

12.9 Circulants and the Discrete Fourier Transform

The spectral theorem of Section 4 gives every normal matrix a unitary of its own, found by solving that matrix's eigenvalue problem. This section does something better for one particular family: it writes down a **single** unitary matrix, before any matrix of the family is named, and that one matrix diagonalizes all of them at once. The family is the circulants, the unitary is built from the n -th roots of unity, and the resulting dictionary between multiplying matrices and multiplying numbers is the discrete Fourier transform.

Two earlier pages pointed here. In Example (Which of these are normal?) (d) a 3×3 matrix whose rows were cyclic shifts of one another turned out to be normal, with the row dot products and the column dot products agreeing entry for entry; Exercise (B2: A circulant, diagonalized) met the same phenomenon again and diagonalized one such matrix by hand. Both times the agreement was described and not explained. It is explained below, and the explanation is the first theorem of the section.

Throughout, the field is $\mathbb{C}$, and **rows, columns, and vector entries are numbered** $0, 1, \dots, n-1$. That is a departure from the rest of the book, and it is worth it: every formula here involves an index read modulo n , and modular arithmetic starts at 0.

12.9.1 Matrices whose rows walk around

Fix $n \geq 1$ and let $\mathbf{S} \in M_n(\mathbb{C})$ be the matrix with $(\mathbf{S})_{jk} = 1$ when $k \equiv j+1 \pmod{n}$ and $(\mathbf{S})_{jk} = 0$ otherwise: a 1 in each position $(j, j+1)$ for $j \leq n-2$, one more in the corner position $(n-1, 0)$, and zeros elsewhere. For $n = 4$,

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Applied to a vector, $\mathbf{S}$ moves every entry one place up and sends the top entry around to the bottom: $\mathbf{S}\mathbf{x} = (x_1, x_2, \dots, x_{n-1}, x_0)$. We call $\mathbf{S}$ the **cyclic shift**. Each row and each column of $\mathbf{S}$ carries exactly one 1, so $\mathbf{S}$ is a permutation matrix (Definition 3.4.2), and shifting n times returns every entry to where it started, so $\mathbf{S}^n = \mathbf{I}$.

Now think of $\mathbf{S}$ as a building block. The powers $\mathbf{I}, \mathbf{S}, \mathbf{S}^2, \dots, \mathbf{S}^{n-1}$ are the n distinct “rotations of the identity”, and a general linear combination of them is the object of this section.

A circulant is a matrix that does not know where the list begins: shift the indices and nothing changes.

Definition 12.9.1. Circulant matrix

Let $n \geq 1$ and let $\mathbf{c} = (c_0, c_1, \dots, c_{n-1}) \in \mathbb{C}^n$. The **circulant** with first row $\mathbf{c}$ is the matrix $\mathbf{C} \in M_n(\mathbb{C})$ whose entries are

$$(\mathbf{C})_{jk} = c_{k-j} \quad (0 \leq j, k \leq n-1),$$

where the subscript $k - j$ is read **modulo** n . The polynomial

$$p_{\mathbf{c}}(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1} \in \mathbb{C}[x]$$

is called the **symbol** of $\mathbf{C}$.

In words: the entry in a given position depends only on how far the column is to the right of the diagonal, counted cyclically. Written out,

$$\mathbf{C} = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix},$$

so each row is the row above it moved one step to the right, with the entry that falls off the right-hand end reappearing on the left. The columns behave the same way going down. Note that the data is one vector of length n , not n^2 numbers: a circulant is a very small object wearing a large costume.

Three examples, simplest first. The vector $\mathbf{c} = (1, 0, \dots, 0)$ gives $\mathbf{C} = \mathbf{I}$, and $\mathbf{c} = (0, 1, 0, \dots, 0)$ gives $\mathbf{C} = \mathbf{S}$; in both cases the definition is checked by reading off $(\mathbf{C})_{jk} = c_{k-j}$, which is 1 exactly when $k = j$, respectively $k = j + 1$, modulo n . The vector $\mathbf{c} = (1, 1, \dots, 1)$ gives the all-ones matrix. The 3×3 matrix of Example (Which of these are normal?) (d) is the circulant with first row $(1, 2, 3)$, and the matrix of Exercise (B2: A circulant, diagonalized) is the circulant with first row $(1, 1, 0)$.

Our running example for the whole section is the 4×4 circulant with first row $(1, 2, 3, 4)$:

$$\mathbf{C} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad p_{\mathbf{c}}(x) = 1 + 2x + 3x^2 + 4x^3.$$

A nonexample by minimal change. Keep the first row $(1, 2, 3, 4)$ but let the entries that fall off the end vanish instead of wrapping around, giving a matrix $\mathbf{B}$ with rows $(1, 2, 3, 4)$, $(0, 1, 2, 3)$, $(0, 0, 1, 2)$, $(0, 0, 0, 1)$. Every diagonal of $\mathbf{B}$ is still constant, and its first row is still $\mathbf{c}$, but $(\mathbf{B})_{10} = 0$ while the definition demands $c_{0-1} = c_3 = 4$. The clause that fails is the one that does all the work here: the subscript must be read modulo n . Matrices with constant diagonals and no wrap-around have their own name and their own theory, and the warning at the end of the section says why they do not fit into this one.

The reason a circulant is worth a name is that it is a polynomial in the shift.

Theorem 12.9.1. Circulants Are Exactly the Polynomials in the Shift

Let $\mathbf{c} \in \mathbb{C}^n$ and let $\mathbf{C}$ be the circulant with first row $\mathbf{c}$, with symbol $p_{\mathbf{c}}$. Then:

- (a) $\mathbf{C} = p_{\mathbf{c}}(\mathbf{S})$; conversely, $q(\mathbf{S})$ is a circulant for every $q \in \mathbb{C}[x]$.
- (b) $\mathbf{C}^*$ is the circulant with first row $(\overline{c_0}, \overline{c_{n-1}}, \overline{c_{n-2}}, \dots, \overline{c_1})$.
- (c) $\mathbf{C}$ is normal, and any two circulants of the same size commute.

Idea. For (a), read off the entries of $\mathbf{S}^m$: it has a 1 in position (j, k) exactly when $k = j + m$ modulo n . So $\mathbf{S}^m$ is the circulant that picks out the coefficient c_m , and summing over m rebuilds $\mathbf{C}$ one coefficient at a time. Once (a) is in hand, (b) and (c) are arithmetic with the single relation $\mathbf{S}^* = \mathbf{S}^{n-1}$: the adjoint of a polynomial in $\mathbf{S}$ is again a polynomial in $\mathbf{S}$, and two polynomials in one matrix always commute (Theorem 0.9.9).

Proof.

- (a) We claim that $(\mathbf{S}^m)_{jk} = 1$ if $k \equiv j + m \pmod{n}$ and $(\mathbf{S}^m)_{jk} = 0$ otherwise, for every $m \geq 0$. For $m = 0$ this says $\mathbf{S}^0 = \mathbf{I}$. Assuming it for m , expand one product: $(\mathbf{S}^{m+1})_{jk} = \sum_i (\mathbf{S})_{ji} (\mathbf{S}^m)_{ik}$, and the only nonzero factor $(\mathbf{S})_{ji}$ sits at $i \equiv j + 1$, so the sum is $(\mathbf{S}^m)_{j+1, k}$, which is 1 exactly when $k \equiv j + 1 + m$. This proves the claim by induction. Hence, for each j and k ,

$$(p_{\mathbf{C}}(\mathbf{S}))_{jk} = \sum_{m=0}^{n-1} c_m (\mathbf{S}^m)_{jk} = c_{k-j},$$

because exactly one value of m in $\{0, \dots, n-1\}$ satisfies $m \equiv k - j \pmod{n}$. So $p_{\mathbf{C}}(\mathbf{S}) = \mathbf{C}$. Conversely, let $q \in \mathbb{C}[x]$. Dividing by $x^n - 1$ (Theorem 6.1.1) gives $q = h \cdot (x^n - 1) + r$ with $\deg r < n$, and $\mathbf{S}^n = \mathbf{I}$ gives $q(\mathbf{S}) = h(\mathbf{S})(\mathbf{S}^n - \mathbf{I}) + r(\mathbf{S}) = r(\mathbf{S})$, which is the circulant whose first row is the coefficient list of r .

