CIRCULANTS (Extract)

Alun Wyn-jones

Last revised in December 2013.

Chapter 1.
Circulants.

1.1 **Introduction** Circulants have been known to humanity since at least the beginning of the nineteenth century when they were revealed in their original manifestation as circulant determinants. Later in the century, matrices were invented and circulants were reinterpreted as matrices. Later still, matrices became part of a new, formal, and more abstract algebra of the Twentieth Century. Circulants could then be viewed as a special kind of algebra, a sub-algebra of the matrix algebra.

However circulants retain their basic simplicity. One can understand circulants, study them, discover things about them, and take delight in them with but a little background in college algebra. The primary goal of this book is to describe circulants in an algebraic context. However, the older forms cannot be ignored, else the theory presented herein would be merely a special case of several modern algebraic theories. Therefore, much of the book is concerned with old problems, especially those parts dealing with the circulant determinant. Consequently, the book oscillates between the point of view of circulants as a commutative algebra, and the concrete point of view of circulants as matrices with emphasis on their determinants..

There are many applications of the theory of circulants. Indeed, many researchers have independently rediscovered circulants for this reason. Sometimes a researcher would be unaware of the name "circulant"; indeed one common alternative name was "cyclic matrix."

The applications are mainly in pure mathematics and technology which mysteriously reflects the abstract - concrete dichotomy of the theory of circulants. For instance, modern telecommunications would be impossible without frequency analysis. With the advent of fast digital computing, the main technique of frequency analysis has become the discrete Fourier transform. This transform is also the single most important transform on circulants, so much so, that much of the theory of circulants can be regarded as the theory of the discrete Fourier transform. Circulants are important in digital encoding; this is a wondrous technology –– it enables devices ranging from computers to music players to recover from errors in transmission and storage of data, and restore the original data. However, the initial impetus to the study of circulants was not technological but rather stemmed from problems in pure mathematics, particularly number theory. Several other applications to pure mathematics have since been discovered. Prof. D. L. Johnson has used circulants to analyze cyclically presented abelian groups. Prof. P. Davies and others have described properties of nested polygons as circulant transformations. Circulant graphs are an important sub-field of graph theory. As we proceed, we shall point out applications of circulants to homogeneous diophantine equations and combinatorial analysis. Finally, towards the end of the book, we shall return to the physical sciences with an application of circulants to the evolution of density fluctuations.

Given such an imposing collection of applications, it may be a pleasant surprise to discover that the most general circulant matrix can be described very simply. Circulant matrices are always square. Here is the most general circulant matrix of order N.

$$\begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{N-1} \\ a_{N-1} & a_0 & a_1 & \dots & a_{N-2} \\ a_{N-2} & a_{N-1} & a_0 & \dots & a_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{pmatrix} \tag{1}$$

We shall explain in a moment why we index the entries in the matrix with the numbers 0 through $N-1$ rather than the more conventional 1 through N. We shall generally reserve the capital letter N to denote the order of circulant matrices.

1.2.1 **How to Construct a Circulant Matrix.** A circulant matrix can be constructed from any sequence of N objects, $a = (a_0, a_1, \dots, a_{N-1})$, say, by the following procedure.

Take the sequence $(a_0, a_1, \dots, a_{N-1})$ as the first row of A. For its second row take the sequence $(a_{N-1}, a_0, a_1, \dots, a_{N-2})$ which is the same sequence but rotated once to the right. For the third row,

take the sequence $(a_{N-2}, a_{N-1}, a_0, \ldots, a_{N-3})$ which is a rotation to the right of the second row. Continue in like manner until all the rows are filled. If the last row were rotated, it would repeat the first row.

We could adopt this contruction as our definition of a circulant matrix by insisting that whatever can be built using the construction is a circulant matrix, and that all circulant matrices can be so constructed. Unfortunately, such heuristic rules do not lend themselves to easy analysis. So, instead, we adopt a formal definition which is easily shown to be equivalent to the construction.

1.2.2 **Definition** Let $A = (a_{i,j})$ be an $N \times N$ matrix. A is a **circulant matrix** if and only if

$$a_{i,j} = a_{k,l} \text{ whenever } j - i \equiv l - k \pmod{N}$$

That is, the value of an entry in a circulant matrix depends only the difference of its column and row position modulo N. You can check that the matrix in (1) above satisfies this criterion.

1.2.3 **Exercise.** Show that this definition guarantees that the matrix is circulant in the sense that it can be constructed from the first row by rotations as described in section 1.2.1.

The construction of §1.2.1 shows that a circulant matrix A is completely defined by any row. For definiteness, we shall always regard A as defined by its top row just as in the construction. Because of its appointed rôle, we shall refer to the top row as the **circulant vector**.

1.3 **Definition**
(i) Let $a = (a_0, a_1, \ldots, a_{N-1})$ then the circulant matrix $A = (a_{i,j})_{i,j}$ where $a_{i,j} = a_{j-i \pmod{N}}$ is denoted by $\mathbf{CIRC}_N(a)$.
(ii) The vector a (the top row of the circulant matrix) is called the **circulant vector**

It is clear that the natural indexing set for entries in $N \times N$ circulant matrices is the set of residues modulo N, that is, remainders after division by N. We shall denote the set of residues modulo N by $\mathbb{Z}_N$, and henceforth, it will be the indexing set for entries in most matrices and vectors. Sometimes, there will be subscripts involving products of residues. So, $\mathbb{Z}_N$ should be regarded as the ring of residues, not just the additive group. The reason we indexed the entries in the circulant matrix from zero to $N-1$ rather than 1 to N is because the remainders modulo N contain 0 but not N.

We shall generally (Chapter §9 is the only exception) assume that the entries in a circulant matrix belong to a commutative ring with identity. If we have no specific ring in mind, we shall usually denote it by R. The order of the circulant matrix, N, should not be a divisor of zero in the ring R else most of the ensuing theory will not work. In fact, R can usually be assumed to be a complex domain, that is, a subring of $\mathbb{C}$, and unless otherwise indicated this can be assumed.

