

CIRCULANTS (Extract)

Alun Wyn-jones

Last revised in December 2013.

Please copy this book for your own reading only. Refers others to this website. Thank You.

CHAPTER 2.
Circulant Matrices

This chapter places the circulants in the context of the familiar matrices. Later chapters will focus on circulants almost to the exclusion of the general matrix ring in which they reside. In keeping with this goal, the circulants in this chapter are matrices whereas in subsequent chapters circulants will usually be viewed as vectors with convolution as the ring product. We shall continue to assume that the underlying ring is an integral domain whose characteristic does not divide the order of any circulant matrix under discussion. We usually denote the circulant order by N , and the underlying ring by R . Thus, we are assuming that $\text{char } R \nmid N$.

We shall describe circulant matrices as a subring of $M_N(R)$, or multiplicatively as a subgroup of $\text{GL}_N(R)$. We shall start the study of ring homomorphisms on circulants by describing all those automorphisms on the circulant matrices over the reals and complex numbers which are given by similarity transformations by real and complex non-singular matrices. It will be shown that the similarity transformations (i.e. the inner-automorphisms of $\text{GL}(R)$) account for all of the linear ring automorphisms on $\text{CIRC}_N(R)$ for $R \subset \mathbb{C}$. We can therefore use the general group formula

$$\text{Inn}_G(H) \approx \text{Norm}_G(H)/C_G(H)$$

to estimate the isomorphism class of the linear automorphisms on circulants. In our case, $G = \text{GL}_N(R)$ and $H = \text{CIRC}_N(R)$. This leads us to calculate the centralizer and normalizer of the circulant matrices in the general linear group.

2.1 The Centralizer of the Circulant Matrices in the General Linear Group.

We begin by determining the centralizer of $\text{CIRC}(R)$. This part proves to be easy to do and easy to state: The largest set of non-singular matrices which commute with the circulants is the circulants themselves.

2.1.1 Proposition Let R be a complex domain. $\text{CIRC}_N(R)$ is its own centralizer in $M_N(R)$.

Proof. Let A commute with every member of $\text{CIRC}_N(R)$, and pick any $C \in \text{CIRC}_N(R)$ which has distinct eigenvalues. Such a circulant matrix certainly exists, for instance, U , the generator of the standard basis for $\text{CIRC}_N(R)$, has the N distinct eigenvalues $1, \zeta, \zeta^2, \dots, \zeta^{N-1}$. Let e_0, e_1, \dots, e_{N-1} be the standard orthonormal basis for the eigenspace. These vectors are also the common eigenvectors of all circulant matrices. (See 1.10.) Multiply e_i by CA .

$$CAe_i = ACE_i = A\lambda_i e_i = \lambda_i Ae_i$$

Therefore, Ae_i is an eigenvector of C with eigenvalue λ_i . Since the eigenvalues $\lambda_0, \lambda_1, \dots, \lambda_{N-1}$ are distinct, this is possible only if Ae_i is a multiple of e_i for every $i \in \mathbb{Z}_N$ which means that A is diagonalized by F . By Theorem 1.5(iii), A must be circulant. \square

The above proof can easily be generalized to give the following corollary.

2.1.2 Corollary Let R be a complex domain, and let $A \in M_N(R)$. Then, A is circulant iff it commutes with any circulant matrix having distinct eigenvalues. In particular, A is circulant iff it commutes with U . \square

This corollary has a rather pretty corollary of its own which is the subject of the next exercise.

2.1.3 Exercise For a complex domain R , define a linear map $\tau : M_n(R) \rightarrow M_n(R)$ by $\tau(M) = n^{-1} \sum_{i=0}^{n-1} U_n^{-i} M U_n^i$. Then, τ is a projection map onto $\text{circ}_n(R)$.

2.1.4 Exercise Let A be an $N \times N$ circulant matrix, let P_σ be the permutation matrix corresponding to a cyclic permutation σ on N objects, and let $\tilde{P}_\sigma = F^{-1} P_\sigma F$ where F is the Fourier matrix. Show that $(A\tilde{P}_\sigma)^N = \det(A)I$ where I is the identity matrix.

2.2 The Shift-circulant Matrices.

The next definition introduces a generalization of circulant matrices called the **shift-circulant** or **s -circulant** matrices. Like a circulant matrix, an s -circulant matrix is completely determined by its first row. Each subsequent row is the prior row rotated $s \pmod{N}$ columns to the right. The amount, s , that each row is rotated is called the **shift** of the matrix. Thus, a matrix of shift 1 is circulant; that is, 1-circulant means circulant.

2.2.1 Definition The Shift-circulant or s -circulant Matrices.

$$(i) \quad K_N(R) := \{A \in M_N(R) \mid \exists s \in \mathbb{Z}_N, A_{i,j} = A_{0,j-si}, \forall i, j \in \mathbb{Z}_N\}.$$

$$(ii) \quad K_N^*(R) := GL_N(R) \cap K_N(R).$$

K_N is the set of all shift-circulant matrices and K_N^* is the set of non-singular shift-circulant matrices.

