CHAPTER

5

Complex Numbers

The shortest path between two truths in the real domain passes through the complex domain.

IN THIS CHAPTER

5.1	Machin's Formula	STEP III 1987
5.2	The Roots-of-Unity Filter	STEP I 1990
5.3	Factorising $z^{2n}+1$	STEP III 2013
5.4	Distances in a Regular Polygon	STEP III 2016
5.5	Rational Angles That Aren't	STEP III 2022
5.6	The Fermat Point	STEP III 2000
5.7	The Complex Exponential as a Map	STEP II 1997
5.8	A Hidden Distance in the Disc	STEP III 2013

Imaginary Numbers, Real Answers

The shortest path between two truths in the real domain passes through the complex domain.

— Jacques Hadamard

In 1545 Gerolamo Cardano published a formula for solving cubic equations, and found something disturbing. For some cubics with three perfectly good real roots, the formula required taking square roots of negative numbers along the way. The square roots cancelled out at the end, leaving real answers, but in the middle of the calculation there were quantities that, by every standard of the time, did not exist. Cardano called the manipulation "as subtle as it is useless". For two centuries, mathematicians used these "imaginary" numbers uneasily, as a trick that happened to work.

It took until the early nineteenth century, when Caspar Wessel, Jean-Robert Argand and Gauss pictured complex numbers as points in a plane, for them to become respectable. Once they did, they turned out to be not a trick but a revelation. Complex numbers explain why real phenomena behave as they do, and they keep providing the shortest route to results about real numbers, exactly as Hadamard said.

Why complex numbers matter

Every polynomial equation has a solution in the complex numbers: the *fundamental theorem of algebra*. So the complex numbers are, in a precise sense, where algebra is complete. They are also the natural language of rotation and oscillation. Multiplying by a complex number of modulus 1 rotates the plane, so anything involving angles, waves or periodic behaviour is simpler with complex numbers. Electrical engineers use them to analyse alternating currents; physicists cannot write down quantum mechanics without them; signal processing, from mobile phones to MRI scanners, runs on the complex Fourier transform. And in pure mathematics, functions of a complex variable turn out to be astonishingly rigid and beautiful: knowing a complex function on a tiny region determines it everywhere.

What you will meet at university

You will use complex numbers from the first week, but the big step comes with *complex analysis*, usually in the second year. Its central objects are *holomorphic* functions, complex functions that can be differentiated, and the theory is full of surprises. Every holomorphic function is automatically infinitely differentiable and equal to its Taylor series. Integrals round closed curves can be evaluated by looking only at the points where the function blows up (Cauchy's residue theorem), which lets you compute real integrals that resist every other method. The theory leads on to conformal maps

(Section 5.7), Riemann surfaces, and the Riemann zeta function that controls the primes (Section 1.2). In parallel, the algebra of roots of unity (Section 5.2, Section 5.3) leads to Galois theory, and complex geometry leads to the non-Euclidean geometry of Section 5.8.

The questions that run through this chapter

Three big questions

1. **How do complex numbers help with real problems?** Why should a detour through the complex plane produce facts about π, binomial coefficients or cosines?
2. **What is special about the roots of unity?** Why do the solutions of $z^n = 1$ keep appearing, in sums, products, polygons and angles?
3. **What does a complex function look like?** How can we picture functions whose inputs and outputs are both points of a plane, and what geometry do they reveal?

The chapter moves from arguments and multiplication (Machin's formula), through the algebra of roots of unity (filters, factorisations, polygon distances and irrational angles), to geometry with complex numbers (the Fermat point) and finally to complex functions as maps (the exponential, and a hidden distance in the unit disc).

Ideas in this chapter

§	Topic	Idea
5.1	Machin's formula	Arguments add when complex numbers multiply
5.2	Every third binomial coefficient	Averaging over roots of unity filters by residue
5.3	Products of cosines	Factorise over the roots, then substitute
5.4	Chords of a regular polygon	Distances are moduli; $z^n - 1 = \prod(z - \omega^k)$
5.5	$\cos^{-1}\frac{1}{3}$ is irrational	De Moivre, parity and a minimal counterexample
5.6	The Fermat point	Rotation is multiplication; good coordinates
5.7	e^w as a map	Picturing complex functions; many-to-one maps
5.8	A hidden distance	An inequality as a triangle inequality in hyperbolic geometry

How to read this chapter

Draw pictures. Almost every result in this chapter has a geometric meaning, and a quick sketch of the complex plane (where the roots of unity are, which way a multiplication rotates, where the grid lines go) often shows you what to do before any algebra. Then do the algebra carefully, because the pictures can mislead about signs and multiples of 2π.

5.1 Machin's Formula

The question

(i) If $z = x + \mathrm{i}y$, with x, y real, show that

$$|x|\cos\alpha + |y|\sin\alpha \leq |z| \leq |x| + |y|$$

for all real α.

(ii) By considering $(5-\mathrm{i})^4(1+\mathrm{i})$, show that

$$\frac{\pi}{4} = 4\tan^{-1}\left(\frac{1}{5}\right) - \tan^{-1}\left(\frac{1}{239}\right).$$

Prove similarly that

$$\frac{\pi}{4} = 3\tan^{-1}\left(\frac{1}{4}\right) + \tan^{-1}\left(\frac{1}{20}\right) + \tan^{-1}\left(\frac{1}{1985}\right).$$

The hook

In 1706 the London mathematician John Machin computed π to 100 decimal places, by hand. Nobody had come close before. His secret was the formula in part (ii), which expresses $\frac{\pi}{4}$ in terms of angles whose tangents are tiny fractions. Tiny fractions make the series $\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots$ converge very fast, and Machin's formula, or variants of it, remained the standard way to compute π for two and a half centuries, until computers and the methods of Section 2.6 took over.

The formula looks like a miracle of trigonometric identities. The question shows it is really a fact about multiplying complex numbers: when you multiply complex numbers, their arguments add. One multiplication of whole-number complex numbers proves the whole thing.

Before you start

Ask yourself What is the argument of $5 - \mathrm{i}$? Of $1 + \mathrm{i}$? What happens to arguments when complex numbers are multiplied?

$\arg(5-\mathrm{i}) = -\tan^{-1}\frac{1}{5}$, a small negative angle, and $\arg(1+\mathrm{i}) = \frac{\pi}{4}$. When you multiply

complex numbers, the moduli multiply and the arguments add (up to multiples of 2π). So

$$\arg\big((5-\mathrm{i})^4(1+\mathrm{i})\big) = \frac{\pi}{4} - 4\tan^{-1}\frac{1}{5} \quad \text{(possibly plus a multiple of } 2\pi\text{)}.$$

If we can compute the product and find its argument another way, we get an identity.

Working it through

STAGE 1 Bounds on the modulus

Ask yourself For the upper bound, square both sides. For the lower bound, think of $|x|\cos\alpha + |y|\sin\alpha$ as a dot product of two vectors.

Upper bound: $(|x|+|y|)^2 = x^2+y^2+2|x||y| \geq x^2+y^2 = |z|^2$, and both sides are non-negative, so $|z| \leq |x|+|y|$.

Lower bound: $|x|\cos\alpha + |y|\sin\alpha$ is the dot product of the vector $(|x|,|y|)$, of length $|z|$, with the unit vector $(\cos\alpha, \sin\alpha)$. A dot product is at most the product of the lengths, so it is at most $|z|$. (Algebraically: $|z|^2 - (|x|\cos\alpha + |y|\sin\alpha)^2 = (|x|\sin\alpha - |y|\cos\alpha)^2 \geq 0$.) ■

STAGE 2 The key multiplication

Ask yourself Compute $(5-\mathrm{i})^2$, then square it, then multiply by $1+\mathrm{i}$.

$(5-\mathrm{i})^2 = 25 - 10\mathrm{i} - 1 = 24 - 10\mathrm{i}$. Then $(24-10\mathrm{i})^2 = 576 - 480\mathrm{i} - 100 = 476 - 480\mathrm{i}$. Finally

$$(476-480\mathrm{i})(1+\mathrm{i}) = 476 + 476\mathrm{i} - 480\mathrm{i} + 480 = 956 - 4\mathrm{i}.$$

Pause and try Now compare arguments. What is the argument of $956-4\mathrm{i}$? Why are you sure there is no extra multiple of 2π?

$956 - 4\mathrm{i} = 4(239-\mathrm{i})$, which has argument $-\tan^{-1}\frac{1}{239}$. So

$$\frac{\pi}{4} - 4\tan^{-1}\frac{1}{5} = -\tan^{-1}\frac{1}{239} + 2k\pi$$

for some integer k. But $0 < \tan^{-1}\frac{1}{5} < \frac{1}{5}$, so the left-hand side lies between $\frac{\pi}{4} - \frac{4}{5}$ and $\frac{\pi}{4}$, that is, between about -0.02 and 0.79; and the right-hand side without the $2k\pi$ lies between -0.005 and 0. Since 2π is much bigger than the gap, $k = 0$. Rearranging,

$$\frac{\pi}{4} = 4\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{239}. \qquad \blacksquare$$

Figure 5.1.1 shows what happened: starting at $45°$, each multiplication by $5-\mathrm{i}$ rotates clockwise by about $11.3°$, and after four of them we have overshot the real axis by just $0.24°$.

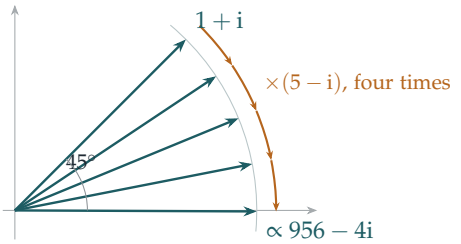

Figure 5.1.1 Multiplying $1+\mathrm{i}$ by $5-\mathrm{i}$ four times. Each multiplication rotates clockwise by $\tan^{-1}\frac{1}{5} \approx 11.31°$ (and stretches, which is ignored here). Four rotations take $45°$ to $-0.24°$, the argument of $956-4\mathrm{i}$, which is $-\tan^{-1}\frac{1}{239}$.

STAGE 3 The second formula

Ask yourself By analogy, which product should you compute this time?

Consider $(4+\mathrm{i})^3(20+\mathrm{i})(1985+\mathrm{i})$, whose argument is $3\tan^{-1}\frac{1}{4}+\tan^{-1}\frac{1}{20}+\tan^{-1}\frac{1}{1985}$. Now $(4+\mathrm{i})^2 = 15+8\mathrm{i}$, so $(4+\mathrm{i})^3 = (15+8\mathrm{i})(4+\mathrm{i}) = 52+47\mathrm{i}$. Then $(52+47\mathrm{i})(20+\mathrm{i}) = 1040+52\mathrm{i}+940\mathrm{i}-47 = 993+992\mathrm{i}$, and

$$(993+992\mathrm{i})(1985+\mathrm{i}) = (1971105-992)+(993+1969120)\mathrm{i} = 1970113+1970113\mathrm{i}.$$

Real part equals imaginary part, so the argument is $\frac{\pi}{4}$. All three angles are small and positive (their sum is less than $3\cdot\frac{1}{4}+\frac{1}{20}+\frac{1}{1985} < 1 < 2\pi$), so there is no multiple of 2π to worry about, and the formula follows. ■

Looking back

- **Arguments add under multiplication.** That single fact turned a product of Gaussian integers into an identity between angles.
- **Mind the multiples of 2π.** Arguments are only defined up to 2π, so each identity needed a size estimate to pin down the multiple.
- **Whole numbers do the work.** All the arithmetic was with integers; the identity for π came out exactly, with no approximation.

