If $\mathcal{A}_1 \subseteq F^{r_1 \times r_1}, \ldots, \mathcal{A}_m \subseteq F^{r_m \times r_m}$ and $\mathcal{B}_1 \subseteq F^{s_1 \times s_1}, \ldots, \mathcal{B}_n \subseteq F^{s_n \times s_n}$ are subsets, all with the same finite cardinality, then*

$$C\left(\bigoplus_{i=1}^{m} \mathcal{A}_i, \bigoplus_{j=1}^{n} \mathcal{B}_j \right) \cong \bigoplus_{i=1}^{m} \bigoplus_{j=1}^{n} C(\mathcal{A}_i, \mathcal{B}_j).$$

Proof. Write $X = (X_{ij})$ as a block matrix where X_{ij} has size $r_i \times s_j$. The condition $X \in C\left(\bigoplus_{i=1}^{m} \mathcal{A}_i, \bigoplus_{j=1}^{n} \mathcal{B}_j\right)$ is equivalent to $X_{ij} \in C(\mathcal{A}_i, \mathcal{B}_j)$ for each i, j. □

Corollary 2.5. *Suppose that $\mathcal{A}_1, \ldots, \mathcal{A}_m$ and $\mathcal{B}_1, \ldots, \mathcal{B}_n$ are as in Lemma 2.4, and suppose that for $i \neq j$, the characteristic polynomials of matrices in $\mathcal{A}_i$ are coprime to the characteristic polynomials of matrices in $\mathcal{B}_j$. Then*

$$C\left(\bigoplus_{i=1}^{m} \mathcal{A}_i, \bigoplus_{j=1}^{n} \mathcal{B}_j \right) \cong \bigoplus_{i=1}^{\min\{m,n\}} C(\mathcal{A}_i, \mathcal{B}_i).$$

Proof. Use Lemma 2.4, and note that $C(\mathcal{A}_i, \mathcal{B}_j) = \{0\}$ for $i \neq j$ by Lemma 2.2. □

Remark 2.6. Let F be a finite field. Standard arguments, for example [6, p. 168], can be used to relate $\dim_{\overline{F}}(C_{\overline{F}}(\mathcal{A}, \mathcal{B}))$ to data computed over F. This remark and Remark 2.10 explain the details. Let $p_1(t), p_2(t), \ldots$ enumerate the (monic) irreducible polynomials over F and write $c_A(t) = \prod_{i \geqslant 1} p_i(t)^{k_i}$ and $c_B(t) = \prod_{i \geqslant 1} p_i(t)^{\ell_i}$, respectively. This gives rise to the A-invariant primary decomposition $F^r = \bigoplus_{i \geqslant 1} \ker(p_i(A)^{k_i})$, and the B-invariant decomposition $F^s = \bigoplus_{i \geqslant 1} \ker(p_i(A)^{\ell_i})$. Let A_i be the restriction of A to $\ker(p_i(A)^{k_i})$ and B_i the restriction of B to $\ker(p_i(A)^{\ell_i})$. Corollary 2.5 shows that $\dim(C(A, B)) = \sum_{i \geqslant 1} \dim(C(A_i, B_i))$. The second ingredient involves partitions and is described in Remark 2.10.

It is straightforward to see that $C(\mathcal{A},\mathcal{B}) = \bigcap_{i\in I} C(A_i,B_i)$ where $C(A_i,B_i)$ means $C(\{A_i\},\{B_i\})$. Recall that a matrix $A \in F^{r\times r}$ is *nilpotent* if and only if $A^r = 0$. We say that A is *α-potent*, where $\alpha \in F$, if $(A-\alpha I)^r = 0$. The following lemma and theorem reduce our deliberations from α-potent matrices to nilpotent matrices. For $\mathcal{A} = \{A_i \mid i \in I\} \subseteq F^{r\times r}$, let $\mathcal{A} - \alpha I_r$ denote the set $\{A_i - \alpha I_r \mid i \in I\}$.