Example.

$$A = \begin{pmatrix} 1 & -2 & 3 & -4 & 5 \\ -4 & 5 & 1 & -2 & 3 \\ -2 & 3 & -4 & 5 & 1 \\ 5 & 1 & -2 & 3 & -4 \\ 3 & -4 & 5 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 3 & -1 & 1 & 4 \\ -1 & 1 & 4 & 0 & 3 \\ 4 & 0 & 3 & -1 & 1 \\ 3 & -1 & 1 & 4 & 0 \\ 1 & 4 & 0 & 3 & -1 \end{pmatrix}$$

$$\therefore AB = \begin{pmatrix} 7 & 25 & -4 & -3 & -4 \\ -4 & 7 & 25 & -4 & -3 \\ -3 & -4 & 7 & 25 & -4 \\ -4 & -3 & -4 & 7 & 25 \\ 25 & -4 & -3 & -4 & 7 \end{pmatrix}, \quad BA = \begin{pmatrix} 7 & -3 & 25 & -4 & -4 \\ -4 & 7 & -3 & 25 & -4 \\ -4 & -4 & 7 & -3 & 25 \\ 25 & -4 & -4 & 7 & -3 \\ -3 & 25 & -4 & -4 & 7 \end{pmatrix}$$

The matrix A is a 2-circulant, B is a 3-circulant, and the products, AB and BA , are both 1-circulants, that is, circulant. As is clear from this example, shift-circulants do not necessarily commute.

2.2.2 Lemma If $A \in K_N^*(R)$ has shift s , then s is coprime to N .

Proof. Suppose $\gcd(s, N) = d > 1$, then the $(N/d)^{\text{th}}$ row equals the 0^{th} row. So the determinant is zero. \square

2.2.3 Proposition $K_N^*(R)$ is a multiplicative group, and is non-abelian for $N > 2$.[†]

Proof. That it is non-abelian is obvious since for $N > 2$, $K_N^* \not\subset \text{CIRC}_N$, and CIRC_N is its own centralizer. The case of $N = 2$ is semi-trivial, since $K_2^* \subset \text{CIRC}_2$.

Let $A, B \in K_N$ with shifts of s, t respectively.

(i) Closure.

$$\begin{aligned} (AB)_{i,k} &= \sum_{j \in \mathbb{Z}_N} a_{j-si} b_{k-tj} \\ &= \sum_{j \in \mathbb{Z}_N} a_{j-s(i+1)} b_{k+st-tj} \quad \text{by } j \rightarrow j-s \\ &= (AB)_{i+1, k+st} \\ \therefore AB &\in K_N \text{ with shift } st \end{aligned}$$

(ii) Inverse

Suppose $A \in K_N^*$ with shift s . By the lemma, s is coprime to N , so $s^{-1} \pmod{N}$ exists. Choose any $B \in K_N^*$ with shift $s^{-1} \pmod{N}$. By part (i), AB has a shift of 1 and so is circulant. $\therefore F$ diagonalizes AB . $\therefore F^{-1}ABF = D$ where D is diagonal. Since A and $B \in K_N^*$, D is non-singular. $\therefore F^{-1}ABFD^{-1} = We$. $\therefore ABFD^{-1}F^{-1} = We$. By Theorem 1.6(iii), $FD^{-1}F^{-1} \in \text{CIRC}_N$ and by part (i), $BFD^{-1}F^{-1} \in K_N$, $\therefore A^{-1} = BFD^{-1}F^{-1} \in K_N$. \square

[†] The assumption that $\text{char } R$ does not divide N is crucial. For example, $K_3^*(\mathbb{Z}_3)$ is abelian.

The above proof can be easily adapted to show that the shift-circulant matrices are multiplicatively closed, that is, form a semigroup. However, the set of shift-circulants is not a ring since they are not additively closed.

The proof showed more than was stated. The implied results are in the next corollary.

2.2.4 Corollary

- (i) If $A, B \in K_N$ have shift of s, t respectively, then AB has a shift of st .
(ii) If $A \in K_N^*$ then $A^{-1} = BC$ where $B \in K_N^*$ with $\text{shift}(B) = s^{-1}$ and is otherwise arbitrary, and C (depending on B) is circulant. \square

2.2.5 Proposition $K_N^*(R) \subset \text{Norm CIRC}_N(R)$, the normalizer of $\text{CIRC}_N(R)$ in $\text{GL}_N(R)$.

Proof. Let $\text{shift}(A) = s$. Let $C \in \text{CIRC}_N$. By the corollary Part (ii), we have

$$\begin{aligned} ACA^{-1} &= ACBC_1 \text{ where } \text{shift}(B) = s^{-1}, \text{shift}(C_1) = 1 \\ \therefore \text{shift}(ACA^{-1}) &= \text{shift}(A)\text{shift}(C)\text{shift}(B)\text{shift}(C_1) \\ &= s \times 1 \times s^{-1} \times 1 \\ &= 1 \quad \square \end{aligned}$$

It is clear by now that the shift is a (multiplicative) group homomorphism from $K_N^*(R) \rightarrow \mathbb{Z}_N^*$. If $p \mid \phi(N)$ there will be a subgroup of order p in \mathbb{Z}_N^* and so its inverse image will be a subgroup in K_N^* . For example, if $N > 2$, the non-singular anticirculant ($\text{shift} = -1$) and circulant matrices form a subgroup of K_N^* .

By Theorem 1.6(iii) the Fourier matrix diagonalizes a shift-circulant only if the shift is 1. Nevertheless, the Fourier matrix does bring the shift-circulant into a form where each row and column has exactly one non-zero element.

2.2.6 Proposition

Let F be the $N \times N$ Fourier matrix, and let $T \in K_N^*$ with $T_{i,j} = t_{j-si}$, say. Then, $(F^{-1}TF)_{i,h} = \delta_{i-sh}\lambda_h(t)$, and in particular, $\det T = \pm\Delta_N(t)$, the circulant determinant on the same vector.

