

Chapter 9

Complex Vectors and Matrices

Real versus Complex

$\mathbf{R}$ = line of all real numbers $-\infty < x < \infty \leftrightarrow \mathbf{C}$ = plane of all complex numbers $z = x + iy$

$|x|$ = absolute value of $x \leftrightarrow |z| = \sqrt{x^2 + y^2} = r$ = absolute value (or modulus) of z
 1 and -1 solve $x^2 = 1 \leftrightarrow z = 1, w, \dots, w^{n-1}$ solve $z^n = 1$ where $w = e^{2\pi i/n}$

The **complex conjugate** of $z = x + iy$ is $\bar{z} = x - iy$. $|z|^2 = x^2 + y^2 = z\bar{z}$ and $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

The **polar form** of $z = x + iy$ is $|z|e^{i\theta} = re^{i\theta} = r \cos \theta + ir \sin \theta$. The angle has $\tan \theta = \frac{y}{x}$.

$\mathbf{R}^n$: vectors with n real components $\leftrightarrow \mathbf{C}^n$: vectors with n complex components

length: $\|x\|^2 = x_1^2 + \dots + x_n^2 \leftrightarrow$ length: $\|z\|^2 = |z_1|^2 + \dots + |z_n|^2$

transpose: $(A^T)_{ij} = A_{ji} \leftrightarrow$ conjugate transpose: $(A^H)_{ij} = \overline{A_{ji}}$

dot product: $x^T y = x_1 y_1 + \dots + x_n y_n \leftrightarrow$ inner product: $u^H v = \bar{u}_1 v_1 + \dots + \bar{u}_n v_n$

reason for A^T : $(Ax)^T y = x^T (A^T y) \leftrightarrow$ reason for A^H : $(Au)^H v = u^H (A^H v)$

orthogonality: $x^T y = 0 \leftrightarrow$ orthogonality: $u^H v = 0$

symmetric matrices: $S = S^T \leftrightarrow$ Hermitian matrices: $S = S^H$

$S = Q\Lambda Q^{-1} = Q\Lambda Q^T$ (real Λ) $\leftrightarrow S = U\Lambda U^{-1} = U\Lambda U^H$ (real Λ)

skew-symmetric matrices: $K^T = -K \leftrightarrow$ skew-Hermitian matrices $K^H = -K$

orthogonal matrices: $Q^T = Q^{-1} \leftrightarrow$ unitary matrices: $U^H = U^{-1}$

orthonormal columns: $Q^T Q = I \leftrightarrow$ orthonormal columns: $U^H U = I$

$(Qx)^T (Qy) = x^T y$ and $\|Qx\| = \|x\| \leftrightarrow (Ux)^H (Uy) = x^H y$ and $\|Uz\| = \|z\|$

A complete presentation of linear algebra must include complex numbers $z = x + iy$. Even when the matrix is real, *the eigenvalues and eigenvectors are often complex*. Example: A 2 by 2 rotation matrix has complex eigenvectors $x = (1, i)$ and $\bar{x} = (1, -i)$. I will summarize Sections 9.1 and 9.2 in these few unforgettable words: When you transpose a vector v or a matrix A , *take the conjugate of every entry (i changes to $-i$)*. Section 9.3 is about the most important complex matrix of all—the *Fourier matrix* F .

9.1 Complex Numbers

Start with the imaginary number i . Everybody knows that $x^2 = -1$ has no real solution. When you square a real number, the answer is never negative. So the world has agreed on a solution called i . (Except that electrical engineers call it j .) Imaginary numbers follow the normal rules of addition and multiplication, with one difference. **Replace i^2 by -1 .**

This section gives the main facts about complex numbers. It is a review for some students and a reference for everyone. Everything comes from $i^2 = -1$ and $e^{2\pi i} = 1$.

A complex number (say $3 + 2i$) **is a real number** (3) **plus an imaginary number** ($2i$). Addition keeps the real and imaginary parts separate. Multiplication uses $i^2 = -1$:

$$\textbf{Add: } (3 + 2i) + (3 + 2i) = 6 + 4i$$

$$\textbf{Multiply: } (3 + 2i)(1 - i) = 3 + 2i - 3i - 2i^2 = 5 - i.$$

If I add $3 + i$ to $1 - i$, the answer is 4. The real numbers $3 + 1$ stay separate from the imaginary numbers $i - i$. We are adding the vectors $(3, 1)$ and $(1, -1)$ to get $(4, 0)$.

The number $(1 + i)^2$ is $1 + i$ times $1 + i$. The rules give the surprising answer $2i$:

$$(1 + i)(1 + i) = 1 + i + i + i^2 = 2i.$$

In the complex plane, $1 + i$ is at an angle of 45° . It is like the vector $(1, 1)$. When we square $1 + i$ to get $2i$, the angle doubles to 90° . If we square again, the answer is $(2i)^2 = -4$. The 90° angle doubled to 180° , the direction of a negative real number.

A real number is just a complex number $z = a + bi$, with zero imaginary part: $b = 0$.

The **real part** is $a = \text{Re}(a + bi)$. The **imaginary part** is $b = \text{Im}(a + bi)$.

The Complex Plane

Complex numbers correspond to points in a plane. Real numbers go along the x axis. Pure imaginary numbers are on the y axis. **The complex number $3 + 2i$ is at the point with coordinates $(3, 2)$.** The number zero, which is $0 + 0i$, is at the origin.

Adding and subtracting complex numbers is like adding and subtracting vectors in the plane. The real component stays separate from the imaginary component. The vectors go head-to-tail as usual. The complex plane $\mathbf{C}^1$ is like the ordinary two-dimensional plane $\mathbf{R}^2$, except that we multiply complex numbers and we didn't multiply vectors.

Now comes an important idea. **The complex conjugate of $3 + 2i$ is $3 - 2i$.** The complex conjugate of $z = 1 - i$ is $\bar{z} = 1 + i$. In general the conjugate of $z = a + bi$ is $\bar{z} = a - bi$. (Some writers use a "bar" on the number and others use a "star": $\bar{z} = z^*$.) The imaginary parts of z and " z bar" have opposite signs. In the complex plane, $\bar{z}$ is the image of z on the other side of the real axis.

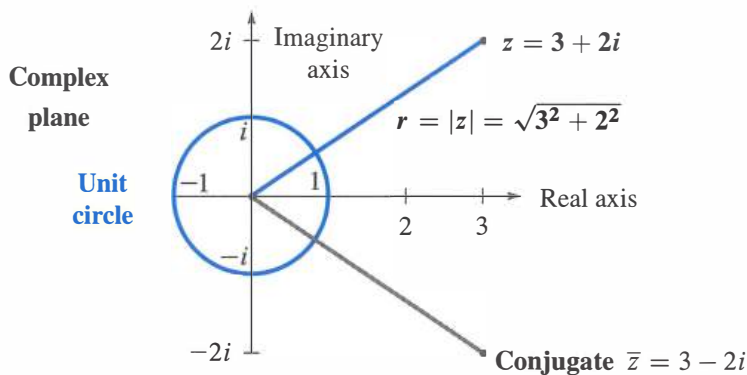


Figure 9.1: The number $z = a + bi$ corresponds to the point (a, b) and the vector $\begin{bmatrix} a \\ b \end{bmatrix}$.

Two useful facts. **When we multiply conjugates $\bar{z}_1$ and $\bar{z}_2$, we get the conjugate of $z_1 z_2$.** And when we add $\bar{z}_1$ and $\bar{z}_2$, we get the conjugate of $z_1 + z_2$:

$$\bar{z}_1 + \bar{z}_2 = (3 - 2i) + (1 + i) = 4 - i. \text{ This is the conjugate of } z_1 + z_2 = 4 + i.$$

$$\bar{z}_1 \times \bar{z}_2 = (3 - 2i) \times (1 + i) = 5 + i. \text{ This is the conjugate of } z_1 \times z_2 = 5 - i.$$

Adding and multiplying is exactly what linear algebra needs. By taking conjugates of $Ax = \lambda x$, when A is real, we have another eigenvalue $\bar{\lambda}$ and its eigenvector $\bar{x}$:

$$\text{Eigenvalues } \lambda \text{ and } \bar{\lambda} \quad \text{If } Ax = \lambda x \text{ and } A \text{ is real then } A\bar{x} = \bar{\lambda}\bar{x}. \quad (1)$$