- (b) $\mathbf{S}$ is a permutation matrix with real entries, so $\mathbf{S}^* = \mathbf{S}^\top = \mathbf{S}^{-1}$ (Lemma 3.4.7 (b)), and $\mathbf{S}^{-1} = \mathbf{S}^{n-1}$ because $\mathbf{S}^n = \mathbf{I}$. Therefore

$$\mathbf{C}^* = \sum_{m=0}^{n-1} \overline{c_m} (\mathbf{S}^*)^m = \sum_{m=0}^{n-1} \overline{c_m} \mathbf{S}^{-m}.$$

For $m = 0$ the term is $\overline{c_0} \mathbf{I}$, and for $1 \leq m \leq n-1$ we have $\mathbf{S}^{-m} = \mathbf{S}^{n-m}$. So the coefficient of $\mathbf{S}^r$ is $\overline{c_0}$ when $r = 0$ and $\overline{c_{n-r}}$ when $1 \leq r \leq n-1$, which is the stated first row.

- (c) By (a) and (b), both $\mathbf{C}$ and $\mathbf{C}^*$ are polynomials in $\mathbf{S}$, and any two polynomials in the same matrix commute (Theorem 0.9.9). Hence $\mathbf{C}^* \mathbf{C} = \mathbf{C} \mathbf{C}^*$. The same argument applies to two circulants $\mathbf{C} = p(\mathbf{S})$ and $\mathbf{C}' = q(\mathbf{S})$. ■

That settles the coincidence noticed in Example (Which of these are normal?) (d) and again in Exercise (B2: A circulant, diagonalized) (a). It is worth seeing it once at the level of entries, because the entries are what those pages compared. For any $\mathbf{A} \in M_n(\mathbb{C})$, writing $\mathbf{a}_0, \dots, \mathbf{a}_{n-1}$ for its columns and $\mathbf{r}_0, \dots, \mathbf{r}_{n-1}$ for its rows, the definition of the matrix product gives

$$(\mathbf{A}^* \mathbf{A})_{jk} = \langle \mathbf{a}_k, \mathbf{a}_j \rangle, \quad (\mathbf{A} \mathbf{A}^*)_{jk} = \langle \mathbf{r}_j, \mathbf{r}_k \rangle.$$

So “normal” says exactly that the table of column inner products and the table of row inner products agree. Now let $\mathbf{A} = \mathbf{C}$ be a circulant, so that $(\mathbf{C})_{ik} = c_{k-i}$. Re-index the first sum by $t = j - i$ and the second by $t = i - k$; as i runs through all residues modulo n , so does t in each case. This gives

$$\langle \mathbf{a}_k, \mathbf{a}_j \rangle = \sum_{t=0}^{n-1} c_{t+k-j} \overline{c_t} = \langle \mathbf{r}_j, \mathbf{r}_k \rangle,$$

all subscripts modulo n . Both tables depend only on $k - j$, and they are literally the same sum. Cyclic shifting is what makes the re-indexing legal, and that is the whole of the coincidence.

Quick check. Which vectors $\mathbf{c} \in \mathbb{C}^n$ give a circulant that is Hermitian? Answer for $n = 4$ with real entries.

Solution. By Theorem 12.9.1 (b), $\mathbf{C}^* = \mathbf{C}$ forces $\overline{c_0} = c_0$ and $\overline{c_{n-m}} = c_m$ for $1 \leq m \leq n - 1$. For $n = 4$ with real entries this reads $c_1 = c_3$, with c_0 and c_2 unconstrained: the symmetric real 4×4 circulants are exactly those whose first row has the form (c_0, c_1, c_2, c_1) .

12.9.2 The eigenvectors of the shift

Everything now follows from one eigenvalue problem, the one for $\mathbf{S}$. Since $\mathbf{S}^n = \mathbf{I}$, any eigenvalue λ satisfies $\lambda^n = 1$, so the candidates are the n -th roots of unity of Example (Roots of Unity). There are n of them and $\mathbf{S}$ is an $n \times n$ matrix, which is encouraging. The eigenvectors are equally explicit.

Fix

$$\omega := e^{2\pi i/n},$$

so that $1, \omega, \omega^2, \dots, \omega^{n-1}$ are the n distinct n -th roots of unity, and $\omega^r = 1$ if and only if n divides r (Example (Roots of Unity)).

Theorem 12.9.2. Eigenvectors of the Cyclic Shift

For $j = 0, 1, \dots, n - 1$ let

$$\mathbf{f}_j := \frac{1}{\sqrt{n}}(1, \omega^j, \omega^{2j}, \dots, \omega^{(n-1)j}) \in \mathbb{C}^n,$$

that is, $(\mathbf{f}_j)_m = \omega^{jm}/\sqrt{n}$. Then $\mathbf{S}\mathbf{f}_j = \omega^j \mathbf{f}_j$. Consequently $\mathbf{S}$ is a unitary matrix with the n distinct eigenvalues $1, \omega, \dots, \omega^{n-1}$, and each eigenspace $E_{\omega^j}(\mathbf{S})$ is the line $\text{span}(\mathbf{f}_j)$.

Idea. The shift moves entry $m + 1$ into slot m . If the entries are the successive powers of a fixed number z , that is the same as multiplying the whole vector by z — provided the wrap-around at the end also multiplies by z , which is exactly the condition $z^n = 1$. So the geometric vectors $(1, z, z^2, \dots)$ with $z^n = 1$ are forced on us, and there are n of them.

Proof. Fix j and let $0 \leq m \leq n - 1$. By the description of $\mathbf{S}$, the m -th entry of $\mathbf{S}\mathbf{f}_j$ is the $(m + 1)$ -st entry of $\mathbf{f}_j$, indices modulo n . For $m < n - 1$ this is

$$\frac{\omega^{j(m+1)}}{\sqrt{n}} = \omega^j \cdot \frac{\omega^{jm}}{\sqrt{n}} = \omega^j (\mathbf{f}_j)_m,$$

and for $m = n - 1$ it is $(\mathbf{f}_j)_0 = 1/\sqrt{n}$, while $\omega^j (\mathbf{f}_j)_{n-1} = \omega^j \omega^{j(n-1)}/\sqrt{n} = \omega^{jn}/\sqrt{n} = 1/\sqrt{n}$ because $\omega^n = 1$. The two agree, so $\mathbf{S}\mathbf{f}_j = \omega^j \mathbf{f}_j$. Each $\mathbf{f}_j$ is nonzero, since its 0-th entry is $1/\sqrt{n}$, so ω^j is an eigenvalue.

The matrix $\mathbf{S}$ is a real permutation matrix, so $\mathbf{S}^\top \mathbf{S} = \mathbf{I}$ by Lemma 3.4.7 (b), and $\mathbf{S}^* = \mathbf{S}^\top$; hence $\mathbf{S}$ is unitary (Definition 11.7.2). The n numbers $1, \omega, \dots, \omega^{n-1}$ are distinct, so $\mathbf{S}$ has n distinct eigenvalues, and the sum of their eigenspaces is direct (Corollary 9.2.3). Each of those n eigenspaces has dimension at least 1 and their dimensions add up to at most $\dim \mathbb{C}^n = n$, so each has dimension exactly 1. Since $\mathbf{f}_j \in E_{\omega^j}(\mathbf{S})$ is nonzero, that eigenspace is $\text{span}(\mathbf{f}_j)$. ■

The vectors $\mathbf{f}_0, \dots, \mathbf{f}_{n-1}$ are called the **Fourier vectors**. They are already known to be pairwise orthogonal: $\mathbf{S}$ is unitary, hence normal, and eigenvectors of a normal matrix for distinct eigenvalues are orthogonal (Corollary 12.3.4). But the direct computation is short, it produces the normalization at the same time, and it is the computation on which the rest of the section runs. So we do it.

