Smith forms of circulant polynomial matrices

Agnese Ilaria Telloni[a], Gerald Williams[b]

[a] *Dipartimento di Ingegneria Industriale e Scienze Matematiche, Università Politecnica delle Marche, Via Brecce Bianche, 60131 Ancona, Italy*
[b] *Department of Mathematical Sciences, University of Essex, Colchester, Essex CO4 3SQ, U.K.*

Abstract

We obtain the Smith normal forms of a class of circulant polynomial matrices (λ-matrices) in terms of their "associated polynomials" when these polynomials do not have repeated roots. We apply this to the case when the associated polynomials are products of cyclotomic polynomials and show that the entries of the Smith normal form are products of cyclotomics.

Keywords: Smith normal form, circulant matrix, polynomial matrix.
2000 MSC: 15A21, 15B05.

1. Introduction

The Smith normal form theorem ("possibly the most important theorem in all of elementary matrix theory" [4]) has a range of applications, such as for solving differential equations and for detecting similarity of matrices and obtaining Frobenius forms. Many algorithms for finding Smith forms of matrices and λ-matrices have been developed, implemented, and their complexities studied. This has been done both from floating point and symbolic computation perspectives – see for example [8],[2],[6],[7]. The multiplicativity of Smith forms of λ-matrices is considered in [5].

A circulant matrix is one in which each row is a cyclic shift of the preceding row by one column. Such matrices are applied in mathematics, statistics, the physical sciences and electronic engineering; some of their applications are described in [1]. Circulant matrices form a class of matrices where attractive, concise, results may be obtained. For instance, the eigenvalues of a circulant matrix C can be stated in terms of the roots of its "representer polynomial", and hence the diagonal form of C may be stated in terms of this polynomial.

In this paper we investigate the Smith normal form for a class of circulant λ-matrices, and show that this can also be stated in terms of the representer polynomial. Our methods are based on those developed by Lin and Phoong [3]

[1] telloni@dipmat.univpm.it
[2] gwill@essex.ac.uk

in their study of "pseudo-circulant matrices" used in the design of digital filters. Our results differ from the Smith form results cited above, in that they give an expression for the Smith form itself, rather than an algorithm for finding it; this further supports Davis's assertion that "practically every matrix theoretic question for circulants may be resolved in 'closed form'" [1, page xi].

2. Preliminaries

We first set out the notation and terminology that we will use throughout the paper; this will be based on [4]. Let R be a principal ideal ring and F a field (often this will be $\mathbb{Q}$) and let $R^{n\times n}$ denote the ring of $n \times n$ matrices over R (where $n \geq 1$). If $R = F[\lambda]$, where λ is an indeterminate, then an element $M = M(\lambda) \in F[\lambda]^{n\times n}$ is a λ-*matrix* over F (these are also known as *polynomial matrices*). An element $M \in R^{n\times n}$ is *unimodular* if its determinant is a unit of R. Thus if $R = F[\lambda]$ the unimodular matrices are those with non-zero constant determinant.

2.1. Smith normal forms

The Smith normal form theorem can be stated as follows (see for example [4, Theorem II.9]).

Theorem 2.1. *Let R be a principal ideal ring and let $M \in R^{n\times n}$. Then there exist unimodular matrices $U, V \in R^{n\times n}$ such that $UMV = S$ where*

$$S = \mathrm{diag}_n(s_1, \ldots, s_r, 0, \ldots, 0),$$

where $r = \mathrm{rank}_R(M)$, $s_1, \ldots, s_r$ are non-zero elements of R and $s_i | s_{i+1}$ for $1 \leq i \leq r-1$.

The matrix S in Theorem 2.1 is called a *Smith normal form (over R)* of M and it is unique up to multiplication of the entries by units in R. This means that for λ-matrices over F, the entries may be taken to be monic polynomials; some authors include this condition as part of the definition of a Smith form – we shall not do so however.

2.2. Circulant λ-matrices

For $\gamma = (a_1, \ldots, a_n) \in R^n$ the *circulant matrix* $\mathrm{circ}_n(\gamma) \in R^{n\times n}$ is the matrix (over R)

$$C = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}.$$

We will be concerned with circulant λ-matrices (over F) so

$$\gamma = \gamma(\lambda) = (a_1(\lambda), \ldots, a_n(\lambda))$$

and

$$C = C(\lambda) = \mathrm{circ}_n(\gamma(\lambda)).$$

We now give some standard derivations which extend the theory of circulant matrices from Section 2, Chapter 3 of [1] to circulant λ-matrices. We keep our notation and terminology close to that used in [1]. Let

$$\pi = \mathrm{circ}_n(0, 1, 0, \ldots, 0)$$

be the *fundamental circulant*. The *representer polynomial* for $C(\lambda)$ (in the indeterminate z) is the polynomial

