

CHAPTER

6

Algebra and Structure

Behind every clever trick in algebra there is a structure, and understanding the structure is worth more than any number of tricks.

IN THIS CHAPTER

- | | | |
|-----|---|---------------|
| 6.1 | Polynomials Whose Roots Are Roots of Unity | STEP III 2010 |
| 6.2 | Gauss's 17-gon | STEP II 1991 |
| 6.3 | A Polynomial for $\sqrt{2} + \sqrt{3} + \sqrt{5}$ | STEP II 2023 |
| 6.4 | Solving the Quartic | STEP III 2017 |
| 6.5 | Differentiation from Axioms | STEP II 1992 |
| 6.6 | Unions of Subgroups | STEP II 1988 |
| 6.7 | The Quaternion Group | STEP III 1993 |
| 6.8 | A Group of Singular Matrices | STEP III 1996 |

From Equations to Structures

Mathematics is the art of giving the same name to different things.

— Henri Poincaré

For most of its history, algebra meant solving equations. The Babylonians solved quadratics four thousand years ago; Italian mathematicians of the sixteenth century solved cubics and quartics; and then, for nearly three hundred years, the quintic resisted every attempt. The resolution, when it came in the 1820s and 1830s, changed algebra completely. Niels Henrik Abel proved that no general formula for the quintic exists, and Évariste Galois explained exactly which equations can be solved, by studying not the equations themselves but the *symmetries of their roots*. To do so he had to invent a new kind of mathematical object: the group.

That shift, from computing answers to studying structures, is the subject of this chapter. Instead of asking “what is the solution?”, modern algebra asks “what kind of object is this, and what operations does it support?” The answers turn out to explain why the old tricks worked, and to settle questions that tricks alone never could.

Why structure matters

Here are three ancient problems that resisted all attempts for two thousand years: with ruler and compasses, double the volume of a cube; trisect an arbitrary angle; construct a regular 7-sided polygon. None of them can be done, and the proofs, found in the nineteenth century, are not clever constructions but arguments about structure: about the *degree* of the field of numbers that ruler and compasses can reach (Section 6.3). The same theory explains why the regular 17-gon *can* be constructed (Section 6.2), and why the quartic can be solved but the quintic cannot (Section 6.4).

Structures also travel. The axioms for a group are satisfied by the symmetries of a crystal, the permutations of the roots of a polynomial, the rotations of three-dimensional space, the moves of Rubik’s cube and the points of an elliptic curve. A theorem proved once from the axioms holds in all of them. That economy, giving “the same name to different things” as Poincaré put it, is what makes abstract algebra so powerful.

What you will meet at university

In the first year you will meet *groups* (sets with an operation satisfying three axioms) and, usually, *rings* and *fields* (sets with addition and multiplication). You will prove Lagrange’s theorem, that the size of a subgroup divides the size of the group, and meet permutations, symmetry groups and matrix groups. In the second and third years these lead to *Galois theory*, which connects fields and groups and answers the

classical questions about equations and constructions; to *representation theory*, which studies groups through matrices; and to *algebraic number theory*, where the ideas of Section 1.7 are put on a firm footing. The last three sections of this chapter are about groups. Groups are no longer on the STEP syllabus (they were removed in 2019), but the questions are self-contained and each section gives the definitions it needs.

The questions that run through this chapter

Three big questions

- 1. **What can be built from what?** Which numbers can be reached from the rationals by square roots, or by solving equations, and how do we prove something cannot be reached?
- 2. **Why do the formulas for equations work?** What is the hidden reason that quadratics, cubics and quartics have formulas, and what changes at degree five?
- 3. **What is gained by abstraction?** What happens when we describe objects only by the rules they obey?