Something special happens when $z = 3 + 2i$ combines with its own complex conjugate $\bar{z} = 3 - 2i$. The result from adding $z + \bar{z}$ or multiplying $z\bar{z}$ is always real:

$$\begin{aligned} z + \bar{z} &= \text{real} & (3 + 2i) + (3 - 2i) &= 6 \quad (\text{real}) \\ z\bar{z} &= \text{real} & (3 + 2i) \times (3 - 2i) &= 9 + 6i - 6i - 4i^2 = 13 \quad (\text{real}). \end{aligned}$$

The sum of $z = a + bi$ and its conjugate $\bar{z} = a - bi$ is the real number $2a$. The product of z times $\bar{z}$ is the real number $a^2 + b^2$:

$$\text{Multiply } z \text{ times } \bar{z} \text{ to get } |z|^2 = r^2 \quad (a + bi)(a - bi) = a^2 + b^2. \quad (2)$$

The next step with complex numbers is $1/z$. How to divide by $a + ib$? The best idea is to multiply first by $\bar{z}/\bar{z} = 1$. That produces $z\bar{z}$ in the denominator, which is $a^2 + b^2$:

$$\frac{1}{a + ib} = \frac{1}{a + ib} \frac{a - ib}{a - ib} = \frac{a - ib}{a^2 + b^2} \quad \frac{1}{3 + 2i} = \frac{1}{3 + 2i} \frac{3 - 2i}{3 - 2i} = \frac{3 - 2i}{13}.$$

In case $a^2 + b^2 = 1$, this says that $(a + ib)^{-1}$ is $a - ib$. **On the unit circle, $1/z$ equals $\bar{z}$.** Later we will say: $1/e^{i\theta}$ is $e^{-i\theta}$. Use distance r and angle θ to multiply and divide.

The Polar Form $re^{i\theta}$

The square root of $a^2 + b^2$ is $|z|$. This is the **absolute value** (or **modulus**) of the number $z = a + ib$. The square root $|z|$ is also written r , because it is the distance from 0 to z . **The real number r in the polar form gives the size of the complex number z :**

The absolute value of $z = a + ib$ is $|z| = \sqrt{a^2 + b^2}$. **This is called r .**

The absolute value of $z = 3 + 2i$ is $|z| = \sqrt{3^2 + 2^2}$. This is $r = \sqrt{13}$.

The other part of the polar form is the angle θ . The angle for $z = 5$ is $\theta = 0$ (because this z is real and positive). The angle for $z = 3i$ is $\pi/2$ radians. The angle for a negative $z = -9$ is π radians. **The angle doubles when the number is squared.** The polar form is excellent for multiplying complex numbers (not good for addition).

When the distance is r and the angle is θ , trigonometry gives the other two sides of the triangle. The real part (along the bottom) is $a = r \cos \theta$. The imaginary part (up or down) is $b = r \sin \theta$. Put those together, and the rectangular form becomes the polar form $re^{i\theta}$.

The number $z = a + ib$ is also $z = r \cos \theta + ir \sin \theta$. This is $re^{i\theta}$

Note: $\cos \theta + i \sin \theta$ has absolute value $r = 1$ because $\cos^2 \theta + \sin^2 \theta = 1$. Thus $\cos \theta + i \sin \theta$ lies on the circle of radius 1—the unit circle.

Example 1 Find r and θ for $z = 1 + i$ and also for the conjugate $\bar{z} = 1 - i$.

Solution The absolute value is the same for z and $\bar{z}$. It is $r = \sqrt{1 + 1} = \sqrt{2}$:

$$|z|^2 = 1^2 + 1^2 = 2 \quad \text{and also} \quad |\bar{z}|^2 = 1^2 + (-1)^2 = 2.$$

The distance from the center is $r = \sqrt{2}$. What about the angle θ ? The number $1 + i$ is at the point $(1, 1)$ in the complex plane. The angle to that point is $\pi/4$ radians or 45° . The cosine is $1/\sqrt{2}$ and the sine is $1/\sqrt{2}$. Combining r and θ brings back $z = 1 + i$:

$$r \cos \theta + ir \sin \theta = \sqrt{2} \left(\frac{1}{\sqrt{2}} \right) + i\sqrt{2} \left(\frac{1}{\sqrt{2}} \right) = 1 + i.$$

The angle to the conjugate $1 - i$ can be positive or negative. We can go to $7\pi/4$ radians which is 315° . Or we can go *backwards through a negative angle*, to $-\pi/4$ radians or -45° . **If z is at angle θ , its conjugate $\bar{z}$ is at $2\pi - \theta$ and also at $-\theta$.**

We can freely add 2π or 4π or -2π to any angle! Those go full circles so the final point is the same. This explains why there are infinitely many choices of θ . Often we select the angle between 0 and 2π . But $-\theta$ is very useful for the conjugate $\bar{z}$. And $1 = e^0 = e^{2\pi i}$.

Powers and Products: Polar Form

Computing $(1 + i)^2$ and $(1 + i)^8$ is quickest in polar form. That form has $r = \sqrt{2}$ and $\theta = \pi/4$ (or 45°). If we square the absolute value to get $r^2 = 2$, and double the angle to get $2\theta = \pi/2$ (or 90°), we have $(1 + i)^2$. For the eighth power we need r^8 and 8θ :

$$(1 + i)^8 \quad r^8 = 2 \cdot 2 \cdot 2 \cdot 2 = 16 \quad \text{and} \quad 8\theta = 8 \cdot \frac{\pi}{4} = 2\pi.$$

This means: $(1 + i)^8$ has absolute value 16 and angle 2π . So $(1 + i)^8 = 16$.

Powers are easy in polar form. So is multiplication of complex numbers.

The n th power of $z = r(\cos \theta + i \sin \theta)$ is $z^n = r^n(\cos n\theta + i \sin n\theta)$. (3)

In that case z multiplies itself. To multiply z times z' , **multiply r 's and add angles**:

$$r(\cos \theta + i \sin \theta) \text{ times } r'(\cos \theta' + i \sin \theta') = rr'(\cos(\theta + \theta') + i \sin(\theta + \theta')). \quad (4)$$

One way to understand this is by trigonometry. Why do we get the double angle 2θ for z^2 ?

$$(\cos \theta + i \sin \theta) \times (\cos \theta + i \sin \theta) = \cos^2 \theta + i^2 \sin^2 \theta + 2i \sin \theta \cos \theta.$$

The real part $\cos^2 \theta - \sin^2 \theta$ is $\cos 2\theta$. The imaginary part $2 \sin \theta \cos \theta$ is $\sin 2\theta$. Those are the “double angle” formulas. They show that θ in z becomes 2θ in z^2 .

There is a second way to understand the rule for z^n . It uses the only amazing formula in this section. Remember that $\cos \theta + i \sin \theta$ has absolute value 1. The cosine is made up of even powers, starting with $1 - \frac{1}{2}\theta^2$. The sine is made up of odd powers, starting with $\theta - \frac{1}{6}\theta^3$. The beautiful fact is that $e^{i\theta}$ combines both of those series into $\cos \theta + i \sin \theta$:

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots \quad \text{becomes} \quad e^{i\theta} = 1 + i\theta + \frac{1}{2}i^2\theta^2 + \frac{1}{6}i^3\theta^3 + \cdots$$

Write -1 for i^2 to see $1 - \frac{1}{2}\theta^2$. **The complex number $e^{i\theta}$ is $\cos \theta + i \sin \theta$:**

Euler's Formula $e^{i\theta} = \cos \theta + i \sin \theta$ gives $z = r \cos \theta + ir \sin \theta = re^{i\theta}$ (5)

The special choice $\theta = 2\pi$ gives $\cos 2\pi + i \sin 2\pi$ which is 1. Somehow the infinite series $e^{2\pi i} = 1 + 2\pi i + \frac{1}{2}(2\pi i)^2 + \cdots$ adds up to 1.

Now multiply $e^{i\theta}$ times $e^{i\theta'}$. Angles add for the same reason that exponents add:

$$e^2 \text{ times } e^3 \text{ is } e^5$$

$$e^{i\theta} \text{ times } e^{i\theta'} \text{ is } e^{2i\theta}$$

$$e^{i\theta} \text{ times } e^{i\theta'} \text{ is } e^{i(\theta+\theta')}$$

The powers $(re^{i\theta})^n$ are equal to $r^n e^{in\theta}$. They stay on the unit circle when $r = 1$ and $r^n = 1$. Then we find n different numbers whose n th powers equal 1:

Set $w = e^{2\pi i/n}$. The n th powers of $1, w, w^2, \dots, w^{n-1}$ all equal 1.

Those are the “ n th roots of 1.” They solve the equation $z^n = 1$. They are equally spaced around the unit circle in Figure 9.2b, where the full 2π is divided by n . Multiply their angles by n to take n th powers. That gives $w^n = e^{2\pi i}$ which is 1. Also $(w^2)^n = e^{4\pi i} = 1$. Each of those numbers, to the n th power, comes around the unit circle to 1.