Lemma 2.7. *If $\mathcal{A} \subseteq F^{r\times r}$, $\mathcal{B} \subseteq F^{s\times s}$ and $\alpha \in F$, then $C(\mathcal{A},\mathcal{B}) = C(\mathcal{A}-\alpha I_r, \mathcal{B}-\alpha I_s)$.*

Proof. For $i \in I$, $A_iX = XB_i$ holds if and only if $(A_i-\alpha I_r)X = X(B_i-\alpha I_s)$. □

A *partition* λ of r, written $\lambda \vdash r$, is a sequence $\lambda = (\lambda_1,\lambda_2,\dots)$ of integers satisfying

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant 0 \quad \text{and} \quad \lambda_1+\lambda_2+\cdots = r.$$

We call λ_i the ith part of λ, and we usually omit parts of size zero. Let N_r be the $r\times r$ nilpotent matrix with all entries 0 except for an entry 1 in position $(i,i+1)$ for $1 \leqslant i < r$. Let $N_\lambda = \bigoplus N_{\lambda_i}$ where $\lambda \vdash r$. Every nilpotent $r\times r$ matrix is conjugate to some N_λ for a unique $\lambda \vdash r$. Furthermore, if an $r\times r$ matrix R has eigenvalues $\rho_1,\dots,\rho_m$ and associated generalized eigenspaces of dimensions $r_1,\dots,r_m$ where $r_1+\cdots+r_m = r$, then R has Jordan form $\bigoplus_{i=1}^m (\rho_i I_{r_i} + N_{\lambda_i})$ where λ_i is a *partition* of r_i (not a part of a partition).

Theorem 2.8. *Suppose $A \in F^{r\times r}$, $B \in F^{s\times s}$ and $\gcd(c_A(t),c_B(t))$ has distinct roots $\zeta_1,\dots,\zeta_m$ in $\overline{F}$. Suppose that the sizes of the Jordan blocks of A associated with the generalized ζ_i-eigenspace of A determine a partition α_i, and the sizes of the Jordan blocks of B associated with the generalized ζ_i-eigenspace of B determine a partition β_i. Then*

$$\dim(C(A,B)) = \sum_{i=1}^m \dim(C(N_{\alpha_i},N_{\beta_i})).$$

Proof. By Lemma 2.3 we may assume that $F = \overline{F}$. Let A_i be the restriction of A to its generalized ζ_i-eigenspace $\{v \mid v(A-\zeta_i I)^k$ for some $k \geqslant 0\}$. Then A_i is ζ_i-potent, and so determines a partition α_i. Similarly, the restriction B_i of B to the ζ_i-eigenspace determines a partition β_i. By Corollary 2.5 and Lemma 2.7, we have

$$\dim(C(A,B)) = \sum_{i=1}^m \dim(C(A_i,B_i)) = \sum_{i=1}^m \dim(C(N_{\alpha_i},N_{\beta_i})). \qquad \square$$

Theorem 2.9. *Given partitions λ of r and μ of s, the dimension of $C(N_\lambda,N_\mu)$ equals*

$$\dim(C(N_\lambda,N_\mu)) = \sum_{i\geqslant 1}\sum_{j\geqslant 1} \min\{\lambda_i,\mu_j\}.$$

Proof. As $\lambda \vdash r$ and $\mu \vdash s$, we have $\sum_{i\geqslant 1}\lambda_i = r$ and $\sum_{j\geqslant 1}\mu_j = s$. Lemma 2.4 shows that $C(N_\lambda,N_\mu) \cong \bigoplus_{i\geqslant 1}\bigoplus_{j\geqslant 1} C(N_{\lambda_i},N_{\mu_j})$. Taking dimensions, it suffices to show $\dim(C(N_{\lambda_i},N_{\mu_j})) = \min\{\lambda_i,\mu_j\}$. This can be shown by solving $N_{\lambda_i}X = XN_{\mu_j}$ for X and counting the number of free variables. Alternatively, F^{λ_i} is a uniserial $\langle N_{\lambda_i}\rangle$-module with 1-dimensional quotient modules, and similarly for F^{λ_j}. As their largest common quotient module is $F^{\min\{\lambda_i,\lambda_j\}}$, we have $\dim(C(N_{\lambda_i},N_{\lambda_j})) = \min\{\lambda_i,\lambda_j\}$. □