What's really going on

Computing π by hand

Why is Machin's formula so good for computation? The series $\tan^{-1}x = x-\frac{x^3}{3}+\frac{x^5}{5}-\cdots$ converges for $|x|\le 1$, but painfully slowly at $x=1$: you need about a thousand terms of $1-\frac{1}{3}+\frac{1}{5}-\cdots$ for three decimal places of $\frac{\pi}{4}$. With $x=\frac{1}{5}$, each term is about 25 times smaller than the last, giving roughly 1.4 decimal digits per term; with $x=\frac{1}{239}$, about 4.8 digits per term. Machin's 100 digits needed only about 70 terms of

the first series and 20 of the second. In 1873 William Shanks used a Machin formula to compute 707 digits, spending years on the arithmetic; in 1944 it was discovered that he had made an error at the 528th place, and everything after it was wrong.

Gaussian integers again

The numbers $a + b\mathrm{i}$ with a, b whole numbers are the *Gaussian integers* of Section 1.6. The reason $5 - \mathrm{i}$ works so well is a factorisation: $239^2 + 1 = 57122 = 2 \times 13^4$, which in the Gaussian integers reads $(239 + \mathrm{i})(239 - \mathrm{i}) = (1 + \mathrm{i})(1 - \mathrm{i})\,(3 + 2\mathrm{i})^4(3 - 2\mathrm{i})^4$, and $5 - \mathrm{i} = (1 - \mathrm{i})(3 + 2\mathrm{i})$ is built from the same Gaussian primes. Finding Machin-type formulas is really a problem about factorising numbers of the form $n^2 + 1$. In 1899 Carl Størmer used exactly this idea to prove that there are only four formulas of the form $\frac{\pi}{4} = a\tan^{-1}\frac{1}{m} + b\tan^{-1}\frac{1}{n}$ with whole numbers a, b, m, n: Machin's, and three found by Euler, Hermann and Hutton.

Rotations

Multiplying by a complex number of modulus 1 is a rotation of the plane. That is why complex numbers are the natural language for anything involving rotation and angles, from the geometry later in this chapter to alternating currents in electrical engineering and the wave functions of quantum mechanics. In Section 4.2 the addition formula for $\tan^{-1}$ came from a clever substitution; here it comes from the multiplication of complex numbers, and the two are the same fact seen from different sides.

The thinking skill

Translate angles into complex multiplication

An identity between sums of angles is an identity between arguments of a product. Compute the product exactly and read off its argument.

Arithmetic with complex numbers is mechanical and exact, while manipulating inverse trigonometric functions is error-prone. Translating from one to the other is often the quickest route to a proof.

Try this next

1. **A cousin.** (From STEP II, 2000.) By considering $(5 - \mathrm{i})^2(1 + \mathrm{i})$, prove that $\tan^{-1}\frac{7}{17} + 2\tan^{-1}\frac{1}{5} = \frac{\pi}{4}$.
2. **Euler's formula.** Show that $\tan^{-1}\frac{1}{2} + \tan^{-1}\frac{1}{3} = \frac{\pi}{4}$ by considering $(2 + \mathrm{i})(3 + \mathrm{i})$.
3. **Hutton's formula.** Show that $\frac{\pi}{4} = 2\tan^{-1}\frac{1}{3} + \tan^{-1}\frac{1}{7}$.

4 **How many terms?** Estimate how many terms of each series in Machin's formula you would need to compute π to 20 decimal places. Then do it for the first few digits with a calculator.

Further reading

- **Jonathan Borwein & Peter Borwein, *Pi and the AGM*** (Wiley). Includes the history of Machin-type formulas.
- **Jörg Arndt & Christoph Haenel, *Pi Unleashed*** (Springer). A lively account of methods for computing π, old and new.
- **Tristan Needham, *Visual Complex Analysis*** (Oxford University Press). The best book for seeing complex numbers geometrically. Chapter 1 is about multiplication as rotation and scaling.

5.2 The Roots-of-Unity Filter

The question

Let $\omega = e^{2\pi i/3}$. Show that $1 + \omega + \omega^2 = 0$ and calculate the modulus and argument of $1 + \omega^2$.

Let n be a positive integer. By evaluating $(1 + \omega^r)^n$ in two ways, taking $r = 1, 2$ and 3, or otherwise, prove that

$$\binom{n}{0} + \binom{n}{3} + \binom{n}{6} + \cdots + \binom{n}{k} = \frac{1}{3}\left(2^n + 2\cos\left(\frac{n\pi}{3}\right)\right),$$

where k is the largest multiple of 3 less than or equal to n. Without using a calculator, evaluate

$$\binom{25}{0} + \binom{25}{3} + \cdots + \binom{25}{24}$$

and

$$\binom{24}{2} + \binom{24}{5} + \cdots + \binom{24}{23}.$$

[$2^{25} = 33554432$.]

The hook

The binomial coefficients in a row of Pascal's triangle add up to 2^n: put $x = 1$ in $(1 + x)^n$. Put $x = -1$ instead and you get the alternating sum, which is 0; adding the two shows that the coefficients in even positions add up to 2^{n-1}, exactly half. So the even-numbered and odd-numbered entries split the total evenly.

What about every *third* coefficient? There is no real number x that can pick out positions $0, 3, 6, \ldots$. But there is a complex one. The cube roots of unity act as a filter: averaged over 1, ω and ω^2, the powers ω^k give 1 when k is a multiple of 3 and 0 otherwise. This idea, the *roots-of-unity filter*, is a small version of the *discrete Fourier transform*, one of the most important algorithms ever devised.

Before you start

Ask yourself ω is a root of $z^3 - 1 = 0$. Factorise $z^3 - 1$. Why does this give $1 + \omega + \omega^2 = 0$?

$z^3 - 1 = (z-1)(z^2 + z + 1)$. Since $\omega^3 = 1$ but $\omega \neq 1$, ω must be a root of the second factor: $\omega^2 + \omega + 1 = 0$. Geometrically, 1, ω and ω^2 are the vertices of an equilateral triangle centred at the origin, and their sum is the centre of mass of the triangle, which is 0.

Working it through

STAGE 1 Moduli and arguments

From $1 + \omega + \omega^2 = 0$, $1 + \omega^2 = -\omega$. Since $\omega = e^{2\pi \mathrm{i}/3}$, $-\omega = e^{\mathrm{i}\pi} e^{2\pi\mathrm{i}/3} = e^{5\pi\mathrm{i}/3} = e^{-\pi\mathrm{i}/3}$. So $1 + \omega^2$ has modulus 1 and argument $-\frac{\pi}{3}$. Similarly $1 + \omega = -\omega^2 = e^{\pi\mathrm{i}/3}$, with modulus 1 and argument $\frac{\pi}{3}$.

STAGE 2 The filter

Ask yourself What is $1 + \omega^k + \omega^{2k}$ when k is a multiple of 3? When it is not?

If $3 \mid k$, then $\omega^k = 1$ and the sum is 3. If not, ω^k is ω or ω^2, and in either case $1 + \omega^k + \omega^{2k} = 1 + \omega + \omega^2 = 0$. So the expression $\frac{1}{3}(1 + \omega^k + \omega^{2k})$ is 1 when $3 \mid k$ and 0 otherwise: it is a detector for multiples of 3.

Pause and try Add up $(1 + \omega^r)^n$ for $r = 1, 2, 3$, expanding each by the binomial theorem. What do you get?

Since $\omega^3 = 1$,

$$(1+\omega)^n + (1+\omega^2)^n + (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} (\omega^k + \omega^{2k} + 1) = 3 \sum_{3 \mid k} \binom{n}{k}.$$

On the other hand, by Stage 1,

$$(1+\omega)^n + (1+\omega^2)^n + 2^n = e^{n\pi\mathrm{i}/3} + e^{-n\pi\mathrm{i}/3} + 2^n = 2\cos\frac{n\pi}{3} + 2^n.$$

Dividing by 3 gives the formula. ∎

STAGE 3 Two calculations

For $n = 25$: $\frac{25\pi}{3} = 8\pi + \frac{\pi}{3}$, so $\cos\frac{25\pi}{3} = \frac{1}{2}$, and

$$\binom{25}{0} + \binom{25}{3} + \cdots + \binom{25}{24} = \frac{33554432 + 1}{3} = 11184811.$$

Ask yourself For the second sum we need positions $k \equiv 2 \pmod 3$. How can you modify the filter to detect $k \equiv 2$ instead of $k \equiv 0$?

Shift the detector: $\frac{1}{3}\left(1+\omega^{k-2}+\omega^{2(k-2)}\right)$ is 1 exactly when $k \equiv 2 \pmod 3$. Since $\omega^{-2}=\omega$ and $\omega^{-4}=\omega^2$, this gives

$$\sum_{k\equiv 2}\binom{n}{k} = \frac{1}{3}\left(2^n+\omega(1+\omega)^n+\omega^2(1+\omega^2)^n\right).$$

With $n=24$, $(1+\omega)^{24}=e^{8\pi i}=1$ and likewise $(1+\omega^2)^{24}=1$, so the bracket is $2^{24}+\omega+\omega^2=2^{24}-1$. With $2^{24}=16777216$,

$$\binom{24}{2}+\binom{24}{5}+\cdots+\binom{24}{23} = \frac{16777215}{3} = 5592405. \qquad \blacksquare$$

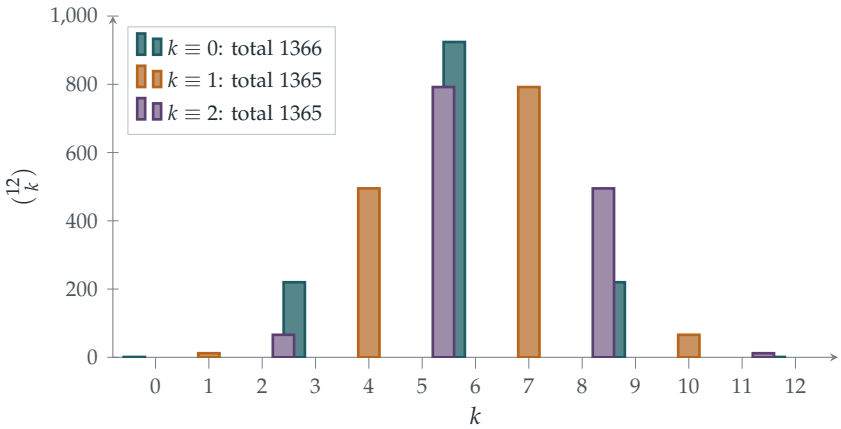

Figure 5.2.1 Row 12 of Pascal's triangle, coloured by the remainder of k on division by 3. The three classes have totals 1366, 1365 and 1365: almost exactly a third each of $2^{12}=4096$. The roots-of-unity filter predicts the small differences from $\frac{4096}{3}$ exactly: $+\frac{2}{3}$, $-\frac{1}{3}$ and $-\frac{1}{3}$.

Looking back

- **Averaging over roots of unity filters by residue.** $\frac{1}{3}\sum_r \omega^{rk}$ is 1 or 0 according as 3 divides k. Summing a polynomial's values at the roots of unity picks out every third coefficient.
- **Shifting the filter.** Multiplying by ω^{-rj} before averaging picks out the coefficients with $k \equiv j$.
- **Complex numbers for real answers.** The answer is a whole number, and the question was about whole numbers, but the route went through the complex plane.