Theorem 12.9.3. The Fourier Matrix Is Unitary

The list $(\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{n-1})$ is an orthonormal basis of $\mathbb{C}^n$. Equivalently, the **Fourier matrix**

$$\mathbf{F} := \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix}, \quad (\mathbf{F})_{mj} = \frac{\omega^{mj}}{\sqrt{n}},$$

whose j -th column is $\mathbf{f}_j$, is unitary: $\mathbf{F}^* \mathbf{F} = \mathbf{F} \mathbf{F}^* = \mathbf{I}$.

Proof. Let $0 \leq j, k \leq n-1$. Since $|\omega| = 1$ we have $\bar{\omega} = \omega^{-1}$, hence $\overline{\omega^{km}} = \omega^{-km}$, and so

$$\langle \mathbf{f}_j, \mathbf{f}_k \rangle = \frac{1}{n} \sum_{m=0}^{n-1} \omega^{jm} \overline{\omega^{km}} = \frac{1}{n} \sum_{m=0}^{n-1} (\omega^{j-k})^m,$$

the inner product being linear in the first slot and conjugate-linear in the second.

Case $j = k$. Then $\omega^{j-k} = \omega^0 = 1$, so every one of the n terms equals 1 and the sum is n . Hence $\langle \mathbf{f}_j, \mathbf{f}_j \rangle = n/n = 1$, and $\|\mathbf{f}_j\| = 1$.

Case $j \neq k$. Put $z := \omega^{j-k}$. Then $z^n = (\omega^n)^{j-k} = 1$. Also $z \neq 1$: by Example (Roots of Unity), $\omega^{j-k} = 1$ would force n to divide $j-k$, and $0 < |j-k| \leq n-1$ makes that impossible. Now use the factorization

$$(z-1)(1+z+z^2+\cdots+z^{n-1}) = z^n - 1 = 0,$$

which is a polynomial identity valid for every complex z . Since $z-1 \neq 0$, we may divide by it, and

$$\sum_{m=0}^{n-1} z^m = 0.$$

Hence $\langle \mathbf{f}_j, \mathbf{f}_k \rangle = 0$.

So the list is orthonormal; it has $n = \dim \mathbb{C}^n$ members and is independent (Theorem 11.1.4),

hence a basis (Theorem 1.6.4). A matrix whose columns form an orthonormal basis of $\mathbb{C}^n$ is unitary (Theorem 11.7.3 (f)), which gives $\mathbf{F}^* \mathbf{F} = \mathbf{F} \mathbf{F}^* = \mathbf{I}$. ■

The second case is where the whole section lives. A sum of n unit complex numbers, spread evenly around the circle, cancels — and it cancels for the cheapest possible reason, that a geometric series with ratio $z \neq 1$ and $z^n = 1$ telescopes to zero.

Example. The Fourier matrix of size four

Write out $\mathbf{F}$ for $n = 4$, and check one orthogonality relation by hand.

Solution. For $n = 4$, $\omega = e^{2\pi i/4} = i$, and the powers cycle $i^0 = 1$, $i^1 = i$, $i^2 = -1$, $i^3 = -i$. So

$$\mathbf{F} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}.$$

Take $j = 1$, $k = 3$. Then $\omega^{j-k} = i^{-2} = -1$, and

$$\langle \mathbf{f}_1, \mathbf{f}_3 \rangle = \frac{1}{4} (1 + (-1) + 1 + (-1)) = 0,$$

matching the geometric-series computation with $z = -1$, for which $1 + z + z^2 + z^3 = 0$. Note that $\mathbf{F}$ is symmetric, $\mathbf{F}^\top = \mathbf{F}$, but **not** Hermitian: $\mathbf{F}^* = \overline{\mathbf{F}} \neq \mathbf{F}$, since $(\mathbf{F})_{12} = i/2$.

12.9.3 One unitary for every circulant

Theorem 12.9.4. Circulant Diagonalization

Let $\mathbf{C} \in M_n(\mathbb{C})$ be the circulant with first row $\mathbf{c}$ and symbol $p_{\mathbf{c}}$. Then every Fourier vector is an eigenvector of $\mathbf{C}$,

$$\mathbf{C} \mathbf{f}_j = p_{\mathbf{c}}(\omega^j) \mathbf{f}_j \quad (j = 0, 1, \dots, n-1),$$

and consequently

$$\mathbf{F}^* \mathbf{C} \mathbf{F} = \text{diag}(p_{\mathbf{c}}(1), p_{\mathbf{c}}(\omega), \dots, p_{\mathbf{c}}(\omega^{n-1})).$$

The same unitary $\mathbf{F}$ works for every circulant of size n , and $\det \mathbf{C} = \prod_{j=0}^{n-1} p_{\mathbf{c}}(\omega^j)$.

Idea. There is nothing left to prove. A circulant is a polynomial in $\mathbf{S}$ by Theorem 12.9.1, and a polynomial in a matrix acts on an eigenvector of that matrix by the same polynomial applied to the eigenvalue. The point is not the proof but the statement: $\mathbf{F}$ was written down without reference to $\mathbf{c}$, so a single change of orthonormal basis handles the entire family at once.

Proof. By [Theorem 12.9.1](#) (a), $\mathbf{C} = p_{\mathbf{c}}(\mathbf{S})$, and by [Theorem 12.9.2](#), $\mathbf{S}\mathbf{f}_j = \omega^j \mathbf{f}_j$. Iterating the latter gives $\mathbf{S}^m \mathbf{f}_j = \omega^{jm} \mathbf{f}_j$ for every $m \geq 0$, so

$$\mathbf{C}\mathbf{f}_j = \sum_{m=0}^{n-1} c_m \mathbf{S}^m \mathbf{f}_j = \left(\sum_{m=0}^{n-1} c_m \omega^{jm} \right) \mathbf{f}_j = p_{\mathbf{c}}(\omega^j) \mathbf{f}_j.$$

The j -th column of $\mathbf{F}$ is $\mathbf{f}_j$, so the j -th column of $\mathbf{C}\mathbf{F}$ is $p_{\mathbf{c}}(\omega^j) \mathbf{f}_j$, which is the j -th column of $\mathbf{F}\mathbf{D}$ for $\mathbf{D} = \text{diag}(p_{\mathbf{c}}(1), \dots, p_{\mathbf{c}}(\omega^{n-1}))$. Hence $\mathbf{C}\mathbf{F} = \mathbf{F}\mathbf{D}$, and multiplying on the left by $\mathbf{F}^* = \mathbf{F}^{-1}$ ([Theorem 12.9.3](#)) gives $\mathbf{F}^* \mathbf{C}\mathbf{F} = \mathbf{D}$. Taking determinants, $\det \mathbf{C} = \det \mathbf{D} = \prod_j p_{\mathbf{c}}(\omega^j)$, since $\det \mathbf{F}^* \det \mathbf{F} = \det(\mathbf{F}^* \mathbf{F}) = 1$ ([Theorem 7.5.4](#)). ■

This re-proves normality from the other side, and closes the account opened in Example (Which of these are normal?) (d): $\mathbf{C} = \mathbf{F}\mathbf{D}\mathbf{F}^*$ with $\mathbf{F}$ unitary and $\mathbf{D}$ diagonal is precisely the factorization that [Corollary 12.4.3](#) reserves for normal matrices. The eigenvalues $p_{\mathbf{c}}(\omega^j)$ need no characteristic polynomial either: they are n evaluations of a polynomial of degree less than n .

Example. The running example, diagonalized

Find the eigenvalues and the determinant of the circulant $\mathbf{C}$ with first row $(1, 2, 3, 4)$, and write down $\mathbf{F}^* \mathbf{C}\mathbf{F}$.