1.4 **Definition** Let R be a (commutative) ring. The set of all $N \times N$ circulant matrices over R will be denoted by $\mathbf{CIRC}_N(R)$. Thus, in terms of Definition 1.3,

$$\mathrm{CIRC}_N(R) \;:=\; \{\mathrm{CIRC}_N(v) \vdash v \in R^N\}$$

1.5 **Examples.**
(i) The first example of a circulant matrix is any matrix of the form cI where c is a scalar and I is the identity matrix. In particular, the identity and zero matrices are circulant.
(ii)

$$\mathrm{CIRC}_3(1,\ 2,\ 3) \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \mathrm{CIRC}_3(4,\ 3,\ 2) \;=\; \begin{pmatrix} 4 & 3 & 2 \\ 2 & 4 & 3 \\ 3 & 2 & 4 \end{pmatrix}$$

(iii)

$$\therefore\ \mathrm{CIRC}_3(1,\ 2,\ 3)\mathrm{CIRC}_3(4,\ 3,\ 2) \ = \ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 4 & 3 & 2 \\ 2 & 4 & 3 \\ 3 & 2 & 4 \end{pmatrix} \ = \ \begin{pmatrix} 17 & 17 & 20 \\ 20 & 17 & 17 \\ 17 & 20 & 17 \end{pmatrix}$$
$$= \ \mathrm{CIRC}_3(17,\ 17,\ 20)$$

In example *(iii)*, the product of the two circulant matrices is itself a circulant matrix. This is a general property of circulants as will be proved shortly.

The beginning for most of the ensuing theory is the next theorem (Theorem 1.6). Even though it is easy to prove, its importance to circulant matrices is fundamental, and so some discussion of the theorem is in order.

The theorem is effectively restricted to circulant matrices having entries in a subring of a field which has an extension containing N^{th} roots of unity. The first statement of the theorem says that there is a matrix (denoted by F) which diagonalizes every circulant matrix. That is, given any circulant matrix C, then $F^{-1}CF$ is a diagonal matrix. This is a highly unusual property. A general matrix need not possess any diagonalizing matrix, so the fact that every circulant matrix can be diagonalized is unusual enough. That the single matrix F suffices to diagonalize all circulants is truly exceptional. If we regard the circulant matrices as linear transformations on a complex vector space, then the first statement says that there is a basis for the vector space in which the circulant matrices transform as a set of diagonal matrices. To look ahead a bit, it is easy to see that the set of all diagonal matrices is closed under matrix multiplication, and is also commutative. This shows that multiplication of circulant matrices is commutative, and suggests that it might also be closed. Since circulant matrices are trivially closed under addition, it already emerges that they actually form a commutative ring. We can show that the circulant matrices are closed under multiplication by showing that the set of diagonalized circulant matrices are closed under multiplication. This is effectively accomplished in the third statement of the theorem which shows that inverse diagonalization (whereby diagonal D is mapped to FDF^{-1}) always produces a circulant matrix. To summarize, the transformation $C \mapsto F^{-1}CF = D$ say, acting on circulant matrices, produces a set of diagonal matrices. The property of being a diagonal matrix is obviously preserved by addition and multiplication. The reverse transformation acting on diagonal matrices, $D \mapsto FDF^{-1} = C$ say, produces a set of circulant matrices.

Now, if a matrix C can be diagonalized, the entries down the main diagonal of its diagonalized form are the eigenvalues of C. The second statement of the theorem merely applies this fact to a general circulant matrix to obtain a formula for its eigenvalues. Later, this formula is interpreted as a map, a linear and multiplicative map, on circulants. That the circulant matrices form a ring implies that this map is actually a ring homomorphism. This homomorphism is the single most important map on circulant matrices.

The proof of the theorem uses a notation which will be used throughout the remainder of the book.

$$\delta_x^N \ := \ \begin{cases} 1 & \text{if } x \equiv 0 \pmod{N} \\ 0 & \text{otherwise} \end{cases}$$

When the order, N, is understood or is clear from the context it will often be omitted. (All special terms such as δ_x and N are listed in the Glossary, as well as general notations and conventions.)

1.6 **The Circulant Diagonalization Theorem** Let R be a subring of a field whose characteristic does not divide N, and let ζ be a primitive N^{th} root of unity, if necessary in an extension, R_ζ, of R. Let F be the matrix with entries $F_{i,j} = \sqrt{N^{-1}}\zeta^{ij}$.

(i) F is a simultaneous, unitary, diagonalizing matrix for $\mathrm{CIRC}_N(R)$.

(ii) Let $A = \mathrm{CIRC}_N(a) \in \mathrm{CIRC}_N(R)$ then the eigenvalues of A are

$$\lambda_j \ = \ \lambda_j(A) \ = \ \sum_{i \in \mathbb{Z}_N} a_i \zeta^{ij}, \qquad \forall j \in \mathbb{Z}_N.$$

(iii) If $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}) \in R_\zeta^N$ then $F\,\mathrm{Diag}(\mu)F^{-1} \in \mathrm{CIRC}_N(R_\zeta)$.

Proof. Let $A = \mathrm{CIRC}_N(a)$ and let F be as stated. Consider $F^\dagger AF$.

$$\begin{aligned}
(F^\dagger AF)_{i,l} &= N^{-1} \sum_{j,k\in\mathbb{Z}_N} \zeta^{-ij} a_{k-j} \zeta^{kl} \\
&= N^{-1} \sum_{j,s\in\mathbb{Z}_N} a_s \zeta^{(s+j)l-ij} \quad (\text{setting } s = k-j) \\
&= N^{-1} \sum_{s\in\mathbb{Z}_N} a_s \zeta^{sl} \sum_{j\in\mathbb{Z}_N} \zeta^{j(l-i)} \\
&= N^{-1} N \delta_{i-l} \sum_{s\in\mathbb{Z}_N} a_s \zeta^{sl} \\
\therefore F^\dagger AF &= \mathrm{Diag}(\lambda_0, \lambda_1, \ldots, \lambda_{N-1}), \quad \text{where } \lambda_j = \sum_{i\in\mathbb{Z}_N} a_i \zeta^{ij}
\end{aligned} \tag{1}$$

To show that F is a diagonalizing matrix, we shall prove that $F^\dagger = F^{-1}$. Equations (1) are valid for any circulant matrix over R, so substitute the identity matrix, I, for A. It is trivial that I is circulant and that $\lambda_j(I) = 1,\ \forall j \in \mathbb{Z}_N$. Therefore, $F^\dagger F = \mathrm{Diag}(1, 1, \ldots, 1) = I$. QED *(i)* and *(ii)*.