What's really going on

The discrete Fourier transform

Replace 3 by any N and ω by $e^{2\pi i/N}$. For a list of numbers $a_0, a_1, \ldots, a_{N-1}$, the numbers

$$\hat{a}_r = \sum_{k=0}^{N-1} a_k \, \omega^{rk}$$

form its *discrete Fourier transform*, and the filter identity $\sum_r \omega^{rk} = N$ or 0 is what lets you recover the a_k from the $\hat{a}_r$. The transform splits a signal into its component frequencies. In 1965 James Cooley and John Tukey published a way to compute it in about $N \log_2 N$ steps instead of N^2, the *fast Fourier transform*; Gauss had found essentially the same idea in 1805 but never published it. The FFT is used in audio and image compression (MP3 and JPEG), in mobile phone signals, in medical scanners and in multiplying enormous numbers, and it is often listed among the most important algorithms of the twentieth century.

Generating functions

The filter is one instance of a bigger idea: a sequence $a_0, a_1, a_2, \ldots$ can be packaged as a polynomial or power series $A(x) = \sum a_k x^k$, its *generating function*, and operations on the sequence become operations on the function. Evaluating at 1 sums the sequence; evaluating at roots of unity picks out residue classes; differentiating weights the terms by k. This turns many counting problems into algebra, and it is a central technique in Chapter 9. The same filter, applied to $(1+x)^n$ with fifth roots of unity, connects to the regular pentagon and the golden ratio; we pursue the algebra of roots of unity in Section 6.1.

The thinking skill

Use symmetry to filter

To pick out the terms of a sum that satisfy a condition, find a function that is 1 *on the terms you want and* 0 *on the rest, built from a symmetric average.*

Averaging over the roots of unity is the classic example. The same principle, "average over a group of symmetries to isolate the part you want", appears throughout algebra, from the roots-of-unity filter to the representation theory of groups.

Try this next

1 **Even positions.** Use the filter with square roots of unity (1 and -1) to show that $\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = 2^{n-1}$.

2 **The other residue.** Find a formula for $\binom{n}{1} + \binom{n}{4} + \binom{n}{7} + \cdots$, and check that the three formulas add up to 2^n.

3 **Every fourth.** Using i, show that $\binom{n}{0} + \binom{n}{4} + \binom{n}{8} + \cdots = \frac{1}{4}\left(2^n + 2^{n/2+1}\cos\frac{n\pi}{4}\right)$.

4 **Divisibility.** How many subsets of $\{1, 2, \ldots, 12\}$ have a number of elements divisible by 3? How many have a sum divisible by 3? (*Hint for the second:* consider $\prod_{j=1}^{12}(1 + x^j)$ at the cube roots of unity.)

Further reading

- **Herbert Wilf, *generatingfunctionology*** (A. K. Peters; free online). A delightful introduction to generating functions; the roots-of-unity filter appears early on.
- **Arthur Benjamin & Jennifer Quinn, *Proofs That Really Count*** (Mathematical Association of America). Binomial identities proved by counting.
- **T. W. Körner, *Fourier Analysis*** (Cambridge University Press). Includes the discrete Fourier transform and the FFT.

5.3 Factorising $z^{2n}+1$

The question

Show that $(z-\mathrm{e}^{\mathrm{i}\theta})(z-\mathrm{e}^{-\mathrm{i}\theta}) = z^2 - 2z\cos\theta + 1$.

Write down the $(2n)$th roots of -1 in the form $\mathrm{e}^{\mathrm{i}\theta}$, where $-\pi < \theta \leq \pi$, and deduce that

$$z^{2n}+1 = \prod_{k=1}^{n}\left(z^2 - 2z\cos\left(\tfrac{(2k-1)\pi}{2n}\right)+1\right).$$

Here, n is a positive integer, and the $\prod$ notation denotes the product.

(i) By substituting $z = \mathrm{i}$ show that, when n is even,

$$\cos\left(\tfrac{\pi}{2n}\right)\cos\left(\tfrac{3\pi}{2n}\right)\cos\left(\tfrac{5\pi}{2n}\right)\cdots\cos\left(\tfrac{(2n-1)\pi}{2n}\right) = (-1)^{\frac{1}{2}n}2^{1-n}.$$

(ii) Show that, when n is odd,

$$\cos^2\left(\tfrac{\pi}{2n}\right)\cos^2\left(\tfrac{3\pi}{2n}\right)\cos^2\left(\tfrac{5\pi}{2n}\right)\cdots\cos^2\left(\tfrac{(n-2)\pi}{2n}\right) = n2^{1-n}.$$

You may use without proof the fact that $1+z^{2n} = (1+z^2)(1-z^2+z^4-\cdots+z^{2n-2})$ when n is odd.

The hook

What is $\cos\frac{\pi}{8}\cos\frac{3\pi}{8}\cos\frac{5\pi}{8}\cos\frac{7\pi}{8}$? Each factor is an ugly number involving nested square roots, but the product is exactly $\frac{1}{8}$. Products of cosines at equally spaced angles keep collapsing to simple numbers like this, and the question explains why. The trick is to recognise the angles as arguments of roots of a polynomial, factorise the polynomial, and then substitute a cleverly chosen value.

Over the complex numbers, every polynomial factorises completely into linear factors. Over the real numbers, the non-real roots come in conjugate pairs, which combine into real quadratic factors. Both facts are on display here, and the second is how $z^{2n}+1$, which has no real roots at all, can be written as a product of real quadratics.

Before you start

Ask yourself Where are the solutions of $z^{2n} = -1$ in the complex plane? Draw them for $n = 2$ and $n = 4$.

They all have modulus 1, and their arguments θ satisfy $2n\theta = \pi + 2m\pi$, so θ is an odd multiple of $\frac{\pi}{2n}$. They form a regular $2n$-gon on the unit circle, rotated so that no vertex lies on the real axis (Figure 5.3.1). They pair up into conjugates $e^{\pm i\theta}$, mirror images in the real axis.

Working it through

STAGE 1 Conjugate pairs give real quadratics

$(z - e^{i\theta})(z - e^{-i\theta}) = z^2 - (e^{i\theta} + e^{-i\theta})z + 1 = z^2 - 2z\cos\theta + 1.$ ■

STAGE 2 The factorisation

The $2n$th roots of -1 with $-\pi < \theta \leq \pi$ are $e^{\pm i\theta_k}$ with

$$\theta_k = \frac{(2k-1)\pi}{2n}, \qquad k = 1, 2, \ldots, n.$$

These are $2n$ distinct numbers, all roots of the degree-$2n$ polynomial $z^{2n} + 1$, so they are all its roots and

$$z^{2n} + 1 = \prod_{k=1}^{n} (z - e^{i\theta_k})(z - e^{-i\theta_k}) = \prod_{k=1}^{n} (z^2 - 2z\cos\theta_k + 1).$$ ■

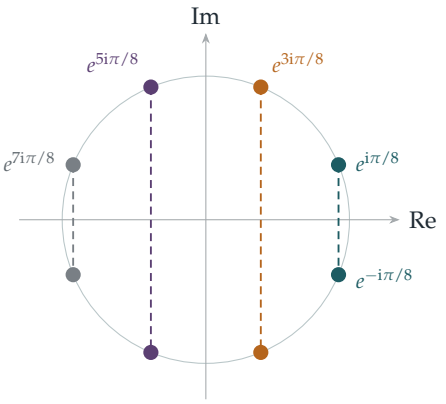

Figure 5.3.1 The eight solutions of $z^8 = -1$ (the case $n = 4$). They pair into four conjugate pairs, joined by dashed lines; each pair gives one real quadratic factor $z^2 - 2z\cos\theta_k + 1$ of $z^8 + 1$.

STAGE 3 Substituting $z = \mathrm{i}$

Ask yourself What is $\mathrm{i}^2 - 2\mathrm{i}\cos\theta + 1$? And what is $\mathrm{i}^{2n}+1$ when n is even?

Each factor becomes

$$\mathrm{i}^2 - 2\mathrm{i}\cos\theta_k + 1 = -2\mathrm{i}\cos\theta_k,$$

so the right-hand side is $(-2\mathrm{i})^n \prod_k \cos\theta_k$. The left-hand side is $\mathrm{i}^{2n}+1 = (-1)^n + 1$, which is 2 when n is even. And for even n,

$$(-2\mathrm{i})^n = 2^n(-\mathrm{i})^n = 2^n\big((-\mathrm{i})^2\big)^{n/2} = 2^n(-1)^{n/2}.$$

So

$$\prod_{k=1}^{n}\cos\theta_k = \frac{2}{2^n(-1)^{n/2}} = (-1)^{\frac{1}{2}n}2^{1-n}. \qquad \blacksquare$$

For $n=4$ this is the product in the hook: $(-1)^2 2^{-3} = \frac{1}{8}$.

STAGE 4 When n is odd

Ask yourself For odd n, one of the angles θ_k equals $\frac{\pi}{2}$. Which one, and why does that make the substitution $z = \mathrm{i}$ useless as it stands?

When $k = \frac{n+1}{2}$, $\theta_k = \frac{n\pi}{2n} = \frac{\pi}{2}$, so that factor is z^2+1, which vanishes at $z=\mathrm{i}$; both sides are 0, and we learn nothing. The way round is to divide out that factor first, which is what the given identity does.

Pause and try Divide both sides of the factorisation by $1+z^2$, substitute $z=\mathrm{i}$, and simplify. You will need to pair up the remaining cosines.

Dividing by the factor z^2+1 (the term $k=\frac{n+1}{2}$),

$$1 - z^2 + z^4 - \cdots + z^{2n-2} = \prod_{k\neq\frac{n+1}{2}}\big(z^2 - 2z\cos\theta_k + 1\big).$$

At $z=\mathrm{i}$, every term on the left is $(-z^2)^j = 1^j = 1$, and there are n terms, so the left-hand side is n. The right-hand side is $(-2\mathrm{i})^{n-1}\prod_{k\neq\frac{n+1}{2}}\cos\theta_k$. Since $n-1$ is even, $(-2\mathrm{i})^{n-1} = 2^{n-1}(-1)^{(n-1)/2}$.

Now pair the angles: $\theta_{n+1-k} = \pi - \theta_k$, so $\cos\theta_{n+1-k} = -\cos\theta_k$. The $n-1$ remaining factors form $\frac{n-1}{2}$ such pairs, with $k = 1, \ldots, \frac{n-1}{2}$, and each pair contributes $-\cos^2\theta_k$. So

$$n = 2^{n-1}(-1)^{(n-1)/2}\cdot(-1)^{(n-1)/2}\prod_{k=1}^{(n-1)/2}\cos^2\theta_k = 2^{n-1}\prod_{k=1}^{(n-1)/2}\cos^2\theta_k,$$

giving $\prod\cos^2\theta_k = n2^{1-n}$. The last factor, with $k=\frac{n-1}{2}$, is $\cos^2\frac{(n-2)\pi}{2n}$, as in the question. $\blacksquare$

Check with $n=3$: $\cos^2\frac{\pi}{6} = \frac{3}{4} = 3\cdot 2^{-2}$.

Looking back

- **Know the roots, know the polynomial.** A polynomial is determined (up to a constant) by its roots, so listing the roots of $z^{2n}+1$ gave its factorisation.
- **Substitute a clever value.** Putting $z=\mathrm{i}$ turned every quadratic factor into a multiple of a cosine.
- **When a substitution gives $0=0$, divide first.** For odd n one factor vanished at $z=\mathrm{i}$, so we removed it before substituting.