Remark 2.10. Suppose $|F| = q$ is finite. Following on from Remark 2.6 it suffices to consider the case where $c_A(t) = p(t)^{r/d}$, $c_B(t) = p(t)^{s/d}$, where $p(t)$ is irreducible over F of degree d. The field $K := F[t]/(p(t))$ has order q^d. In this case the structure of the modules $F^r = K^{r/d}$ and $F^s = K^{s/d}$ is determined by partitions $\lambda \vdash r/d$ and $\mu \vdash s/d$. It turns out that A is conjugate (see below) to $N_{\lambda,p} := \mathrm{diag}(N_{\lambda_1,p}, N_{\lambda_2,p}, \dots) \in F^{r\times r}$ where

$$N_{m,p} = \begin{pmatrix} C(p) & I & & \\ & \ddots & \ddots & \\ & & C(p) & I \\ & & & C(p) \end{pmatrix} \in F^{dm\times dm}$$

and $C(p) \in F^{d\times d}$ is the companion matrix of $p(t)$. Now $C(p)$ is conjugate in $\mathrm{GL}(d, K)$ to $\mathrm{diag}(\zeta_1, \dots, \zeta_d)$ where $\zeta_1, \dots, \zeta_d$ are the (distinct) roots of $p(t)$ in K. It follows from Theorems 2.9 and 2.8 that

$$\dim(C(A, B)) = \dim(C(N_{\lambda,p}, N_{\mu,p})) = d\sum_{i\geqslant 1}\sum_{j\geqslant 1}\min\{\lambda_i, \mu_j\}. \tag{2.1}$$

As an example, suppose A is cyclic and $c_A(t) = p(t)^3$ where $d = \deg(p) = 3$. In this case $r = 9$ and $\lambda = (3)$. Write $p(t) = t^3 + p_2t^2 + p_1t + p_0 = (t - \zeta_1)(t - \zeta_2)(t - \zeta_3)$. Then A is conjugate in $\mathrm{GL}(9, F)$ by [5] to

$$\begin{pmatrix} C(p) & N & 0 \\ 0 & C(p) & N \\ 0 & 0 & C(p) \end{pmatrix}$$

where

$$C(p) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -p_0 & -p_1 & -p_2 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As $p(t)$ is separable, [5, Theorem 1] implies that A is conjugate in $\mathrm{GL}(9, F)$ to

$$\begin{pmatrix} C(p) & I & 0 \\ 0 & C(p) & I \\ 0 & 0 & C(p) \end{pmatrix}.$$

Hence A is conjugate in $\mathrm{GL}(9, K)$ to

$$\begin{pmatrix} D(\zeta_1) & 0 & 0 \\ 0 & D(\zeta_2) & 0 \\ 0 & 0 & D(\zeta_3) \end{pmatrix}$$

where

$$D(\zeta) = \begin{pmatrix} \zeta & 1 & 0 \\ 0 & \zeta & 1 \\ 0 & 0 & \zeta \end{pmatrix}.$$

This explains the factor of $d = \deg(p(t))$ in equation (2.1) and relates the generalized Jordan form of A over F to the Jordan form of A over K.

3 Conjugate partitions

In this section we simplify the formula in Theorem 2.9 for $\dim(C(N_\lambda, N_\mu))$. We prove an identity in Lemma 3.2 involving partitions which replaces multiple sums by a single sum. In order to state the simpler dimension formula we need to define 'conjugate partitions'. The *conjugate* of $\lambda \vdash r$ is the partition $\lambda' = (\lambda'_1, \lambda'_2, \dots)$ of r whose parts satisfy $\lambda'_i = |\{j \mid \lambda_j \geqslant i\}|$, for each i. The Young diagram of λ', is obtained from that of λ by swapping rows and columns as shown in Figure 1.

Figure 1: Young diagrams for $\lambda = (5, 3, 3, 1)$ and $\lambda' = (4, 3, 3, 1, 1)$.

For the following result, note that the number of nonzero λ_i is λ'_1, and $r = \sum_{i=1}^{\lambda'_1} \lambda_i$.

Theorem 3.1. *Given partitions λ of r and μ of s, the dimension of $C(N_\lambda, N_\mu)$ equals*

$$\dim(C(N_\lambda, N_\mu)) = \sum_{i \geqslant 1} \lambda'_i \mu'_i = \sum_{i=1}^{\min\{\lambda_1, \mu_1\}} \lambda'_i \mu'_i.$$

To prove Theorem 3.1 we need a technical lemma which we have not been able to find in the literature, see [4]. Lemma 3.2 below says $\sum_{i \geqslant 1} \lambda_i = \sum_{i \geqslant 1} \lambda'_i$ when $k = 1$. We only need the case $k = 2$ for the proof of Theorem 3.1, however, the argument for $k > 2$ is not much harder.