$$p_{\gamma(\lambda)}(z) = \sum_{i=1}^{n} a_i(\lambda) z^{i-1}$$

and we observe that

$$p_{\gamma(\lambda)}(\pi) = \mathrm{circ}_n(\gamma(\lambda)).$$

We shall call the polynomial

$$f(\lambda) = p_{\gamma(\lambda)}(1) = \sum_{i=1}^{n} a_i(\lambda)$$

the *associated polynomial* for $C(\lambda)$. We define

$$\Omega = \mathrm{diag}_n(1, \zeta_n, \ldots, \zeta_n^{n-1})$$

where (as throughout this paper) ζ_n denotes $e^{2\pi\sqrt{-1}/n}$, and we let $\mathcal{F}$ be the Fourier matrix of order n (i.e., it is the $n \times n$ matrix whose (i,j)th matrix is $\zeta_n^{-(i-1)(j-1)}/\sqrt{n}$). By Theorem 3.2.1 of [1], we have $\pi = \mathcal{F}^*\Omega\mathcal{F}$, where $\mathcal{F}^*$ is the conjugate transpose of $\mathcal{F}$. Therefore

$$C(\lambda) = p_{\gamma(\lambda)}(\pi) = p_{\gamma(\lambda)}(\mathcal{F}^*\Omega\mathcal{F}) = \mathcal{F}^* p_{\gamma(\lambda)}(\Omega)\mathcal{F} = \mathcal{F}^*\Lambda(\lambda)\mathcal{F}$$

where

$$\Lambda(\lambda) = \mathrm{diag}_n(p_{\gamma(\lambda)}(1), p_{\gamma(\lambda)}(\zeta_n), \ldots, p_{\gamma(\lambda)}(\zeta_n^{n-1})). \tag{1}$$

In summary, we have the λ-matrix version of Theorem 3.2.2 of [1]:

Theorem 2.2. *Let $C(\lambda), \Lambda(\lambda)$ and $\Omega, \pi, \mathcal{F}$ be the λ-matrices and matrices defined above. Then $\Lambda(\lambda)$ is a diagonal form for $C(\lambda)$ with diagonalizing matrices $\mathcal{F}^*$, $\mathcal{F}$:*

$$C(\lambda) = \mathcal{F}^*\Lambda(\lambda)\mathcal{F}.$$

Taking determinants (and noting $\mathcal{F}^*\mathcal{F} = I$) we get

$$\det(C(\lambda)) = \prod_{j=1}^{n} p_{\gamma(\lambda)}(\zeta_n^{j-1}). \tag{2}$$

The same argument as [1, pages 75–76] shows that we may express this as a resultant

$$\prod_{j=1}^{n} p_{\gamma(\lambda)}(\zeta_n^{j-1}) = \mathrm{Res}(z^n - 1, p_{\gamma(\lambda)}(z)).$$

Further, we have

$$\mathrm{Res}(z^n - 1, p_{\gamma(\lambda)}(z)) = a_n(\lambda)^n \prod_{i=1}^{n-1} (\mu_i(\lambda)^n - 1)$$

where $\mu_1(\lambda), \ldots, \mu_{n-1}(\lambda) \in F[\lambda]$ are the roots of $p_{\gamma(\lambda)}(z)$. Summarizing, we have the following corollary to Theorem 2.2 (which is the analogue of the formula obtained in [1, page 76]).

Corollary 2.3. *In the above notation* $\det(C(\lambda)) = a_n(\lambda)^n \prod_{i=1}^{n-1}(\mu_i(\lambda)^n - 1)$.

We record another corollary of Theorem 2.2:

Corollary 2.4. *In the above notation* $\mathrm{rank}_{F[\lambda]}(C(\lambda)) = \mathrm{rank}_{F[\lambda]}(\Lambda(\lambda))$.

3. Smith forms of circulant λ-matrices

In this section we consider circulant λ-matrices

$$C(\lambda) = \mathrm{circ}_n(c_1(\lambda^n), \lambda c_2(\lambda^n), \ldots, \lambda^{n-1} c_n(\lambda^n)) \in \mathbb{Q}[\lambda]^{n \times n}. \tag{3}$$

We obtain a Smith form for $C(\lambda)$ in terms of the roots of the associated polynomial $f(\lambda) = \sum_{i=1}^{n} \lambda^{i-1} c_i(\lambda^n)$. Our methods apply to a slightly wider class of circulant λ-matrices than these; we will comment on this at the end of the section.

Our methods are based on those developed in [3]. Following the terminology of that paper, for $n \geq 2$ we say that $\theta, \phi \in \mathbb{C}$ are *congruous with respect to* n if $\theta^n = \phi^n$ and we say that a set of complex numbers is *congruous with respect to* n if all pairs of its elements are.

Proposition 3.1. *Let* $0 \leq j \leq n-1$. *The representer polynomial* $p_{\gamma(\lambda)}(z)$ *and the associated polynomial* $f(\lambda)$ *of a circulant* λ*-matrix of the form (3) satisfy* $p_{\gamma(\lambda)}(\zeta_n^j) = f(\lambda \zeta_n^j)$.