Proof.

$$\begin{aligned} (F^{-1}TF)_{i,h} &= N^{-1} \sum_{j,k \in \mathbb{Z}_N} \zeta^{-ij} t_{k-sj} \zeta^{kh} \\ &= N^{-1} \sum_{j \in \mathbb{Z}_N} \sum_{m \in \mathbb{Z}_N} t_m \zeta^{(m+sj)h-ij} \quad \text{substituting } m = k - sj \\ &= N^{-1} \sum_{m \in \mathbb{Z}_N} t_m \zeta^{mh} \sum_{j \in \mathbb{Z}_N} \zeta^{j(sh-i)} \\ &= \delta_{i-sh} \sum_{m \in \mathbb{Z}_N} t_m \zeta^{mh} \\ &= \delta_{i-sh} \lambda_h(t) \end{aligned}$$

Since s is coprime to N , δ_{i-sh} has a single 1 in every column and row, and so $\lambda_h(t)$ occurs exactly once and shares no column or row with another. The determinant of T is therefore the product of $\lambda_h(t)$ to within sign. \square

Of course, $\det T = \pm\Delta(t)$ is an easy consequence of the fact that T is just a row permutation of $\mathbf{circ}(t)$.

A matrix such as $F^{-1}TF$ in the proposition which has only one non-zero entry in every row and column is called a **monomial** matrix or a **PD-** matrix (for Permutation Diagonal). Note however, that the $F^{-1}TF$ matrix in the proposition is not a completely general PD-matrix: the non-zero entry on each row occurs a fixed displacement from the row above. It will transpire that the full normalizer of the circulants is in fact FPF^\dagger where P is the set of PD-matrices (see Proposition 2.5.1).

One might imagine from the proposition that the shift-circulant shares not just its determinant but also its eigenvalues with the circulant having the same top row. This is not the case. Indeed, the eigenvalues of the shift-circulants of order N are not typically in the ring $R(\zeta_N)$. For a simple example, take the vector $t = (1, 0, 0, 0, 0)$ with a shift of 2.

$$T := \begin{pmatrix} 1, & 0, & 0, & 0, & 0 \\ 0, & 0, & 1, & 0, & 0 \\ 0, & 0, & 0, & 0, & 1 \\ 0, & 1, & 0, & 0, & 0 \\ 0, & 0, & 0, & 1, & 0 \end{pmatrix}$$

One easily verifies that $T^4 = I$, and that this is the minimum polynomial for T . Hence, the eigenvalues of T are the fourth roots of unity whereas the eigenvalues of $\text{CIRC}(1, 0, 0, 0, 0)$, which is the 5×5 identity matrix are all 1.

2.4 Normalizer of CIRC_N

We have seen in Proposition 2.2.5 that the shift-circulants are a subgroup of the normalizer group of the circulants. They are not however the entire normalizer subgroup.

2.4.1 Assumptions and Notation.

(i) To characterize the full normalizer group of the circulants, we shall need a notation for the general permutation matrix. We shall denote it by P_σ where σ is any permutation on the symbols $\{0, 1, 2, \dots, N-1\}$. We denote the set of all such permutations, the full symmetric group on N objects, by \mathcal{S}_N .

For all $\sigma \in \mathcal{S}_N$, let P_σ be the permutation matrix given by $(P_\sigma)_{i,j} = \delta_{i-\sigma(j)}$.

Let u_0, u_1, \dots, u_{N-1} be the unit coordinate vectors. Then, $P_\sigma : u_k \mapsto u_{\sigma(k)}$ which shows that P is a group homomorphism. That is, $P_\sigma P_\tau = P_{\sigma\tau}$.

(ii) In the next few lemmas and propositions, R and S are to represent rings which are subsumed by a larger ring (all integral domains with characteristic not dividing N). The containment by a larger ring guarantees that the ring operations on mixed elements from R and S make sense. This slightly strange requirement allows us to generalize theorems on the normalizer of the circulants to answer natural questions such as, for example, ‘‘What is the normalizer within the complex matrices of the real circulants.’’

(iii) We shall be discussing the normalizer in $\text{GL}_N(S)$ of $\text{circ}_N(R)$ where R and S , as we discussed, are integral domains subsumed in a larger domain.. The standard notation for the normalizer would be the unwieldy $\text{Norm}_{\text{GL}_N(S)} \text{CIRC}_N(R)$. We shall simplify this to $\text{Norm}_S \text{CIRC}_N(R)$.

(iv) To reduce clutter in expressions involving similarity transformations with the Fourier matrix F , we shall denote $F^{-1}AF$ by \hat{A} . Thus, if A is circulant, $\hat{A} = \text{Diag}(\lambda_0, \lambda_1, \dots, \lambda_{N-1})$ where $\lambda_i = \lambda_i(A)$. Thus, for example, $\hat{U} = \text{Diag}(1, \zeta, \zeta^2, \dots, \zeta^{N-1})$.

Proposition 2.2.6 showed that if T is a shift-circulant, then \hat{T} is a PD-matrix (one with a single non-zero entry in every row and column). We also know that the shift-circulants normalize the circulants. So it is no surprise to find that the full normalizer of the circulants is in fact all matrices similar to monomial matrices under similarity transformation by F . This is proved in Proposition 2.5.1 below.

Recall that our standard basis for the circulant matrices are powers of the matrix $U = \text{CIRC}(0, 1, 0, \dots, 0)$. The next lemma shows that D is in the normalizer of the circulants if and only if U^D is circulant. ■

2.4.2 **Lemma** $D \in \text{Norm}_S \text{CIRC}_N(R) \Leftrightarrow D^{-1}UD \in \text{CIRC}_N(R)$.