(iii) Let $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}) \in R_\zeta^N$ and let $A = (a_{i,j}) = F\,\mathrm{Diag}(\mu)F^{-1}$.

$$a_{i,l} \;=\; N^{-1} \sum_{j,k\in\mathbb{Z}_N} \zeta^{ij} \mu_j \delta_{j-k} \zeta^{-kl} \;=\; N^{-1} \sum_{j\in\mathbb{Z}_N} \zeta^{-j(l-i)} \mu_j$$

which therefore depends only on $l - i \pmod N$ and so, by definition 1.3, A is a circulant matrix. □

We should emphasize that Theorem 1.6 applies only to integral domains whose characteristic does not divide the order of the circulant. So fundamental is the theorem that all ensuing development will be restricted to such rings, and indeed large sections will be further restricted to subrings of the complex numbers.

The matrix F is called the **Fourier matrix**. When the order of the matrix is not clear, we shall denote the Fourier matrix of order $n \times n$ by F_n. Thus, $(F_n)_{i,j} = \sqrt{n^{-1}}\zeta_n^{ij}$ where ζ_n is a primitive n^{th} root of unity. For a given n, there are $\phi(n)$ primitive roots of unity where $\phi(n)$ is the Euler (or totient) function. The choice of ζ_n is arbitrary. In the case of a complex domain, we shall take $\zeta_n = e^{2\pi\iota/n}$. In the case of rings of finite characteristic we just assume that one primitive root can be singled out.

The symbol ζ will be reserved throughout for a primitive root of unity, often an N^{th} primitive root of unity. The symbol λ will be reserved for the eigenvalues of a circulant matrix (though when several circulant matrices are present we might use other Greek letters such as μ). The eigenvalues of a general matrix are not usually presented to us in any natural order. Of course, we can impose an order on them, but this is usually arbitrary. However, once a primitive root of unity has been chosen, the eigenvalues of circulant matrices have a natural order, namely, $\lambda_0, \lambda_1, \ldots, \lambda_{N-1}$, which is the order given in the theorem. We can therefore regard λ as a function acting on a circulant matrix yielding a sequence of N complex numbers --the N eigenvalues of the circulant matrix. Thus, λ is seen to be a map from the set of circulant matrices into an N-dimensional complex vector space. This is formally stated in the next definition.

1.7 **Definition** The following definitions apply to $\mathrm{CIRC}_N(R)$ where R is an integral domain whose characteristic does not divide N.

(i) R_ζ and $R(\zeta)$ will denote same thing, namely, the smallest ring containing both R and a primitive N^{th} root of unity, ζ. Thus, if R contains a primitive N^{th} root of unity, then $R_\zeta = R$

(ii) Let $A = \mathrm{CIRC}_N(a_0, a_1, \ldots, a_{N-1})$. For every $j \in \mathbb{Z}_N$, define the map $\lambda_j : \mathrm{CIRC}_N(R) \to R_\zeta$ by

$$\lambda_j(A) \;:=\; \sum_{i\in\mathbb{Z}_N} a_i \zeta^{ij}$$

$\lambda_j(A)$ is the j^{th} eigenvalue of the circulant matrix A.

(iii) Now define $\lambda(A)$ to be the vector of all the eigenvalues in their natural order.

$$\lambda(A) \;:=\; (\lambda_0(A),\; \lambda_1(A),\; \ldots,\; \lambda_{N-1}(A))$$

In formal notation, λ is the map $\quad \lambda_0 \oplus \lambda_1 \oplus \cdots \oplus \lambda_{N-1} \;:\; \mathrm{CIRC}_N(R) \;\rightarrow\; R_\zeta^N$.

With these notations, $\lambda_i(A)$ and $\lambda(A)_i$ denote the same thing: the i^{th} component of $\lambda(A)$ which is the i^{th} eigenvalue of A.

(iv) We shall denote the image of $\mathrm{CIRC}_N(R)$ under λ by $\Lambda_{N,\zeta}(R)$. Thus, λ maps $\mathrm{CIRC}_N(R)$ onto $\Lambda_{N,\zeta}(R)$. We shall almost always omit the ζ in the subscript and write just $\Lambda_N(R)$, or even $\Lambda(R)$.

1.8 **Corollaries of Theorem** 1.6

The theme of the corollaries which follow is that the map λ is a ring homomorphism. However, before showing that λ is a ring homomorphism it is wise to first ensure that the range of λ is a ring.

There is always considerable leeway in choosing the range of a map. Of course, the range must contain the image, which in our case is $\Lambda(R)$, but is otherwise arbitrary. For circulant matrices with complex entries, it would therefore seem natural to take the range of λ as the N-dimensional vector space $\mathbb{C}^N$. However, most circulants of interest have real, algebraic numbers, rational numbers, or integer entries, and not general complex numbers. Let us take the set of integral circulant matrices, $\mathrm{CIRC}(\mathbb{Z})$, as an example. To take $\mathbb{C}^N$ as the range of λ acting on $\mathrm{CIRC}(\mathbb{Z})$ misses all the subtleties of cyclotomic and rational integer arithmetic which makes the set $\mathrm{CIRC}(\mathbb{Z})$ interesting in the first place. So why not take the range as the image itself? One practical reason is that it is not at all easy to characterize the set $\Lambda_N(\mathbb{Z})$. Instead, we could try the set $\mathbb{Z} \oplus \mathbb{Z}_\zeta^{N-1}$ (the direct sum of $\mathbb{Z}$ and $\mathbb{Z}_\zeta$ with itself $N-1$ times). The first component reflects that fact that λ_0 is just the sum of the circulant vector and so is always in the base ring, in this case $\mathbb{Z}$. But this choice is too *ad hoc*. This set does indeed contain $\Lambda_N(\mathbb{Z})$, but it is not homogenous -- the first component differs from all the others. Instead, we prefer to take the even simpler set $\mathbb{Z}_\zeta^N$ as our range. As a range, $\mathbb{Z}_\zeta^N$ has the great advantage that it can be homogenously extended to the vector space, $\mathbb{Q}_\zeta^N$. Lastly, and most importantly, $\mathbb{Q}_\zeta^N$ is the minimum vector space on which the Fourier matrix acts as a linear transform in the case $R = \mathbb{Z}$.