These n roots of 1 are the key numbers for signal processing. The Discrete Fourier Transform uses $w = e^{2\pi i/n}$ and its powers. Section 9.3 shows how to decompose a vector (a signal) into n frequencies by the Fast Fourier Transform.

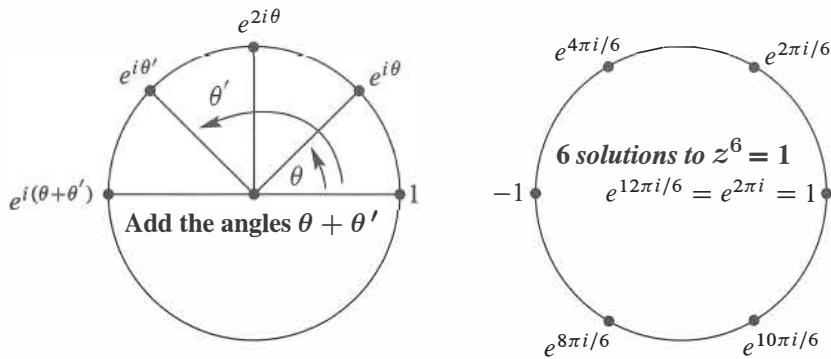


Figure 9.2: (a) $e^{i\theta}$ times $e^{i\theta'}$ is $e^{i(\theta+\theta')}$. (b) The n th power of $e^{2\pi i/n}$ is $e^{2\pi i} = 1$.

■ REVIEW OF THE KEY IDEAS ■

1. Adding $a + ib$ to $c + id$ is like adding $(a, b) + (c, d)$. Use $i^2 = -1$ to multiply.
2. The conjugate of $z = a + bi = re^{i\theta}$ is $\bar{z} = z^* = a - bi = re^{-i\theta}$.
3. z times $\bar{z}$ is $re^{i\theta}$ times $re^{-i\theta}$. This is $r^2 = |z|^2 = a^2 + b^2$ (real).
4. Powers and products are easy in polar form $z = re^{i\theta}$. Multiply r 's and add θ 's.

Problem Set 9.1

Questions 1–8 are about operations on complex numbers.

- Add and multiply each pair of complex numbers:
 (a) $2 + i, 2 - i$ (b) $-1 + i, -1 + i$ (c) $\cos \theta + i \sin \theta, \cos \theta - i \sin \theta$
- Locate these points on the complex plane. Simplify them if necessary:
 (a) $2 + i$ (b) $(2 + i)^2$ (c) $\frac{1}{2+i}$ (d) $|2 + i|$
- Find the absolute value $r = |z|$ of these four numbers. If θ is the angle for $6 - 8i$, what are the angles for the other three numbers?
 (a) $6 - 8i$ (b) $(6 - 8i)^2$ (c) $\frac{1}{6-8i}$ (d) $(6 + 8i)^2$
- If $|z| = 2$ and $|w| = 3$ then $|z \times w| = \underline{\hspace{1cm}}$ and $|z + w| \leq \underline{\hspace{1cm}}$ and $|z/w| = \underline{\hspace{1cm}}$ and $|z - w| \leq \underline{\hspace{1cm}}$.
- Find $a + ib$ for the numbers at angles $30^\circ, 60^\circ, 90^\circ, 120^\circ$ on the unit circle. If w is the number at 30° , check that w^2 is at 60° . What power of w equals 1?
- If $z = r \cos \theta + ir \sin \theta$ then $1/z$ has absolute value $\underline{\hspace{1cm}}$ and angle $\underline{\hspace{1cm}}$. Its polar form is $\underline{\hspace{1cm}}$. Multiply $z \times 1/z$ to get 1.
- The complex multiplication $M = (a + bi)(c + di)$ is a 2 by 2 real multiplication

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \\ \end{bmatrix}.$$

The right side contains the real and imaginary parts of M . Test $M = (1 + 3i)(1 - 3i)$.

- $A = A_1 + iA_2$ is a complex n by n matrix and $\mathbf{b} = \mathbf{b}_1 + i\mathbf{b}_2$ is a complex vector. The solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x}_1 + i\mathbf{x}_2$. Write $A\mathbf{x} = \mathbf{b}$ as a real system of size $2n$:

$$\begin{array}{l} \text{Complex } n \text{ by } n \\ \text{Real } 2n \text{ by } 2n \end{array} \quad \begin{bmatrix} \\ \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}.$$

Questions 9–16 are about the conjugate $\bar{z} = a - ib = re^{-i\theta} = z^*$.

- Write down the complex conjugate of each number by changing i to $-i$:
 (a) $2 - i$ (b) $(2 - i)(1 - i)$ (c) $e^{i\pi/2}$ (which is i)
 (d) $e^{i\pi} = -1$ (e) $\frac{1+i}{1-i}$ (which is also i) (f) $i^{103} = \underline{\hspace{1cm}}$.
- The sum $z + \bar{z}$ is always $\underline{\hspace{1cm}}$. The difference $z - \bar{z}$ is always $\underline{\hspace{1cm}}$. Assume $z \neq 0$. The product $z \times \bar{z}$ is always $\underline{\hspace{1cm}}$. The ratio $z/\bar{z}$ has absolute value $\underline{\hspace{1cm}}$.

- 11 For a real matrix, the conjugate of $Ax = \lambda x$ is $A\bar{x} = \bar{\lambda}\bar{x}$. This proves two things: $\bar{\lambda}$ is another eigenvalue and $\bar{x}$ is its eigenvector. Find the eigenvalues $\lambda, \bar{\lambda}$ and eigenvectors $x, \bar{x}$ of $A = \begin{bmatrix} a & b \\ c & -b \end{bmatrix}$.

- 12 The eigenvalues of a real 2 by 2 matrix come from the quadratic formula:

$$\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

gives the two eigenvalues $\lambda = \left[a + d \pm \sqrt{(a + d)^2 - 4(ad - bc)} \right] / 2$.

- (a) If $a = b = d = 1$, the eigenvalues are complex when c is _____.
 (b) What are the eigenvalues when $ad = bc$?
- 13 In Problem 12 the eigenvalues are not real when $(\text{trace})^2 = (a + d)^2$ is smaller than _____. Show that the λ 's are real when $bc > 0$.
- 14 A real skew-symmetric matrix ($A^T = -A$) has pure imaginary eigenvalues. First proof: If $Ax = \lambda x$ then block multiplication gives

$$\begin{bmatrix} 0 & A \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ ix \end{bmatrix} = i\lambda \begin{bmatrix} x \\ ix \end{bmatrix}.$$

This block matrix is symmetric. Its eigenvalues must be _____! So λ is _____.

Questions 15–22 are about the form $re^{i\theta}$ of the complex number $r \cos \theta + ir \sin \theta$.

- 15 Write these numbers in Euler's form $re^{i\theta}$. Then square each number:
 (a) $1 + \sqrt{3}i$ (b) $\cos 2\theta + i \sin 2\theta$ (c) $-7i$ (d) $5 - 5i$.
- 16 (A favorite) Find the absolute value and the angle for $z = \sin \theta + i \cos \theta$ (careful). Locate this z in the complex plane. Multiply z by $\cos \theta + i \sin \theta$ to get _____.
- 17 Draw all eight solutions of $z^8 = 1$ in the complex plane. What is the rectangular form $a + ib$ of the root $z = \bar{w} = \exp(-2\pi i/8)$?
- 18 Locate the cube roots of 1 in the complex plane. Locate the cube roots of -1 . Together these are the sixth roots of _____.
- 19 By comparing $e^{3i\theta} = \cos 3\theta + i \sin 3\theta$ with $(e^{i\theta})^3 = (\cos \theta + i \sin \theta)^3$, find the "triple angle" formulas for $\cos 3\theta$ and $\sin 3\theta$ in terms of $\cos \theta$ and $\sin \theta$.
- 20 Suppose the conjugate $\bar{z}$ is equal to the reciprocal $1/z$. What are all possible z 's?
- 21 (a) Why do e^i and i^e both have absolute value 1?
 (b) In the complex plane put stars near the points e^i and i^e .
 (c) The number i^e could be $(e^{i\pi/2})^e$ or $(e^{5i\pi/2})^e$. Are those equal?
- 22 Draw the paths of these numbers from $t = 0$ to $t = 2\pi$ in the complex plane:
 (a) e^{it} (b) $e^{(-1+i)t} = e^{-t}e^{it}$ (c) $(-1)^t = e^{t\pi i}$.