Proof. $D^{-1}UD \in \text{CIRC}_N(R) \Leftrightarrow D^{-1}U^r D \in \text{CIRC}_N(R) \Leftrightarrow D^{-1} \left(\sum_{r \in \mathbb{Z}_N} a_r U^r \right) D \in \text{CIRC}_N(R)$. □

2.5.1 Proposition A necessary and sufficient condition for $T \in \text{Norm}_S \text{CIRC}_N(R_\zeta)$ is that $\widehat{T} = P_\sigma \text{Diag}(\theta)$ for some permutation $\sigma \in \mathcal{S}_N$ and some diagonal matrix $\text{Diag}(\theta_0, \theta_1, \dots, \theta_{N-1}) \in \text{GL}_N(S)$.

Proof. By the lemma, the proposition need only be proved for U .

Assume $T \in \text{Norm}_S \text{CIRC}_N(R)$, then $T^{-1}UT = C$ for some circulant matrix C . $\therefore UT e_i = TC e_i = \lambda_i(C)T e_i$. So $T e_i$ is an eigenvector of U . But, the eigenvectors of U are $\{e_0, e_1, \dots, e_{N-1}\}$. Hence, T maps the set of eigenvectors $\{e_0, e_1, \dots, e_{N-1}\}$ to itself, possibly with scalar multipliers in S . Since T is non-singular, T must permute $\{e_0, e_1, \dots, e_{N-1}\}$ within scalar multipliers. Let the permutation be σ , and let the scalar multipliers be $\{\theta_0, \theta_1, \dots, \theta_{N-1}\} \subset S$. Then, for all j ,

$$T e_j = \theta_j e_{\sigma(j)} \quad (1)$$

This says that T maps the eigenspace, columnar, basis vector e_j to θ_j times the columnar basis vector $e_{\sigma(j)}$. So the matrix representation of T in the eigenspace basis, that is \widehat{T} , is given by

$$\begin{aligned} (\widehat{T})_{i,j} &= (\theta_j e_{\sigma(j)})_i = \theta_j e_{i-\sigma(j)} \\ \therefore \widehat{T} &= P_\sigma \text{Diag}(\theta_0, \theta_1, \dots, \theta_{N-1}) \end{aligned}$$

QED Necessity.

Let $\widehat{T} = P_\sigma \text{Diag}(\theta)$. Reversing the above proof of sufficiency, we arrive back at equation (1). Thence,

$$\begin{aligned} \therefore T^{-1}UT e_j &= T^{-1}\theta_j U e_{\sigma(j)} = T^{-1}\theta_j \zeta^{\sigma(j)} e_{\sigma(j)} = \theta_j \zeta^{\sigma(j)} T^{-1} e_{\sigma(j)} \\ &= \theta_j \zeta^{\sigma(j)} \theta_j^{-1} e_j \quad \text{applying (1) again} \\ \therefore T^{-1}UT e_j &= \zeta^{\sigma(j)} e_j \end{aligned} \quad (2)$$

Therefore, e_j is an eigenvector of $T^{-1}UT$, $\forall j$, and so, the matrix constructed from these eigenvectors, one whose every j^{th} column is e_j , is a diagonalizing matrix for $T^{-1}UT$. But this diagonalizing matrix is F . Therefore, by Theorem 1.5(iii), $T^{-1}UT \in \text{CIRC}_N(R_\zeta)$.

QED Sufficiency. \square

The PD matrices have interesting algebraic properties beyond those discussed here. For more details on PD-matrices and their use in investigating s -circulants see [Dav3]. The matrices of interest to us are not immediately the PD-matrices but the PD inversely transformed by the Fourier matrix, $FPDF^{-1}$. Hence, the following definition.

2.5.2 Definition For any permutation $\sigma \in \mathcal{S}_N$, and vector $\theta \in \mathbb{C}^N$, define $\widehat{P}(\sigma, \theta) := FP_\sigma \text{Diag}(\theta)F^{-1}$. In the special case when $\text{Diag}(\theta)$ is the identity matrix, we shall write \widehat{P}_σ . Thus, $\widehat{P}_\sigma = FP_\sigma F^{-1}$.

We shall show that the precise values appearing in the diagonal matrix are irrelevant to automorphisms on circulants; hence, in typifying automorphisms, we can actually take the diagonal matrix to be the identity.

Note that the proof of Proposition 2.5.1 fails to characterize the normalizer of $\mathbf{circ}_N(R)$ because the matrix $T^{-1}UT$ in general has entries in R_ζ not R . However, Proposition 2.5.1 does determine the normalizer of any subset of $\mathbf{circ}(R_\zeta)$ which is defined by the value of the determinant such as for instance, the non-singular complex circulant matrices. This is because the determinant is unaffected by similarity transformations. That is, $\det A = \det(T^{-1}AT)$. This gives us an easy corollary.

2.5.3 Corollary Let $\widehat{P}(\mathcal{S}_N, R^*)$ be the set of all $\widehat{P}(\sigma, \theta)$ where $\sigma \in \mathcal{S}_N$, and $\theta \in R_*^N$ where $R_* = R - \{0\} \subset \mathbb{C}$ -- i.e. θ has no zero components. Then,

- (i) $\widehat{P}(\mathcal{S}_N, R^*)$ is the normalizer in GL_N of $\text{GL}_N(\mathbb{C}) \cap \text{CIRC}_N(\mathbb{C})$ and $\text{SL}_N(\mathbb{C}) \cap \text{CIRC}_N(\mathbb{C})$, and
- (ii) $\widehat{P}(\mathcal{S}_N, R^*)$ is a multiplicative group.