Lemma 3.2. *If $\lambda, \mu, \dots, \omega$ are partitions and $\lambda', \mu', \dots, \omega'$ are their conjugates, then*

$$\sum_{i=1}^{\lambda'_1} \sum_{j=1}^{\mu'_1} \cdots \sum_{k=1}^{\omega'_1} \min\{\lambda_i, \mu_j, \dots, \omega_k\} = \sum_{i=1}^{\min\{\lambda_1, \mu_1, \dots, \omega_1\}} \lambda'_i \mu'_i \cdots \omega'_i. \tag{3.1}$$

Proof. By permuting the partitions $\lambda, \mu, \dots, \omega$ if necessary, we can assume that

$$\lambda_1 \leqslant \mu_1 \leqslant \cdots \leqslant \omega_1.$$

If $\lambda_1 = 0$, then $\lambda \vdash 0$ and both sides of (3.1) are zero. If $\lambda_1 = 1$, then

$$\mathrm{LHS}(1) = \sum_{i=1}^{\lambda'_1} \sum_{j=1}^{\mu'_1} \cdots \sum_{k=1}^{\omega'_1} 1 = \lambda'_1 \mu'_1 \cdots \omega'_1 = \sum_{i=1}^{\min\{\lambda_1, \mu_1, \dots, \omega_1\}} \lambda'_i \mu'_i \cdots \omega'_i = \mathrm{RHS}(1).$$

Suppose now that $\lambda_1 > 1$. We use induction on λ_1. Let $\widehat{\lambda}$ be the partition of $(\sum_{i \geqslant 1} \lambda_i) - \lambda'_1$ obtained by deleting the first column of the Young diagram of λ. Since $1 < \mu_1 \leqslant \cdots \leqslant \omega_1$, we define $\widehat{\mu}, \dots, \widehat{\omega}$ similarly. It is clear that $\widehat{\lambda}_i = \lambda_i - 1$ for $1 \leqslant i \leqslant \lambda'_1$ and $\widehat{\lambda}'_i = \lambda'_{i+1}$ for $i \geqslant 1$, and similarly for $\widehat{\mu}, \dots, \widehat{\omega}$. As $\widehat{\lambda}_1 < \lambda_1$, induction shows

$$\sum_{i=1}^{\widehat{\lambda}'_1} \sum_{j=1}^{\widehat{\mu}'_1} \cdots \sum_{k=1}^{\widehat{\omega}'_1} \min\{\widehat{\lambda}_i, \widehat{\mu}_j, \dots, \widehat{\omega}_k\} = \sum_{i=1}^{\min\{\widehat{\lambda}_1, \widehat{\mu_1}, \dots, \widehat{\omega}_1\}} \widehat{\lambda}'_i \widehat{\mu}'_i \cdots \widehat{\omega}'_i.$$

Note that $\widehat{\lambda}_i = 0$ for each $i \in [\widehat{\lambda}'_1 + 1, \lambda'_1]$ since $\widehat{\lambda}'_1 = \lambda'_2$, so the upper limit $\widehat{\lambda}'_1$ of the sum $\sum_{i=1}^{\widehat{\lambda}'_1}$ can be replaced by λ'_1. Similarly, the upper limits $\widehat{\mu}'_1, \ldots, \widehat{\omega}'_1$ can be replaced by $\mu'_1, \ldots, \omega'_1$. Hence, since $\widehat{\lambda}_i = \lambda'_i - 1, \ldots, \widehat{\omega}_i = \omega'_i - 1$, we have

$$\sum_{i=1}^{\lambda'_1} \sum_{j=1}^{\mu'_1} \cdots \sum_{k=1}^{\omega'_1} \min\{\lambda_i - 1, \mu_j - 1, \ldots, \omega_k - 1\} = \sum_{i=1}^{\min\{\lambda_1-1, \mu_1-1, \ldots, \omega_1-1\}} \lambda'_{i+1}\mu'_{i+1} \cdots \omega'_{i+1}.$$