9.2 Hermitian and Unitary Matrices

The main message of this section can be presented in one sentence: **When you transpose a complex vector z or matrix A , take the complex conjugate too.** Don't stop at z^T or A^T . Reverse the signs of all imaginary parts. From a column vector with $z_j = a_j + ib_j$, the good row vector $\bar{z}^T$ is the *conjugate transpose* with components $a_j - ib_j$:

$$\text{Conjugate transpose} \quad \bar{z}^T = [\bar{z}_1 \ \cdots \ \bar{z}_n] = [a_1 - ib_1 \ \cdots \ a_n - ib_n]. \quad (1)$$

Here is one reason to go to $\bar{z}$. The length squared of a real vector is $x_1^2 + \cdots + x_n^2$. The length squared of a complex vector is *not* $z_1^2 + \cdots + z_n^2$. With that wrong definition, the length of $(1, i)$ would be $1^2 + i^2 = 0$. A nonzero vector would have zero length—not good. Other vectors would have complex lengths. Instead of $(a + bi)^2$ we want $a^2 + b^2$, the *absolute value squared*. This is $(a + bi)$ times $(a - bi)$.

For each component we want z_j times $\bar{z}_j$, which is $|z_j|^2 = a_j^2 + b_j^2$. That comes when the components of z multiply the components of $\bar{z}$:

$$\text{Length squared} \quad [\bar{z}_1 \ \cdots \ \bar{z}_n] \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = |z_1|^2 + \cdots + |z_n|^2. \quad \text{This is } \bar{z}^T z = \|z\|^2. \quad (2)$$

Now the squared length of $(1, i)$ is $1^2 + |i|^2 = 2$. The length is $\sqrt{2}$. The squared length of $(1 + i, 1 - i)$ is 4. The only vectors with zero length are zero vectors.

The length $\|z\|$ is the square root of $\bar{z}^T z = z^H z = |z_1|^2 + \cdots + |z_n|^2$

Before going further we replace two symbols by one symbol. Instead of a bar for the conjugate and T for the transpose, we just use a superscript H. Thus $\bar{z}^T = z^H$. This is “ z Hermitian,” the *conjugate transpose* of z . The new word is pronounced “Hermeeeshan.” The new symbol applies also to matrices: The conjugate transpose of a matrix A is A^H .

Another popular notation is A^* . The MATLAB transpose command $'$ automatically takes complex conjugates (z' is $z^H = \bar{z}^T$ and A' is $A^H = \bar{A}^T$).

$$A^H \text{ is “A Hermitian”} \quad \text{If } A = \begin{bmatrix} 1 & i \\ 0 & 1 + i \end{bmatrix} \quad \text{then } A^H = \begin{bmatrix} 1 & 0 \\ -i & 1 - i \end{bmatrix}$$

Complex Inner Products

For real vectors, the length squared is $x^T x$ —the *inner product of x with itself*. For complex vectors, the length squared is $z^H z$. It will be very desirable if $z^H z$ is the inner product of z with itself. To make that happen, the complex inner product should use the conjugate transpose (not just the transpose). This has no effect on real vectors.

DEFINITION The inner product of real or complex vectors u and v is $u^H v$:

$$u^H v = [\bar{u}_1 \ \cdots \ \bar{u}_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \bar{u}_1 v_1 + \cdots + \bar{u}_n v_n. \quad (3)$$

With complex vectors, $u^H v$ is different from $v^H u$. *The order of the vectors is now important.* In fact $v^H u = \bar{v}_1 u_1 + \cdots + \bar{v}_n u_n$ is the complex conjugate of $u^H v$. We have to put up with a few inconveniences for the greater good.

Example 1 The inner product of $u = \begin{bmatrix} 1 \\ i \end{bmatrix}$ with $v = \begin{bmatrix} i \\ 1 \end{bmatrix}$ is $[1 \ -i] \begin{bmatrix} i \\ 1 \end{bmatrix} = 0$.

Example 1 is surprising. Those vectors $(1, i)$ and $(i, 1)$ don't look perpendicular. But they are. *A zero inner product still means that the (complex) vectors are orthogonal.* Similarly the vector $(1, i)$ is orthogonal to the vector $(1, -i)$. Their inner product is $1 - 1 = 0$. We are correctly getting zero for the inner product—where we would be incorrectly getting zero for the length of $(1, i)$ if we forgot to take the conjugate.

Note We have chosen to conjugate the first vector u . Some authors choose the second vector v . Their complex inner product would be $u^T \bar{v}$. I think it is a free choice.

The inner product of Au with v equals the inner product of u with $A^H v$:

$$A^H \text{ is also called the "adjoint" of } A \quad (Au)^H v = u^H (A^H v). \quad (4)$$

The conjugate of Au is $\overline{Au}$. Transposing $\overline{Au}$ gives $\bar{u}^T \bar{A}^T$ as usual. This is $u^H A^H$. Everything that should work, does work. The rule for H comes from the rule for T . We constantly use the fact that $(a - ib)(c - id)$ is the conjugate of $(a + ib)(c + id)$.

The conjugate transpose of AB is $(AB)^H = B^H A^H$.

Hermitian Matrices $S = S^H$

Among real matrices, *symmetric matrices* form the most important special class: $S = S^T$. They have real eigenvalues and the orthogonal eigenvectors in an orthogonal matrix Q . Every real symmetric matrix can be written as $S = Q\Lambda Q^{-1}$ and also as $S = Q\Lambda Q^T$ (because $Q^{-1} = Q^T$). All this follows from $S^T = S$, when S is real.

Among complex matrices, the special class contains the **Hermitian matrices**: $S = S^H$. The condition on the entries is $s_{ij} = \overline{s_{ji}}$. In this case we say that “ S is Hermitian.” Every real symmetric matrix is Hermitian, because taking its conjugate has no effect. The next matrix is also Hermitian, $S = S^H$:

Example 2 $S = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}$ The main diagonal must be real since $s_{ii} = \overline{s_{ii}}$.
Across it are conjugates $3+3i$ and $3-3i$.

This example will illustrate the three crucial properties of all Hermitian matrices.

If $S = S^H$ and z is any real or complex column vector, the number $z^H S z$ is real.

Quick proof: $z^H S z$ is certainly 1 by 1. Take its conjugate transpose:

$$(z^H S z)^H = z^H S^H (z^H)^H \text{ which is } z^H S z \text{ again.}$$

So the number $z^H S z$ equals its conjugate and must be real. Here is that “energy” $z^H S z$:

$$\begin{bmatrix} \bar{z}_1 & \bar{z}_2 \end{bmatrix} \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \underbrace{2\bar{z}_1 z_1 + 5\bar{z}_2 z_2}_{\text{diagonal}} + \underbrace{(3-3i)\bar{z}_1 z_2 + (3+3i)z_1 \bar{z}_2}_{\text{off-diagonal}}.$$

The terms $2|z_1|^2$ and $5|z_2|^2$ from the diagonal are both real. The off-diagonal terms are conjugates of each other—so their sum is real. (The imaginary parts cancel when we add.) The whole expression $z^H S z$ is real, and this will make λ real.

Every eigenvalue of a Hermitian matrix is real.

Proof Suppose $Sz = \lambda z$. Multiply both sides by z^H to get $z^H S z = \lambda z^H z$. On the left side, $z^H S z$ is real. On the right side, $z^H z$ is the length squared, real and positive. So the ratio $\lambda = z^H S z / z^H z$ is a real number. Q.E.D.

The example above has eigenvalues $\lambda = 8$ and $\lambda = -1$, real because $S = S^H$:

$$\begin{vmatrix} 2-\lambda & 3-3i \\ 3+3i & 5-\lambda \end{vmatrix} = \lambda^2 - 7\lambda + 10 - |3+3i|^2 \\ = \lambda^2 - 7\lambda + 10 - 18 = (\lambda - 8)(\lambda + 1).$$

The eigenvectors of a Hermitian matrix are orthogonal (when they correspond to different eigenvalues). If $Sz = \lambda z$ and $Sy = \beta y$ then $y^H z = 0$.

Proof Multiply $Sz = \lambda z$ on the left by y^H . Multiply $y^H S^H = \beta y^H$ on the right by z :

$$y^H S z = \lambda y^H z \quad \text{and} \quad y^H S^H z = \beta y^H z. \quad (5)$$

The left sides are equal so $\lambda y^H z = \beta y^H z$. Then $y^H z$ must be zero.