What's really going on

The fundamental theorem of algebra

That $z^{2n}+1$ can be factorised completely into linear factors over $\mathbb{C}$ is a special case of the *fundamental theorem of algebra*: every non-constant polynomial with complex coefficients has a complex root, and so factorises completely into linear factors. Gauss gave the first widely accepted proof in his doctoral thesis of 1799, and returned to the theorem several times with new proofs. A consequence is that every real polynomial factorises into real linear and quadratic factors, because non-real roots come in conjugate pairs, exactly as in Stage 1. The theorem is "algebraic" in name but analytic in substance: every proof uses some property of the real numbers, such as continuity, and you will see one of the slickest proofs in a first course on complex analysis.

Products of sines and cosines

The method gives a whole family of identities. Using z^n-1 in place of $z^{2n}+1$ leads to

$$\prod_{k=1}^{n-1}\sin\frac{k\pi}{n}=\frac{n}{2^{n-1}},$$

which is proved, in its geometric form, in Section 5.4. Identities like these appear in the analysis of vibrating systems (the frequencies of n masses joined by springs involve $\sin\frac{k\pi}{2(n+1)}$) and in counting problems (the number of spanning trees of some graphs, and the number of domino tilings of a rectangle, are given by products of cosines). Kasteleyn's 1961 formula for domino tilings of an $m\times n$ rectangle is a double product of expressions in $\cos^2\frac{j\pi}{m+1}$ and $\cos^2\frac{k\pi}{n+1}$.

The thinking skill

From roots to products

To evaluate a product of trigonometric values at equally spaced angles, recognise them as coming from the roots of a simple polynomial, factorise, and substitute a convenient value.

The choice of value is the art: here $z = \mathrm{i}$ turned each quadratic factor into a cosine. Other choices (such as $z = 1$ or $z = -1$) give other identities.

Try this next

1 **A heptagonal product.** Use part (ii) with $n = 7$ to show that $\cos\frac{\pi}{14}\cos\frac{3\pi}{14}\cos\frac{5\pi}{14} = \frac{\sqrt{7}}{8}$. Then show that $\cos\frac{\pi}{7}\cos\frac{2\pi}{7}\cos\frac{3\pi}{7} = \frac{1}{8}$. (*Hint:* relate the angles using $\cos\theta = \sin(\frac{\pi}{2}-\theta)$, or use the identity $\sin 2x = 2\sin x\cos x$ repeatedly.)

2 **Another substitution.** Substitute $z = 1$ in the factorisation to show that

$$\prod_{k=1}^{n} \sin\frac{(2k-1)\pi}{4n} = 2^{\frac{1}{2}-n}.$$

3 **Factorise z^6+1.** Write z^6+1 as a product of three real quadratics, and check your answer by multiplying out.

4 **Sines.** By factorising $z^{2n}-1$ similarly and dividing by z^2-1, show that

$$\prod_{k=1}^{n-1} \sin\frac{k\pi}{2n} = \frac{\sqrt{n}}{2^{n-1}}.$$

Further reading

- **Tristan Needham, *Visual Complex Analysis*** (Oxford University Press). Chapter 2 on roots of unity and polynomials, with beautiful pictures.
- **Benjamin Fine & Gerhard Rosenberger, *The Fundamental Theorem of Algebra*** (Springer). Several proofs of the theorem, from analysis, algebra and topology, written for undergraduates.
- **Paul Nahin, *An Imaginary Tale*** (Princeton University Press). A history of $\sqrt{-1}$, including roots of unity and their uses.

5.4 Distances in a Regular Polygon

The question

Let $\omega = e^{2\pi i/n}$, where n is a positive integer. Show that, for any complex number z,

$$(z-1)(z-\omega)\cdots(z-\omega^{n-1}) = z^n - 1.$$

The points $X_0, X_1, \ldots, X_{n-1}$ lie on a circle with centre O and radius 1, and are the vertices of a regular polygon.

(i) The point P is equidistant from X_0 and X_1. Show that, if n is even,

$$|PX_0| \times |PX_1| \times \cdots \times |PX_{n-1}| = |OP|^n + 1,$$

where $|PX_k|$ denotes the distance from P to X_k.

Give the corresponding result when n is odd. (There are two cases to consider.)

(ii) Show that

$$|X_0X_1| \times |X_0X_2| \times \cdots \times |X_0X_{n-1}| = n.$$

The hook

Draw a regular heptagon in a circle of radius 1. From one vertex, draw the six chords to the other vertices and measure them. Multiply the six lengths together. The answer is exactly 7. For a regular 100-gon, the product of the 99 chords from one vertex is exactly 100.

Why should anything so clean be true? The chord lengths themselves are messy ($2\sin\frac{\pi}{7}$, $2\sin\frac{2\pi}{7}$, and so on). The question shows that the result is a geometric shadow of a simple algebraic identity, the factorisation of $z^n - 1$ over the roots of unity, and that distances in the plane are moduli of complex numbers.

Before you start

? **Ask yourself** Put the polygon's vertices at the nth roots of unity, $X_k = \omega^k$. If P is the point z, what is $|PX_k|$ in terms of z? So what is the product of all the distances?

$|PX_k| = |z - \omega^k|$, and the product of all of them is $|(z-1)(z-\omega)\cdots(z-\omega^{n-1})|$, which the first part of the question says is $|z^n - 1|$. A product of n distances has become a single modulus. Everything else is choosing z cleverly.

Working it through

STAGE 1 The factorisation

Each ω^k satisfies $(\omega^k)^n = (\omega^n)^k = 1$, so $1, \omega, \ldots, \omega^{n-1}$ are roots of $z^n - 1$. They are distinct (their arguments $\frac{2\pi k}{n}$ are different, between 0 and 2π), so they are all n roots of this degree-n polynomial with leading coefficient 1, and $z^n - 1 = \prod_{k=0}^{n-1}(z - \omega^k)$. ■

STAGE 2 A point on the perpendicular bisector

Rotate the picture so that $X_k = \omega^k$. The points equidistant from $X_0 = 1$ and $X_1 = \omega$ form the perpendicular bisector of X_0X_1, which is the line through O at angle $\frac{\pi}{n}$.

Ask yourself So P is $z = te^{i\pi/n}$ for some real t (positive or negative), with $|OP| = |t|$. What is z^n?

$z^n = t^n e^{i\pi} = -t^n$. So

$$|PX_0| \times \cdots \times |PX_{n-1}| = |z^n - 1| = |-t^n - 1| = |t^n + 1|.$$

If n is even, $t^n = |t|^n \geq 0$, so this is $|OP|^n + 1$. ■

Pause and try What happens when n is odd? Consider $t > 0$ and $t < 0$ separately, and say what each case means geometrically.

If n is odd, t^n has the same sign as t.

- If $t \geq 0$, so P is on the same side of O as the midpoint of X_0X_1, the product is $|OP|^n + 1$.
- If $t < 0$, so P is on the opposite side of O (towards the vertex opposite the side X_0X_1), then $t^n = -|OP|^n$ and the product is $\left|1 - |OP|^n\right|$.

For odd n, the line through O perpendicular to X_0X_1 passes through a vertex on the far side; in the second case, the product vanishes when P is that vertex ($|OP| = 1$), as it must, since one of the distances is then zero.

STAGE 3 Chords from a vertex

Ask yourself Now $P = X_0$, and one of the distances, $|X_0X_0|$, is zero, so $|z^n - 1| = 0$ is no use. What polynomial has the roots $\omega, \omega^2, \ldots, \omega^{n-1}$ but not 1?

Dividing $z^n - 1$ by $z - 1$,

$$(z-\omega)(z-\omega^2)\cdots(z-\omega^{n-1}) = \frac{z^n-1}{z-1} = 1 + z + z^2 + \cdots + z^{n-1}.$$

The right-hand side is a polynomial, so this holds for every z, including $z = 1$, where it equals n. Taking moduli at $z = 1$,

$$|X_0X_1| \times |X_0X_2| \times \cdots \times |X_0X_{n-1}| = |1-\omega|\,|1-\omega^2|\cdots|1-\omega^{n-1}| = n. \qquad \blacksquare$$

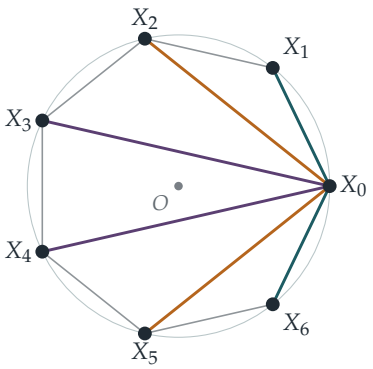

Figure 5.4.1 A regular heptagon inscribed in a circle of radius 1. The six chords from X_0 have lengths $2\sin\frac{k\pi}{7}$, about 0.868 (teal), 1.564 (amber) and 1.950 (plum), each occurring twice. Their product is exactly 7.

Looking back

- **Distances are moduli.** $|PX_k| = |z - \omega^k|$, so a product of distances is the modulus of a product, and the product had a closed form.
- **Use symmetry to choose coordinates.** Rotating so that the vertices are the roots of unity, and using the symmetry line through O, made z^n easy to compute.
- **Cancel the zero before substituting.** As in Section 5.3, dividing out the factor that vanished at the point of interest rescued the substitution.

What's really going on

Cotes's theorem

Part (i) is a theorem of Roger Cotes, the brilliant young Cambridge mathematician who edited the second edition of Newton's *Principia* and died in 1716 at the age of thirty-three (Newton said "if he had lived we might have known something"). Cotes found that if P lies on a ray from O through a vertex of a regular n-gon, then the product of the distances from P to the vertices is $\big||OP|^n - 1\big|$, and if P lies on a

ray through the midpoint of a side, it is $|OP|^n + 1$. He used this, without complex numbers, to factorise $x^n \pm 1$ into real quadratics, which he needed for integrating rational functions. The theorem was published after his death, in 1722.

A product of sines

Since $|1 - \omega^k| = |e^{-\mathrm{i}\pi k/n} - e^{\mathrm{i}\pi k/n}| = 2\sin\frac{k\pi}{n}$, part (ii) says

$$\prod_{k=1}^{n-1} \sin\frac{k\pi}{n} = \frac{n}{2^{n-1}}.$$

This is the companion of the cosine products in Section 5.3. A related product gives the *discriminant* of $z^n - 1$: the product of the squared distances between all pairs of distinct nth roots of unity is n^n. Discriminants measure how far apart the roots of a polynomial are, and they are central to the algebra of polynomials and of number fields, which we take up in Chapter 6.

Electrostatics and equal spacing

There is also a physical reading. Place n equal electric charges on a circle; the configuration that minimises the potential energy (with the logarithmic potential of two-dimensional electrostatics) is the regular polygon, and the energy involves exactly $\sum \ln|X_jX_k|$, the logarithm of a product of distances. Points that are "as spread out as possible" are important in numerical analysis (Section 3.6, where the best interpolation points come from a regular polygon via $x = \cos\theta$) and in the theory of random matrices.

The thinking skill

Turn geometry into algebra with complex numbers

Points in the plane are complex numbers, distances are moduli, and regular polygons are roots of unity. A product of distances becomes the modulus of a polynomial.

The translation lets algebraic identities, like the factorisation of $z^n - 1$, prove geometric facts that would be very hard to see directly.