Re-indexing the right sum, and using $\sum_{i=1}^{\lambda'_1} \sum_{j=1}^{\mu'_1} \cdots \sum_{k=1}^{\omega'_1} (-1) = -\lambda'_1\mu'_1 \cdots \omega'_1$ gives

$$-\lambda'_1\mu'_1 \cdots \omega'_1 + \sum_{i=1}^{\lambda'_1} \sum_{j=1}^{\mu'_1} \cdots \sum_{k=1}^{\omega'_1} \min\{\lambda_i, \mu_j, \ldots, \omega_k\} = \sum_{i=2}^{\min\{\lambda_1, \mu_1, \ldots, \omega_1\}} \lambda'_i\mu'_i \cdots \omega'_i.$$

Adding $\lambda'_1\mu'_1 \cdots \omega'_1$ to both sides completes the inductive proof of (3.1). □

Proof of Theorem 3.1. Apply Theorem 2.9 and Lemma 3.2 with $k = 2$. □

4 Minimum distances

In Section 2 a formula is given for $k := \dim(C(\mathcal{A}, B))$; where we suppress mention of the field F in our notation. In this section we choose $\mathcal{A}$ and $\mathcal{B}$ to maximize the value of the minimum distance $d := d(C(\mathcal{A}, \mathcal{B}))$ as a function of k. We focus on the case when $|\mathcal{A}| = |\mathcal{B}| = 1$. The action of $\mathrm{GL}(r, F) \times \mathrm{GL}(s, F)$ of $C(A, B)$ fixes $k = \dim(C(A, B))$ but can change d wildly, e.g. from 1 to rs as setting $k = 1$ in Theorem 4.3 illustrates.

Let E_{ij} denote the $r \times s$ matrix with all entries 0, except the (i, j) entry which is 1.

Lemma 4.1. *Suppose $r, s, k \in \mathbb{Z}$ where $1 \leqslant k \leqslant \min\{r, s\}$, and suppose F is a field with $|F| \geqslant k + \min\{1, r-k\} + \min\{1, s-k\}$. Fix pairwise distinct scalars $\zeta_1, \ldots, \zeta_k, \alpha, \beta \in F$ and set*

$$A_0 := \mathrm{diag}(\zeta_1, \ldots, \zeta_k, \alpha, \ldots, \alpha) \in F^{r\times r} \quad \text{and}$$
$$B_0 := \mathrm{diag}(\zeta_1, \ldots, \zeta_k, \beta, \ldots, \beta) \in F^{s\times s}.$$

Then $C(A_0, B_0) = \langle E_{11}, \ldots, E_{kk}\rangle$ has dimension k and minimum distance 1.

Proof. Note first that if $k = \min\{r, s\}$, then A_0 has no αs, or B_0 has no βs. Thus the assumption $|F| \geqslant k + \min\{1, r-k\} + \min\{1, s-k\}$ ensures that distinct scalars $\zeta_1, \ldots, \zeta_k, \alpha, \beta$ in F exist. Using a direct calculation of $C(A_0, B_0)$, or Corollary 2.5, shows that $C(A_0, B_0) = \langle E_{11}, \ldots, E_{kk}\rangle$. Since $d(0, E_{11}) = 1$, we have $d(C(A_0, B_0)) = 1$. □

We now seek $R \in \mathrm{GL}(r, F)$ and $S \in \mathrm{GL}(s, F)$ such that $R^{-1}\langle E_{11}, \ldots, E_{kk}\rangle S$ has large minimum distance. For brevity, we write $T := R^{-1}$.

Denote the ith row of a matrix A by A_{i*} and its jth column by A_{*j}.

Lemma 4.2. *Suppose $r, s, k \in \mathbb{Z}$ where $k \leqslant \min\{r, s\}$. Fix $S \in F^{s\times s}$ and $T \in F^{r\times r}$ and define $X^{(1)}, \ldots, X^{(k)} \in F^{r\times s}$ by $X^{(\ell)} = T_{*\ell}S_{\ell*}$ for $1 \leqslant \ell \leqslant k$. Then $TE_{\ell\ell}S = X^{(\ell)}$ for $1 \leqslant \ell \leqslant k$.*

Proof. Suppose δ_{ij} is 1 if $i = j$ and 0 otherwise. Then the (i, j) entry of $E_{\ell\ell}$ is $\delta_{i\ell}\delta_{\ell j}$. The (i', j') entry of $T_{*\ell}S_{\ell *}$ is $t_{i'\ell}s_{\ell j'}$. This agrees with the (i', j') entry of $TE_{\ell\ell}S$, namely

$$\sum_{i=1}^{r}\sum_{j=1}^{s} t_{i'i}\delta_{i\ell}\delta_{\ell j}s_{jj'} = t_{i'\ell}s_{\ell j'}. \qquad \square$$