The eigenvectors are orthogonal in our example with $\lambda = 8$ and $\beta = -1$:

$$(S - 8I)z = \begin{bmatrix} -6 & 3-3i \\ 3+3i & -3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad z = \begin{bmatrix} 1 \\ 1+i \end{bmatrix}$$

$$(S + I)y = \begin{bmatrix} 3 & 3-3i \\ 3+3i & 6 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 1-i \\ -1 \end{bmatrix}.$$

Orthogonal eigenvectors $y^H z = [1+i \quad -1] \begin{bmatrix} 1 \\ 1+i \end{bmatrix} = 0.$

These eigenvectors have squared length $1^2 + 1^2 + 1^2 = 3$. After division by $\sqrt{3}$ they are unit vectors. They were orthogonal, now they are **orthonormal**. They go into the columns of the *eigenvector matrix* X , which diagonalizes S .

When S is real and symmetric, X is Q —an orthogonal matrix. Now S is complex and Hermitian. Its eigenvectors are complex and orthonormal. **The eigenvector matrix X is like Q , but complex: $Q^H Q = I$.** We assign Q a new name “unitary” but still call it Q .

Unitary Matrices

A **unitary matrix** Q is a (complex) square matrix that has **orthonormal columns**.

Unitary matrix that diagonalizes S : $Q = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}$

This Q is also a Hermitian matrix. I didn’t expect that! The example is almost too perfect. We will see that the eigenvalues of this Q must be 1 and -1 .

The matrix test for real orthonormal columns was $Q^T Q = I$. The zero inner products appear off the diagonal. In the complex case, Q^T becomes Q^H . The columns show themselves as orthonormal when Q^H multiplies Q . The inner products fill up $Q^H Q = I$:

Every matrix Q with orthonormal columns has $Q^H Q = I$.

If Q is square, it is a unitary matrix. Then $Q^H = Q^{-1}$.

Suppose Q (with orthonormal columns) multiplies any z . The vector length stays the same, because $z^H Q^H Q z = z^H z$. If z is an eigenvector of Q we learn something more: **The eigenvalues of unitary (and orthogonal) matrices Q all have absolute value $|\lambda| = 1$.**

If Q is unitary then $\|Qz\| = \|z\|$. Therefore $Qz = \lambda z$ leads to $|\lambda| = 1$.

Our 2 by 2 example is both Hermitian ($Q = Q^H$) and unitary ($Q^{-1} = Q^H$). That means real eigenvalues and it means $|\lambda| = 1$. A real number with $|\lambda| = 1$ has only two possibilities: **The eigenvalues are 1 or -1 .** The trace of Q is zero so $\lambda = 1$ and $\lambda = -1$.

Example 3 The 3 by 3 *Fourier matrix* is in Figure 9.3. Is it Hermitian? Is it unitary? F_3 is certainly symmetric. It equals its transpose. But it doesn't equal its conjugate transpose—it is *not Hermitian*. If you change i to $-i$, you get a different matrix.

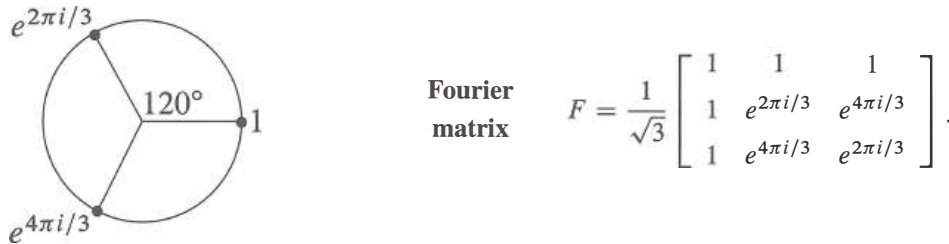


Figure 9.3: The cube roots of 1 go into the Fourier matrix $F = F_3$.

Is F unitary? *Yes*. The squared length of every column is $\frac{1}{3}(1 + 1 + 1)$ (unit vector). The first column is orthogonal to the second column because $1 + e^{2\pi i/3} + e^{4\pi i/3} = 0$. This is the sum of the three numbers marked in Figure 9.3.

Notice the symmetry of the figure. If you rotate it by 120° , the three points are in the same position. Therefore their sum S also stays in the same position! The only possible sum in the same position after 120° rotation is $S = 0$.

Is column 2 of F orthogonal to column 3? Their dot product looks like

$$\frac{1}{3}(1 + e^{6\pi i/3} + e^{6\pi i/3}) = \frac{1}{3}(1 + 1 + 1).$$

This is not zero. The answer is wrong because we forgot to take complex conjugates. The complex inner product uses H not T :

$$\begin{aligned} (\text{column 2})^H (\text{column 3}) &= \frac{1}{3}(1 \cdot 1 + e^{-2\pi i/3} e^{4\pi i/3} + e^{-4\pi i/3} e^{2\pi i/3}) \\ &= \frac{1}{3}(1 + e^{2\pi i/3} + e^{-2\pi i/3}) = 0. \end{aligned}$$

So we do have orthogonality. **Conclusion:** F is a unitary matrix.

The next section will study the n by n Fourier matrices. Among all complex unitary matrices, these are the most important. When we multiply a vector by F , we are computing its **Discrete Fourier Transform**. When we multiply by F^{-1} , we are computing the *inverse transform*. The special property of unitary matrices is that $F^{-1} = F^H$. The inverse transform only differs by changing i to $-i$:

Change i to $-i$
$$F^{-1} = F^H = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{-2\pi i/3} & e^{-4\pi i/3} \\ 1 & e^{-4\pi i/3} & e^{-2\pi i/3} \end{bmatrix}.$$

Everyone who works with F recognizes its value. The last section of this chapter will bring together Fourier analysis and complex numbers and linear algebra.

Problem Set 9.2

- Find the lengths of $\mathbf{u} = (1 + i, 1 - i, 1 + 2i)$ and $\mathbf{v} = (i, i, i)$. Find $\mathbf{u}^H \mathbf{v}$ and $\mathbf{v}^H \mathbf{u}$.
- Compute $A^H A$ and AA^H . Those are both _____ matrices:

$$A = \begin{bmatrix} i & 1 & i \\ 1 & i & i \end{bmatrix}.$$

- Solve $A\mathbf{z} = \mathbf{0}$ to find a vector $\mathbf{z}$ in the nullspace of A in Problem 2. Show that $\mathbf{z}$ is orthogonal to the columns of A^H . Show that $\mathbf{z}$ is *not* orthogonal to the columns of A^T . **The good row space is no longer $C(A^T)$. Now it is $C(A^H)$.**
- Problem 3 indicates that the four fundamental subspaces are $C(A)$ and $N(A)$ and _____ and _____. Their dimensions are still r and $n - r$ and r and $m - r$. They are still orthogonal subspaces. *The symbol H takes the place of T .*
- Prove that $A^H A$ is always a Hermitian matrix.
 - If $A\mathbf{z} = \mathbf{0}$ then $A^H A\mathbf{z} = \mathbf{0}$. If $A^H A\mathbf{z} = \mathbf{0}$, multiply by $\mathbf{z}^H$ to prove that $A\mathbf{z} = \mathbf{0}$. The nullspaces of A and $A^H A$ are _____. Therefore $A^H A$ is an invertible Hermitian matrix when the nullspace of A contains only $\mathbf{z} = \mathbf{0}$.
- True or false (give a reason if true or a counterexample if false):
 - If A is a real matrix then $A + iI$ is invertible.
 - If S is a Hermitian matrix then $S + iI$ is invertible.
 - If Q is a unitary matrix then $Q + iI$ is invertible.
- When you multiply a Hermitian matrix by a real number c , is cS still Hermitian? Show that iS is skew-Hermitian when S is Hermitian. The 3 by 3 Hermitian matrices are a subspace provided the “scalars” are real numbers.
- Which classes of matrices does P belong to: invertible, Hermitian, unitary?

$$P = \begin{bmatrix} 0 & i & 0 \\ 0 & 0 & i \\ i & 0 & 0 \end{bmatrix}.$$

Compute P^2 , P^3 , and P^{100} . What are the eigenvalues of P ?

- Find the unit eigenvectors of P in Problem 8, and put them into the columns of a unitary matrix Q . What property of P makes these eigenvectors orthogonal?
- Write down the 3 by 3 circulant matrix $C = 2I + 5P$. It has the same eigenvectors as P in Problem 8. Find its eigenvalues.
- If Q and U are unitary matrices, show that Q^{-1} is unitary and also QU is unitary. Start from $Q^H Q = I$ and $U^H U = I$.