Proof.

$$p_{\gamma(\lambda)}(z) = \sum_{i=1}^{n} \lambda^{i-1} c_i(\lambda^n) z^{i-1} = \sum_{i=1}^{n} c_i(\lambda^n)(\lambda z)^{i-1}$$

so

$$p_{\gamma(\lambda\zeta_n^j)}(z) = \sum_{i=1}^{n} c_i(\lambda^n)(\lambda(\zeta_n^j z))^{i-1} = p_{\gamma(\lambda)}(\zeta_n^j z)$$

and in particular $f(\lambda\zeta_n^j) = p_{\gamma(\lambda\zeta_n^j)}(1) = p_{\gamma(\lambda)}(\zeta_n^j)$. □

We require the following proposition, the proof of which is straightforward.

Proposition 3.2. *Let $\{\alpha_1, \ldots, \alpha_q\}$ be a set of roots of a polynomial in $\mathbb{Q}[\lambda]$ that is congruous with respect to $n \geq 2$. Then, for $j = 1, \ldots, q$ we have $\alpha_j = \alpha_1 \zeta_n^{n_j}$ for some distinct $0 \leq n_j \leq n-1$.*

Our key lemma is:

Lemma 3.3. *Let $C(\lambda) \in \mathbb{Q}[\lambda]^{n \times n}$ be of the form (3) and let $f(\lambda) \in \mathbb{Q}[\lambda]$ be its associated polynomial. Let α_i be a root of f. Let Σ be a congruous set of roots of f (with respect to n) that contains α_i such that Σ cannot be properly contained in any other set of congruous roots of f (with respect to n).*

If $S(\lambda) = \mathrm{diag}_n(s_1(\lambda), \ldots, s_n(\lambda))$ is a Smith form for $C(\lambda)$ then $(\lambda^n - \alpha_i^n)$ is a factor of $s_j(\lambda)$ if and only if $n - |\Sigma| < j \leq n$.

Proof. By Proposition 3.1 and (1) we have that

$$\Lambda(\lambda) = \mathrm{diag}_n(f(\lambda), f(\lambda\zeta_n), \ldots, f(\lambda\zeta_n^{n-1})).$$

Therefore for each $0 \leq \ell \leq n-1$ we have

$$\Lambda(\alpha_i\zeta_n^\ell) = \mathrm{diag}_n(f(\alpha_i\zeta_n^\ell), f(\alpha_i\zeta_n^{\ell+1}), \ldots, f(\alpha_i\zeta_n^{\ell+n-1}))$$

and so the set of diagonal terms of $\Lambda(\alpha_i\zeta_n^\ell)$ is

$$\{f(\alpha_i\zeta_n^\ell), f(\alpha_i\zeta_n^{\ell+1}), \ldots, f(\alpha_i\zeta_n^{\ell+n-1})\} = \{f(\alpha_i), f(\alpha_i\zeta_n), \ldots, f(\alpha_i\zeta_n^{n-1})\}.$$

By Proposition 3.2 we may write $\Sigma = \{\alpha_i\zeta_n^{n_1}, \ldots, \alpha_i\zeta_n^{n_{|\Sigma|}}\}$ for distinct $0 \leq n_1, \ldots, n_{|\Sigma|} \leq n-1$. Thus $f(\alpha_i\zeta_n^{n_1}) = \ldots = f(\alpha_i\zeta_n^{n_{|\Sigma|}}) = 0$ so at least $|\Sigma|$ of the diagonal entries of $\Lambda(\alpha_i\zeta_n^\ell)$ are zero. No other diagonal entries can be zero, for otherwise Σ could be expanded to form a set of congruous roots of f that contains Σ, a contradiction. Thus $\mathrm{rank}_{\mathbb{Q}[\alpha_i\zeta_n^\ell]}(\Lambda(\alpha_i\zeta_n^\ell)) = n - |\Sigma|$ and by Corollary 2.4 we have

$$\mathrm{rank}_{\mathbb{Q}[\alpha_i\zeta_n^\ell]}(C(\alpha_i\zeta_n^\ell)) = n - |\Sigma|.$$

Therefore by Theorem 2.1

$$\mathrm{rank}_{\mathbb{Q}[\alpha_i\zeta_n^\ell]}(S(\alpha_i\zeta_n^\ell)) = n - |\Sigma|$$

so $(\lambda - \alpha_i\zeta_n^\ell)$ divides exactly $|\Sigma|$ of the diagonal entries of $S(\lambda)$. That is, $(\lambda - \alpha_i\zeta_n^\ell)$ divides $s_j(\lambda)$ if and only if $n - |\Sigma| < j \leq n$. Since this is true for each $0 \leq \ell \leq n-1$ we have that

$$\prod_{\ell=0}^{n-1}(\lambda - \alpha_i\zeta_n^\ell) = (\lambda^n - \alpha_i^n)$$