Proof. Part (i) is just Proposition 2.5.1 with Definition 2.5.2, and part (ii) follows from (i) by fact that the normalizer of any set is a group. \square

The group product in $\widehat{P}(\mathcal{S}_N, R)$ is not quite as simple as the composition of permutations. There is a twist in the product whenever the matrix on the right of the product is not circulant. Hence, even an abelian group of permutations does not define an abelian subgroup of $\widehat{P}(\mathcal{S}_N, R)$. This is shown next.

2.5.4 Proposition

- (i) $\widehat{P}(\sigma, \theta)\widehat{P}(\tau, \phi) = \widehat{P}(\sigma\tau, \beta)$ where $\beta_i = \theta_{\tau(i)}\phi_i$.
(ii) $\widehat{P}(\sigma, \theta)$ is unitary iff $|\theta_i| = 1, \forall i$.

Proof.

(i) We can determine the product rule by calculating the product $F^{-1}\widehat{P}(\sigma, \theta)\widehat{P}(\tau, \phi)F$ which by definition equals $P_\sigma \text{Diag}(\theta)P_\tau \text{Diag}(\phi)$.

$$\begin{aligned} (P_\sigma \text{Diag}(\theta)P_\tau \text{Diag}(\phi))_{i,m} &= \sum_{j,k,l} \delta_{i-\sigma(j)}\delta_{j-k}\theta_k\delta_{k-\tau(l)}\delta_{l-m}\phi_m \\ &= \sum_k \delta_{i-\sigma(k)}\theta_k\delta_{k-\tau(m)}\phi_m \\ &= \delta_{i-\sigma\tau(m)}\theta_{\tau(m)}\phi_m \\ &= (P_{\sigma\tau} \text{Diag}(\beta))_{i,m} \quad \text{where } \beta_m = \theta_{\tau(m)}\phi_m \text{ as required.} \end{aligned}$$

QED (i).

- (ii) $P_\sigma^\dagger = P_\sigma^{-1}$ for all permutation matrices P_σ . Hence, $\widehat{P}(\sigma, \theta)^\dagger = \widehat{P}(\sigma^{-1}, \bar{\theta}) = \widehat{P}(\sigma, \theta)^{-1}$ iff $\theta\bar{\theta} = 1$.
□

Let \widehat{P} be any member of $\widehat{P}(\mathcal{S}_N, \mathbb{C})$, and let $C = \widehat{P}^{-1}U\widehat{P}$. Then, C , of course, is circulant. As in the proof of Proposition 2.5.1, we see that $U\widehat{P}_\sigma e_i = \lambda_i(C)\widehat{P}_\sigma e_i$ which shows that the eigenvalues of C is a permutation of the eigenvalues of U . Since $P_\sigma e_i = e_{\sigma(i)}$, the permutation of the eigenvalues of U is none other than σ . One can easily generalize this to any matrix of the form $C = \widehat{P}^{-1}B\widehat{P}$ where B is any circulant matrix. If we do so we will find that the eigenvalues of C are again just of those of B permuted by σ . This strongly suggests that the automorphism $B \rightarrow \widehat{P}^{-1}B\widehat{P}$ where $\widehat{P} = \widehat{P}(\sigma, \theta)$ is independent of θ .

For any non-singular matrix A , let $\iota A : M_N \rightarrow M_N$ denote the similarity transformation,

$$\iota A : X \mapsto A^{-1}XA$$

2.5.5 **Lemma** Let $\widehat{P} = \widehat{P}(\sigma, \theta)$ then the automorphism $\iota\widehat{P} : \text{CIRC}_N \rightarrow \text{CIRC}_N$ is independent of θ .

Proof. Let $C = (\iota\widehat{P})(U)$. From equation (2) and the above remarks, $\lambda(C)$ and hence C is independent of θ . So, $(\iota\widehat{P})(U)$ is independent of θ , and so is U . Hence, $(\iota\widehat{P})(U^i)$ is independent of θ . Therefore, $\iota\widehat{P}$ is independent of θ on CIRC_N . □

However, the vector θ is not wholly arbitrary; it may not contain a zero component otherwise $\widehat{P}(\sigma, \theta)$ would be singular.

Because of Lemma 2.5.5 and Proposition 2.5.4, in dealing with inner-automorphism, $\iota\widehat{P}$, we can take θ as the vector $(1, 1, \dots, 1)$ which is tantamount to setting the diagonal matrix in the PD-matrix to be the identity. That is, we can always take $\widehat{P}(\sigma, \theta) = \widehat{P}_\sigma$ in similarity transformations on circulants. Hence, the automorphisms of $\text{CIRC}_N(R)$ induced by inner-automorphisms of $\text{GL}_N(R)$ are just permutations of the eigenvalues. The question now is whether distinct permutations lead to distinct automorphisms. But, the answer to this question is obvious. Any subalgebra of $\text{CIRC}(R)$ which contains a matrix with N distinct eigenvalues must be transformed non-trivially by \widehat{P}_σ unless σ is the identity permutation. This finally gives us the complete characterization of the similarity-induced automorphisms on circulant matrices over rings which contain ζ .

2.5.6 Proposition The group of distinct automorphisms of $\text{CIRC}_N(R_\zeta)$ generated by inner automorphisms of $\text{GL}_N(R_\zeta)$ is identical to the group of permutation matrices acting on $\lambda\text{CIRC}_N(R_\zeta)$. In particular, the group is isomorphic to S_N .

2.6 The Linear Automorphisms of $\text{CIRC}_N(R_\zeta)$.

We have seen that all automorphisms on circulants over R_ζ arising from similarity transforms are essentially just permutations of the circulants' eigenvalues. This must apply to any circulant ring. The only difference when $\zeta \notin R$ will be that some eigenvalue permutations will map a circulant into a circulant matrix with some entries not in R . It is an opportune moment to consider what type of automorphisms we might be interested in, regardless of the base ring. We shall argue that we are interested only in linear homomorphisms, and in linear automorphisms in particular.