Solution. Here $n = 4$, $\omega = i$, and $p_{\mathbf{c}}(x) = 1 + 2x + 3x^2 + 4x^3$. Evaluate at the four fourth roots of unity, using $i^2 = -1$ and $i^3 = -i$:

$$\begin{aligned} p_{\mathbf{c}}(1) &= 1 + 2 + 3 + 4 = 10, \\ p_{\mathbf{c}}(i) &= 1 + 2i - 3 - 4i = -2 - 2i, \\ p_{\mathbf{c}}(-1) &= 1 - 2 + 3 - 4 = -2, \\ p_{\mathbf{c}}(-i) &= 1 - 2i - 3 + 4i = -2 + 2i. \end{aligned}$$

So $\mathbf{F}^* \mathbf{C}\mathbf{F} = \text{diag}(10, -2 - 2i, -2, -2 + 2i)$ with $\mathbf{F}$ as in [Example \(The Fourier matrix of size four\)](#), and

$$\det \mathbf{C} = 10 \cdot (-2 - 2i) \cdot (-2) \cdot (-2 + 2i) = -20(4 + 4) = -160,$$

using $(-2 - 2i)(-2 + 2i) = 4 + 4 = 8$. As a check on the first eigenvalue: every row of $\mathbf{C}$ sums to 10, so $\mathbf{C}\mathbf{1} = 10 \cdot \mathbf{1}$, and $\mathbf{f}_0 = \frac{1}{2}\mathbf{1}$.

Quick check. The matrix of [Example \(The running example, diagonalized\)](#) is real, and two of its eigenvalues are not. Why must those two be complex conjugates of each other, and which two are they in general?

Solution. If every c_m is real then $\overline{p_{\mathbf{c}}(z)} = p_{\mathbf{c}}(\bar{z})$ for every z . Taking $z = \omega^j$ and using

$$\overline{\omega^j} = \omega^{-j} = \omega^{n-j},$$

$$\overline{p_{\mathbf{c}}(\omega^j)} = p_{\mathbf{c}}(\omega^{n-j}),$$

so the eigenvalue attached to $\mathbf{f}_j$ and the one attached to $\mathbf{f}_{n-j}$ are conjugate. Here $n = 4$ and the pair is $j = 1, 3$: $\overline{-2 - 2i} = -2 + 2i$. The eigenvalues for $j = 0$ and, when n is even, $j = n/2$ are their own conjugates, hence real — which is why 10 and -2 came out real. (This is also Theorem 9.2.5, seen here with the eigenvectors named.)

Once every circulant is $\mathbf{FDF}^*$, the family closes up on itself.

Corollary 12.9.5. Circulants Form a Commutative Algebra

Let $\mathcal{C}_n \subseteq M_n(\mathbb{C})$ be the set of circulants of size n . Then:

- (a) $\mathcal{C}_n$ is a subspace of $M_n(\mathbb{C})$ of dimension n , with basis $\mathbf{I}, \mathbf{S}, \dots, \mathbf{S}^{n-1}$;
- (b) $\mathcal{C}_n$ is closed under products, and any two of its members commute;
- (c) $\mathbf{A} \in \mathcal{C}_n$ if and only if $\mathbf{F}^* \mathbf{A} \mathbf{F}$ is diagonal;
- (d) a circulant $\mathbf{C}$ is invertible if and only if $p_{\mathbf{c}}(\omega^j) \neq 0$ for every j , and then $\mathbf{C}^{-1}$ is again a circulant.

Proof.

- (a) The map $\mathbf{c} \mapsto p_{\mathbf{c}}(\mathbf{S})$ from $\mathbb{C}^n$ to $M_n(\mathbb{C})$ is linear and, by Theorem 12.9.1 (a), its image is exactly $\mathcal{C}_n$. It is injective, because the first row of $p_{\mathbf{c}}(\mathbf{S})$ is $\mathbf{c}$. So $\mathcal{C}_n$ is a subspace of dimension n , and the images $\mathbf{I}, \mathbf{S}, \dots, \mathbf{S}^{n-1}$ of the standard basis form a basis of it.
- (b) If $\mathbf{C} = p(\mathbf{S})$ and $\mathbf{C}' = q(\mathbf{S})$, then $\mathbf{C}\mathbf{C}' = (pq)(\mathbf{S})$ (Theorem 0.9.9), which is a circulant by Theorem 12.9.1 (a). Commutativity is part (c) of that theorem.
- (c) $(\Rightarrow)$ is Theorem 12.9.4. $(\Leftarrow)$ Suppose $\mathbf{F}^* \mathbf{A} \mathbf{F} = \mathbf{D} = \text{diag}(d_0, \dots, d_{n-1})$. The n numbers $1, \omega, \dots, \omega^{n-1}$ are distinct, so by Lagrange interpolation (Theorem 6.6.1) there is $q \in \mathbb{C}[x]$ with $q(\omega^j) = d_j$ for every j . By Theorem 12.9.4 applied to the circulant $q(\mathbf{S})$, which has symbol the remainder of q on division by $x^n - 1$,

$$\mathbf{F}^* q(\mathbf{S}) \mathbf{F} = \text{diag}(q(1), \dots, q(\omega^{n-1})) = \mathbf{D}.$$

Multiplying both sides by $\mathbf{F}$ on the left and $\mathbf{F}^*$ on the right gives $q(\mathbf{S}) = \mathbf{A}$, so $\mathbf{A}$ is a circulant.

- (d) From $\mathbf{C} = \mathbf{FDF}^*$, $\det \mathbf{C} = \det \mathbf{D}$, so $\mathbf{C}$ is invertible exactly when no diagonal entry $p_{\mathbf{c}}(\omega^j)$ vanishes (Theorem 7.5.2). In that case $\mathbf{C}^{-1} = \mathbf{F}\mathbf{D}^{-1}\mathbf{F}^*$, and $\mathbf{F}^* \mathbf{C}^{-1} \mathbf{F} = \mathbf{D}^{-1}$ is diagonal, so $\mathbf{C}^{-1}$ is a circulant by (c). ■

Part (c) is the sharpest statement of the section: being a circulant is not a statement about the pattern of the entries at all, but a statement about which orthonormal basis diagonalizes the matrix. The pattern was only ever a way of recognizing it.

12.9.4 Multiplying circulants is convolution

A circulant is its first row. So the product of two circulants must be computable from the two first rows, and it is worth knowing what that computation is.

Given $\mathbf{c}, \mathbf{d} \in \mathbb{C}^n$, define their **cyclic convolution** $\mathbf{c} * \mathbf{d} \in \mathbb{C}^n$ by

$$(\mathbf{c} * \mathbf{d})_r := \sum_{l=0}^{n-1} c_l d_{r-l} \quad (r = 0, \dots, n-1),$$

subscripts modulo n . In words: to get entry r , slide the reversed list $\mathbf{d}$ along $\mathbf{c}$ so that the indices add up to r , and add the products. It is the multiplication of polynomials with the exponents read modulo n , and the next theorem says so.

Theorem 12.9.6. Convolution Theorem

Let $\mathbf{c}, \mathbf{d} \in \mathbb{C}^n$ with circulants $\mathbf{C}, \mathbf{C}'$ and symbols $p_{\mathbf{c}}, p_{\mathbf{d}}$. Then:

- (a) $\mathbf{C}\mathbf{C}'$ is the circulant with first row $\mathbf{c} * \mathbf{d}$;
- (b) for every j , $p_{\mathbf{c} * \mathbf{d}}(\omega^j) = p_{\mathbf{c}}(\omega^j) p_{\mathbf{d}}(\omega^j)$.

In words: convolving the first rows corresponds to multiplying the eigenvalue lists entry by entry.

Idea. Both parts are the identity $p_{\mathbf{c}} p_{\mathbf{d}} \equiv p_{\mathbf{c} * \mathbf{d}} \pmod{x^n - 1}$, looked at twice. Multiplying the two symbols produces a polynomial of degree up to $2n - 2$; reducing the exponents modulo n folds the coefficients back into n slots, and the fold is precisely the convolution sum. Evaluating at a root of $x^n - 1$ makes the reduction invisible, which is (b); substituting $\mathbf{S}$, for which $\mathbf{S}^n = \mathbf{I}$, makes it invisible as well, which is (a).