divides $s_j(\lambda)$ if and only if $n - |\Sigma| < j \leq n$. □

Theorem A. *Let $C(\lambda) \in \mathbb{Q}[\lambda]^{n\times n}$ be of the form (3) and suppose that its associated polynomial $f(\lambda) \in \mathbb{Q}[\lambda]$ does not have repeated roots. Suppose that the sets A_i ($i \in I$) form a partition of the set of roots of f such that each set A_i is congruous with respect to $n \geq 2$ and such that there is no such partition into fewer sets. For each $i \in I$ let α_i denote one element of A_i. Then there exist unimodular matrices $U(\lambda), V(\lambda) \in \mathbb{Q}[\lambda]^{n\times n}$ such that $U(\lambda)C(\lambda)V(\lambda) = S(\lambda)$ where*

$$S(\lambda) = \mathrm{diag}_n(s_1(\lambda), \ldots, s_n(\lambda))$$

where for each $j = 1, \ldots, n$

$$s_j(\lambda) = \prod_{\substack{i\in I,\\ |A_i|>n-j}} (\lambda^n - \alpha_i^n).$$

The matrix $S(\lambda)$ is a Smith form (over $\mathbb{Q}[\lambda]$) of $C(\lambda)$.

Proof. By Theorem 2.1 there exist unimodular matrices $U(\lambda), V(\lambda) \in \mathbb{Q}[\lambda]^{n\times n}$ such that $U(\lambda)C(\lambda)V(\lambda) = S(\lambda)$ where $S(\lambda) = \mathrm{diag}_n(s_1(\lambda), \ldots, s_n(\lambda))$ and $s_j | s_{j+1}$ for all $1 \leq j \leq n-1$. The condition on the partition ensures that no congruous set A_i can be properly contained in any other congruous set of roots of f (with respect to n). Therefore Lemma 3.3 implies that the entry $s_j(\lambda)$ ($j = 1, \ldots, n$) of $S(\lambda)$ is

$$\begin{aligned} s_j(\lambda) &= h_j(\lambda) \prod_{\substack{f(\alpha)=0,\\ \alpha^n=\alpha_i^n,\\ i\in I,\\ |A_i|>n-j}} (\lambda^n - \alpha^n) \\ &= h_j(\lambda) \prod_{\substack{i\in I,\\ |A_i|>n-j}} (\lambda^n - \alpha_i^n) \end{aligned} \tag{4}$$

It remains to show that each $h_j(\lambda)$ is a unit, so may be taken to be 1. We have

$$\begin{aligned} s_1(\lambda)\ldots s_n(\lambda) &= \det(S(\lambda)) \\ &= c\cdot\det(C(\lambda)) \qquad \text{for some } c \in \mathbb{Q} \text{ by Theorem 2.1} \\ &= c\prod_{j=1}^{n} p_{\gamma(\lambda)}(\zeta_n^{j-1}) \qquad \text{by (2)} \\ &= c\prod_{j=1}^{n} f(\lambda\zeta_n^{j-1}) \qquad \text{by Proposition 3.1.} \end{aligned} \tag{5}$$

Let $k\lambda^L$ ($k \in \mathbb{Q}$) be the leading term of $f(\lambda)$ so, since f does not have repeated roots, we have that $L = \sum_{i\in I} |A_i|$. Then $f(\lambda) = k\prod_{l=1}^{L}(\lambda - \theta_l)$, where θ_l,

$l = 1, \ldots, L$, are the roots of $f(\lambda)$ and we have

$$\begin{aligned}\prod_{j=1}^{n} f(\lambda\zeta_n^{j-1}) &= k^n \prod_{l=1}^{L} (\lambda - \theta_l)(\lambda\zeta_n - \theta_l)\ldots(\lambda\zeta_n^{n-1} - \theta_l) \\ &= k^n \prod_{l=1}^{L} (-1)^n(\lambda^n - \theta_l^n) \\ &= (-1)^{Ln} k^n \prod_{l=1}^{L} (\lambda^n - \theta_l^n)\end{aligned}$$

which has degree nL so by (5) we have

$$\deg(s_1(\lambda)\ldots s_n(\lambda)) = nL. \tag{6}$$

Now by (4) we have

$$\deg(s_1(\lambda)\ldots s_n(\lambda)) = \deg(h_1(\lambda)\ldots h_n(\lambda)) + \deg(g(\lambda)) \tag{7}$$

where

$$g(\lambda) = \prod_{j=1}^{n} \prod_{\substack{i\in I,\\ |A_i|>n-j}} (\lambda^n - \alpha_i^n).$$

This gives that for $i \in I$, $(\lambda^n - \alpha_i^n)$ is a factor of $g(\lambda)$ of multiplicity $|A_i|$ so $\deg(g(\lambda)) = n\sum_{i\in I}|A_i| = nL$. Then (6),(7) imply $\deg(h_1(\lambda)\ldots h_n(\lambda)) = 0$, so for every $j = 1, \ldots, n$ we have that $h_j(\lambda)$ is a constant, and therefore a unit. □

Note that in general the associated polynomial does not determine the Smith form of a λ-matrix. For example $\mathrm{circ}_3(1, \lambda, \lambda^2)$ has a Smith form $\mathrm{diag}_3(1, \lambda^3 - 1, \lambda^3 - 1)$ whereas $\mathrm{circ}_3(1, \lambda + \lambda^2, 0)$ has a Smith form $\mathrm{diag}_3(1, 1, \lambda^6 + 3\lambda^5 + 3\lambda^4 + \lambda^3 + 1)$. The hypothesis that $C(\lambda)$ is of the form (3) may be replaced by the condition that the associated polynomial $f(\lambda\zeta_n^j) = p_{\gamma(\lambda)}(\zeta_n^j)$ for all $0 \leq j \leq n-1$. Such a weakening of hypothesis appears to be somewhat marginal, however.