More generally, we take the range of λ acting on $\mathrm{CIRC}(R)$ to be the set R_ζ^N. $\Lambda(R)$ is a subset of the set R_ζ^N. The set R_ζ^N is made into a ring by defining addition and multiplication componentwise. Thus, for $x, y \in R_\zeta^N$,

$$x + y \;:=\; (x_0 + y_0,\; x_1 + y_1,\; \ldots,\; x_{N-1} + y_{N-1})$$
$$xy \;:=\; (x_0 y_0,\; x_1 y_1,\; \ldots,\; x_{N-1} y_{N-1})$$

We can go further and allow scalar multiplication: for all $c \in R_\zeta$, $cx = (cx_0,\; cx_1,\; \ldots,\; cx_{N-1})$. This makes R_ζ^N into an algebra over R_ζ. If R_ζ is a field, then R_ζ^N will be a vector space over R_ζ. If R is not a field, we can still find a range for λ which is a vector space by taking instead the minimum field which contains R. Let Q be this field. Q is called the **quotient field** of R and can always be constructed when R is an integral domain (see [Kap1] or [Lang]). When Q exists, Q_ζ becomes a field extension of Q, and is therefore the quotient field of R_ζ. Thus, given a domain R we can construct a vector space Q_ζ^N which contains $\Lambda(R)$ which in a sense, is the smallest vector space which contains $\Lambda(R)$.

Usually, we take R_ζ^N as the range of λ. When it is important that the range be a vector space then we extend the range to Q_ζ^N. In either case, we refer to the range of λ as the **circulant eigenspace** or just the **eigenspace** of $\mathrm{CIRC}_N(R)$ (strictly it is the "eigenvaluespace", but this is too big a mouthful). In this nomenclature, an element of Q_ζ^N is an "eigenspace vector." The reader should beware of confusing this with the space of eigenvectors which is the space spanned by the common eigenvectors of the circulant matrices. (The circulant eigenvectors are the columns of the Fourier matrix.)

Let us now get to the corollaries of Theorem 1.6. In these corollaries, R is an integral domain, and Q is a field whose characteristic does not divide N, the order of the circulants under discussion.

1.8.1 **Corollary** $\lambda : \mathrm{CIRC}_N(Q_\zeta) \to Q_\zeta^N$ is an algebra isomorphism.

Proof. The map λ is a non-singular, linear map between vector spaces of the same dimension over the same field. Hence, λ is a vector space isomorphism.

By parts *(i)* and *(ii)* of the theorem, λ is equivalent to the map $A \mapsto \mathrm{Diag}^{-1}F^{-1}AF$ where Diag^{-1} is an inverse diagonal map which maps a matrix to the vector of its diagonal entries. The similarity transform $\alpha_F : A \mapsto F^{-1}AF$ is certainly a matrix ring isomorphism. It remains to prove that Diag^{-1} is a ring isomorphism when restricted to the set $\alpha_F(\mathrm{CIRC}_N(Q))$. But this is a set of diagonal matrices by the theorem. It is easy to verify that the map Diag when restricted to diagonal matrices is indeed a ring isomorphism provided addition and multiplication are taken componentwise in its domain. □

1.8.2 **Corollary** $\mathrm{CIRC}_N(Q_\zeta)$ is a commutative algebra over Q_ζ.

Proof. Q_ζ^N with componentwise addition and multiplication is a commutative algebra over Q_ζ. Apply the inverse eigenvalue map, λ^{-1}, and we see that $\mathrm{CIRC}_N(Q_\zeta) = \lambda^{-1}Q_\zeta^N$ is a commutative algebra over Q_ζ with matrix addition and multiplication. □

1.8.3 **Corollary** $\mathrm{CIRC}_N(R)$ is a commutative algebra over R.

Proof. Let $A, B \in \mathrm{CIRC}_N(R)$, and let $c \in R$. Then, clearly, $A + B$, AB, and cA are matrices with entries in R. By the previous corollary, these matrices are also in $\mathrm{CIRC}_N(Q_\zeta)$ where Q is the quotient field for R. Hence, they are in $\mathrm{CIRC}_N(R)$. □

1.8.4 **Corollary** $\lambda : \mathrm{CIRC}_N(R) \to \Lambda_N(R)$ is a ring isomorphism. □

One last, but important, corollary of Corollary 1.8.1 is the existence of the inverse map λ^{-1}.

1.8.5 **Corollary** Let $\mu \in \Lambda_N(R)$. Then, $\lambda^{-1}(\mu) = A$ is a circulant matrix given by

$$A_{i,j} \;=\; \frac{1}{N}\sum_{k=0}^{N-1} \mu_k \zeta^{(i-j)k}$$

Proof. One can verify the formula by direct substitution into the formula for λ. □

1.9 Circulant Vectors.

Recall that the circulant vector of a circulant matrix is simply the top row of the matrix. The question now naturally arises whether the circulant vectors also form a commutative algebra. Specifically, we ask:

> Can we define addition and multiplication on vectors in R^N so that the map
> $\mathrm{CIRC} : R^N \to \mathrm{CIRC}_N(R)$ is a ring isomorphism?

If we construct the circulant matrix with top row $a_0, a_1, \ldots, a_{N-1}$ and add it to the circulant matrix whose top row is $b_0, b_1, \ldots, b_{N-1}$, we should quite obviously get the circulant matrix whose top row is $a_0{+}b_0,\ a_1{+}b_1,\ \ldots,\ a_{N-1}{+}b_{N-1}$. So, we expect that addition be componentwise addition.