Theorem 4.3. *Suppose $r, s, k \in \mathbb{Z}$ where $1 \leqslant k \leqslant \min\{r, s\}$, and suppose F is a field with $|F| \geqslant k + 2$. Then there exist $A \in F^{r\times r}$ and $B \in F^{s\times s}$ such that the linear code $C(A, B)$ has dimension k and minimum distance $d = \lfloor r/k \rfloor s$.*

Proof. By Lemma 4.1 there exist diagonal matrices $A_0 \in F^{r\times r}$ and $B_0 \in F^{s\times s}$ such that $C(A_0, B_0) = \langle E_{11}, \ldots, E_{kk}\rangle$ has dimension k. We seek invertible matrices $R \in F^{r\times r}$ and $S \in F^{s\times s}$ such that $A = A_0^R$ and $B = B_0^S$ give $C(A, B) = \langle E_{11}, \ldots, E_{kk}\rangle^{(R,S)}$ with minimum distance $d = \lfloor r/k \rfloor s$. Let $X^{(\ell)} = R^{-1}E_{\ell\ell}S$ for $1 \leqslant \ell \leqslant k$. The $r \times s$ matrices $X^{(\ell)}$, $1 \leqslant \ell \leqslant k$, will have a form which makes it clear that $d = \lfloor r/k \rfloor s$.

First, we partition the set $\{1, \ldots, r\}$ of rows into the following k subsets:

$$I_1 = \left\{1, \ldots, \left\lfloor \frac{r}{k} \right\rfloor\right\}, I_2 = \left\{\left\lfloor \frac{r}{k} \right\rfloor + 1, \ldots, 2\left\lfloor \frac{r}{k} \right\rfloor\right\}, \ldots,$$
$$I_k = \left\{(k-1)\left\lfloor \frac{r}{k} \right\rfloor + 1, \ldots, r\right\}.$$

Choose the ith row of the matrix $X^{(\ell)}$ to be zero if $i \notin I_\ell$, and to be a vector with all s entries nonzero otherwise. Since $\lfloor \frac{r}{k} \rfloor = |I_\ell| \leqslant |I_k|$ for $\ell < k$, it follows that

$$d(0, X^{(\ell)}) = \sum_{i\in I_\ell} s = |I_\ell|s \geqslant \left\lfloor \frac{r}{k} \right\rfloor s \qquad \text{for } 1 \leqslant \ell \leqslant k$$

with equality if $\ell < k$. The choice of these matrices is such that for each nonzero X in the span $\langle X^{(1)}, \ldots, X^{(k)}\rangle$ we also have $d(0, X) \geqslant d(0, X^{(\ell)})$ for some ℓ, and hence $\langle X^{(1)}, \ldots, X^{(k)}\rangle$ has minimum distance $d = \lfloor r/k \rfloor s$.

It is well known that if the first few rows of a square matrix are linearly independent, then the remaining rows can be chosen so that the matrix is invertible. A similar remark holds if the first few columns are linearly independent. Our construction uses k linearly independent $1 \times s$ row vectors $u_1, \ldots, u_k$ which give the first k rows of $S \in \mathrm{GL}(s, F)$, and k linearly independent $r \times 1$ column vectors $v^{(1)}, \ldots, v^{(k)}$ which give the first k columns of $R^{-1} \in \mathrm{GL}(r, F)$. The pair (R, S) will be used to construct A and B.

Henceforth suppose that $1 \leqslant \ell \leqslant k$. Since $|F| \geqslant 3$, we may choose $\gamma \in F \setminus \{1, 1 - s\}$. Let J be the $s \times s$ matrix with all entries 1. Then the $s \times s$ matrix $S' = (\gamma - 1)I + J$ is invertible as $\det(S') = (\gamma - 1)^{s-1}(\gamma + s - 1)$ is nonzero. Let $u_\ell = (1, \ldots, 1, \gamma, 1, \ldots, 1)$ be the ℓth row of S'. Since $u_1, \ldots, u_k$ are linearly independent, there exists an invertible matrix $S \in \mathrm{GL}(s, F)$ with $S_{\ell *} = u_\ell$. Of course $S = S'$ is one possibility. Similarly, let $v^{(\ell)}$ be the $r \times 1$ column vector

$$v_i^{(\ell)} = \begin{cases} 1 & \text{if } i \in I_\ell, \\ 0 & \text{if } i \notin I_\ell. \end{cases}$$