4. Circulant λ-matrices associated with products of cyclotomics

In this section we consider circulant λ-matrices of the form (3) where the roots of the associated polynomial are roots of unity and are not repeated. That is, for associated polynomials of the form

$$f(\lambda) = \prod_{i=1}^{r} \Phi_{d_i}(\lambda) \tag{8}$$

where $r \geq 1, d_i \geq 1$, $d_i \neq d_j$ for $i \neq j$. If we let $m = \mathrm{lcm}(d_1, \ldots, d_r)$ then we may write (8) as

$$f(\lambda) = \prod_{d|m} \Phi_d(\lambda)^{\delta(d)} \tag{9}$$

where

$$\delta(d) = \begin{cases} 1 & \text{if } d \in \{d_1, \ldots, d_r\}, \\ 0 & \text{if } d \notin \{d_1, \ldots, d_r\}, \end{cases}$$

and clearly we can write (9) in the form (8). Examples of λ-matrices of this form are

$$\mathrm{circ}_n(1, \lambda, \lambda^2, \ldots, \lambda^{m-1}, 0, \ldots, 0) \in \mathbb{Q}[\lambda]^{n \times n}, \tag{10}$$

where $n \geq m$, and their associated polynomials are

$$1 + \lambda + \lambda^2 + \ldots + \lambda^{m-1} = \prod_{\substack{d|m \\ d \neq 1}} \Phi_d(\lambda).$$

These appear in Problem 28, page 82 of [1], which asks for a proof that the determinant is equal to

$$(-1)^{(m,n)-1} \frac{(\lambda^{nm/(m,n)} - 1)^{(m,n)}}{\lambda^n - 1}$$

(attributing the result to Oystein Ore). Theorem B will calculate the Smith form of λ-matrices of the form (3) with associated polynomial of the form (9) and show that its diagonal entries are products of cyclotomics in λ^n; Corollary C will apply it to the λ-matrices (10).

By (9) we see that the set of roots of f is

$$B = \{\zeta_d^i \mid i = 1, \ldots, d-1, (i,d) = 1, d|m, \delta(d) = 1\}. \tag{11}$$

Now

$$f(\lambda) | \prod_{d|m} \Phi_d = (\lambda^m - 1)$$

and $(\lambda^m - 1)$ has roots ζ_m^i, $i = 0, \ldots, m-1$, so the set of roots of f is

$$B = \{\zeta_m^i \mid i = 0, \ldots, m-1, \Phi_d(\zeta_m^i) = 0 \text{ for some } d|m \text{ with } \delta(d) = 1\}.$$

With this in mind, we set $M = m/(m,n), N = n/(m,n)$ and for each $\mu = 0, \ldots, M-1$ we define the sets

$$B_\mu = \{\zeta_m^{\mu + jM} \mid j = 0, \ldots, (m,n) - 1, \Phi_d(\zeta_m^{\mu+jM}) = 0 \text{ for some } d|m, \delta(d) = 1\} \tag{12}$$

and it is clear that these form a partition of the set B. Furthermore, given $\mu = 0, \ldots, M-1$, if $\theta \in B_\mu$ then

$$\theta^n = \zeta_m^{n\mu+jMn} = \zeta_m^{n\mu} \cdot (\zeta_m^m)^{jn/(m,n)} = \zeta_m^{n\mu} \tag{13}$$

so each set B_μ ($\mu = 0, \ldots, M-1$) is congruous with respect to n.

Also, if $\zeta_m^i \in B$ is congruous to an element of B_μ then $\zeta_m^i \in B_\mu$. To see this note that $\zeta_m^i \in B$ can be written $\zeta_m^i = \zeta_m^{\mu'+jM}$ for some $\mu' \in \{0, \ldots, M-1\}$, $j \in \{0, \ldots, (m,n)-1\}$, so $(\zeta_m^i)^n = \zeta_m^{n\mu}$ if and only if $\zeta_m^{n\mu'} = \zeta_m^{n\mu}$ or equivalently $\zeta_M^{N\mu'} = \zeta_M^{N\mu}$ or $\zeta_M^{N(\mu'-\mu)} = 1$ or $(\mu'-\mu) \equiv 0 \bmod M$, ie $\mu' = \mu$ so $\zeta_m^i = \zeta_m^{\mu+jM} \in B_\mu$. Thus the set B_μ cannot be properly contained in any other set of congruous roots of f (with respect to n).