However, to agree with the matrix multiplication in $\mathrm{CIRC}_N(R)$, the multiplication of two circulant vectors a and b must be taken as the **convolution** $a * b$ which is defined by

$$(a * b)_i \;:=\; \sum_{j\in\mathbb{Z}_N} a_j b_{i-j} \qquad \forall a, b \in R^N,\ \forall i \in \mathbb{Z}_N$$

In the analysis of power series products, this is sometimes called the Cauchy product, but we shall always call it the convolution.

So the answer to our question is to take addition of circulant vectors as componentwise addition, and multiplication of circulant vectors as the convolution. We shall verify in Proposition 1.9.2 that these operations together with the usual scalar multiplication of vectors makes the circulant vectors into a commutative algebra, and, as required, it makes CIRC into an algebra isomorphism.

To distinguish the set of circulant vectors from the direct sum R^N, we shall denote the set of circulant N-vectors over the ring R by $\mathbf{circ}_N(R)$. Formally,

1.9.1 **Definition** $\mathbf{circ}_N(R)$ is the set of N-tuples over the ring R endowed with componentwise addition, scalar multiplication, and convolution as a product. It is called the **circulant space** of dimension N over the ring R. When R is an integral domain, $\mathbf{circ}(R)$ is sometimes called the **circulant vector space**.

We can now regard CIRC as a map, $\mathrm{CIRC} : \mathbf{circ}_N(R) \to \mathrm{CIRC}_N(R)$.

1.9.2 **Proposition** $\mathbf{circ}_N(R)$ is a ring, and $\mathrm{CIRC} : \mathbf{circ}_N(R) \to \mathrm{CIRC}_N(R)$ is a ring isomorphism.

Proof. We shall first prove that CIRC is an additive and multiplicative bijection. That $\mathbf{circ}_N(R)$ is a ring then follows trivially.

The map CIRC is a bijection because given any vector the algorithm of §1.2.1 shows how to construct the circulant matrix, and given any circulant matrix, its first row is the circulant vector. So it remains to show that CIRC is preserves sums and products.

Although the additivity of CIRC is quite obvious, we nevertheless present a proof since in this chapter the reader is still becoming acquainted with the terminology.

Let a and b be circulant vectors in $\mathbf{circ}_N(R)$.

$$(\mathrm{CIRC}(a+b))_{i,j} \;=\; \big((a+b)_{j-i}\big)_{i,j} \;=\; \big((a_{j-i}+b_{j-i}\big)_{i,j} \;=\; (\mathrm{CIRC}(a)+\mathrm{CIRC}(b))_{i,j}$$

Also $\mathrm{CIRC}(0) = 0$. This shows that CIRC is an additive map.

It is easy to see that the multiplicative identity in $\mathbf{circ}_N(R)$ (that is the identity of the convolution operator) is the vector $(1, 0, 0, \ldots, 0)$, and CIRC maps this vector to the identity matrix. Hence, CIRC maps the identity to the identity.

Again let a and b be circulant vectors in $\mathbf{circ}_N(R)$ and let $A = \mathrm{CIRC}(a)$, and $B = \mathrm{CIRC}(b)$ be their corresponding circulant matrices. We need to show that $\mathrm{CIRC}(a * b) = AB$.

$$(AB)_{i,j} \;=\; \left(\sum_{k\in\mathbb{Z}_N} A_{i,k}B_{k,j}\right)_{i,j} \;=\; \left(\sum_{k\in\mathbb{Z}_N} a_{k-i}b_{j-k}\right)_{i,j}$$

The top row in the matrix on the right is found by setting $i = 0$. It is the vector

$$\left(\sum_{k\in\mathbb{Z}_N} a_k b_{j-k}\right)_{j} \;=\; (a*b)_j \;\square$$

1.9.3 **Corollary** Let R be a complex domain then $\mathbf{circ}_N(R)$ is a commutative algebra over R.

Proof. $\mathrm{CIRC}_N(R)$ is a commutative algebra with identity. Now CIRC^{-1} is a ring isomorphism and is trivially linear over R. That is, $\mathrm{CIRC}^{-1}(rA) = r\mathrm{CIRC}^{-1}(A)$ for all ring elements r, and circulant matrices, A. Therefore, $\mathbf{circ}_N(R)$ must also be an R-algebra. $\square$

This corollary is true when R is any commutative ring as can easily be proved by direct calculation.

For various reasons it is awkward to use the "$*$" symbol to denote the ring product in circulant vector spaces. For one thing, it is an extra symbol that would appear repeatedly throughout the remainder of the text. For another, repeated convolutions such as $a * a * a$, appear often which we would like to write as a^3. Lastly, the symbol $\mathbf{circ}_N(R)$ means the ring R^N specifically with convolution. For these reasons, the asterisk will be dropped and the ring product in $\mathbf{circ}_N(R)$ will be denoted henceforth by juxtaposition. If this can lead to confusion, the asterisk will be reintroduced but only to remind the reader that the product in $\mathbf{circ}(R)$ is convolution, and not componentwise multiplication.

The connection between a circulant vector and its corresponding circulant matrix is unusually close: either one can very easily be converted to the other, and this conversion is an algebra isomorphism. Therefore, almost any question regarding one can be settled by reference to the other. It is our intention to use whichever provides the most convenient approach for a particular purpose. When we wish to be indiscrimate as to

whether we mean a circulant vector or its circulant matrix, we shall refer to one, both, or either simply as a circulant and we shall refer to the set of circulants as **circulant space**.

So far we have an eigenvalue map defined on circulant matrices but the equivalent map on circulant vectors is the notationally cumbersome $\lambda \circ \mathrm{CIRC} : \mathbf{circ}_N(R) \to R_\zeta^N$. In the spirit of keeping $\mathbf{circ}_N(R)$ and $\mathrm{CIRC}_N(R)$ on par, we shall allow λ to act on $\mathbf{circ}_N(R)$ directly. Thus, $\lambda : \mathbf{circ}_N(R) \to R_\zeta^N$ and is given by

$$\lambda(a) \;=\; \sum_{j \in \mathbb{Z}_N} a_j \zeta^{ij}$$

There will be little danger of confusion between the two uses of λ because we always use capital letters for matrices and lower-case letters for vectors, but if necessary, we can refer to one or the other by $\lambda|\mathbf{circ}$ or $\lambda|\mathrm{CIRC}$.