Proof. Expand the product of the symbols and collect exponents modulo n :

$$p_{\mathbf{c}}(x) p_{\mathbf{d}}(x) = \sum_{l=0}^{n-1} \sum_{m=0}^{n-1} c_l d_m x^{l+m}.$$

For each $r \in \{0, \dots, n-1\}$, the pairs (l, m) with $l + m \equiv r \pmod{n}$ are exactly those with $m \equiv r - l$, one value of m for each l . Hence, working modulo $x^n - 1$ and using $x^{l+m} \equiv x^{(l+m) \bmod n}$,

$$p_{\mathbf{c}}(x) p_{\mathbf{d}}(x) \equiv \sum_{r=0}^{n-1} \left(\sum_{l=0}^{n-1} c_l d_{r-l} \right) x^r = p_{\mathbf{c} * \mathbf{d}}(x) \pmod{x^n - 1}.$$

- (a) Substitute $\mathbf{S}$. Since $\mathbf{S}^n = \mathbf{I}$, the multiple of $x^n - 1$ contributes nothing, so

$$\mathbf{C}\mathbf{C}' = p_{\mathbf{c}}(\mathbf{S}) p_{\mathbf{d}}(\mathbf{S}) = p_{\mathbf{c} * \mathbf{d}}(\mathbf{S}),$$

the first equality by Theorem 12.9.1 (a) and Theorem 0.9.9. By Theorem 12.9.1 (a) again, $p_{\mathbf{c} * \mathbf{d}}(\mathbf{S})$ is the circulant with first row $\mathbf{c} * \mathbf{d}$, the coefficient list having length n .

- (b) Substitute $x = \omega^j$, a root of $x^n - 1$, so that the multiple of $x^n - 1$ again contributes nothing. This gives $p_{\mathbf{c}}(\omega^j) p_{\mathbf{d}}(\omega^j) = p_{\mathbf{c} * \mathbf{d}}(\omega^j)$.



Part (b) deserves its own sentence, because it is the reason the discrete Fourier transform is used at all. Define the **discrete Fourier transform** of $\mathbf{c} \in \mathbb{C}^n$ to be the vector of eigenvalues,

$$\widehat{\mathbf{c}} := (p_{\mathbf{c}}(1), p_{\mathbf{c}}(\omega), \dots, p_{\mathbf{c}}(\omega^{n-1})) = \sqrt{n} \mathbf{F} \mathbf{c},$$

the second expression because the j -th entry of $\mathbf{F} \mathbf{c}$ is $n^{-1/2} \sum_m \omega^{jm} c_m$. Then (b) reads

$$(\widehat{\mathbf{c} * \mathbf{d}})_j = \widehat{\mathbf{c}}_j \widehat{\mathbf{d}}_j \quad (j = 0, \dots, n-1) :$$

the transform turns convolution into entrywise multiplication. Convolution is a tangled sum of n^2 products; after one unitary change of basis it is n ordinary multiplications. Everything the transform buys comes from this line.

Example. Convolution in size four

Let $\mathbf{c} = (1, 2, 3, 4)$ and $\mathbf{d} = (1, 1, 0, 0)$. Compute $\mathbf{c} * \mathbf{d}$, and verify [Theorem 12.9.6](#) (b) at all four roots of unity.

Solution. Since $d_0 = d_1 = 1$ and $d_2 = d_3 = 0$, the sum $\sum_l c_l d_{r-l}$ keeps only the terms with $r - l \equiv 0$ or 1 , that is $l = r$ and $l = r - 1$. So $(\mathbf{c} * \mathbf{d})_r = c_r + c_{r-1}$, subscripts modulo 4:

$$\mathbf{c} * \mathbf{d} = (1 + 4, 2 + 1, 3 + 2, 4 + 3) = (5, 3, 5, 7).$$

The three symbols are $p_{\mathbf{c}}(x) = 1 + 2x + 3x^2 + 4x^3$, $p_{\mathbf{d}}(x) = 1 + x$, and $p_{\mathbf{c} * \mathbf{d}}(x) = 5 + 3x + 5x^2 + 7x^3$. Evaluating each at $1, i, -1, -i$, with $p_{\mathbf{c}}$ already computed in [Example \(The running example, diagonalized\)](#):

$$\begin{array}{ll} j = 0 : & 10 \cdot 2 = 20, & 5 + 3 + 5 + 7 = 20; \\ j = 1 : & (-2 - 2i)(1 + i) = -4i, & 5 + 3i - 5 - 7i = -4i; \\ j = 2 : & (-2) \cdot 0 = 0, & 5 - 3 + 5 - 7 = 0; \\ j = 3 : & (-2 + 2i)(1 - i) = 4i, & 5 - 3i - 5 + 7i = 4i. \end{array}$$

All four agree. For the middle line, $(-2 - 2i)(1 + i) = -2 - 2i - 2i - 2i^2 = -4i$.

12.9.5 Solving a circulant system, and a word about speed

If $\mathbf{C} \mathbf{x} = \mathbf{b}$ with $\mathbf{C}$ an invertible circulant, then in the Fourier basis the system falls apart into n scalar equations. Writing $\mathbf{C} = \mathbf{F} \mathbf{D} \mathbf{F}^*$ and setting $\mathbf{y} = \mathbf{F}^* \mathbf{x}$, $\mathbf{g} = \mathbf{F}^* \mathbf{b}$, the system becomes $\mathbf{D} \mathbf{y} = \mathbf{g}$, that is $y_j = g_j / \lambda_j$ with $\lambda_j = p_{\mathbf{c}}(\omega^j)$. Then $\mathbf{x} = \mathbf{F} \mathbf{y}$. Three steps: transform, divide, transform back.

Example. A circulant system

Solve $\mathbf{C} \mathbf{x} = \mathbf{b}$ for the circulant $\mathbf{C}$ with first row $(1, 2, 3, 4)$ and $\mathbf{b} = (1, 1, -1, -1)$.

Solution. The eigenvalues are $\lambda = (10, -2 - 2i, -2, -2 + 2i)$ by [Example \(The running example, diagonalized\)](#). First transform $\mathbf{b}$: the j -th entry of $\mathbf{F}^*\mathbf{b}$ is $\frac{1}{2}\sum_m \omega^{jm}b_m = \frac{1}{2}\sum_m (-i)^{jm}b_m$, so

$$\begin{aligned} g_0 &= \frac{1}{2}(1 + 1 - 1 - 1) = 0, \\ g_1 &= \frac{1}{2}(1 + (-i) - (-1) - (i)) = 1 - i, \\ g_2 &= \frac{1}{2}(1 - 1 - 1 + 1) = 0, \\ g_3 &= \frac{1}{2}(1 + i + 1 + i) = 1 + i. \end{aligned}$$

Now divide. The two zero entries stay zero, and

$$y_1 = \frac{1 - i}{-2 - 2i} = \frac{(1 - i)(-2 + 2i)}{8} = \frac{4i}{8} = \frac{i}{2}, \quad y_3 = \overline{y_1} = -\frac{i}{2},$$

where the middle step multiplies numerator and denominator by $\overline{-2 - 2i} = -2 + 2i$, whose product with $-2 - 2i$ is 8. Finally $\mathbf{x} = \mathbf{F}\mathbf{y}$, whose m -th entry is $\frac{1}{2}(i^m y_1 + i^{3m} y_3) = \frac{1}{2} \cdot \frac{i}{2}(i^m - (-i)^m)$:

$$\mathbf{x} = \left(0, -\frac{1}{2}, 0, \frac{1}{2}\right).$$

Check directly: the second row of $\mathbf{C}$ is $(4, 1, 2, 3)$, and $-\frac{1}{2} + \frac{3}{2} = 1 = b_1$; the third row is $(3, 4, 1, 2)$, and $-2 + 1 = -1 = b_2$. The answer is real, as it must be: $\mathbf{C}$ and $\mathbf{b}$ are real, so $\mathbf{x} = \mathbf{C}^{-1}\mathbf{b}$ is real, and this shows up in the transform as the conjugate symmetry $y_3 = \overline{y_1}$.