As $v^{(1)}, \ldots, v^{(k)}$ are linearly independent, there exists an $r \times r$ invertible matrix, which we call R^{-1}, whose first k columns are $v^{(1)}, \ldots, v^{(k)}$. Lemma 4.2 shows that $R^{-1}E_{\ell\ell}S = X^{(\ell)}$ for $1 \leqslant \ell \leqslant k$. Hence $C(A_0^R, B_0^S) = C(A_0, B_0)^{(R,S)} = \langle X^{(1)}, \ldots, X^{(k)}\rangle$ has minimum distance $\lfloor r/k \rfloor s$ as desired. $\square$

Corollary 4.4. *If* $|F| \geqslant \min\{r,s\}+2$*, then there exist matrices* $A \in F^{r\times r}$ *and* $B \in F^{s\times s}$ *such that* $C(A,B)$ *has dimension* $\min\{r,s\}$ *and minimum distance* $\max\{r,s\}$.

Proof. Since $AX = XB$ if and only if $X^TA^T = B^TX^T$ we see that $C(B^T,A^T)$ equals $C(A,B)^T$. Because $C(A,B)$ and $C(A,B)^T$ have the same dimension and minimum distance, we may assume that $r \leqslant s$. If $|F| \geqslant r+2$, then applying Theorem 4.3 with $k = r$ gives the desired result. □

Remark 4.5. Suitable matrices A and B in Theorem 4.3 are constructed by first choosing the diagonal matrices A_0 and B_0 in Lemma 4.1, and then taking $A = R^{-1}A_0R$ and $B = S^{-1}B_0S$ where R and S are constructed in the proof of Theorem 4.3.

It is desirable for a code to have both a high rate, *viz.* $R = k/n$, and a high distance d. Can the product Rd be a constant for intertwining codes? By setting $r = s = k$ in Theorem 4.3 we obtain a rate of $R = 1/r$ and a distance of $d = r$, so the answer is affirmative. It is natural to ask how the maximum value of Rd for an intertwining code depends on (r,s,F)? *We wonder whether there is a sequence* $C_1, C_2, \ldots$ *of intertwining codes over a field* F *with parameters* $[r_is_i, k_i, d_i]$ *for which* $R_id_i = \frac{k_id_i}{r_is_i}$ *approaches infinity.*

5 Upper and lower bounds for $\dim_F(C(\mathcal{A},\mathcal{B}))$

Denote that rank and nullity of $A \in F^{r\times r}$ by $\mathrm{Rk}(A)$ and $\mathrm{Null}(A)$, respectively. Note that $\mathrm{Rk}(A) + \mathrm{Null}(A) = r$ and $\mathrm{Null}(N_\lambda) = \lambda'_1$. In this section we bound $k = \dim(C(A,B))$ in terms of the rank and nullity of A and B. If $\lambda \vdash r$ and $\mu \vdash s$, Theorem 2.9 implies that

$$\lambda'_1\mu'_1 \leqslant \sum_{i\geqslant 1}\lambda'_i\mu'_i = \dim(C(N_\lambda,N_\mu)) \leqslant \left(\sum_{i\geqslant 1}\lambda'_i\right)\left(\sum_{j\geqslant 1}\mu'_j\right) = rs. \qquad (5.1)$$

View $A \in F^{r\times r}$ as acting on an r-dimensional vector space over the algebraic closure $\overline{F}$. Let the α-eigenspace, and the generalized α-eigenspace, of A have dimensions $k_{A,\alpha}$ and $m_{A,\alpha}$, respectively. Then $c_A(t) = \prod(t-\alpha)^{m_{A,\alpha}}$ where $m_{A,\alpha} \neq 0$ for finitely many $\alpha \in \overline{F}$ and $0 \leqslant k_{A,\alpha} \leqslant m_{A,\alpha}$. The following result generalizes [2, Theorems 2.8 and 4.7].