For each $\mu = 0, \ldots, M-1$, let $\chi(\mu)$ denote any element of B_μ. We collect the sets B_μ according to the 'primitivity' of $\chi(\mu)^n$. That is, since $\chi(\mu)^n = \zeta_m^{n\mu}$ is a primitive $m/(n\mu, m)$'th root of unity, we collect the sets B_μ according to the value of $m/(n\mu, m)$, or equivalently according to the value of $(n\mu, m)$. Now

$$\begin{aligned}
\{(n\mu, m) \mid \mu = 0, \ldots, M-1\} &= \{(N(m,n)\mu, M(m,n)) \mid \mu = 0, \ldots, M-1\} \\
&= \{(m,n)(N\mu, M) \mid \mu = 0, \ldots, M-1\} \\
&= \{(m,n)(\mu, M) \mid \mu = 0, \ldots, M-1\} \\
&= \{(m,n)D \mid D|M\}. \qquad (14)
\end{aligned}$$

That is, the possible values of $(n\mu, m)/(m,n)$ are the divisors D of M. We partition the set $\{0, \ldots, M-1\}$ according to these values. For each $D|M$, defining

$$\mathcal{C}_D = \{\mu \mid \mu = 0, \ldots, M-1, (n\mu, m)/(m,n) = D\}$$

gives such a partition, i.e.

$$\cup_{D|M} \mathcal{C}_D = \{\mu \mid \mu = 0, \ldots, M-1\}. \tag{15}$$

Proposition 4.1. $\prod_{\mu \in \mathcal{C}_D} (\lambda^n - \chi(\mu)^n) = \Phi_{M/D}(\lambda^n)$.

Proof.

$$\begin{aligned}
\prod_{\mu \in \mathcal{C}_D} (\lambda^n - \chi(\mu)^n) &= \prod_{\substack{\mu=0,\ldots,M-1 \\ (n\mu,m)=(m,n)D}} (\lambda^n - \zeta_m^{n\mu}) \\
&= \prod_{\substack{\mu=0,\ldots,M-1 \\ (N\mu,M)=D}} (\lambda^n - \zeta_m^{N\mu(m,n)}) \\
&= \prod_{\substack{\mu=0,\ldots,M-1 \\ (N\mu,M)=D}} (\lambda^n - \zeta_M^{N\mu}) \\
&= \prod_{\substack{\nu=0,\ldots,M-1 \\ (\nu,M)=D}} (\lambda^n - \zeta_M^{\nu}) \qquad \text{where } \nu = N\mu \bmod M
\end{aligned}$$

and the result follows. □

Proposition 4.2. *Let $D|M$. If $\mu, \mu' \in \mathcal{C}_D$ then $|B_\mu| = |B_{\mu'}|$.*

Proof. Let $\mu, \mu' \in \mathcal{C}_D$. Then we have $(n\mu, m)/(m, n) = D = (n\mu', m)/(m, n)$, so $(n\mu, m) = (n\mu', m)$. Hence there exists $1 \leq t \leq m$, $(t, m) = 1$ such that $t(n\mu) \equiv (n\mu') \bmod m$. Therefore $t(N\mu) \equiv N\mu' \bmod M$ and thus $t\mu \equiv \mu' \bmod M$ (since $(M, N) = 1$), so $t\mu = \mu' + rM$ for some $r \in \mathbb{Z}$. If $\zeta_m^{\mu+jM} \in B_\mu$ then $\zeta_m^{(\mu+jM)t} = \zeta_m^{\mu t+jMt} = \zeta_m^{\mu'+j'M}$ where $j' = r + jt \bmod m$. Now since $(t, m) = 1$ we have $(\mu + jM, m) = (\mu t + jtM, m) = (\mu' + rM + jtM, m) = (\mu' + j'M, m)$, i.e.

$$(\mu' + j'M, m) = (\mu + jM, m). \tag{16}$$

Now since $\zeta_m^{\mu+jM} \in B_\mu$ we have $\Phi_d(\zeta_m^{\mu+jM}) = 0$ for some $d|m$ with $\delta(d) = 1$. That is, $\zeta_m^{\mu+jM}$ is a primitive d'th root of unity (where $d|m$ with $\delta(d) = 1$). By (16) we have that $\zeta_m^{\mu'+j'M}$ is also a primitive d'th root of unity so $\zeta_m^{\mu'+j'M} \in B_{\mu'}$. Therefore we may define $\iota : B_\mu \to B_{\mu'}$ by $\iota(\theta) = \theta^t$. Then $\zeta_m^{\mu'+j'M} = \iota(\zeta_m^{\mu+jM})$ so ι is onto; we now show that it is one-to-one.

Let $\theta_1 = \zeta_m^{\mu+j_1M}, \theta_2 = \zeta_m^{\mu+j_2M} \in B_\mu$ $(j_1, j_2 \in \{1, \ldots, (m, n) - 1\})$. Then $\iota(\theta_1) = \iota(\theta_2)$ if and only if $\zeta_m^{(j_1-j_2)Mt} = 1$. This occurs if and only if $(j_1 - j_2)Mt \equiv 0 \bmod m$, which occurs if and only if $(j_1 - j_2)M \equiv 0 \bmod m$ (since $(m, t) = 1$), which occurs if and only if $(m, n)|(j_1 - j_2)$, i.e. $j_1 = j_2$. Therefore ι is a bijection, and the result follows. □

Therefore, for each $D|M$ we may set

$$b_D = |B_\mu| \tag{17}$$

for any $\mu \in \mathcal{C}_D$. We can give a self contained expression for b_D, which we shall use in the statement of Theorem B.