We are now in a position to appreciate our choice of range for λ. The map $\lambda|\mathbf{circ}_N(Q_\zeta)$ is a vector space endomorphism, $\lambda : Q_\zeta^N \to Q_\zeta^N$, and as such must have a matrix representation. The matrix is obviously $(\zeta^{ij})_{i,j} = \sqrt{N}F$. That is, $\lambda(a) = \sqrt{N}Fa$ where F is the Fourier matrix. The map $\lambda|\mathbf{circ}(\mathbb{C})$ when regarded as a vector space map on $\mathbb{C}^N$ is better known as the **discrete Fourier transform**. (See for instance, [Fla] or [Dav1].) As a simple application of this point of view of circulants as vectors we shall restate Corollary 1.8.5 in more natural language.

1.9.4 Corollary (The Fourier Inversion Formula)

Let $\mu \in \Lambda_N(R)$, then $\lambda = \lambda|_{\mathbf{circ}}$ has an inverse given by $a = \lambda^{-1}(\mu) \in \mathbf{circ}_N(R)$ where

$$a_i \;=\; \frac{1}{N} \sum_{j=0}^{N-1} \mu_j \zeta^{-ij} \qquad \square$$

1.10 Standard Bases for Circulant Space and Eigenspace.

The columns of the Fourier matrix are common eigenvectors of all the circulant matrices. These eigenvectors are

$$e_i \;:=\; \sqrt{N^{-1}}(1, \zeta^i, \zeta^{2i}, \ldots, \zeta^{(N-1)i}) \quad \text{for } i = 0, 1, 2, \ldots, N-1$$

Since the Fourier matrix is unitary, these vectors form an orthonormal basis for the eigenspace R_ζ^N. We shall adopt $e_0, e_1, \ldots, e_{N-1}$ as the standard basis for the eigenspace. But, what should be the standard basis for the circulant space? The natural choice would be circulant vectors which are mapped by λ to the standard basis for R_ζ^N. This basis would be $\lambda^{-1}(e_0), \lambda^{-1}(e_1), \ldots, \lambda^{-1}(e_{N-1})$. Although this would be a perfectly logical choice, there is better. For each $i \in \mathbb{Z}_N$, define $\bar{e}_i = \sqrt{N}e_i$. Hence,

$$\bar{e}_i \;=\; \left(1, \zeta^i, \zeta^{2i}, \ldots, \zeta^{(N-1)i}\right)$$

The set $\{\bar{e}_0, \bar{e}_1, \ldots, \bar{e}_{N-1}\}$ is multiplicatively closed. Recall that the product in the eigenspace is taken componentwise. So, $\bar{e}_i\bar{e}_j = \bar{e}_{i+j}$, and $\bar{e}_1^i = \bar{e}_i$. Of course, all subscripts are residues modulo N. Let $\vec{b}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbf{circ}_N(R)$ where the 1 occurs in position i. It is very easy to check that $\lambda(\vec{b}_i) = (1, \zeta^i, \zeta^{2i}, \ldots, \zeta^{(N-1)i}) = \bar{e}_i$. Since λ is a ring isomorphism, the set of vectors $\{\vec{b}_0, \vec{b}_1, \ldots, \vec{b}_{N-1}\}$ must also be multiplicatively closed under convolution. Indeed, setting $\mathrm{u} := \vec{b}_1$, we see that $\vec{b}_i = \mathrm{u}^i$, the i^{th} power of u. Since the set $\{\bar{e}_0, \bar{e}_1, \ldots, \bar{e}_{N-1}\}$ is a basis for the eigenspace, the set $\{\mathrm{u}^0, \mathrm{u}, \mathrm{u}^2, \ldots, \mathrm{u}^{N-1}\}$ must be a basis for the circulant vectors. But this is obvious since the vectors u^i are just the usual unit orthonormal basis for R^N.

For these reasons, the best choice for a standard basis for the circulant vectors is the set of powers under convolution of the the circulant vector u, namely, $\{\mathrm{u}^0, \mathrm{u}, \mathrm{u}^2, \ldots, \mathrm{u}^{N-1}\}$. To simplify notation, we shall always identify the zeroth power of u with the identity of the base ring R. Thus, $\mathrm{u}^0 = 1 \in R$. In terms of the basis $1, \mathrm{u}, \mathrm{u}^2, \ldots, \mathrm{u}^{N-1}$, an arbitrary circulant vector $a = (a_0, a_1, \ldots, a_{N-1})$ has the expansion

$$a \;=\; a_0 + a_1\mathrm{u} + a_2\mathrm{u}^2 + \ldots + a_i\mathrm{u}^i + \ldots + a_{N-1}\mathrm{u}^{N-1}$$

Define $\mathrm{U} = \mathrm{CIRC}_N(\mathrm{u})$. Then,

$$\mathrm{U} \;=\; \mathrm{CIRC}_N(0,1,0,\ldots,0) \;=\; \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

By the above, the powers of the U matrix form a natural basis for the circulant matrices. Thus, an arbitrary circulant matrix $A = \mathrm{CIRC}_N(a_0, a_1, \ldots, a_{N-1})$ can be expanded as

$$A \;=\; a_0 I + a_1\mathrm{U} + a_2\mathrm{U}^2 + \ldots + a_i\mathrm{U}^i + \ldots + a_{N-1}\mathrm{U}^{N-1}$$

where I is the $N \times N$ identity matrix.