First we define what we mean by linear. Informally, a homomorphism on $\text{CIRC}_N(R)$ is linear when the homomorphism is a trivial homomorphism on the base ring, R . If the homomorphism maps $\text{CIRC}(R)$ to another algebra over the ring R then the definition of linear is simple: α is linear iff $\alpha(rA) = r\alpha(A)$, $\forall A \in \text{CIRC}(R)$, $\forall r \in R$. Homomorphisms which have a non-trivial action on the base ring are certainly of interest to general ring theory. But, in a study of circulants, as opposed to general ring theory, the description of the linear circulant homomorphisms must be paramount, and so the actions of homomorphisms on general rings R are left to other texts.

The formal definition of linear depends on the following fact: Given R is a subring of S and that $A(R)$ is an R -algebra, then there exists an S -algebra $A(S)$ which contains $A(R)$. Informally, this can be seen by assuming that $A(R)$ has an R -basis $\{v_1, v_2, \dots, v_n\}$, say. Then, $A(S)$ consists of linear sums $s_1v_1 + \dots + s_nv_n$ where $s_i \in S$. (Formally, $A(S)$ is identified with the tensor product $S \otimes_R A(R)$.) In the case of circulant algebras, $\text{CIRC}_N(S)$ is defined by Definition 1.2.2 for all rings S , and indeed, it is easy to see that $\text{CIRC}_N(S)$ can be constructed as explained above from $\text{CIRC}_N(R)$, for instance, by taking the standard basis $\{I, U, U^2, \dots, U^{N-1}\}$ for $\text{CIRC}_N(R)$ and turning it into an S -basis for $\text{CIRC}_N(S)$.

2.6.1 Definition Let α be a ring homomorphism from $\text{CIRC}_N(R)$ to an R -algebra, $A(R)$.

- (i) α is said to be R -linear if $\alpha(ra) = r\alpha(a)$ for all $r \in R$ and all $a \in \text{CIRC}_N(R)$.
- (ii) If R is a subring of S then α is said to be S -linear if there exists an extension $\bar{\alpha}$ of α which is S -linear, and maps $\text{CIRC}_N(S)$ to the S -algebra $A(S)$.
- (iii) α is said to be **linear** if α is S -linear for all rings S containing (or equal to) R .

An R -linear ring homomorphism on $\text{CIRC}_N(R)$ is an algebra homomorphism on the algebra of circulant matrices over the base ring R . If the homomorphism maps the identity to the identity then it must map rI to $r1_A$ for all $r \in R$ where 1_A is the identity in $A(R)$. Conversely, any homomorphism which maps every rI to $r1_A$ must be R -linear.

Suppose $\alpha : \text{CIRC}_N(R) \rightarrow A(R)$ is an algebra homomorphism, that is, an R -linear ring homomorphism. Let R be a subring of S . Then, α can be extended to $\bar{\alpha} : \text{CIRC}_N(S) \rightarrow A(S)$ by defining

$$\bar{\alpha} \left(\sum_{i \in \mathbb{Z}_N} a_i U^i \right) := \sum_{i \in \mathbb{Z}_N} a_i \alpha(U)^i$$

From this one sees first that $\alpha : \text{CIRC}_N(R) \rightarrow A(R)$ is R -linear iff it is linear, and secondly that a linear map is completely specified by its action on U . A good example of an R -linear homomorphism on $\text{CIRC}_N(R)$ is the eigenvalue map, λ .

It is quite easy to invent homomorphisms on $\text{CIRC}_N(R)$ which are not R -linear. For instance, let $R = \mathbb{Q}(\sqrt{2})$ and let α be any linear automorphism on $\text{CIRC}_N(R)$. Define β to be the field automorphism on R given by $\beta(r + s\sqrt{2}) = r - s\sqrt{2}$, $\forall r, s \in \mathbb{Q}$. Finally, define $\gamma : \text{CIRC}_N(R) \rightarrow \text{CIRC}_N(R)$ by $\gamma\text{CIRC}_N(a_0, a_1, \dots, a_{N-1}) = \alpha\text{CIRC}_N(\beta(a_0), \beta(a_1), \dots, \beta(a_{N-1}))$. The map γ is a ring automorphism but is not linear. For instance, take α as the identity map, then $\gamma((1 + \sqrt{2})C) = (1 - \sqrt{2})\gamma C$.

The proposition which follows essentially proves that the similarity transformations which we have already characterized for $R = \mathbb{C}$ account for all the linear automorphisms on $\text{CIRC}_N(R)$. In this proposition, we introduce a notation which will be used frequently. Given any circulant automorphism α , let $\tilde{\alpha}$ denote the map $\lambda\alpha\lambda^{-1}$. The map $\tilde{\alpha}$ is the automorphism on the eigenspace, $\tilde{\alpha} : R_\zeta^N \rightarrow R_\zeta^N$ which agrees with the homomorphism α on the circulant space.

2.6.2 Theorem Let R be a complex domain. Let $\alpha : \text{CIRC}_N(R) \rightarrow \text{CIRC}_N(R)$ be a linear ring automorphism. Then, α is a similarity transformation, and $\tilde{\alpha} : \lambda \mapsto P_\sigma\lambda$ for some permutation $\sigma \in \mathcal{S}_N$.