Counted honestly, the recipe above saves something, but not yet very much. Forming $\mathbf{F}^*\mathbf{b}$ as a matrix–vector product costs about n^2 multiplications, and so does forming $\mathbf{F}\mathbf{y}$, so the whole solve costs on the order of n^2 . That beats general elimination, which is on the order of n^3 , but it is no better than multiplying $\mathbf{b}$ by $\mathbf{C}^{-1}$ once that inverse has been found. What changes the picture is that $\mathbf{F}$ does not have to be applied as a general matrix. The powers of ω repeat, and splitting the sum $\sum_m \omega^{jm}b_m$ into its even and odd m reduces one transform of length n to two of length $n/2$; iterating brings the cost down to about $n \log n$. That algorithm is the **fast Fourier transform**. It proves no new theorem: the matrix $\mathbf{F}$, the factorization $\mathbf{C} = \mathbf{F}\mathbf{D}\mathbf{F}^*$ and everything above are unchanged, and the fast Fourier transform is only a way of applying $\mathbf{F}$ quickly. Chapter 24 develops it, together with the cost accounting that makes “quickly” precise.

Warning. A **Toeplitz** matrix — one with constant diagonals, $(\mathbf{A})_{jk} = a_{k-j}$ with the subscript **not** reduced modulo n — is not a circulant, and $\mathbf{F}$ does not diagonalize it. Take

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Every diagonal of $\mathbf{A}$ is constant, so $\mathbf{A}$ is Toeplitz, but the circulant with first row $(0, 1, 0)$ is $\mathbf{S}$, whose second row is $(0, 0, 1)$, not $(1, 0, 1)$. And $\mathbf{F}$ fails at the very first column: $\mathbf{f}_0 = \frac{1}{\sqrt{3}}(1, 1, 1)$ and $\mathbf{A}\mathbf{f}_0 = \frac{1}{\sqrt{3}}(1, 2, 1)$, which is not a multiple of $\mathbf{f}_0$, so $\mathbf{f}_0$ is not an eigenvector and $\mathbf{F}^*\mathbf{A}\mathbf{F}$ is not diagonal. Normality is not the issue: $\mathbf{A}$ is symmetric, hence normal, and Theorem 12.5.1 does hand it an orthonormal eigenbasis. It is simply a different basis, with the irrational eigenvalues $0, \pm\sqrt{2}$ in place of evaluations of a polynomial at roots of unity. What the wrap-around buys is not diagonalizability; it is the *same* diagonalizing basis for everybody.

12.9.6 Exercises

12.9.6.1 A. Check your understanding

Exercise. A1

- (a) Define the circulant with first row $\mathbf{c} = (c_0, \dots, c_{n-1})$, and write down its symbol.
- (b) State the eigenvalues and the eigenvectors of the cyclic shift $\mathbf{S}$ on $\mathbb{C}^n$.
- (c) Give the eigenvalues of the circulant with first row $\mathbf{c}$, and say how many multiplications are needed to compute all of them from $\mathbf{c}$ by direct evaluation.
- (d) Determine whether the following are true, with a reason. (i) Every Toeplitz matrix is a circulant. (ii) Every circulant is normal. (iii) The Fourier matrix $\mathbf{F}$ is Hermitian.
- (e) Explain in one sentence why two circulants of the same size commute, without computing any entries.

Solution.

- (a) The matrix $\mathbf{C} \in M_n(\mathbb{C})$ with $(\mathbf{C})_{jk} = c_{k-j}$, subscripts modulo n and indices running from 0 to $n-1$ (Definition 12.9.1). Its symbol is $p_{\mathbf{c}}(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1}$.
- (b) With $\omega = e^{2\pi i/n}$: the eigenvalues are $1, \omega, \dots, \omega^{n-1}$, all distinct, and ω^j has the one-dimensional eigenspace spanned by $\mathbf{f}_j = n^{-1/2}(1, \omega^j, \dots, \omega^{(n-1)j})$ (Theorem 12.9.2).
- (c) The eigenvalues are $p_{\mathbf{c}}(\omega^j)$ for $j = 0, \dots, n-1$ (Theorem 12.9.4). Each evaluation is a sum of n products, so all n of them cost about n^2 multiplications done directly.
- (d) (i) False: the 3×3 matrix in the warning above has constant diagonals but is not a circulant. (ii) True, by Theorem 12.9.1 (c); equivalently by Corollary 12.4.3, since $\mathbf{C} = \mathbf{F}\mathbf{D}\mathbf{F}^*$. (iii) False. $\mathbf{F}$ is symmetric, but $\mathbf{F}^* = \overline{\mathbf{F}}$, and for $n = 4$ the entry $(\mathbf{F})_{12} = i/2$ is not real (Example (The Fourier matrix of size four)). $\mathbf{F}$ is Hermitian only for $n = 1, 2$, where all its entries are real.
- (e) Both are polynomials in the single matrix $\mathbf{S}$ (Theorem 12.9.1 (a)), and two polynomials in one matrix always commute (Theorem 0.9.9).

12.9.6.2 B. Practice

Exercise. B1: Eigenvalues of small circulants

For each vector $\mathbf{c}$, write down the symbol, compute all eigenvalues of the circulant with first row $\mathbf{c}$, and give its determinant.

- (a) $\mathbf{c} = (1, 1, 1)$, $n = 3$.
- (b) $\mathbf{c} = (0, 1, 0, 0)$, $n = 4$.
- (c) $\mathbf{c} = (2, -1, 0, -1)$, $n = 4$.

Solution.

- (a) $p_{\mathbf{c}}(x) = 1 + x + x^2$. With $\omega = e^{2\pi i/3}$, the eigenvalues are $p_{\mathbf{c}}(1) = 3$ and $p_{\mathbf{c}}(\omega^j) = 1 + \omega^j + \omega^{2j}$ for $j = 1, 2$. For $j \neq 0$, $z = \omega^j$ satisfies $z \neq 1$ and $z^3 = 1$, so $1 + z + z^2 = (z^3 - 1)/(z - 1) = 0$ — the geometric-series computation of Theorem 12.9.3 again. The

eigenvalues are $3, 0, 0$ and $\det \mathbf{C} = 0$. (This matrix is the all-ones matrix, of rank 1.)
 (b) $p_{\mathbf{C}}(x) = x$, so $\mathbf{C} = \mathbf{S}$. The eigenvalues are $\omega^j = i^j$, namely $1, i, -1, -i$, and

$$\det \mathbf{S} = 1 \cdot i \cdot (-1) \cdot (-i) = (i \cdot (-i))(1 \cdot (-1)) = -1.$$

This agrees with the sign of the permutation: $\mathbf{S}$ permutes the coordinates in one 4-cycle, which is odd.

(c) $p_{\mathbf{C}}(x) = 2 - x - x^3$. Then $p_{\mathbf{C}}(1) = 0$; $p_{\mathbf{C}}(i) = 2 - i - i^3 = 2 - i + i = 2$; $p_{\mathbf{C}}(-1) = 2 + 1 + 1 = 4$; $p_{\mathbf{C}}(-i) = 2 + i - i = 2$. The eigenvalues are $0, 2, 4, 2$, and $\det \mathbf{C} = 0$. The kernel contains $\mathbf{f}_0$, that is the all-ones vector, as one sees directly: every row of $\mathbf{C}$ sums to 0.

Exercise. B2: Solve a circulant system by transforming

Let $\mathbf{C}$ be the 4×4 circulant with first row $(3, 1, 0, 1)$ and let $\mathbf{b} = (1, 0, 0, 0)$. Compute the eigenvalues of $\mathbf{C}$, and solve $\mathbf{C}\mathbf{x} = \mathbf{b}$ by transforming, dividing, and transforming back. Check your answer against the first two rows of $\mathbf{C}$.