Calculations involving the standard bases are merely the familiar polynomial additions and multiplications, but with all powers of u and U taken modulo N. For example, let $a, b \in \mathbf{circ}_5(\mathbb{Z})$ with $a = (1, 0, -3, 2, -1)$, and $b = (3, -1, 2, 0, 4)$. Then,

$$\begin{aligned} a \;&=\; 1 - 3\mathrm{u}^2 + 2\mathrm{u}^3 - \mathrm{u}^4 \\ b \;&=\; 3 - \mathrm{u} + 2\mathrm{u}^2 + 4\mathrm{u}^4 \\ \therefore\; ab \;&=\; \left(1 - 3\mathrm{u}^2 + 2\mathrm{u}^3 - \mathrm{u}^4\right)\left(3 - \mathrm{u} + 2\mathrm{u}^2 + 4\mathrm{u}^4\right) \\ &=\; 3 - \mathrm{u} - 7\mathrm{u}^2 + 9\mathrm{u}^3 - 7\mathrm{u}^4 + 5\mathrm{u}^5 - 14\mathrm{u}^6 + 8\mathrm{u}^7 - 4\mathrm{u}^8 \\ &=\; 8 - 15\mathrm{u} + \mathrm{u}^2 + 5\mathrm{u}^3 - 7\mathrm{u}^4 \end{aligned}$$

If the example had been the circulant matrices $A = \mathrm{CIRC}(1, 0, -3, 2, -1)$, and $B = \mathrm{CIRC}(3, -1, 2, 0, 4)$ rather than their circulant vectors, then the same calculation with U substituted throughout for u would have shown that $AB = \mathrm{CIRC}(8, -15, 1, 5, -7)$.

1.10.1 **The Representer Polynomial.** Given a circulant $a = \sum_{i=0}^{N-1} a_i\mathrm{u}^i$, define a polynomial $\vartheta(a)$ by $\vartheta(a)(x) := \sum_{i=0}^{N-1} a_i x^i$. It is clear that $\vartheta(a)(x)$ evaluates to the circulant vector $\mathbf{circ}(a)$ at $x = \mathrm{u}$. That is, $\vartheta(a)(\mathrm{u}) = a$. Similarly, $\vartheta(a)(\mathrm{U}) = A$, the circulant matrix. The polynomial $\vartheta(a)$ is called **representer polynomial** † for the circulant a. The above examples show that we can do calculations on circulants using their representer polynomials, and evaluate the resulting polynomial at u to obtain the result of the calculations on circulants.

To summarize,

1.10.2 **Definition**

(i) Let $\mathrm{u} := (0, 1, 0 \ldots, 0) \in \mathbf{circ}_N(R)$, and let $\mathrm{U} := \mathrm{CIRC}_N(0, 1, 0, \ldots, 0)$.
Then, $\{1, \mathrm{u}, \mathrm{u}^2, \ldots, \mathrm{u}^{N-1}\}$ is the standard orthonormal basis for $\mathbf{circ}_N(R)$.
and $\{I, \mathrm{U}, \mathrm{U}^2, \ldots, \mathrm{U}^{N-1}\}$ is the standard orthonormal basis for $\mathrm{CIRC}_N(R)$.

(ii) Let $e_i := \sqrt{N^{-1}}(1, \zeta^i, \zeta^{2i}, \ldots, \zeta^{(N-1)i}) = \sqrt{N^{-1}}\lambda(\mathrm{u}^i) = F\mathrm{u}^i$.
Then, $(e_0, e_1, \ldots, e_{N-1})$ is the standard orthonormal basis for R_ζ^N.

(iii) Let a be the circulant $\displaystyle\sum_{i=0}^{N-1} a_i\mathrm{u}^i$, then its representer polynomial is $\vartheta(a)(x) = \displaystyle\sum_{i=0}^{N-1} a_i x^i$.

† Some authors call it the Hall polynomial, e.g. [Ham], [Lam].

1.11 The Circulant Determinant.

One of the central questions on circulant matrices is the value of their determinants. Indeed, the study of circulants began as a study of their determinants. There are several formulæ for the determinant, each of which has its advantages and disadvantages. Because of frequent need to refer to the determinant, we give it its own symbol.

1.11.1 Definition Let $a \in R^N$ for some (commutative) ring R. Define

$$\Delta_N(a) \;:=\; \det \mathrm{CIRC}_N(a)$$

The last corollary of Theorem 1.6 is a formula for the determinant.

1.11.2 Corollary Let R_ζ be a complex domain, then $\Delta_N(a) = \prod_{j\in\mathbb{Z}_N} \sum_{i\in\mathbb{Z}_N} a_i\zeta^{ij}$

Proof. The determinant is the product of the eigenvalues. □

The above formula for the circulant determinant is an old result, as is the next formula, (for instance, see [Muir1].) but the formula in the next theorem is no less remarkable for that.

1.11.3 Theorem (The Resultant Formula) †

Let $a \in \mathbf{circ}_N(R)$ where R is an integral domain and let $A(x) = \sum_{i=0}^{N-1} a_i x^i \in R[x]$ be the representer polynomial for a of degree d with roots $\alpha_1, \alpha_2, \ldots, \alpha_d$ if necessary in some extension of R. Then,

$$\Delta_N(a) \quad = \quad a_d^N (-1)^{d(N-1)} \prod_{i=1}^{d} (1 - \alpha_i^N)$$

Proof. By Corollary 1.11.2,

$$\Delta_N \quad = \quad \prod_{i=0}^{N-1} \Big(\sum_{j=0}^{d} a_j \zeta^{ij} \Big) \quad = \quad \prod_{i=0}^{N-1} A(\zeta^i)$$

Decompose A into its linear factors, $A(x) = a_d \prod_{j=1}^{d} (x - \alpha_j)$, then

$$\begin{aligned}
\Delta_N \quad &= \quad a_d^N \prod_{i=0}^{N-1} \prod_{j=1}^{d} (\zeta^i - \alpha_j) \\
&= \quad a_d^N \prod_{j=1}^{d} \prod_{i=0}^{N-1} (\zeta^i - \alpha_j) \\
&= \quad a_d^N \zeta^{\frac{1}{2}dN(N-1)} \prod_{j=1}^{d} \prod_{i=0}^{N-1} (1 - \zeta^{-i}\alpha_j) \\
&= \quad a_d^N (-1)^{d(N-1)} \prod_{j=1}^{d} \prod_{i=0}^{N-1} (1 - \zeta^{-i}\alpha_j)
\end{aligned}$$

The product $\prod_i (1 - \zeta^{-i}\alpha_j)$ can be evaluated. Consider the polynomial $f(x) = x^N - \alpha_j^N$. It has N roots given by $\{\alpha_j \zeta^{-i} \vdash i \in \mathbb{Z}_N\}$. Therefore, the product is $x^N - \alpha_j^N$ evaluated at $x = 1$. Substituting $x = 1$ gives the equation in the theorem statement. □

† This theorem is attributed to M.A.Stern by [FG].