Ideas in this chapter

§	Topic	Idea
6.1	Cyclotomic polynomials	Classify roots of unity by order
6.2	The regular 17-gon	Gauss periods: sums and products give quadratics
6.3	Minimal polynomials	Eliminate radicals; degrees multiply
6.4	Solving the quartic	Symmetric functions; the resolvent cubic
6.5	Differentiation from axioms	An operation determined by its rules
6.6	Unions of subgroups	Arguing from the group axioms
6.7	The quaternion group	Representing an operation by matrices
6.8	A group of singular matrices	Reading definitions literally

How to read this chapter

Many of these questions can be done by brute force: expand everything, compare coefficients, check every case. Do that first if you need to. Then look back and ask what made the brute force work, and whether there is a structural reason behind it. The structural reason is usually shorter, always more illuminating, and it is what you will need when the brute force becomes impossible.

6.1 Polynomials Whose Roots Are Roots of Unity

The question

For any given positive integer n , a number a (which may be complex) is said to be a *primitive n th root of unity* if $a^n = 1$ and there is no integer m such that $0 < m < n$ and $a^m = 1$. Write down the two primitive 4th roots of unity.

Let $C_n(x)$ be the polynomial such that the roots of the equation $C_n(x) = 0$ are the primitive n th roots of unity, the coefficient of the highest power of x is one and the equation has no repeated roots. Show that $C_4(x) = x^2 + 1$.

- (i) Find $C_1(x)$, $C_2(x)$, $C_3(x)$, $C_5(x)$ and $C_6(x)$, giving your answers as unfactorised polynomials.
- (ii) Find the value of n for which $C_n(x) = x^4 + 1$.
- (iii) Given that p is prime, find an expression for $C_p(x)$, giving your answer as an unfactorised polynomial.
- (iv) Prove that there are no positive integers q , r and s such that $C_q(x) \equiv C_r(x) C_s(x)$.

The hook

The n th roots of unity, the solutions of $z^n = 1$, sit at the vertices of a regular n -gon. But they are not all alike. Among the 12th roots of unity, for instance, 1 is already a first root of unity, -1 is a square root of unity, i is a fourth root, and only four of the twelve, the *primitive* ones, need all twelve steps to get back to 1.

Sorting the roots of unity by the number of steps they need, and building a polynomial from each class, produces the *cyclotomic polynomials* ("circle-dividing", from the Greek). They are the basic building blocks of $z^n - 1$, in the same way that primes are the building blocks of whole numbers, and they lie behind Gauss's construction of the regular 17-gon (Section 6.2), the roots-of-unity filter (Section 5.2) and a good deal of modern number theory.

Before you start



Ask yourself Every n th root of unity satisfies $z^n = 1$. But the smallest positive power that gives 1 might be less than n . For each of the four 4th roots of unity, 1, i , -1 , $-i$, what is that smallest power?

$1^1 = 1$, $(-1)^2 = 1$, and $i^4 = (-i)^4 = 1$ with no smaller power working. So the smallest powers are 1, 4, 2, 4. This smallest power is called the *order* of the root of unity, and the primitive n th roots of unity are exactly those of order n . The key fact for the whole question is that *every root of unity has exactly one order*.

Working it through

STAGE 1 The first few

The primitive 4th roots are i and $-i$, so $C_4(x) = (x - i)(x + i) = x^2 + 1$. ■



Ask yourself Which roots of unity have order 1, 2, 3, 5 and 6? (Draw the 6th roots of unity and sort them by order.)

Order 1: just 1, so $C_1(x) = x - 1$. Order 2: just -1 , so $C_2(x) = x + 1$.

Order 3: the cube roots of unity other than 1, that is, the roots of $\frac{x^3-1}{x-1}$. So $C_3(x) = x^2 + x + 1$.

Order 5: the fifth roots other than 1, which are the roots of $\frac{x^5-1}{x-1}$, so $C_5(x) = x^4 + x^3 + x^2 + x + 1$.

Order 6: of the sixth roots $e^{k\pi i/3}$, $k = 0, \dots, 5$, those with $k = 0, 3$ have orders 1, 2, and $k = 2, 4$ have order 3. That leaves $e^{\pm\pi i/3}$, so

$$C_6(x) = (x - e^{i\pi/3})(x - e^{-i\pi/3}) = x^2 - 2x \cos \frac{\pi}{3} + 1 = x^2 - x + 1. \quad \blacksquare$$

STAGE 2 Which n gives $x^4 + 1$?