Solution. The symbol is $p_{\mathbf{C}}(x) = 3 + x + x^3$, and with $\omega = i$,

$$p_{\mathbf{C}}(1) = 5, \quad p_{\mathbf{C}}(i) = 3 + i - i = 3, \quad p_{\mathbf{C}}(-1) = 1, \quad p_{\mathbf{C}}(-i) = 3.$$

All four are nonzero, so $\mathbf{C}$ is invertible (Corollary 12.9.5 (d)), with $\det \mathbf{C} = 5 \cdot 3 \cdot 1 \cdot 3 = 45$.

Transform: $(\mathbf{F}^* \mathbf{b})_j = \frac{1}{2} \sum_m \overline{\omega^{jm}} b_m = \frac{1}{2}$ for every j , since $\mathbf{b} = \mathbf{e}_0$. Divide: $y_j = \frac{1}{2} \lambda_j^{-1}$, giving $\mathbf{y} = \frac{1}{2}(\frac{1}{5}, \frac{1}{3}, 1, \frac{1}{3})$. Transform back: the m -th entry of $\mathbf{x} = \mathbf{F}\mathbf{y}$ is

$$x_m = \frac{1}{2} \sum_{j=0}^3 i^{mj} y_j = \frac{1}{4} \left(\frac{1}{5} + \frac{i^m}{3} + (-1)^m + \frac{i^{3m}}{3} \right).$$

For $m = 1$ and $m = 3$ the two middle terms cancel, since $i^m + i^{3m} = i^m + \overline{i^m}$ is 0 when $i^m = \pm i$; the bracket is $\frac{1}{5} - 1$, so $x_1 = x_3 = -\frac{1}{5}$. For $m = 0$ the bracket is $\frac{1}{5} + \frac{1}{3} + 1 + \frac{1}{3} = \frac{28}{15}$, so $x_0 = \frac{7}{15}$. For $m = 2$ it is $\frac{1}{5} - \frac{1}{3} + 1 - \frac{1}{3} = \frac{8}{15}$, so $x_2 = \frac{2}{15}$. Hence

$$\mathbf{x} = \left(\frac{7}{15}, -\frac{1}{5}, \frac{2}{15}, -\frac{1}{5} \right).$$

Check: row 0 of $\mathbf{C}$ is $(3, 1, 0, 1)$ and $3 \cdot \frac{7}{15} - \frac{1}{5} - \frac{1}{5} = \frac{21-3-3}{15} = 1$; row 1 is $(1, 3, 1, 0)$ and $\frac{7}{15} - \frac{3}{5} + \frac{2}{15} = \frac{7-9+2}{15} = 0$.

Exercise. B3: The convolution theorem in an example

Let $\mathbf{c} = (1, 0, 2, 0)$ and $\mathbf{d} = (0, 1, 0, 3)$ in $\mathbb{C}^4$.

- Compute $\mathbf{c} * \mathbf{d}$.
- Verify $p_{\mathbf{c} * \mathbf{d}}(\omega^j) = p_{\mathbf{C}}(\omega^j) p_{\mathbf{D}}(\omega^j)$ for $j = 0, 1, 2, 3$.
- Hence write down the first row of the product of the two circulants, without multiplying any matrices.

Solution.

- (a) Only $c_0 = 1$ and $c_2 = 2$ are nonzero, so $(\mathbf{c} * \mathbf{d})_r = d_r + 2d_{r-2}$, subscripts modulo 4. With $\mathbf{d} = (0, 1, 0, 3)$:

$$\mathbf{c} * \mathbf{d} = (0 + 0, 1 + 6, 0 + 0, 3 + 2) = (0, 7, 0, 5).$$

- (b) The symbols are $p_{\mathbf{c}}(x) = 1 + 2x^2$, $p_{\mathbf{d}}(x) = x + 3x^3$, and $p_{\mathbf{c}*\mathbf{d}}(x) = 7x + 5x^3$. With $\omega = i$:

$$\begin{aligned} j = 0 : & \quad 3 \cdot 4 = 12, & 7 + 5 = 12; \\ j = 1 : & \quad (-1)(-2i) = 2i, & 7i - 5i = 2i; \\ j = 2 : & \quad 3 \cdot (-4) = -12, & -7 - 5 = -12; \\ j = 3 : & \quad (-1)(2i) = -2i, & -7i + 5i = -2i. \end{aligned}$$

Here $p_{\mathbf{c}}(i) = 1 + 2i^2 = -1$ and $p_{\mathbf{d}}(i) = i + 3i^3 = i - 3i = -2i$, and similarly at the other roots. All four rows agree.

- (c) By [Theorem 12.9.6](#) (a), the product is the circulant with first row $\mathbf{c} * \mathbf{d} = (0, 7, 0, 5)$.

12.9.6.3 C. Going deeper

Exercise. C1: What commutes with the shift

Let $\mathbf{A} \in M_n(\mathbb{C})$ satisfy $\mathbf{AS} = \mathbf{SA}$. Prove that $\mathbf{A}$ is a circulant. *Hint: consider the eigenspaces of $\mathbf{S}$.*

Solution. By [Theorem 12.9.2](#), $\mathbf{S}$ has the n distinct eigenvalues ω^j , and $E_{\omega^j}(\mathbf{S}) = \text{span}(\mathbf{f}_j)$ is one-dimensional. Since $\mathbf{A}$ commutes with $\mathbf{S}$, each eigenspace of $\mathbf{S}$ is invariant under $\mathbf{A}$ (Theorem 9.9.1). So $\mathbf{A}\mathbf{f}_j \in \text{span}(\mathbf{f}_j)$, that is $\mathbf{A}\mathbf{f}_j = d_j\mathbf{f}_j$ for some $d_j \in \mathbb{C}$.

Therefore the j -th column of $\mathbf{A}\mathbf{F}$ is $d_j\mathbf{f}_j$, which is the j -th column of $\mathbf{F}\mathbf{D}$ for $\mathbf{D} = \text{diag}(d_0, \dots, d_{n-1})$. Hence $\mathbf{A}\mathbf{F} = \mathbf{F}\mathbf{D}$ and $\mathbf{F}^*\mathbf{A}\mathbf{F} = \mathbf{D}$ is diagonal, so $\mathbf{A}$ is a circulant by [Corollary 12.9.5](#) (c). This proves the converse of [Theorem 12.9.1](#) (c) for the single matrix $\mathbf{S}$: the commutant of the cyclic shift is exactly the algebra of circulants.

Exercise. C2: Determinant and invertibility

Let $\mathbf{C}$ be the circulant with first row $\mathbf{c} \in \mathbb{C}^n$ and symbol $p_{\mathbf{c}}$.

- Prove that $\det \mathbf{C} = \prod_{j=0}^{n-1} p_{\mathbf{c}}(\omega^j)$, and use it to compute the determinant of the $n \times n$ all-ones matrix for every $n \geq 2$.
- Prove that $\mathbf{C}$ is invertible if and only if $p_{\mathbf{c}}$ and $x^n - 1$ are coprime in $\mathbb{C}[x]$.
- Give an example of a 4×4 circulant with integer entries, none of them 0, whose determinant is 0.

Solution.

- (a) This is the last claim of [Theorem 12.9.4](#): from $\mathbf{C} = \mathbf{F}\mathbf{D}\mathbf{F}^*$ and $\det(\mathbf{F}^*)\det(\mathbf{F}) = \det(\mathbf{F}^*\mathbf{F}) = 1$, multiplicativity of the determinant (Theorem 7.5.4) gives $\det \mathbf{C} = \det \mathbf{D}$,

which is the product of the diagonal entries $p_{\mathbf{c}}(\omega^j)$.

For the all-ones matrix, $\mathbf{c} = (1, 1, \dots, 1)$ and $p_{\mathbf{c}}(x) = 1 + x + \dots + x^{n-1}$. Then $p_{\mathbf{c}}(1) = n$, while for $1 \leq j \leq n-1$ the number $z = \omega^j$ satisfies $z^n = 1$ and $z \neq 1$, so $p_{\mathbf{c}}(z) = (z^n - 1)/(z - 1) = 0$. Hence $\det \mathbf{C} = n \cdot 0 \cdots 0 = 0$ for every $n \geq 2$, as it must be, since all the rows are equal.