Proof. The assumption of linearity implies that α can be extended through linearity to a ring homomorphism $\bar{\alpha} : \text{CIRC}_N(Q) \rightarrow \text{CIRC}_N(Q)$ where Q is a subfield of \mathbb{C} which includes R and ζ . We shall drop the bar over $\bar{\alpha}$ and regard α as the map on $\mathbf{circ}_N(Q)$.

We shall temporarily regard the circulants as vectors with convolution as the ring product, and, as usual, componentwise multiplication as the ring product in the eigenspace. The homomorphism $\alpha : \mathbf{circ}_N(Q) \rightarrow \mathbf{circ}_N(Q)$ induces the homomorphism $\tilde{\alpha} : Q^N \rightarrow Q^N$ on the eigenspace. Since α is linear so is $\tilde{\alpha}$. Therefore $\tilde{\alpha}$ is a vector space map and so can be represented as a matrix transformation. Let the map be $\tilde{\alpha} : \lambda \mapsto M\lambda$ where $M \in M_N(Q)$. Since this is also a multiplicative map, for all $\lambda, \mu \in Q^N$, we must have

$$M(\lambda\mu) = (M\lambda)(M\mu)$$

where the vector product is componentwise multiplication. Writing this equation out in terms of components, we get

$$\sum_j m_{i,j}\lambda_j\mu_j = \sum_j m_{i,j}\lambda_j \sum_k m_{i,k}\mu_k = \sum_{j,k} m_{i,j}m_{i,k}\lambda_j\mu_k \quad (8)$$

where all summations are over the set \mathbb{Z}_N .

Since $\mu \in Q^N$ is arbitrary, we can pick $\mu_j = \delta_{j-s}$ for any $s \in \mathbb{Z}_N$. With this setting, equation (8) becomes

$$\begin{aligned} \sum_j m_{i,j}\lambda_j\delta_{j-s} &= m_{i,s}\lambda_s = \sum_{j,k} m_{i,j}m_{i,k}\lambda_j\delta_{s-k} = m_{i,s} \sum_j m_{i,j}\lambda_j \\ \therefore m_{i,s}\lambda_s &= m_{i,s} \sum_j m_{i,j}\lambda_j \end{aligned}$$

Suppose $m_{i,s} \neq 0$. Cancelling $m_{i,s}$, we get the equation

$$\lambda_s = \sum_j m_{i,j}\lambda_j = m_{i,s}\lambda_s + \sum_{j \neq s} m_{i,j}\lambda_j$$

Since λ is also arbitrary, we can pick $\lambda_i = \delta_{i-s}$ which forces $m_{i,s} = 1$. Therefore, $m_{i,s} = 0$ or 1 .

Now pick $\mu_j = 1, \forall j$, then equation (8) gives

$$\sum_j m_{i,j}\lambda_j = \sum_{j,k} m_{i,j}m_{i,k}\lambda_j = \left(\sum_k m_{i,k} \right) \left(\sum_j m_{i,j}\lambda_j \right)$$

Again, λ is arbitrary, so we can deduce that either M is the zero matrix or $\sum_k m_{i,k} = 1$. But the first possibility contradicts the bijectivity of α . Therefore, the sum of every column in M equals 1. This condition together with the condition $m_{i,j} = 0$ or 1 , and the fact that α , hence $\tilde{\alpha}$ are bijective implies that M is a permutation matrix. That is, $\tilde{\alpha}$ permutes the components of the eigenvalue vector. The proposition now follows by Proposition 2.5.6. \square

2.6.3 Corollary The group of linear ring automorphisms on $\text{CIRC}_N(\mathbb{C})$ is induced by the group of permutations of the circulant eigenvalues. \square

2.7 The Linear Automorphisms of $\text{CIRC}_N(\mathbb{R})$.

We shall now characterize Norm $\text{CIRC}(R)$ where R no longer necessarily contains ζ .

2.7.1 Proposition $\widehat{P}_\sigma \in \text{Norm CIRC}_N(R)$ iff $\widehat{P}_\sigma \in M_N(N^{-1}R)$.

Proof. Let $B(k) := \widehat{P}_\sigma^{-1}U^k\widehat{P}_\sigma$. By equation (6) in Proposition 2.5.1, the eigenvalues of $B(k)$ are $\lambda_i = \zeta^{k\sigma(i)}$. So, $\widehat{B}_{i,j}(k) = \zeta^{k\sigma(i)}\delta_{i-j}$. Applying the inverse eigenvalue map, λ^{-1} ,

$$NB_{i,l}(k) = \sum_{j,k} \zeta^{ij} \zeta^{k\sigma(j)} \delta_{j-k} \zeta^{-kl} = \sum_j \zeta^{j(i-l)+k\sigma(j)} \quad (1)$$

On the other hand, $\widehat{P}_\sigma = FP_\sigma F^{-1}$.

$$\therefore (\widehat{P}_\sigma)_{i,l} = N^{-1} \sum_{j,k} \zeta^{ij} \delta_{\sigma(j)-k} \zeta^{kl} = N^{-1} \sum_j \zeta^{ij+l\sigma(j)} \quad (2)$$

Therefore,

$$\begin{aligned} P_\sigma \in \text{Norm CIRC}_N(R) &\Leftrightarrow B(k) \in \text{CIRC}_N(R), \quad \forall k \in \mathbb{Z}_N \\ &\Leftrightarrow N \sum_j \zeta^{jr+k\sigma(j)} \in R, \quad \forall r, k \in R, \quad \text{by (1)} \\ &\Leftrightarrow N\widehat{P}_\sigma \in M_N(R), \quad \text{by (2)} \quad \square \end{aligned}$$

Hence, if $F \subset E$ are fields, then the restriction of an inner automorphism of $\text{CIRC}(E)$ to $\text{CIRC}(F)$ is an inner automorphism on $\text{CIRC}(F)$ - - we get no more automorphisms on $\text{CIRC}(F)$ by extending the field to E .