Try this next

1. **Check small cases.** Verify part (ii) directly for a square ($n = 4$) and an equilateral triangle ($n = 3$).
2. **Through a vertex.** Show that if P lies on the ray from O through X_0, then $|PX_0| \times \cdots \times |PX_{n-1}| = \big||OP|^n - 1\big|$.

3 **The average distance.** Using part (ii) and the AM–GM inequality of Section 3.2, show that the average length of the $n-1$ chords from X_0 is at least $n^{1/(n-1)}$. Find the actual average, $\frac{1}{n-1}\sum_k 2\sin\frac{k\pi}{n}$, for $n=6$, and compare.

4 **All the chords.** HARDERShow that the product of all $\binom{n}{2}$ chord lengths of a regular n-gon inscribed in a unit circle is $n^{n/2}$.

Further reading

- **Tristan Needham, *Visual Complex Analysis*** (Oxford University Press). Geometry with complex numbers, beautifully illustrated.
- **Ross Honsberger, *Mathematical Gems*** (Mathematical Association of America). Several elegant problems about regular polygons.
- **Ronald Gowing, *Roger Cotes: Natural Philosopher*** (Cambridge University Press). A biography of Cotes and his mathematics.

5.5 Rational Angles That Aren't

The question

(i) Use De Moivre's theorem to prove that for any positive integer $k > 1$,

$$\sin(k\theta) = \sin\theta\cos^{k-1}\theta\left(k - \binom{k}{3}(\sec^2\theta - 1) + \binom{k}{5}(\sec^2\theta - 1)^2 - \cdots\right)$$

and find a similar expression for $\cos(k\theta)$.

(ii) Let $\theta = \cos^{-1}\left(\frac{1}{a}\right)$, where θ is measured in degrees, and a is an odd integer greater than 1.

Suppose that there is a positive integer k such that $\sin(k\theta) = 0$ and $\sin(m\theta) \neq 0$ for all integers m with $0 < m < k$.

Show that it would be necessary to have k even and $\cos(\frac{1}{2}k\theta) = 0$.

Deduce that θ is irrational.

(iii) Show that if $\phi = \cot^{-1}\left(\frac{1}{b}\right)$, where ϕ is measured in degrees, and b is an even integer greater than 1, then ϕ is irrational.

The hook

The angles you meet in school geometry are usually whole numbers of degrees: $30°$, $45°$, $60°$. Their cosines are numbers like $\frac{\sqrt{3}}{2}$ and $\frac{1}{2}$. What about going the other way? If the cosine is a simple fraction, such as $\frac{1}{3}$, is the angle a nice number of degrees?

The angle $\cos^{-1}\frac{1}{3} \approx 70.528779\ldots°$ appears all over science. It is the angle between two faces of a regular tetrahedron, and its supplement, $109.47°$, is the angle between the bonds in a methane molecule. This question proves that it is not a rational number of degrees: not $70\frac{1}{2}$, not $\frac{1411}{20}$, not any fraction at all. Its decimal expansion never repeats. This seemingly obscure fact solved one of the famous problems that David Hilbert posed in 1900.

Before you start

Ask yourself Suppose θ were rational, say $\theta = \frac{p}{q}$ degrees. Find a positive integer k with $\sin(k\theta) = 0$.

Take $k = 180q$: then $k\theta = 180p$ degrees, a whole number of half-turns, so $\sin(k\theta) = 0$. So rational angles always have some multiple that lands on a multiple of 180°. To prove θ is irrational, it is enough to show that *no* multiple $k\theta$ has zero sine. Part (ii) does this by taking the smallest such k, if there is one, and deriving a contradiction: the well-ordering principle of Section 1.3 again.

Working it through

STAGE 1 Multiple-angle formulas

Ask yourself Write $\cos k\theta + \mathrm{i}\sin k\theta = (\cos\theta + \mathrm{i}\sin\theta)^k$, and take out a factor of $\cos^k\theta$. What is left inside the bracket?

$(\cos\theta + \mathrm{i}\sin\theta)^k = \cos^k\theta\,(1 + \mathrm{i}\tan\theta)^k = \cos^k\theta \sum_{j=0}^{k} \binom{k}{j} \mathrm{i}^j \tan^j\theta$. The imaginary part comes from odd $j = 2m+1$, where $\mathrm{i}^j = \mathrm{i}(-1)^m$:

$$\begin{aligned}\sin k\theta &= \cos^k\theta \sum_m (-1)^m \binom{k}{2m+1} \tan^{2m+1}\theta \\ &= \sin\theta\cos^{k-1}\theta \sum_m (-1)^m \binom{k}{2m+1} \tan^{2m}\theta.\end{aligned}$$

Since $\tan^2\theta = \sec^2\theta - 1$, this is the required formula. ■ The real part, from even $j = 2m$, gives

$$\cos k\theta = \cos^k\theta \left(1 - \binom{k}{2}(\sec^2\theta - 1) + \binom{k}{4}(\sec^2\theta - 1)^2 - \cdots\right).$$

STAGE 2 Parity

Now $\sec\theta = a$, an odd integer, so $\sec^2\theta - 1 = a^2 - 1$ is an *even* integer. And θ is strictly between 0° and 90°, so $\sin\theta \neq 0$ and $\cos\theta \neq 0$.

Ask yourself If $\sin(k\theta) = 0$, the bracket $k - \binom{k}{3}(a^2-1) + \cdots$ must be zero. Every term after the first is a multiple of $a^2 - 1$. What does that tell you about k?

The bracket is k plus a sum of multiples of the even number $a^2 - 1$. If k were odd, the bracket would be odd, and an odd integer cannot be 0. So k must be even.

Then $\sin(k\theta) = 2\sin(\frac{1}{2}k\theta)\cos(\frac{1}{2}k\theta) = 0$. But $\frac{1}{2}k$ is a positive integer smaller than k, so by the minimality of k, $\sin(\frac{1}{2}k\theta) \neq 0$. Hence $\cos(\frac{1}{2}k\theta) = 0$. ■

STAGE 3 The contradiction

Pause and try Apply the cosine formula with $m = \frac{1}{2}k$. Why can $\cos(m\theta)$ never be zero?

By Stage 1, $\cos(m\theta) = a^{-m}\left(1 - \binom{m}{2}(a^2-1) + \binom{m}{4}(a^2-1)^2 - \cdots\right)$. The bracket is 1 plus multiples of the even number $a^2 - 1$, so it is odd, and in particular not zero. So $\cos(m\theta) \neq 0$, contradicting Stage 2.

Therefore there is no positive integer k with $\sin(k\theta) = 0$. (If there were any, there would be a smallest one, and we have just shown that the smallest one cannot exist.) By the observation in *Before you start*, θ cannot be a rational number of degrees. ■

STAGE 4 The cotangent version

Ask yourself If $\cot\phi = \frac{1}{b}$, what is $\sec^2\phi - 1$? Is it even?

$\tan\phi = b$, so $\sec^2\phi - 1 = \tan^2\phi = b^2$, which is even (indeed a multiple of 4) since b is even. Also $0^\circ < \phi < 90^\circ$, so $\sin\phi$ and $\cos\phi$ are non-zero. Exactly the same argument now works: if k is the smallest positive integer with $\sin(k\phi) = 0$, the bracket $k - \binom{k}{3}b^2 + \cdots$ must vanish, forcing k even; then $\cos(\frac{1}{2}k\phi) = 0$; but the cosine bracket $1 - \binom{m}{2}b^2 + \cdots$ is odd. Contradiction, so ϕ is irrational. ■

For example, $\cot^{-1}\frac{1}{2} = \tan^{-1}2 \approx 63.43^\circ$ is irrational.

Looking back

- **Rational angles have periodic multiples.** That reformulation turned irrationality into a statement about $\sin(k\theta)$ for all k.
- **Parity arguments.** Working modulo 2 showed that certain integers were odd and therefore not zero. Simple, but decisive.
- **The smallest counterexample.** Assuming a smallest k and showing a smaller one must exist (here $\frac{1}{2}k$) is the well-ordering argument of Chapter 1.

What's really going on

Irrational rotations

If θ is a rational number of degrees, repeatedly rotating by θ eventually returns you to where you started. If θ is irrational, you never return, and something remarkable happens: the points $k\theta$ (taken modulo 360°) spread out *evenly* around the circle, coming arbitrarily close to every angle. This is the *equidistribution theorem*, proved independently by Hermann Weyl, Wacław Sierpiński and Piers Bohl around 1910. Figure 5.5.1 shows the first 40 multiples of $\cos^{-1}\frac{1}{3}$: they are already filling the circle, and none ever repeats.

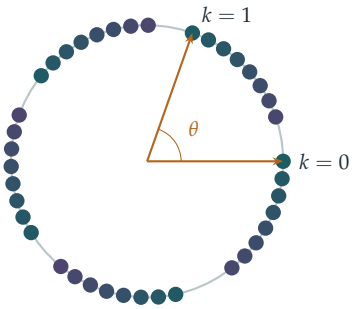

Figure 5.5.1 The points at angles $k\theta$, for $k = 0, 1, \ldots, 39$, where $\theta = \cos^{-1}\frac{1}{3} \approx 70.53^\circ$ (shading from teal to plum as k increases). Because θ is irrational, no point is ever repeated, and the points spread evenly around the whole circle.

Hilbert's third problem

In 1900 David Hilbert listed twenty-three problems for the new century. The third asked: if two polyhedra have the same volume, can one always be cut into finitely many polyhedral pieces and reassembled into the other? (For polygons in the plane the answer is yes: this is the Wallace–Bolyai–Gerwien theorem.) Within a year, Hilbert's student Max Dehn showed the answer is *no*: a regular tetrahedron cannot be cut up and reassembled into a cube of the same volume. Dehn's proof attaches to each polyhedron a quantity built from its edge lengths and dihedral angles, now called the *Dehn invariant*, which is unchanged by cutting and reassembling; for a cube it is zero, and for a regular tetrahedron it is non-zero precisely because the dihedral angle $\cos^{-1}\frac{1}{3}$ is not a rational multiple of 180°. That is exactly what this question proves.

Niven's theorem

How far does this go? In 1956 Ivan Niven proved that if θ is a rational number of degrees and $\cos\theta$ is rational, then $\cos\theta$ is one of $0, \pm\frac{1}{2}, \pm 1$. So the only "nice" angles with rational cosines are 0°, 60°, 90°, 120°, 180° and their relatives. The proof uses the

fact that $2\cos\theta$ is a root of a monic polynomial with integer coefficients, an *algebraic integer*, the idea mentioned at the end of Section 1.3.

The thinking skill

Parity as a proof of non-vanishing

To show an integer expression is never zero, show it is always odd: reduce it modulo 2.

Combined with a minimal counterexample, this simple idea proved a deep fact. Reducing modulo a small number to rule out an equation is one of the most common moves in number theory, as in Section 1.5.

Try this next

1. **The tetrahedron.** Show that the angle between two faces of a regular tetrahedron is $\cos^{-1}\frac{1}{3}$, and that the angle between two bonds in a methane molecule (from the centre of a regular tetrahedron to two vertices) is $\cos^{-1}\left(-\frac{1}{3}\right)$.

2. **The 3–4–5 triangle.** Show that the smallest angle of a 3–4–5 triangle is an irrational number of degrees. (*Hint:* it is $\cot^{-1}\frac{4}{3}$; adapt the argument, or consider powers of the Gaussian integer $4+3\mathrm{i}$ and show none is real.)