- (b) By (a), $\det \mathbf{C} \neq 0$ if and only if $p_{\mathbf{c}}(\omega^j) \neq 0$ for every j , that is, if and only if $p_{\mathbf{c}}$ and $x^n - 1$ have no common root, the roots of $x^n - 1$ being exactly $1, \omega, \dots, \omega^{n-1}$ (Example (Roots of Unity)). Over $\mathbb{C}$, having no common root is the same as having no common nonconstant factor: such a factor would have a root by Theorem 6.5.1, and that root would be common to both; conversely a common root α makes $x - \alpha$ a common factor. So invertibility of $\mathbf{C}$ is equivalent to $p_{\mathbf{c}}$ and $x^n - 1$ being coprime (Definition 6.2.3). Combined with Theorem 7.5.2, this is the statement asked for.

- (c) Take $\mathbf{c} = (1, 2, 1, 2)$, so

$$\mathbf{C} = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \end{pmatrix}, \quad p_{\mathbf{c}}(x) = 1 + 2x + x^2 + 2x^3.$$

Here $p_{\mathbf{c}}(i) = 1 + 2i - 1 - 2i = 0$, so $\det \mathbf{C} = 0$ by (a). The reason is visible without any of the theory — rows 0 and 2 are equal — and the Fourier picture names a kernel vector: $\mathbf{C}\mathbf{f}_1 = \mathbf{0}$, and since $\mathbf{C}$ is real, the real and imaginary parts of $\mathbf{f}_1$ are in the kernel too, giving $\mathbf{C}(1, 0, -1, 0) = \mathbf{0}$.

Exercise. C3: The real orthogonal form of a real circulant

Let $\mathbf{C} \in M_4(\mathbb{R})$ be the circulant with first row $(1, 2, 3, 4)$, with eigenvalues $10, -2 - 2i, -2, -2 + 2i$ as in [Example \(The running example, diagonalized\)](#). Put

$$\mathbf{g} := \sqrt{2} \operatorname{Re} \mathbf{f}_1, \quad \mathbf{h} := \sqrt{2} \operatorname{Im} \mathbf{f}_1,$$

the real and imaginary parts taken entrywise.

- Write $\mathbf{g}$ and $\mathbf{h}$ out, and show that $(\mathbf{f}_0, \mathbf{f}_2, \mathbf{g}, \mathbf{h})$ is an orthonormal basis of $\mathbb{R}^4$.
- Writing $a + bi := -2 + 2i$, prove that $\mathbf{C}\mathbf{g} = a\mathbf{g} + b\mathbf{h}$ and $\mathbf{C}\mathbf{h} = -b\mathbf{g} + a\mathbf{h}$.
- Hence write down the matrix of $\mathbf{C}$ in the ordered basis of (a), and identify it as an instance of Theorem 12.5.3.

Hint for (b): apply $\mathbf{C}$ to $\mathbf{f}_3 = (\mathbf{g} - i\mathbf{h})/\sqrt{2}$ and compare real and imaginary parts.

Solution.

- (a) With $\omega = i$, $\mathbf{f}_1 = \frac{1}{2}(1, i, -1, -i)$, so

$$\mathbf{g} = \frac{1}{\sqrt{2}}(1, 0, -1, 0), \quad \mathbf{h} = \frac{1}{\sqrt{2}}(0, 1, 0, -1).$$

Also $\mathbf{f}_0 = \frac{1}{2}(1, 1, 1, 1)$ and $\mathbf{f}_2 = \frac{1}{2}(1, -1, 1, -1)$ are real, so all four vectors lie in $\mathbb{R}^4$. Each has norm 1: for instance $\|\mathbf{g}\|^2 = \frac{1}{2}(1 + 1) = 1$ and $\|\mathbf{f}_0\|^2 = \frac{1}{4} \cdot 4 = 1$. They are pairwise orthogonal for the real dot product. First $\mathbf{g} \cdot \mathbf{h} = 0$, because the nonzero

entries sit in different slots, and $\mathbf{f}_0 \cdot \mathbf{f}_2 = \frac{1}{4}(1 - 1 + 1 - 1) = 0$. The four remaining pairings are

$$\begin{aligned}\mathbf{f}_0 \cdot \mathbf{g} &= \frac{1}{2\sqrt{2}}(1 - 1) = 0, & \mathbf{f}_2 \cdot \mathbf{g} &= \frac{1}{2\sqrt{2}}(1 - 1) = 0, \\ \mathbf{f}_0 \cdot \mathbf{h} &= \frac{1}{2\sqrt{2}}(1 - 1) = 0, & \mathbf{f}_2 \cdot \mathbf{h} &= \frac{1}{2\sqrt{2}}(-1 + 1) = 0.\end{aligned}$$

Four orthonormal vectors are independent (Theorem 11.1.4) and there are $\dim \mathbb{R}^4 = 4$ of them, so they form a basis (Theorem 1.6.4).

- (b) From $\mathbf{f}_1 = \frac{1}{\sqrt{2}}(\mathbf{g} + i\mathbf{h})$ and $\mathbf{f}_3 = \overline{\mathbf{f}_1}$ we get $\mathbf{f}_3 = \frac{1}{\sqrt{2}}(\mathbf{g} - i\mathbf{h})$. Its eigenvalue is $p_{\mathbf{c}}(\omega^3) = p_{\mathbf{c}}(-i) = -2 + 2i = a + bi$, so $a = -2$ and $b = 2$, and $\mathbf{C}\mathbf{f}_3 = (a + bi)\mathbf{f}_3$ reads

$$\frac{1}{\sqrt{2}}(\mathbf{C}\mathbf{g} - i\mathbf{C}\mathbf{h}) = \frac{1}{\sqrt{2}}(a + bi)(\mathbf{g} - i\mathbf{h}) = \frac{1}{\sqrt{2}}((a\mathbf{g} + b\mathbf{h}) + i(b\mathbf{g} - a\mathbf{h})).$$

Both $\mathbf{C}\mathbf{g}$ and $\mathbf{C}\mathbf{h}$ are real vectors, because $\mathbf{C}$, $\mathbf{g}$, and $\mathbf{h}$ are real. Comparing real parts gives $\mathbf{C}\mathbf{g} = a\mathbf{g} + b\mathbf{h}$, and comparing imaginary parts gives $-\mathbf{C}\mathbf{h} = b\mathbf{g} - a\mathbf{h}$, that is $\mathbf{C}\mathbf{h} = -b\mathbf{g} + a\mathbf{h}$. With $a = -2$ and $b = 2$,

$$\mathbf{C}\mathbf{g} = -2\mathbf{g} + 2\mathbf{h}, \quad \mathbf{C}\mathbf{h} = -2\mathbf{g} - 2\mathbf{h}.$$

A direct check: $\sqrt{2}\mathbf{C}\mathbf{g} = \mathbf{C}(1, 0, -1, 0) = (1 - 3, 4 - 2, 3 - 1, 2 - 4) = (-2, 2, 2, -2)$, while $\sqrt{2}(-2\mathbf{g} + 2\mathbf{h}) = (-2, 2, 2, -2)$.

- (c) Reading the coordinates of the images off (b), the matrix in the ordered basis $(\mathbf{f}_0, \mathbf{f}_2, \mathbf{g}, \mathbf{h})$ is

$$\begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 2 & -2 \end{pmatrix} = (10) \oplus (-2) \oplus \mathbf{\Lambda}(-2 + 2i),$$

since $\mathbf{\Lambda}(a + bi)$ has rows $(a, -b)$ and (b, a) . This is exactly the shape promised by Theorem 12.5.3 for a real normal operator: two 1×1 blocks for the real eigenvalues, and one rotation-scaling block for the conjugate pair, with $b = 2 > 0$. The change of basis is orthogonal, because the basis is orthonormal.