- 12 How do you know that the determinant of every Hermitian matrix is real?
- 13 The matrix $A^H A$ is not only Hermitian but also positive definite, when the columns of A are independent. Proof: $z^H A^H A z$ is positive if z is nonzero because _____.
- 14 Diagonalize these Hermitian matrices to reach $S = Q \Lambda Q^H$:

$$S = \begin{bmatrix} 0 & 1-i \\ i+1 & 1 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 2 & 1+i \\ i-1 & 3 \end{bmatrix}.$$

- 15 Diagonalize this skew-Hermitian matrix to reach $K = Q \Lambda Q^H$. All λ 's are _____:

$$K = \begin{bmatrix} 0 & -1+i \\ 1+i & i \end{bmatrix}.$$

- 16 Diagonalize this orthogonal matrix to reach $U = Q \Lambda Q^H$. Now all λ 's are _____:

$$U = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

- 17 Diagonalize this unitary matrix to reach $U = Q \Lambda Q^H$. Again all λ 's are _____:

$$U = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}.$$

- 18 If $v_1, \dots, v_n$ is an orthonormal basis for $\mathbf{C}^n$, the matrix with those columns is a _____ matrix. Show that any vector z equals $(v_1^H z)v_1 + \dots + (v_n^H z)v_n$.
- 19 $v = (1, i, 1)$, $w = (i, 1, 0)$ and $z = \underline{\hspace{1cm}}$ are an orthogonal basis for _____.
- 20 If $S = A + iB$ is a Hermitian matrix, are its real and imaginary parts symmetric?
- 21 The (complex) dimension of $\mathbf{C}^n$ is _____. Find a non-real basis for $\mathbf{C}^n$.
- 22 Describe all 1 by 1 and 2 by 2 Hermitian matrices and unitary matrices.
- 23 How are the eigenvalues of A^H related to the eigenvalues of the square matrix A ?
- 24 If $u^H u = 1$ show that $I - 2uu^H$ is Hermitian and also unitary. The rank-one matrix uu^H is the projection onto what line in $\mathbf{C}^n$?
- 25 If $A + iB$ is a unitary matrix (A and B are real) show that $Q = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$ is an orthogonal matrix.
- 26 If $A + iB$ is Hermitian (A and B are real) show that $\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$ is symmetric.
- 27 Prove that the inverse of a Hermitian matrix is also Hermitian (transpose $S^{-1}S = I$).
- 28 A matrix with orthonormal eigenvectors has the form $N = Q \Lambda Q^{-1} = Q \Lambda Q^H$. Prove that $NN^H = N^H N$. These N are exactly the **normal matrices**. Examples are Hermitian, skew-Hermitian, and unitary matrices. Construct a 2 by 2 normal matrix from $Q \Lambda Q^H$ by choosing complex eigenvalues in Λ .

9.3 The Fast Fourier Transform

Many applications of linear algebra take time to develop. It is not easy to explain them in an hour. The teacher and the author must choose between completing the theory and adding new applications. Often the theory wins, but this section is an exception. It explains the most valuable numerical algorithm in the last century.

We want to multiply quickly by F and F^{-1} , the Fourier matrix and its inverse. This is achieved by the Fast Fourier Transform. An ordinary product Fc uses n^2 multiplications (F has n^2 entries). The FFT needs only n times $\frac{1}{2} \log_2 n$. We will see how.

The FFT has revolutionized signal processing. Whole industries are speeded up by this one idea. Electrical engineers are the first to know the difference—they take your Fourier transform as they meet you (if you are a function). Fourier’s idea is to represent f as a sum of harmonics $c_k e^{ikx}$. The function is seen in *frequency space* through the coefficients c_k , instead of *physical space* through its values $f(x)$. The passage backward and forward between c ’s and f ’s is by the Fourier transform. Fast passage is by the FFT.

Roots of Unity and the Fourier Matrix

Quadratic equations have two roots (or one repeated root). Equations of degree n have n roots (counting repetitions). This is the Fundamental Theorem of Algebra, and to make it true we must allow complex roots. This section is about the very special equation $z^n = 1$. The solutions z are the “ n th roots of unity.” They are n evenly spaced points around the unit circle in the complex plane.

Figure 9.4 shows the eight solutions to $z^8 = 1$. Their spacing is $\frac{1}{8}(360^\circ) = 45^\circ$. The first root is at 45° or $\theta = 2\pi/8$ radians. **It is the complex number** $w = e^{i\theta} = e^{i2\pi/8}$. We call this number w_8 to emphasize that it is an 8th root. You could write it in terms of $\cos \frac{2\pi}{8}$ and $\sin \frac{2\pi}{8}$, but don’t do it. The seven other 8th roots are $w^2, w^3, \dots, w^7$, going around the circle. Powers of w are best in polar form, because we work only with the angles $\frac{2\pi}{8}, \frac{4\pi}{8}, \dots, \frac{16\pi}{8} = 2\pi$. Those 8 angles in degrees are $45^\circ, 90^\circ, 135^\circ, \dots, 360^\circ$.

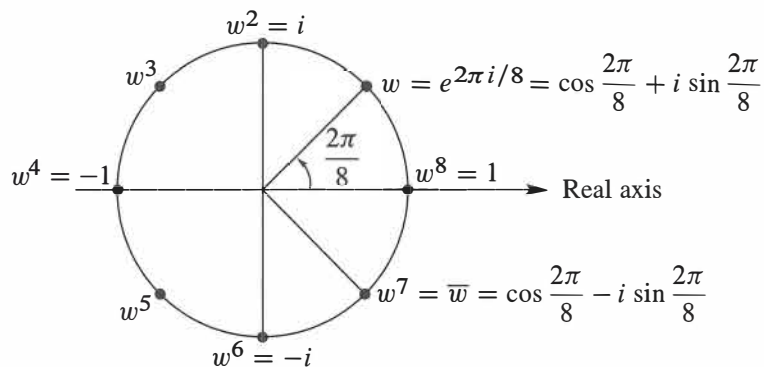


Figure 9.4: The eight solutions to $z^8 = 1$ are $1, w, w^2, \dots, w^7$ with $w = (1 + i)/\sqrt{2}$.

The fourth roots of 1 are also in the figure. They are $i, -1, -i, 1$. The angle is now $2\pi/4$ or 90° . The first root $w_4 = e^{2\pi i/4}$ is nothing but i . Even the square roots of 1 are seen, with $w_2 = e^{i2\pi/2} = -1$. Do not despise those square roots 1 and -1 . The idea behind the FFT is to go from an **8 by 8** Fourier matrix (containing powers of w_8) to the **4 by 4** matrix below (with powers of $w_4 = i$). The same idea goes from 4 to 2. By exploiting the connections of F_8 down to F_4 and up to F_{16} (and beyond), the FFT makes multiplication by F_{1024} very quick.

We describe the *Fourier matrix*, first for $n = 4$. Its rows contain powers of 1 and w and w^2 and w^3 . These are the fourth roots of 1, and their powers come in a special order.

$$\begin{array}{l} \text{Fourier} \\ \text{matrix} \\ n = 4 \\ w = i \end{array} \quad F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 \\ 1 & w^2 & w^4 & w^6 \\ 1 & w^3 & w^6 & w^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix}.$$

The matrix is symmetric ($F = F^T$). It is *not* Hermitian. Its main diagonal is not real. But $\frac{1}{2}F$ is a **unitary matrix**, which means that $(\frac{1}{2}F^H)(\frac{1}{2}F) = I$:

The columns of F give $F^H F = 4I$. Its inverse is $\frac{1}{4} F^H$ which is $F^{-1} = \frac{1}{4} \overline{F}$.

The inverse changes from $w = i$ to $\overline{w} = -i$. That takes us from F to $\overline{F}$. When the Fast Fourier Transform gives a quick way to multiply by F , it does the same for $\overline{F}$ and F^{-1} .

Every column has length $\sqrt{n}$. So the unitary matrices are $Q = F/\sqrt{n}$ and $Q^{-1} = \overline{F}/\sqrt{n}$. We avoid $\sqrt{n}$ and just use F and $F^{-1} = \overline{F}/n$. The main point is to multiply F times c_0, c_1, c_2, c_3 :

$$\begin{array}{l} \text{4-point} \\ \text{Fourier} \\ \text{series} \end{array} \quad \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = F \mathbf{c} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 \\ 1 & w^2 & w^4 & w^6 \\ 1 & w^3 & w^6 & w^9 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}. \quad (1)$$

The input is four complex coefficients c_0, c_1, c_2, c_3 . The output is four function values y_0, y_1, y_2, y_3 . The first output $y_0 = c_0 + c_1 + c_2 + c_3$ is the value of the Fourier series $\sum c_k e^{ikx}$ at $x = 0$. The second output is the value of that series $\sum c_k e^{ikx}$ at $x = 2\pi/4$:

$$y_1 = c_0 + c_1 e^{i2\pi/4} + c_2 e^{i4\pi/4} + c_3 e^{i6\pi/4} = c_0 + c_1 w + c_2 w^2 + c_3 w^3.$$

The third and fourth outputs y_2 and y_3 are the values of $\sum c_k e^{ikx}$ at $x = 4\pi/4$ and $x = 6\pi/4$. These are *finite* Fourier series! They contain $n = 4$ terms and they are evaluated at $n = 4$ points. Those points $x = 0, 2\pi/4, 4\pi/4, 6\pi/4$ are equally spaced.