3. **Why odd a?** What goes wrong in the argument when $a=2$? Is $\cos^{-1}\frac{1}{2}$ rational?

4. **Niven for small cases.** Find all whole numbers of degrees θ between 0° and 90° for which $\cos\theta$ is rational, and check that your list agrees with Niven's theorem.

Further reading

- **Ivan Niven, *Irrational Numbers*** (Mathematical Association of America). Includes Niven's theorem on rational values of trigonometric functions.
- **Martin Aigner & Günter Ziegler, *Proofs from THE BOOK***, chapter on Hilbert's third problem. A beautiful account of Dehn's solution.
- **Benjamin Yandell, *The Honors Class: Hilbert's Problems and Their Solvers*** (A. K. Peters). The stories behind all twenty-three problems.

5.6 The Fermat Point

The question

Given that $\alpha = e^{i\pi/3}$, prove that $1 + \alpha^2 = \alpha$.

A triangle in the Argand plane has vertices A, B, and C represented by the complex numbers p, $q\alpha^2$ and $-r\alpha$ respectively, where p, q and r are positive real numbers. Sketch the triangle ABC.

Three equilateral triangles ABL, BCM and CAN (each lettered clockwise) are erected on sides AB, BC and CA respectively. Show that the complex number representing N is $(1-\alpha)p - \alpha^2 r$ and find similar expressions for the complex numbers representing L and M.

Show that lines LC, MA and NB all meet at the origin, and that these three line segments have the common length $p + q + r$.

The hook

Three villages want to build a road network connecting all of them, as short as possible. Where should the junction go? In about 1640 Pierre de Fermat posed the problem of finding the point P inside a triangle that minimises $PA + PB + PC$, and Evangelista Torricelli solved it: it is the point from which the three sides are seen at angles of $120°$.

This question uses complex numbers to prove a beautiful construction of that point. Build an equilateral triangle outwards on each side of the triangle, and join each new apex to the opposite vertex of the original. The three lines meet at a single point, the *Fermat point*, and all three have the same length, which is the minimum total road length. With complex numbers and one identity about $e^{i\pi/3}$, the whole thing is a few lines of algebra.

Before you start

Ask yourself Where are A, B and C? What are their arguments? What angles do OA, OB and OC make with each other?

$A = p$ lies on the positive real axis (argument $0°$). $B = q\alpha^2 = qe^{2\pi i/3}$ has argument $120°$. $C = -r\alpha = re^{i\pi}e^{i\pi/3}$ has argument $240°$. So the origin is a point inside the triangle from which the three vertices are seen at $120°$ to one another, at distances p, q and r. The origin *is* the Fermat point, by construction; the question shows that the

equilateral-triangle construction finds it.

Working it through

STAGE 1 The key identity

$\alpha^3 = e^{i\pi} = -1$, so α is a root of $z^3+1 = (z+1)(z^2-z+1)$. Since $\alpha \neq -1$, $\alpha^2 - \alpha + 1 = 0$, that is, $1+\alpha^2 = \alpha$. ■

Two consequences will be useful: $1-\alpha = -\alpha^2$, and $\alpha - \alpha^2 = 1$. Also $1-\alpha = \frac{1}{2} - \frac{\sqrt{3}}{2}i = e^{-i\pi/3}$, the complex conjugate of α.

STAGE 2 Equilateral triangles as rotations

Ask yourself Triangle CAN is equilateral, lettered clockwise. If you stand at C, how do you get from A to N? Express $N-C$ in terms of $A-C$.

Going clockwise round C, A, N, the point N is obtained by rotating A about C through $60°$ clockwise. Multiplying by $e^{-i\pi/3} = 1-\alpha$ rotates by $60°$ clockwise, so

$$\begin{aligned} N = C + (1-\alpha)(A-C) &= -r\alpha + (1-\alpha)(p+r\alpha) \\ &= (1-\alpha)p + r(\alpha - \alpha^2 - \alpha) = (1-\alpha)p - \alpha^2 r. \end{aligned}$$ ■

(Rotating clockwise rather than anticlockwise puts N on the far side of CA from the origin, outside the triangle.)

Pause and try Find L and M in the same way, and simplify using the identities from Stage 1.

Similarly $L = A + (1-\alpha)(B-A)$ and $M = B + (1-\alpha)(C-B)$:

$$\begin{aligned} L &= p + (1-\alpha)(q\alpha^2 - p) = \alpha p + (\alpha^2 - \alpha^3)q = \alpha p + (\alpha^2+1)q = \alpha(p+q), \\ M &= q\alpha^2 + (1-\alpha)(-r\alpha - q\alpha^2) = q\alpha^3 + r(\alpha^2 - \alpha) = -q - r. \end{aligned}$$

And using $1-\alpha = -\alpha^2$, $N = -\alpha^2 p - \alpha^2 r = -\alpha^2(p+r)$.

STAGE 3 Three lines through the origin

Ask yourself Compare $L = \alpha(p+q)$ with $C = -r\alpha$. What do you notice? Do the same for M and A, and for N and B.

L is a positive multiple of α and C is a negative multiple of α. So L and C lie on opposite sides of O on the same line through O, and the segment LC passes through the origin, with length $|L-C| = |\alpha|(p+q+r) = p+q+r$.

Likewise $M = -(q+r)$ and $A = p$ lie on the real axis on opposite sides of O, so MA passes through O with length $p+q+r$. And $N = -\alpha^2(p+r)$, $B = q\alpha^2$ lie on the line

through O in the direction α^2, on opposite sides, so NB passes through O with length $p+q+r$. ■

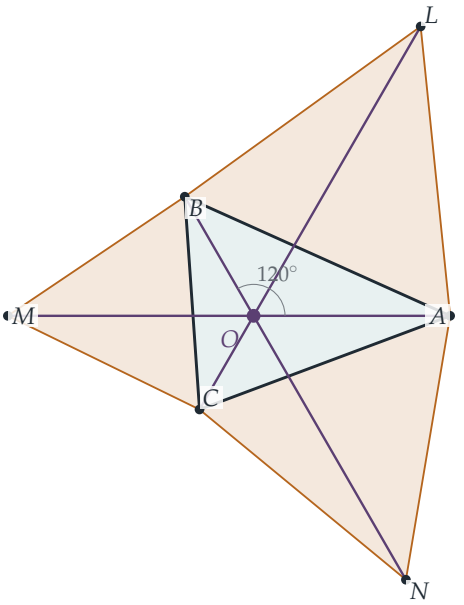

Figure 5.6.1 The Fermat point, with $p = 2$, $q = 1.4$, $r = 1.1$. Equilateral triangles (amber) are built outwards on the sides of ABC. The lines LC, MA and NB (plum) all pass through the origin O, where the sides of ABC subtend $120°$, and each has length $p+q+r = 4.5$.

Looking back

- **Rotation is multiplication.** Building an equilateral triangle meant rotating one vertex about another by $60°$, which is multiplication by $e^{\pm i\pi/3}$.
- **Clever coordinates.** Placing the Fermat point at the origin, with the vertices along the directions 1, α^2, $-\alpha$, made every expression collapse.
- **One identity, used repeatedly.** Every simplification came from $\alpha^2 - \alpha + 1 = 0$ and $\alpha^3 = -1$.

What's really going on

Why the Fermat point minimises distance

Why should $PA + PB + PC$ be smallest at the point where the sides subtend $120°$? Here is a lovely argument. Rotate the whole picture about B by $60°$, sending A to L' (the apex of an equilateral triangle on AB) and P to P'. Triangle BPP' is equilateral,

so $PB = PP'$, and $PA = P'L'$. Then $PA + PB + PC = L'P' + P'P + PC$, the length of a path from L' to C. That is at least $|L'C|$, with equality exactly when the path is straight, which happens when the angles at P are $120°$. So the minimum total length is $|LC| = p + q + r$, as found in the question. (This works provided every angle of the triangle is less than $120°$; otherwise the minimum is at the obtuse vertex.)

Soap films and Steiner trees

Dip a frame with three pegs (representing A, B, C) into soapy water: the soap film, minimising its surface tension energy, forms three walls meeting at $120°$ at the Fermat point. For more points the shortest connecting network is a *Steiner tree*, with extra junctions where three edges meet at $120°$. Finding Steiner trees is computationally hard in general (it is an NP-hard problem), but they matter in practice for designing networks of roads, pipelines and the wiring on computer chips.

Napoleon's theorem

The same configuration contains another gem: the *centres* of the three equilateral triangles on the sides of any triangle form an equilateral triangle. This is traditionally credited to Napoleon Bonaparte (who was a keen amateur mathematician), though there is no evidence he discovered it. It too has a short proof with complex numbers, using $1 + \alpha^2 = \alpha$.

The thinking skill

Choose coordinates that fit the problem

Put the most important point at the origin and line things up with the symmetry of the problem. A good choice can make the algebra almost trivial.

Here the origin was the Fermat point, and the three vertices lay along the cube roots of -1 (up to sign). With any other coordinates the same calculation would be much messier.

Try this next

1. **Check the lettering.** Show that if the equilateral triangles were lettered anticlockwise (built inwards), the three lines would generally not pass through O.

2. **Napoleon's theorem.** Show that the centre of the equilateral triangle ABL is $\frac{1}{3}(A + B + L)$. Using the expressions above, show that the three centres form an equilateral triangle.

3 **The minimum.** With $p = 2$, $q = 1.4$, $r = 1.1$, compute $PA + PB + PC$ for P at each vertex and at the centroid $\frac{1}{3}(A + B + C)$, and check that all exceed 4.5.

4 **An equilateral test.** Show that z_1, z_2, z_3 form an equilateral triangle (in either orientation) if and only if $z_1^2 + z_2^2 + z_3^2 = z_1z_2 + z_2z_3 + z_3z_1$.

Further reading

- **Richard Courant & Herbert Robbins, *What Is Mathematics?*** (Oxford University Press). Chapter 7 discusses the Fermat point, Steiner trees and soap films.
- **Liang-shin Hahn, *Complex Numbers and Geometry*** (Mathematical Association of America). Many geometric theorems proved with complex numbers, at exactly this level.
- **Tristan Needham, *Visual Complex Analysis*** (Oxford University Press). Chapter 1 includes Napoleon's theorem and other geometric results.

5.7 The Complex Exponential as a Map

The question

The complex numbers $w = u + \mathrm{i}v$ and $z = x + \mathrm{i}y$ are related by the equation

$$z = (\cos v + \mathrm{i}\sin v)\mathrm{e}^u.$$

Find all w which correspond to $z = \mathrm{i}\,\mathrm{e}$.

Find the loci in the x–y plane corresponding to the lines $u =$ constant in the u–v plane. Find also the loci corresponding to the lines $v =$ constant. Illustrate your answers with clearly labelled sketches.

Identify two subsets W_1 and W_2 of the u–v plane each of which is in one-to-one correspondence with the first quadrant $\{(x, y) : x > 0, y > 0\}$ of the x–y plane. Identify also two subsets W_3 and W_4 each of which is in one-to-one correspondence with the set $\{z : 0 < |z| < 1\}$.

[**NB** 'one-to-one' means here that to each value of w there is only one corresponding value of z, and vice-versa.]

The hook

You know the exponential function as a curve, $y = e^x$. But a function of a complex variable cannot be drawn as a curve: its input is a point of a plane and so is its output. The way to picture it is as a *map*, a transformation that moves every point of one plane to a point of another, stretching and bending as it goes.