One easy application of the above proposition is to show that the linear automorphism group of $\text{CIRC}_N(\mathbb{R})$ is identical to $\{\widehat{P}_\sigma \mid \sigma(-i) \equiv -\sigma(i) \pmod{N}\}$. These are the permutations of the eigenvalues which are odd functions on \mathbb{Z}_N . From this we get the next theorem.

2.7.2 Theorem The number of distinct linear automorphisms of $\text{CIRC}_N(\mathbb{R})$ generated by $\text{GL}_N(\mathbb{C})$ is $(\frac{1}{2}(N-1))!$ for N odd and $2(\frac{1}{2}N-1)!$ for N even. The group of these automorphisms is isomorphic to the subgroup J_N of \mathcal{S}_N consisting of all those permutations which satisfy $\sigma(-i) = -\sigma(i)$, for all $i \in \mathbb{Z}_N$.

Proof. As discussed above, the set $\{\widehat{P}_\sigma \mid \sigma \in J_N\}$ is the normalizer of $\text{CIRC}_N(\mathbb{R})$. Since $\alpha_{\widehat{P}_\sigma} = \alpha_{\widehat{P}_\tau} \Leftrightarrow \sigma = \tau$, the automorphism group of Norm $\text{CIRC}_N(\mathbb{R})$ is isomorphic to J_N .

Case I N odd.

The permutation σ which is an odd function on \mathbb{Z}_N is fully specified when one specifies $\sigma(0), \sigma(1), \dots, \sigma(\frac{1}{2}(N-1))$. The remaining values are completely determined by previous choices and the constraint $\sigma(-i) = -\sigma(i)$. The number of free choices is therefore $(\frac{1}{2}(N-1))!$ **QED** Case I.

Case II N even.

In this case, both $\sigma(0)$ and $\sigma(\frac{1}{2}N)$ are constrained to the set $\{0, \frac{1}{2}N\}$. Consequently, there are two choices for $\sigma(0)$ and this determines the value of $\sigma(\frac{1}{2}N)$.

As in case I, the values of σ on $1, 2, \dots, \frac{1}{2}N-1$ is free and thereafter all values are determined. Therefore, the number of free choices is $2 \times (\frac{1}{2}N-1)!$

Lastly, since all linear automorphisms are similarity transforms, it follows that the above characterizes all linear ring automorphisms on $\text{CIRC}_N(\mathbb{R})$. \square

The theorem can be stated more intuitively. Every linear automorphism of $\text{CIRC}(\mathbb{R})$ is a permutation of the eigenvalues which commutes with complex conjugation.

According to Lemma 2.7.1 if $\widehat{P}_\sigma \in \text{Norm CIRC}_N(\mathbb{R})$ then \widehat{P}_σ is a real matrix. Consequently, all the automorphisms of $\text{CIRC}_N(\mathbb{R})$ generated by $\text{GL}_N(\mathbb{C})$ are also generated by $\text{GL}_N(\mathbb{R})$. We get no extra automorphisms by extending the field. Since J_N is a subgroup of \mathcal{S}_N , it follows that $\text{Inn}(\text{CIRC}_N(\mathbb{R}))$ is a subgroup of $\text{Inn}(\text{CIRC}_N(\mathbb{C}))$. However, except for $N \leq 2$, $\text{Inn}(\text{CIRC}_N(\mathbb{R}))$ is not normal in $\text{Inn}(\text{CIRC}_N(\mathbb{C}))$. One can for instance easily show that J_N is not normal in \mathcal{S}_N when $N > 2$.

2.8 The Galois Group and Linear Automorphism of $\text{CIRC}_N(R)$.

We give a preview here of the rôle played by the Galois group of cyclotomic field extensions in subsequent sections. (See Appendix A for a brief overview of cyclotomic theory).

Let R be a complex domain, and let Q be its quotient field. Then, Q_ζ is a cyclotomic extension of Q . Let G be the Galois group for this extension. Concretely, G is the set of field automorphisms of Q_ζ which leave Q fixed. The orbits of G acting on Q_ζ is a partitioning of Q_ζ . Similarity transformation by the matrix \widehat{P}_σ is an automorphism of $\text{CIRC}_N(R)$ if and only if the map $\zeta^i \rightarrow \zeta^{\sigma(i)}$ maps G -orbits into G -orbits. (Note that the roots of unity is a union of G -orbits.)

For example, take $R = \mathbb{R}$, then $Q = \mathbb{R}$ also, and the Galois group is generated by the map $x + iy \leftrightarrow x - iy$. Hence, orbits under G consists of the singleton sets of all real numbers and all sets of pairs of conjugate complex numbers. Let $\sigma(i) = j$. By the condition given above for \widehat{P}_σ to be an automorphism on $\text{CIRC}_N(\mathbb{R})$, if ζ^i is not real then $\sigma(-i)$ must equal $-j$. This gives the constraint stated in Lemma 2.7.1 part (i) except when $i = 0$ or $1/2N$. But, in these latter two cases, $i = -i$ (in \mathbb{Z}_N) anyway. Another important case is when $R = \mathbb{Q}$. Automorphisms which preserve rationality must be permutation of the eigenvalues which map eigenvalues into conjugate eigenvalues. Thus, λ_i can be mapped to any λ_{hi} where $h \in \mathbb{Z}_N^*$. There will be more on this in later chapters.