The next point would be $x = 8\pi/4$ which is 2π . Then the series is back to y_0 , because $e^{2\pi i}$ is the same as $e^0 = 1$. Everything cycles around with period 4. In this world $2 + 2$ is 0 because $(w^2)(w^2) = w^0 = 1$. We follow the convention that **j and k go from 0 to $n - 1$** (instead of 1 to n). The “zeroth row” and “zeroth column” of F contain all ones.

The n by n Fourier matrix contains powers of $w = e^{2\pi i/n}$:

$$F_n \mathbf{c} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{n-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)^2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix} = \mathbf{y}. \quad (2)$$

F_n is symmetric but not Hermitian. **Its columns are orthogonal**, and $F_n \overline{F}_n = nI$. Then F_n^{-1} is $\overline{F}_n/n$. The inverse contains powers of $\overline{w}_n = e^{-2\pi i/n}$. Look at the pattern in F :

The entry in row j , column k is w^{jk} . Row zero and column zero contain $w^0 = 1$.

When we multiply $\mathbf{c}$ by F_n , we sum the series at n points. When we multiply $\mathbf{y}$ by F_n^{-1} , we find the coefficients $\mathbf{c}$ from the function values $\mathbf{y}$. In MATLAB that command is $\mathbf{c} = \text{fft}(\mathbf{y})$. The matrix F passes from “frequency space” to “physical space.”

Important note. Many authors prefer to work with $\omega = e^{-2\pi i/N}$, which is the *complex conjugate* of our w . (They often use the Greek omega, and I will do that to keep the two options separate.) With this choice, their DFT matrix contains powers of ω not w . It is $\overline{F}$, the conjugate of our F . $\overline{F}$ goes from physical space to frequency space.

$\overline{F}$ is a completely reasonable choice! MATLAB uses $\omega = e^{-2\pi i/N}$. The DFT matrix $\text{fft}(\text{eye}(N))$ contains powers of this number $\omega = \overline{w}$. **The Fourier matrix F with w 's reconstructs $\mathbf{y}$ from $\mathbf{c}$. The matrix $\overline{F}$ with ω 's computes Fourier coefficients as $\text{fft}(\mathbf{y})$.**

Also important. When a function $f(x)$ has period 2π , and we change x to $e^{i\theta}$, the function is defined around the unit circle (where $z = e^{i\theta}$). The Discrete Fourier Transform is the same as interpolation. Find the polynomial $p(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}$ that matches n values $f_0, \dots, f_{n-1}$:

Interpolation Find $c_0, \dots, c_{n-1}$ so that $p(z) = f$ at n points $z = 1, \dots, w^{n-1}$

The Fourier matrix is the Vandermonde matrix for interpolation at those n special points.

One Step of the Fast Fourier Transform

We want to multiply F times $\mathbf{c}$ as quickly as possible. Normally a matrix times a vector takes n^2 separate multiplications—the matrix has n^2 entries. You might think it is impossible to do better. (If the matrix has zero entries then multiplications can be skipped. But the Fourier matrix has no zeros!) By using the special pattern w^{jk} for its entries, F can be factored in a way that produces many zeros. This is the **FFT**.

The key idea is to connect F_n with the half-size Fourier matrix $F_{n/2}$. Assume that n is a power of 2 (say $n = 2^{10} = 1024$). We will connect F_{1024} to **two copies of F_{512}** .

When $n = 4$, the key is in the relation between F_4 and two copies of F_2 :

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} F_2 & \\ & F_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & & \\ 1 & i^2 & & \\ & & 1 & 1 \\ & & 1 & i^2 \end{bmatrix}.$$

On the left is F_4 , with no zeros. On the right is a matrix that is half zero. The work is cut in half. But wait, those matrices are not the same. We need two sparse and simple matrices to complete the FFT factorization:

$$\text{Factors for FFT} \quad F_4 = \begin{bmatrix} 1 & & 1 & \\ & 1 & & i \\ 1 & & -1 & \\ & 1 & & -i \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ 1 & i^2 & & \\ & & 1 & 1 \\ & & 1 & i^2 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}. \quad (3)$$

The last matrix is a permutation. It puts the even c 's (c_0 and c_2) ahead of the odd c 's (c_1 and c_3). The middle matrix performs half-size transforms F_2 and F_2 on the even c 's and odd c 's. The matrix at the left combines the two half-size outputs—in a way that produces the correct full-size output $\mathbf{y} = F_4 \mathbf{c}$.

The same idea applies when $n = 1024$ and $m = \frac{1}{2}n = 512$. The number w is $e^{2\pi i/1024}$. It is at the angle $\theta = 2\pi/1024$ on the unit circle. The Fourier matrix F_{1024} is full of powers of w . The first stage of the FFT is the great factorization discovered by Cooley and Tukey (and foreshadowed in 1805 by Gauss):

$$F_{1024} = \begin{bmatrix} I_{512} & D_{512} \\ I_{512} & -D_{512} \end{bmatrix} \begin{bmatrix} F_{512} & \\ & F_{512} \end{bmatrix} \begin{bmatrix} \text{even-odd} \\ \text{permutation} \end{bmatrix}. \quad (4)$$

I_{512} is the identity matrix. D_{512} is the diagonal matrix with entries $(1, w, \dots, w^{511})$. The two copies of F_{512} are what we expected. Don't forget that they use the 512th root of unity (which is nothing but w^2 !!) The permutation matrix separates the incoming vector $\mathbf{c}$ into its even and odd parts $\mathbf{c}' = (c_0, c_2, \dots, c_{1022})$ and $\mathbf{c}'' = (c_1, c_3, \dots, c_{1023})$.

Here are the algebra formulas which say the same thing as that factorization of F_{1024} :

(One step of the FFT) Set $m = \frac{1}{2}n$. The first m and last m components of $\mathbf{y} = F_n \mathbf{c}$ combine the half-size transforms $\mathbf{y}' = F_m \mathbf{c}'$ and $\mathbf{y}'' = F_m \mathbf{c}''$. Equation (4) shows this step from n to $m = n/2$ as $I\mathbf{y}' + D\mathbf{y}''$ and $I\mathbf{y}' - D\mathbf{y}''$:

$$\begin{aligned} y_j &= y'_j + (w_n)^j y''_j, & j &= 0, \dots, m-1 \\ y_{j+m} &= y'_j - (w_n)^j y''_j, & j &= 0, \dots, m-1. \end{aligned} \quad (5)$$

Split $\mathbf{c}$ into $\mathbf{c}'$ and $\mathbf{c}''$, transform them by F_m into $\mathbf{y}'$ and $\mathbf{y}''$, then (5) reconstructs $\mathbf{y}$.

Those formulas come from separating $c_0, \dots, c_{n-1}$ into even c_{2k} and odd c_{2k+1} : w is w_n .

$$\mathbf{y} = F\mathbf{c} \quad y_j = \sum_{k=0}^{n-1} w^{jk} c_k = \sum_{k=0}^{m-1} w^{2jk} c_{2k} + \sum_{k=0}^{m-1} w^{j(2k+1)} c_{2k+1} \quad \text{with } m = \frac{1}{2}n. \quad (6)$$

The even c 's go into $c' = (c_0, c_2, \dots)$ and the odd c 's go into $c'' = (c_1, c_3, \dots)$. Then come the transforms $F_m c'$ and $F_m c''$. **The key is $w_n^2 = w_m$.** This gives $w_n^{2jk} = w_m^{jk}$.

$$\text{Rewrite (6)} \quad y_j = \sum (w_m)^{jk} c'_k + (w_n)^j \sum (w_m)^{jk} c''_k = y'_j + (w_n)^j y''_j. \quad (7)$$

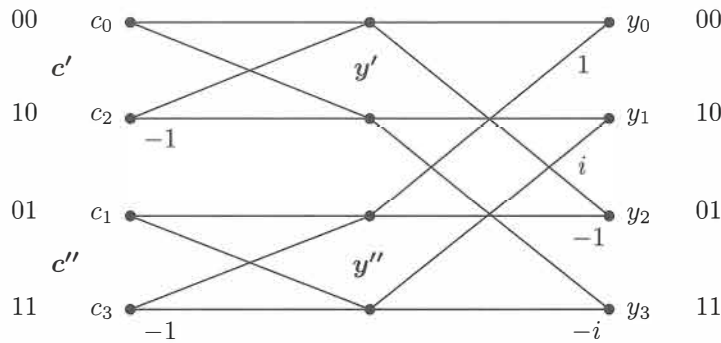
For $j \geq m$, the minus sign in (5) comes from factoring out $(w_n)^m = -1$ from $(w_n)^j$.