Proposition 4.3. *Let $D|M$. Let $\mu = 0, \ldots, M - 1$ satisfy $(n\mu, m) = (m, n)D$ (such a μ exists by (14)). Then*

$$b_D = |\{j \mid j = 0, \ldots, (m, n) - 1, m/(\mu + jM, m) = d \text{ for some } d|m, \delta(d) = 1\}|.$$

Proof. We have that $\mu \in \mathcal{C}_D$ so the result follows from (12),(17). □

Proposition 4.4. *For each $j = 1, \ldots, n$*

$$\prod_{\substack{\mu=0,\ldots,M-1 \\ |B_\mu|>n-j}} (\lambda^n - \chi(\mu)^n) = \prod_{D|M} \Phi_{M/D}(\lambda^n)^{\Delta(D,j)}$$

where $\Delta(D, j) = \begin{cases} 0 & \text{if } b_D \leq n - j, \\ 1 & \text{if } b_D > n - j. \end{cases}$

Proof.

$$\begin{aligned}\prod_{\substack{\mu=0,\dots,M-1\\ |B_\mu|>n-j}}(\lambda^n-\chi(\mu)^n) &= \prod_{D|M}\prod_{\substack{\mu\in\mathcal{C}_D\\ |B_\mu|>n-j}}(\lambda^n-\chi(\mu)^n) \qquad \text{by (15)}\\ &= \prod_{\substack{D|M\\ b_D>n-j}}\prod_{\mu\in\mathcal{C}_D}(\lambda^n-\chi(\mu)^n) \qquad \text{by (17)}\\ &= \prod_{\substack{D|M\\ b_D>n-j}}\Phi_{M/D}(\lambda^n) \qquad \text{by Proposition 4.1}\\ &= \prod_{D|M}\Phi_{M/D}(\lambda^n)^{\Delta(D,j)}.\end{aligned}$$

□

Theorem B. *Let $m,n\in\mathbb{N}$, set $M=m/(m,n)$ and suppose*

$$C(\lambda)=\mathrm{circ}_n(c_1(\lambda^n),\lambda c_2(\lambda^n),\dots,\lambda^{n-1}c_n(\lambda^n))\in\mathbb{Q}[\lambda]^{n\times n}$$

has associated polynomial

$$f(\lambda)=\prod_{d|m}\Phi_d(\lambda)^{\delta(d)}$$

where $\delta(d)\in\{0,1\}$ $(d|m)$. Each $D|M$ may be written $D=(n\mu,m)/(m,n)$ for some $\mu=0,\dots,M-1$; for each such D define

$$b_D=|\{j\mid j=0,\dots,(m,n)-1, m/(\mu+jM,m)=d \text{ for some } d|m,\delta(d)=1\},$$

and for each such D and $1\le j\le n$ define $\Delta(D,j)$ by

$$\Delta(D,j)=\begin{cases}0 & \text{if } b_D\le n-j,\\ 1 & \text{if } b_D>n-j.\end{cases}$$

Then there exist unimodular matrices $U(\lambda),V(\lambda)\in\mathbb{Q}[\lambda]^{n\times n}$ such that

$$U(\lambda)C(\lambda)V(\lambda)=S(\lambda)$$

where $S(\lambda)=\mathrm{diag}_n(s_1(\lambda),\dots,s_n(\lambda))$where

$$s_j(\lambda)=\prod_{D|M}\Phi_{M/D}(\lambda^n)^{\Delta(D,j)}$$

and $S(\lambda)$ is a Smith form (over $\mathbb{Q}[\lambda]$) of $C(\lambda)$.

Proof. Recall that the sets B_μ $(\mu=0,\dots,M-1)$ defined at (12) form a partition of the set B of roots of f defined at (11) and each set B_μ is congruous with respect to n by (13). Further, no set B_μ can be properly contained in any other

set of congruous roots of f (with respect to n) so there is no such partition into fewer sets. Thus by Theorem A there exist unimodular matrices $U(\lambda), V(\lambda) \in \mathbb{Q}[\lambda]^{n\times n}$ such that $S(\lambda) = U(\lambda)C(\lambda)V(\lambda)$ is a Smith form (over $\mathbb{Q}[\lambda]$) of $C(\lambda)$, where

$$S(\lambda) = \mathrm{diag}_n(s_1(\lambda), \ldots, s_n(\lambda))$$

where for each $j = 1, \ldots, n$

$$s_j(\lambda) = \prod_{\substack{\mu=0,\ldots,M-1 \\ |B_\mu|>n-j}} (\lambda^n - \chi(\mu)^n)$$