There is another way of stating the theorem when the circulant has integer components. Suppose $R = \mathbb{Z}$ and that A is monic, that is, $a_d = 1$. Let the roots of A be $\alpha_1, \alpha_2, \ldots, \alpha_d$. The roots are a multiset of conjugate algebraic integers. Consequently, so is the multiset of their N^{th} powers, $\alpha_1^N, \alpha_2^N, \ldots, \alpha_d^N$, and this multiset generates a sub-domain of $\mathbb{Z}(\alpha_1, \alpha_2, \ldots, \alpha_d)$. Hence, the polynomial $A_N(x) = \prod_{i=1}^{d}(x - \alpha_i^N)$ is a polynomial in $\mathbb{Z}[x]$, and the formula of the theorem can be written as $\pm\Delta = A_N(1) =$ sum of coefficients of the polynomial A_N.

1.11.4 Circulant Determinant Expansions for $N \leq 6$.

The above formulas can be used to deduce the following expansions for the circulant determinant. (For $N \geq 7$ expansions are more easily obtained from formulæ which will be derived in Chapter 10.)

$$\Delta_1(a_0) = a_0$$

$$\Delta_2(a_0, a_1) = \begin{vmatrix} a_0 & a_1 \\ a_1 & a_0 \end{vmatrix} = a_0^2 - a_1^2 = (a_0 + a_1)(a_0 - a_1)$$

$$\begin{aligned} \Delta_3(a_0, a_1, a_2) &= \begin{vmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{vmatrix} \\ &= a_0^3 + a_1^3 + a_2^3 - 3a_0a_1a_2 = (a_0 + a_1 + a_2)(a_0^2 + a_1^2 + a_2^2 - a_0a_1 - a_1a_2 - a_2a_0) \end{aligned}$$

$$\begin{aligned} \Delta_4(a_0, a_1, a_2, a_3) &= \begin{vmatrix} a_0 & a_1 & a_2 & a_3 \\ a_3 & a_0 & a_1 & a_2 \\ a_2 & a_3 & a_0 & a_1 \\ a_1 & a_2 & a_3 & a_0 \end{vmatrix} \\ &= a_0^4 - a_1^4 + a_2^4 - a_3^4 - 4(a_0^2a_1a_3 - a_0a_1^2a_2 + a_1a_2^2a_3 - a_0a_2a_3^2) - 2(a_0^2a_2^2 - a_1^2a_3^2) \\ &= (a_0 + a_1 + a_2 + a_3)(a_0 - a_1 + a_2 - a_3)\left((a_0 - a_2)^2 + (a_1 - a_3)^2\right) \end{aligned}$$

$$\begin{aligned} \Delta_5(a_0, a_1, a_2, a_3, a_4) &= \sum_i a_i^5 + 5\sum_i (a_i^2a_{i+1}^2a_{i+3} + a_i^2a_{i+1}a_{i+2}^2 - a_i^3a_{i+2}a_{i+3} - a_i^3a_{i+1}a_{i+4}) - 5a_0a_1a_2a_3a_4 \\ &= (a_0 + a_1 + a_2 + a_3 + a_4)P(a), \quad \text{where } P(a) \in \mathbb{Z}[a_0, a_1, \ldots, a_4] \end{aligned}$$

$$\begin{aligned} \Delta_6(a_0, a_1, a_2, a_3, a_4, a_5) &= \sum_i (-1)^i \Big\{ a_i^6 + 2a_i^3a_{i+2}^3 - 3a_i^4a_{i+3}^2 \\ &\qquad +6(a_i^3a_{i+1}^2a_{i+4} + a_i^3a_{i+2}a_{i+5}^2 - a_i^4a_{i+1}a_{i+5} - a_i^4a_{i+2}a_{i+4}) \\ &\qquad +12(a_i^3a_{i+1}a_{i+2}a_{i+3} + a_i^3a_{i+3}a_{i+4}a_{i+5}) \\ &\qquad -9a_i^2a_{i+1}^2a_{i+2}^2 \\ &\qquad -18a_i^2a_{i+2}^2a_{i+3}a_{i+5} \\ &\qquad \Big\} \\ &\qquad +9\big(a_0^2a_2^2a_4^2 - a_1^2a_3^2a_5^2\big) \\ &= L_0L_3L_2L_1 \end{aligned}$$

where

$$\begin{aligned} L_0 &= a_0 + a_1 + a_2 + a_3 + a_4 + a_5 = \lambda_0(a_0, a_1, a_2, a_3, a_4, a_5) \\ L_3 &= a_0 - a_1 + a_2 - a_3 + a_4 - a_5 = \lambda_3(a_0, a_1, a_2, a_3, a_4, a_5) \\ L_2 &= (a_0 + a_3)^2 + (a_1 + a_4)^2 + (a_2 + a_5)^2 - (a_0 + a_3)(a_1 + a_4) - (a_1 + a_4)(a_2 + a_5) - (a_2 + a_5)(a_0 + a_3) \\ &= \lambda_2\lambda_4 \\ L_1 &= -2\,a_0a_3 - 2\,a_1a_4 - 2\,a_2a_5 \quad - a_0a_2 - a_1a_3 - a_2a_4 - a_3a_5 - a_4a_0 - a_5a_1 \\ &\quad + a_0a_1 + a_1a_2 + a_2a_3 + a_3a_4 + a_4a_5 + a_5a_0 \quad + a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2 + a_0^2 \\ &= \lambda_1\lambda_5 \end{aligned}$$