MATLAB easily separates even c 's from odd c 's and multiplies by w_n^j . We use `conj(F)` or equivalently MATLAB's inverse transform `ifft`, because `fft` is based on $\omega = \bar{w} = e^{-2\pi i/n}$. Problem 16 shows that F and `conj(F)` are linked by permuting rows.

FFT step from n to $n/2$ in MATLAB	$y' = \text{ifft}(c(0:2:n-2)) * n/2;$ $y'' = \text{ifft}(c(1:2:n-1)) * n/2;$ $d = w.^{(0:n/2-1)'};$ $y = [y' + d.*y''; y' - d.*y''];$
---	--

The flow graph shows c' and c'' going through the half-size F_2 . Those steps are called “*butterflies*,” from their shape. Then the outputs y' and y'' are combined (multiplying y'' by $1, i$ from D and also by $-1, -i$ from $-D$) to produce $y = F_4 c$.

This reduction from F_n to two F_m 's almost cuts the work in half—you see the zeros in the matrix factorization. That reduction is good but not great. The full idea of the **FFT** is much more powerful. It saves much more than half the time.



The Full FFT by Recursion

If you have read this far, you probably guessed what comes next. We reduced F_n to $F_{n/2}$. **Keep going to $F_{n/4}$.** Every F_{512} leads to F_{256} . Then 256 leads to 128. *That is recursion.*

Recursion is a basic principle of many fast algorithms. Here is step 2 with four copies of F_{256} and D (256 powers of ω_{512}). Evens of evens $c_0, c_4, c_8, \dots$ come first:

$$\begin{bmatrix} F_{512} & & \\ & F_{512} & \\ & & F_{512} \\ & & & F_{512} \end{bmatrix} = \begin{bmatrix} I & D & & \\ I & -D & & \\ & & I & D \\ & & I & -D \end{bmatrix} \begin{bmatrix} F & & & \\ & F & & \\ & & F & \\ & & & F \end{bmatrix} \begin{bmatrix} \text{pick } 0, 4, 8, \dots \\ \text{pick } 2, 6, 10, \dots \\ \text{pick } 1, 5, 9, \dots \\ \text{pick } 3, 7, 11, \dots \end{bmatrix}.$$

We will count the individual multiplications, to see how much is saved. Before the **FFT** was invented, the count was the usual $n^2 = (1024)^2$. This is about a million multiplications. I am not saying that they take a long time. The cost becomes large when we have many, many transforms to do—which is typical. Then the saving by the FFT is also large:

The final count for size $n = 2^\ell$ is reduced from n^2 to $\frac{1}{2}n\ell$.

The number 1024 is 2^{10} , so $\ell = 10$. The original count of $(1024)^2$ is reduced to $(5)(1024)$. The saving is a factor of 200. A million is reduced to five thousand. That is why the FFT has revolutionized signal processing.

Here is the reasoning behind $\frac{1}{2}n\ell$. There are ℓ levels, going from $n = 2^\ell$ down to $n = 1$. Each level has $n/2$ multiplications from the diagonal D 's, to reassemble the half-size outputs from the lower level. This yields the final count $\frac{1}{2}n\ell$, which is $\frac{1}{2}n \log_2 n$.

One last note about this remarkable algorithm. There is an amazing rule for the order that the c 's enter the FFT, after all the even-odd permutations. Write the numbers 0 to $n - 1$ in binary (like 00, 01, 10, 11 for $n = 4$). Reverse the order of those digits: 00, 10, 01, 11. That gives the **bit-reversed order 0, 2, 1, 3** with evens before odds (See Problem 17.) The complete picture shows the c 's in bit-reversed order, the $\ell = \log_2 n$ steps of the recursion, and the final output $y_0, \dots, y_{n-1}$ which is F_n times c .

The chapter ends with that very fundamental idea, a matrix multiplying a vector.

Problem Set 9.3

- 1 Multiply the three matrices in equation (3) and compare with F . In which six entries do you need to know that $i^2 = -1$?
- 2 Invert the three factors in equation (3) to find a fast factorization of F^{-1} .
- 3 F is symmetric. So transpose equation (3) to find a new Fast Fourier Transform!
- 4 All entries in the factorization of F_6 involve powers of $w_6 =$ sixth root of 1:

$$F_6 = \begin{bmatrix} I & D \\ I & -D \end{bmatrix} \begin{bmatrix} F_3 & \\ & F_3 \end{bmatrix} \begin{bmatrix} P \\ \end{bmatrix}.$$

Write down these matrices with 1, w_6 , w_6^2 in D and $w_3 = w_6^2$ in F_3 . Multiply!

- 5 If $v = (1, 0, 0, 0)$ and $w = (1, 1, 1, 1)$, show that $Fv = w$ and $Fw = 4v$. Therefore $F^{-1}w = v$ and $F^{-1}v = \underline{\hspace{1cm}}$.
- 6 What is F^2 and what is F^4 for the 4 by 4 Fourier matrix?
- 7 Put the vector $c = (1, 0, 1, 0)$ through the three steps of the FFT to find $y = Fc$. Do the same for $c = (0, 1, 0, 1)$.
- 8 Compute $y = F_8c$ by the three FFT steps for $c = (1, 0, 1, 0, 1, 0, 1, 0)$. Repeat the computation for $c = (0, 1, 0, 1, 0, 1, 0, 1)$.

- 9 If $w = e^{2\pi i/64}$ then w^2 and $\sqrt{w}$ are among the _____ and _____ roots of 1.
- 10 (a) Draw all the sixth roots of 1 on the unit circle. Prove they add to zero.
(b) What are the three cube roots of 1? Do they also add to zero?
- 11 The columns of the Fourier matrix F are the *eigenvectors* of the cyclic permutation P (see Section 8.3). Multiply PF to find the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{bmatrix}.$$

This is $PF = F\Lambda$ or $P = F\Lambda F^{-1}$. The eigenvector matrix (usually X) is F .

- 12 The equation $\det(P - \lambda I) = 0$ is $\lambda^4 = 1$. This shows again that the eigenvalues are $\lambda = ______$. Which permutation P has eigenvalues = cube roots of 1?
- 13 (a) Two eigenvectors of C are $(1, 1, 1, 1)$ and $(1, i, i^2, i^3)$. Find the eigenvalues e .

$$\begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = e_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad C \begin{bmatrix} 1 \\ i \\ i^2 \\ i^3 \end{bmatrix} = e_2 \begin{bmatrix} 1 \\ i \\ i^2 \\ i^3 \end{bmatrix}.$$

- (b) $P = F\Lambda F^{-1}$ immediately gives $P^2 = F\Lambda^2 F^{-1}$ and $P^3 = F\Lambda^3 F^{-1}$. Then $C = c_0 I + c_1 P + c_2 P^2 + c_3 P^3 = F(c_0 I + c_1 \Lambda + c_2 \Lambda^2 + c_3 \Lambda^3)F^{-1} = \mathbf{F} \mathbf{E} \mathbf{F}^{-1}$. That matrix E in parentheses is diagonal. It contains the _____ of C .
- 14 Find the eigenvalues of the “periodic” $-1, 2, -1$ matrix from $E = 2I - \Lambda - \Lambda^3$, with the eigenvalues of P in Λ . The -1 ’s in the corners make this matrix periodic:

$$C = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix} \quad \text{has } c_0 = 2, c_1 = -1, c_2 = 0, c_3 = -1.$$

- 15 **Fast convolution = Fast multiplication by C :** To multiply C times a vector x , we can multiply $F(E(F^{-1}x))$ instead. The direct way uses n^2 separate multiplications. Knowing E and F , the second way uses only $n \log_2 n + n$ multiplications. How many of those come from E , how many from F , and how many from F^{-1} ?
- 16 **Notice.** Why is row i of $\overline{F}$ the same as row $N - i$ of F (numbered 0 to $N - 1$)?
- 17 What is the *bit-reversed order* of the numbers $0, 1, \dots, 7$? Write them all in binary (base 2) as 000, 001, $\dots$, 111 and reverse each order. The 8 numbers are now _____.