where $\chi(\mu)$ denotes one element from B_μ so Proposition 4.4 implies

$$s_j(\lambda) = \prod_{D|M} \Phi_{M/D}(\lambda^n)^{\Delta(D,j)}.$$

□

Corollary C. *Let $C(\lambda) = \mathrm{circ}_n(1, \lambda, \lambda^2, \cdots, \lambda^{m-1}, 0, \cdots, 0) \in \mathbb{Q}[\lambda]^{n\times n}$ and set $M = m/(m,n)$. Then there exist unimodular matrices $U(\lambda), V(\lambda) \in \mathbb{Q}[\lambda]^{n\times n}$ such that $U(\lambda)C(\lambda)V(\lambda) = S(\lambda)$ where*

$$S(\lambda) = \mathrm{diag}_n(\underbrace{1, \cdots, 1}_{n-(m,n)}, \sum_{j=0}^{M-1} \lambda^{jn}, \underbrace{\lambda^{nM}-1, \cdots, \lambda^{nM}-1}_{(m,n)-1})$$

is a Smith form for $C(\lambda)$.

Proof. The associated polynomial is

$$f(\lambda) = 1 + \lambda + \lambda^2 + \ldots + \lambda^{m-1} = \frac{\lambda^m - 1}{\lambda - 1} = \prod_{d|m} \Phi_d(\lambda)^{\delta(d)}$$

where $\delta(1) = 0$, $\delta(d) = 1$ $(d|m,\ d \neq 1)$. For $D|M$ with $D = (n\mu, m)/(m,n)$ $(\mu = 0, \ldots, M-1)$ we have

$$\begin{aligned} b_D &= |\{j \mid j = 0, \ldots, (m,n)-1, m/(\mu + jM, m) = d \text{ for some } d|m, d \neq 1\}| \\ &= |\{j \mid j = 0, \ldots, (m,n)-1, m/(\mu + jM, m) \neq 1\}| \\ &= |\{j \mid j = 0, \ldots, (m,n)-1, \mu \neq 0 \text{ or } j \neq 0\}| \\ &= \begin{cases} (m,n) & \text{if } \mu \neq 0, \\ (m,n)-1 & \text{if } \mu = 0. \end{cases} \end{aligned}$$

Now $\mu = 0$ if and only if $D = M$ so $b_D = (m,n)$ if $D \neq M$ and $b_D = (m,n) - 1$ if $D = M$. That is, $\Delta(D, j) = 0$ if $1 \leq j \leq n - (m,n)$ or if $j = n - (m,n) + 1$

and $D = M$, and $\Delta(D, j) = 1$ otherwise, so Theorem B gives

$$\begin{aligned} s_j(\lambda) &= \prod_{D|M} \Phi_{M/D}(\lambda^n)^{\Delta(D,j)} \\ &= \begin{cases} 1 & \text{if } j \le n - (m,n), \\ \prod_{\substack{D|M \\ D \neq M}} \Phi_{M/D}(\lambda^n) & \text{if } j = n - (m,n) + 1, \\ \prod_{D|M} \Phi_{M/D}(\lambda^n) & \text{if } j > n - (m,n) + 1, \end{cases} \\ &= \begin{cases} 1 & \text{if } j \le n - (m,n), \\ \sum_{j=0}^{M-1} (\lambda^n)^j & \text{if } j = n - (m,n) + 1, \\ (\lambda^n)^M - 1 & \text{if } j > n - (m,n) + 1. \end{cases} \end{aligned}$$

□

Acknowledgements

We thank Yuan-Pei Lin for helpful correspondence relating to this work. The second named author would like to thank the Dipartimento di Matematica at the Università di Modena e Reggio Emilia for its hospitality during a research visit in 2009. This research was partially supported by a London Mathematical Society Scheme 4 grant and the project "L'Oréal Italia for Women and Science".

References

[1] Philip J. Davis. *Circulant matrices. 2nd ed.* New York, NY: AMS Chelsea Publishing, 1994.

[2] R. Kannan. Solving systems of linear equations over polynomials. *Theor. Comput. Sci.*, 39:69–88, 1985.

[3] Yuan-Pei Lin and See-May Phoong. Smith form of FIR pseudocirculants. *IEEE Signal Processing Letters*, 9(8), 2002.

[4] Morris Newman. *Integral matrices.* Pure and Applied Mathematics, 45. New York-London: Academic Press XVII, 1972.

[5] V.M. Petrichkovich. Semiscalar equivalence and the Smith normal form of polynomial matrices. *J. Sov. Math.*, 66(1):2030–2033, 1987.

[6] Arne Storjohann and George Labahn. A fast Las Vegas algorithm for computing the Smith normal form of a polynomial matrix. *Linear Algebra Appl.*, 253:155–173, 1997.

[7] Gilles Villard. Generalized subresultants for computing the Smith normal form of polynomial matrices. *J. Symb. Comput.*, 20(3):269–286, 1995.

[8] Jon Wilkening and Jia Yu. A local construction of the Smith normal form of a matrix polynomial. *J. Symb. Comput.*, 46(1):1–22, 2011.