The formula in the question is $z = e^u(\cos v + \mathrm{i}\sin v) = e^{u+\mathrm{i}v} = e^w$. Seen as a map, the complex exponential does something remarkable: it rolls up a horizontal strip of the plane into a whole plane (minus the origin), wrapping vertical lines round circles and laying horizontal lines out as rays. And because it wraps round, it is not one-to-one: infinitely many points go to each point. This is why the complex logarithm is "multi-valued", one of the first surprises of complex analysis.

Before you start

Ask yourself What are the modulus and argument of $z = e^u(\cos v + \mathrm{i}\sin v)$ in terms of u and v?

$|z| = e^u$ and $\arg z = v$ (up to multiples of 2π). So the real part u of w controls the distance from the origin, exponentially, and the imaginary part v controls the direction. That one observation answers almost every part of the question.

Working it through

STAGE 1 Solving $e^w = \mathrm{i}e$

Ask yourself What are the modulus and argument of $\mathrm{i}e$?

$|\mathrm{i}e| = e$ and $\arg(\mathrm{i}e) = \frac{\pi}{2}$. So we need $e^u = e$, giving $u = 1$, and $v = \frac{\pi}{2} + 2k\pi$ for any integer k. The solutions are

$$w = 1 + \mathrm{i}\left(\tfrac{\pi}{2} + 2k\pi\right), \qquad k \in \mathbb{Z}.$$

There are infinitely many, spaced 2π apart vertically.

STAGE 2 The images of grid lines

On a vertical line $u = c$, $|z| = e^c$ is fixed while v varies: the image is the *circle* of radius e^c centred at the origin, traced round infinitely often as v runs over all real numbers.

On a horizontal line $v = d$, $\arg z = d$ is fixed while $|z| = e^u$ takes every positive value: the image is the open *ray* (half-line) from the origin at angle d, not including the origin itself. Lines $v = d$ and $v = d + 2\pi$ give the same ray.

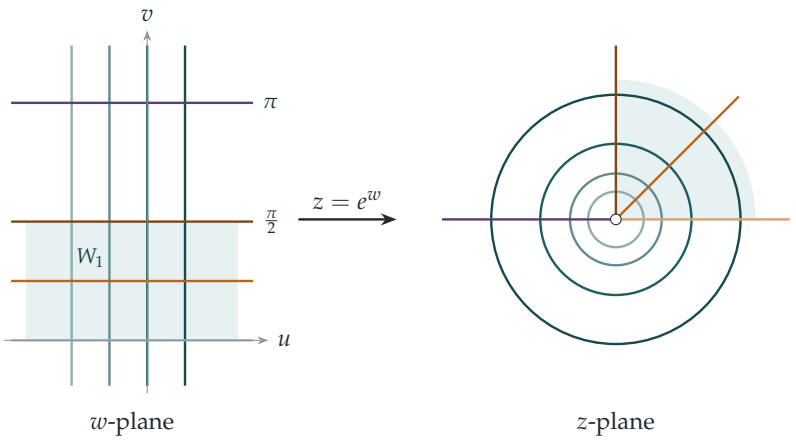

Figure 5.7.1 The complex exponential as a map. Vertical lines $u = c$ in the w-plane (teal) go to circles of radius e^c; horizontal lines $v = d$ (amber, plum) go to rays at angle d. The shaded strip $W_1 = \{0 < v < \frac{\pi}{2}\}$ goes exactly onto the first quadrant. Notice that the grid lines meet at right angles in both pictures: the map preserves angles. The origin (open circle) is never reached.

STAGE 3 One-to-one pieces

Ask yourself Which points w map into the first quadrant $x > 0$, $y > 0$? Why is that set too big to be in one-to-one correspondence with the quadrant?

z is in the open first quadrant exactly when $\arg z$ is strictly between 0 and $\frac{\pi}{2}$ (modulo 2π), with any modulus. So w must satisfy $0 < v < \frac{\pi}{2}$ *or* $2\pi < v < \frac{5\pi}{2}$ or $\ldots$: infinitely many horizontal strips, each of which maps onto the whole quadrant. Any one strip is in one-to-one correspondence with it, since within a strip u and v are determined uniquely by $|z|$ and $\arg z$. So we can take

$$W_1 = \{u + iv : 0 < v < \tfrac{\pi}{2}\}, \qquad W_2 = \{u + iv : 2\pi < v < \tfrac{5\pi}{2}\}.$$

Pause and try Now find two sets in the w-plane in one-to-one correspondence with the punctured disc $0 < |z| < 1$.

$|z| < 1$ means $e^u < 1$, that is $u < 0$; and every argument must be covered exactly once, so v should range over an interval of length exactly 2π, including one end but not the other. For example,

$$W_3 = \{u + iv : u < 0,\ 0 \le v < 2\pi\}, \qquad W_4 = \{u + iv : u < 0,\ -\pi < v \le \pi\}.$$

These are half-strips with one edge included and one excluded. (If both edges were included, the points on the positive real axis would be hit twice; if neither, they would be missed.) ■

Looking back

- **Modulus and argument separate the variables.** $|e^w| = e^{\operatorname{Re} w}$ and $\arg e^w = \operatorname{Im} w$, so the real and imaginary parts of w control distance and direction independently.
- **Picture a complex function as a map.** Tracking where grid lines go gives a clear picture of a function that cannot be drawn as a graph.
- **Periodicity destroys injectivity.** $e^{w+2\pi\mathrm{i}} = e^w$, so the map is many-to-one, and must be restricted to a strip of height 2π to have an inverse.

What's really going on

The complex logarithm

On the real line, ln is the inverse of exp. In the complex plane, $e^w = z$ has infinitely many solutions w for each $z \neq 0$, differing by multiples of $2\pi\mathrm{i}$, so "$\log z$" is not a single number. To get a genuine function you must choose a *branch*, which means choosing a strip like W_3 or W_4: the *principal branch* takes $-\pi < \arg z \leq \pi$. Every branch has a jump somewhere, along a ray from the origin called a *branch cut*. Bernhard Riemann's resolution, in his thesis of 1851, was to stop insisting that the logarithm lives on a plane: it lives on an infinite spiral staircase, a *Riemann surface*, on which it is single-valued and smooth. Riemann surfaces are now central to geometry, number theory and string theory.

Angles are preserved

Look again at Figure 5.7.1. The vertical and horizontal grid lines meet at right angles, and so do the circles and rays they become. This is not a coincidence: wherever the derivative of a complex function is non-zero, the function preserves the angles between curves. Such maps are called *conformal*. Locally, a conformal map is just a rotation and a scaling, which is exactly what multiplication by the complex number $f'(z)$ does. Conformal maps are used to solve problems in fluid flow, electrostatics and heat conduction, by transforming a complicated region into a simple one where the answer is known.

Mercator's map

Here is a surprise connecting this section with Section 4.8. Gerardus Mercator's map of 1569 was designed so that a course of constant compass bearing is a straight line, which requires preserving angles. If you first project the globe stereographically onto a plane (from the north pole), then take the complex logarithm, the result is exactly Mercator's projection: meridians become vertical lines and lines of latitude become horizontal lines, just as rays and circles become horizontal and vertical lines under log. The exponential map in this section, run backwards, is the mathematics of the navigator's chart.

The thinking skill

Visualise functions as transformations

When a function cannot be drawn as a graph, draw what it does: follow a grid of simple curves and see where they go.

For complex functions, the images of horizontal and vertical lines, or of circles and rays, usually tell the whole story. This way of thinking, about what a map does to shapes, runs through all of geometry and analysis.

Try this next

1. **Solving equations.** Find all complex w with $e^w = -1$, and all w with $e^w = 1 + \mathrm{i}$.
2. **The image of a square.** Find and sketch the image under $z = e^w$ of the square with corners 0, 1, $1 + \mathrm{i}$ and i.
3. **Right angles.** Show directly that the circle $|z| = e^c$ and the ray $\arg z = d$ meet at right angles, and explain why this is consistent with the lines $u = c$ and $v = d$ being perpendicular.
4. **Squaring.** Find the images of the lines $x = c$ and $y = d$ under the map $z \mapsto z^2$. Show that they are parabolas, and that they meet at right angles.

Further reading

- **Tristan Needham, *Visual Complex Analysis*** (Oxford University Press). Chapters 2 and 4 treat the exponential, the logarithm and conformal maps with superb pictures.
- **Hilary Priestley, *Introduction to Complex Analysis*** (Oxford University Press). A clear first undergraduate text.
- **Ian Stewart & David Tall, *Complex Analysis*** (Cambridge University Press). Another excellent undergraduate introduction, with an emphasis on geometry.

5.8 A Hidden Distance in the Disc

The question

Let z and w be complex numbers. Use a diagram to show that $|z-w| \leq |z|+|w|$.

For any complex numbers z and w, E is defined by

$$E = zw^* + z^*w + 2|zw|.$$

(i) Show that $|z-w|^2 = \left(|z|+|w|\right)^2 - E$, and deduce that E is real and non-negative.

(ii) Show that $|1-zw^*|^2 = \left(1+|zw|\right)^2 - E$.

Hence show that, if both $|z|>1$ and $|w|>1$, then

$$\frac{|z-w|}{|1-zw^*|} \leq \frac{|z|+|w|}{1+|zw|}.$$

Does this inequality also hold if both $|z|<1$ and $|w|<1$?

The hook

At first sight this is an exercise in manipulating moduli and conjugates, ending in a strange-looking inequality. Why would anyone care about $\frac{|z-w|}{|1-zw^*|}$?

Because it is a distance. Not the ordinary distance between z and w, but the distance in a different geometry, one in which the unit disc is an entire infinite universe, straight lines are circular arcs, and the angles of a triangle add up to less than 180°. This is *hyperbolic geometry*, discovered in the nineteenth century after two thousand years of attempts to prove Euclid's parallel postulate, and the inequality in this question turns out to be its triangle inequality, in disguise. Even more surprisingly, the right-hand side is the formula for adding velocities in special relativity, which we met in Section 3.8.

Before you start

Ask yourself How does $|z-w| \leq |z|+|w|$ follow from a diagram? (Think of z and $-w$ as vectors.)

$z-w = z+(-w)$. Draw z as an arrow from the origin, then $-w$ as an arrow from the

tip of z. The arrow from the origin to the final point is $z - w$, and the three arrows form a triangle with sides $|z|$, $|w|$ and $|z - w|$. One side of a triangle is at most the sum of the other two.

Also recall that $|a|^2 = aa^*$ for any complex number a. That converts every modulus in the question into algebra.

Working it through

STAGE 1 Expanding the moduli

Ask yourself Expand $|z - w|^2 = (z - w)(z^* - w^*)$. Compare with $(|z| + |w|)^2$.

$|z - w|^2 = zz^* - zw^* - z^*w + ww^* = |z|^2 + |w|^2 - zw^* - z^*w$. Meanwhile $(|z| + |w|)^2 - E = |z|^2 + 2|z||w| + |w|^2 - zw^* - z^*w - 2|zw|$, and since $|z||w| = |zw|$ this is the same. ■

$E = (zw^*) + (zw^*)^* + 2|zw| = 2\,\mathrm{Re}(zw^*) + 2|zw|$, which is real. And $E = (|z| + |w|)^2 - |z - w|^2 \geq 0$ by the triangle inequality. ■

STAGE 2 The second identity

$|1 - zw^*|^2 = (1 - zw^*)(1 - z^*w) = 1 - zw^* - z^*w + |z|^2|w|^2$, while $(1 + |zw|)^2 - E = 1 + 2|zw| + |zw|^2 - zw^* - z^*w - 2|zw|$, which is the same. ■

STAGE 3 The inequality

Ask yourself Write $A = (|z| + |w|)^2$ and $B = (1 + |zw|)^2$. The inequality, squared, says $\frac{A-E}{B-E} \leq \frac{A}{B}$. When is that true?