Theorem 5.1. *If* $A \in F^{r\times r}$ *and* $B \in F^{s\times s}$*, then*

(a)
$$\sum k_{A,\alpha}k_{B,\alpha} \leqslant \dim(C(A,B)) \leqslant \sum m_{A,\alpha}m_{B,\alpha}, \quad \textit{and}$$

(b)
$$(r-\mathrm{Rk}(A))(s-\mathrm{Rk}(B)) \leqslant \dim(C(A,B)) \leqslant (r-\mathrm{Rk}(A))(s-\mathrm{Rk}(B)) + \mathrm{Rk}(A)\,\mathrm{Rk}(B).$$

Proof. Part (a) follows immediately from Theorem 2.8 and (5.1).

(b) The lower bound follows from part (a) since $r - \mathrm{Rk}(A) = \mathrm{Null}(A) = k_{A,0}$. For the upper bound, note that A is similar to a diagonal direct sum $N_\lambda \oplus A'$ where N_λ is nilpotent of size $m_{0,A}$ and A' is invertible of size $r - m_{0,A}$. Similarly, B is similar to

$N_\mu \oplus B'$ where N_μ is nilpotent of size $m_{0,B}$ and B' is invertible of size $s - m_{0,B}$. It follows from Theorem 2.8 that $\dim(C(A,B)) = \dim(C(N_\lambda, N_\mu)) + \dim(C(A',B'))$. Further by Theorem 2.9 $\dim(C(N_\lambda, N_\mu)) = \sum_{i\geqslant 1} \lambda'_i\mu'_i$ where, as usual, λ' and μ' denote conjugate partitions. We use the observation:

$$\text{if} \quad 0 \leqslant x \leqslant a \quad \text{and} \quad 0 \leqslant y \leqslant b, \quad \text{then} \quad (a-x)(b-y) + xy \leqslant ab \tag{5.2}$$

to show that

$$\begin{aligned}\dim(C(A,B)) &= \lambda'_1\mu'_1 + \sum_{i\geqslant 2} \lambda'_i\mu'_i + \dim(C(A',B')) \\ &\leqslant \lambda'_1\mu'_1 + (m_{0,A} - \lambda'_1)(m_{0,B} - \mu'_1) + (r - m_{0,A})(s - m_{0,B}) \\ &\leqslant \lambda'_1\mu'_1 + (r - \lambda'_1)(s - \mu'_1).\end{aligned}$$

The result follows since

$$\lambda'_1 = \mathrm{Null}(N_\lambda) = \mathrm{Null}(A) = r - \mathrm{Rk}(A) \quad \text{and} \quad \mu'_1 = s - \mathrm{Rk}(B). \qquad \square$$

The Singleton bound $d + k \leqslant n + 1$ implies that if d is close to $n = rs$, then k is small, and the lower bound of Theorem 5.1(b) implies that A or B has high rank. Setting $k = 1$ in Theorem 4.3, shows that this bound is attained for intertwining codes.

The code $C(A,B)$ is the row nullspace of $A^T \otimes I_s + I_r \otimes B$ and the column nullspace of $A \otimes I_s + I_r \otimes B^T$ where ^{T} denotes transpose.

References

[1] A. Alahmadi, S. Alamoudi, S. Karadeniz, B. Yildiz, C. Praeger and P. Solé, Centraliser codes, *Linear Algebra Appl.* **463** (2014), 68–77, doi:10.1016/j.laa.2014.08.024.

[2] A. Alahmadi, S. P. Glasby and C. E. Praeger, On the dimension of twisted centralizer codes, *Finite Fields Appl.* **48** (2017), 43–59, doi:10.1016/j.ffa.2017.07.005.

[3] A. Alahmadi, S. P. Glasby, C. E. Praeger, P. Solé and B. Yildiz, Twisted centralizer codes, *Linear Algebra Appl.* **524** (2017), 235–249, doi:10.1016/j.laa.2017.03.011.

[4] S. P. Glasby, Lemmas involving two partitions of integers, MathOverflow, 2017, https://mathoverflow.net/q/258722.

[5] D. W. Robinson, Classroom notes: the generalized jordan canonical form, *Amer. Math. Monthly* **77** (1970), 392–395, doi:10.2307/2316152.

[6] R. Stong, Some asymptotic results on finite vector spaces, *Adv. in Appl. Math.* **9** (1988), 167–199, doi:10.1016/0196-8858(88)90012-7.