Both sides of the required inequality are non-negative, so it is equivalent to its square, $\frac{A-E}{B-E} \leq \frac{A}{B}$. Assuming $B - E = |1 - zw^*|^2 > 0$, multiply out:

$$(A - E)B \leq A(B - E) \iff -EB \leq -EA \iff E(B - A) \geq 0.$$

Since $E \geq 0$, this holds whenever $B \geq A$, that is, $1 + |zw| \geq |z| + |w|$.

Pause and try Show that $1 + |z||w| - |z| - |w|$ factorises. When is it non-negative?

$1 + |z||w| - |z| - |w| = (1 - |z|)(1 - |w|)$. If $|z| > 1$ and $|w| > 1$, both factors are negative and the product is positive. So $B \geq A$ and the inequality holds. ■

If instead $|z| < 1$ and $|w| < 1$, both factors are positive, the product is again positive, and the same argument works: **yes**, the inequality holds in this case too. (In both cases $|zw^*| \neq 1$, so $1 - zw^* \neq 0$ and the left-hand side is defined.) It is the mixed case, one

inside the unit circle and one outside, where it can fail.

Looking back

- $|a|^2 = aa^*$. Every identity came from expanding products of conjugates. This is the standard way to handle moduli algebraically.
- **Reduce to the sign of one quantity.** After squaring and cross-multiplying, the whole inequality depended on the sign of $(1 - |z|)(1 - |w|)$.
- **Check other cases.** The final question asked about $|z|, |w| < 1$; the same factorisation answered it immediately.

What's really going on

Non-Euclidean geometry

For two thousand years mathematicians tried to prove Euclid's fifth postulate (through a point not on a line, there is exactly one parallel to the line) from his other axioms. In the 1820s and 1830s Nikolai Lobachevsky, János Bolyai and (privately) Gauss realised that it cannot be done: there is a perfectly consistent geometry in which there are infinitely many parallels. In 1882 Henri Poincaré found a beautiful model of it inside the unit disc. Points are points of the open disc; "straight lines" are diameters and circular arcs that meet the boundary circle at right angles; and the distance between z and w is

$$d(z, w) = 2 \tanh^{-1} \rho(z, w), \qquad \text{where} \qquad \rho(z, w) = \frac{|z - w|}{|1 - zw^*|}.$$

As a point approaches the boundary circle, its distance from the centre tends to infinity: the disc is a whole infinite plane, seen through a distorting lens.

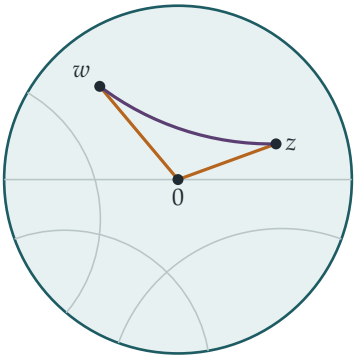

Figure 5.8.1 The Poincaré disc. Hyperbolic "straight lines" are diameters or arcs of circles meeting the boundary at right angles (grey). The hyperbolic triangle with vertices 0, z and w has two sides along diameters (amber) and one along an arc (plum). The inequality in the question says that the plum side is no longer, in hyperbolic distance, than the other two put together.

The triangle inequality in disguise

Here is what the question proved, for $|z|, |w| < 1$. Write $|z| = \tanh a$ and $|w| = \tanh b$, so that the hyperbolic distances from 0 to z and w are $2a$ and $2b$. Then by the addition formula for tanh from Section 3.8,

$$\frac{|z| + |w|}{1 + |zw|} = \frac{\tanh a + \tanh b}{1 + \tanh a \tanh b} = \tanh(a + b).$$

So the inequality says $\tanh\left(\frac{1}{2}d(z, w)\right) \leq \tanh(a + b)$, that is,

$$d(z, w) \leq 2a + 2b = d(z, 0) + d(0, w).$$

This is the triangle inequality for hyperbolic distance (Figure 5.8.1). The relativistic velocity-addition formula and hyperbolic geometry are secretly the same thing: the set of possible velocities in special relativity, with the right notion of distance, *is* a hyperbolic space. We return to this in Section 7.6.

Symmetries of the disc

Why is ρ the right quantity? The maps $z \mapsto \dfrac{z - a}{1 - a^* z}$, for $|a| < 1$, send the unit disc to itself, preserve the hyperbolic distance, and move a to the centre. They are the hyperbolic analogues of translations, and $\rho(z, w)$ is exactly $|\text{image of } z|$ when w is moved to the centre. These *Möbius transformations* are a central topic in complex analysis. M. C. Escher's *Circle Limit* woodcuts, inspired by a diagram of the mathematician H. S. M. Coxeter, show tilings of the Poincaré disc by congruent fish and angels, which shrink towards the edge because hyperbolic distances blow up there.

The thinking skill

Ask what an expression means

When an algebraic expression keeps appearing, ask whether it has a meaning: a distance, an angle, a probability, a symmetry.

Here a strange ratio of moduli turned out to be a distance in another geometry, and a strange inequality turned out to be its triangle inequality. Recognising the meaning explains why the result is true and where it leads.

Try this next

1. **A numerical check.** Take $z = \frac{1}{2}$ and $w = -\frac{1}{2}$. Compute both sides of the inequality. Then compute the hyperbolic distances $d(z, 0)$, $d(0, w)$ and $d(z, w)$ and check the triangle inequality is an equality here. Why?

2. **The mixed case.** Find z, w with $|z| < 1 < |w|$ for which the inequality fails.

3. **An isometry.** Let $f(z) = \dfrac{z - a}{1 - a^* z}$ with $|a| < 1$. Show that $|f(z)| = 1$ whenever $|z| = 1$, and that $f(a) = 0$. Show that $\rho(f(z), f(w)) = \rho(z, w)$.

4. **Infinity is near.** Show that the hyperbolic distance from 0 to r (with $0 < r < 1$) is $\ln \frac{1+r}{1-r}$, and that it tends to infinity as $r \to 1$. How close to the boundary must a point be to be hyperbolic distance 10 from the centre?

Further reading

- **Caroline Series, "Hyperbolic geometry" lecture notes** (University of Warwick, freely available). A clear first course, starting from the Poincaré disc.
- **Jeremy Gray, *Ideas of Space*** (Oxford University Press). The history of non-Euclidean geometry from Euclid to Einstein.
- **Tristan Needham, *Visual Complex Analysis*** (Oxford University Press). Chapter 6 covers non-Euclidean geometry and Möbius transformations, with beautiful pictures.

What We Have Learned

We began with Cardano's unsettling discovery that real answers can require imaginary steps, and with three questions: how complex numbers help with real problems, what is special about roots of unity, and what complex functions look like. Here is what the chapter's eight questions showed.

1. How do complex numbers help with real problems?

Again and again, a fact about real numbers had a short proof through the complex plane.

- **Angles become products.** Machin's formula for π is the statement that $(5-\mathrm{i})^4(1+\mathrm{i})$ is a real multiple of $239-\mathrm{i}$, checked with whole-number arithmetic (Section 5.1).
- **Counting becomes averaging.** Sums of every third binomial coefficient came from evaluating $(1+x)^n$ at the cube roots of unity (Section 5.2).
- **Trigonometric products become polynomial values.** Products of cosines at equally spaced angles came from factorising $z^{2n}+1$ and putting $z=\mathrm{i}$ (Section 5.3).
- **Geometry becomes algebra.** Distances became moduli, so the product of chords from a vertex of a regular n-gon is $|1+1+\cdots+1|=n$ (Section 5.4); rotations became multiplications, so the Fermat point construction collapsed to the identity $1+\alpha^2=\alpha$ (Section 5.6).
- **Irrationality from De Moivre.** Expanding $(\cos\theta+\mathrm{i}\sin\theta)^k$ and arguing about parity showed that $\cos^{-1}\frac{1}{3}$ is not a rational number of degrees, the key fact in Dehn's solution of Hilbert's third problem (Section 5.5).

2. What is special about the roots of unity?

The nth roots of unity are the vertices of a regular n-gon, they are the roots of the simplest possible polynomial z^n-1, and they form a group under multiplication. Those three facts together explain their ubiquity. Because they are evenly spread round a circle, they sum to zero, which gives the filter of Section 5.2 and, in general, the discrete Fourier transform. Because they are the roots of z^n-1, products over them are values of simple polynomials (Section 5.3, Section 5.4). And because they are closed under multiplication, powers of a root of unity cycle round, which is exactly what fails for an angle like $\cos^{-1}\frac{1}{3}$ (Section 5.5). In Chapter 6 roots of unity become the heart of algebra, leading towards Galois theory.

3. What does a complex function look like?

In Section 5.7 we pictured e^w as a map that rolls a strip into a plane, sending vertical lines to circles and horizontal lines to rays, and wrapping round infinitely often, so that the logarithm is many-valued. We saw that it preserves angles, the first example of a *conformal* map, and that it underlies Mercator's chart. In Section 5.8 a strange inequality turned out to be the triangle inequality of hyperbolic geometry, where the unit disc is an infinite plane, straight lines are circular arcs, and distances combine by the relativistic velocity-addition formula.

4. A toolkit for complex numbers

Method	Where you used it
Arguments add under multiplication	Machin's formula (5.1)
Checking multiples of 2π	Machin's formula (5.1)
Averaging over roots of unity	Every third coefficient (5.2)
Factorising over roots, then substituting	Cosine products (5.3); chords (5.4)
Dividing out a vanishing factor	5.3; 5.4
De Moivre's theorem	Multiple-angle formulas (5.5)
Parity and minimal counterexamples	Irrational angles (5.5)
Rotation as multiplication	Equilateral triangles (5.6)
Mapping grids	The exponential map (5.7)
$\lvert a\rvert^2 = aa^*$	The Poincaré disc inequality (5.8)

Where next

- Roots of unity drive Chapter 6, including the cyclotomic polynomials of Section 6.1 and Gauss's construction of the regular 17-gon in Section 6.2.
- The connection between hyperbolic geometry, tanh and relativity is developed in Section 7.6.
- Gaussian integers, which appeared in Machin's formula, were the subject of Section 1.6.

If you want to go further before university

- **Tristan Needham, *Visual Complex Analysis***: the most beautiful book on complex numbers ever written, full of pictures. Start with Chapter 1.
- **Paul Nahin, *An Imaginary Tale***: the history of $\sqrt{-1}$, from Cardano to modern physics.
- **Liang-shin Hahn, *Complex Numbers and Geometry***: Euclidean geometry with complex numbers, at exactly the level of this chapter.
- **Hilary Priestley, *Introduction to Complex Analysis***: for a first look at the univer-

sity course.

Cardano thought his imaginary quantities were as useless as they were subtle. Five centuries later, it is hard to find a corner of mathematics or physics where they do not appear. The lesson of this chapter is Hadamard's: when a problem about real numbers is hard, try stepping off the real line. The complex plane is bigger, and in it, many things that look complicated turn out to be simple.