The roots of $x^4 + 1 = 0$ satisfy $x^4 = -1$, so $x^8 = 1$ but $x^4 \neq 1$. Their order divides 8 but not 4, so it is 8. Conversely, a root of order 8 satisfies $x^8 = 1$, $x^4 \neq 1$, so $x^4 = -1$. There are four primitive 8th roots and four roots of $x^4 + 1$, so $n = 8$. ■

STAGE 3 Primes

If p is prime, the order of a p th root of unity divides p , so it is 1 or p . Every p th root except 1 is therefore primitive, and

$$C_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \dots + x + 1. \quad \blacksquare$$

STAGE 4 No cyclotomic polynomial is a product of two others

Pause and try Suppose $C_q = C_r C_s$. Take a root of C_r . What is its order? What does the identity say about its order? What does that force?

Every root of C_r (there is at least one, since $e^{2\pi i/r}$ is primitive) is a root of $C_r C_s = C_q$, so it has order q . But it also has order r , and a root of unity has only one order, so $r = q$. Then $C_q = C_q C_s$, and since C_q is not the zero polynomial, $C_s = 1$. That is impossible, because C_s has at least one root, $e^{2\pi i/s}$. So no such q, r, s exist. ■

Looking back

- **Sort by order.** Every root of unity has a unique order, and grouping by order splits the roots of $z^n - 1$ into disjoint classes.
- **Define a polynomial by its roots.** Each C_n was pinned down by saying exactly which numbers are its roots.
- **Uniqueness gives impossibility.** Part (iv) needed only one idea: a root cannot have two different orders.

What's really going on**Factorising $x^n - 1$**

Every n th root of unity has some order d that divides n , and every d th root of unity of order d is an n th root of unity when $d \mid n$. So the roots of $x^n - 1$ split into classes by order, giving

$$x^n - 1 = \prod_{d \mid n} C_d(x).$$

For example $x^6 - 1 = C_1 C_2 C_3 C_6 = (x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1)$. Figure 6.1.1 shows the twelve 12th roots of unity sorted in this way. The number of primitive n th roots, which is the degree of C_n , is the number of k between 1 and n with no common factor with n : *Euler's totient function* $\varphi(n)$. Comparing degrees gives the identity $\sum_{d \mid n} \varphi(d) = n$.

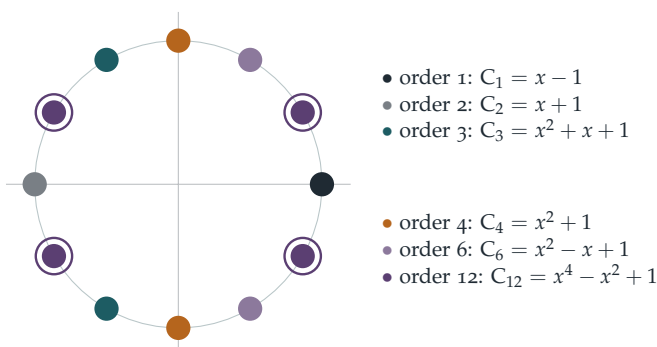


Figure 6.1.1 The twelve 12th roots of unity, coloured by order. The four primitive 12th roots (ringed) are the roots of C_{12} . Together the classes give $x^{12} - 1 = C_1 C_2 C_3 C_4 C_6 C_{12}$, with degrees $1 + 1 + 2 + 2 + 2 + 4 = 12$.

Integer coefficients, and a surprise

It is not obvious from the definition, but every cyclotomic polynomial has *integer* coefficients. (Use the factorisation of $x^n - 1$ and induction: C_n is $x^n - 1$ divided by a product of earlier cyclotomic polynomials, and long division of integer polynomials by monic integer polynomials keeps the coefficients integers.) For a long time all known examples had coefficients 0, 1 and -1 . The first exception is C_{105} , which contains the term $-2x^7$, found by Migotti in 1883; $105 = 3 \times 5 \times 7$ is the smallest number with three distinct odd prime factors. In fact the coefficients of cyclotomic polynomials can be made as large as you like.

Irreducibility

Part (iv) shows that no cyclotomic polynomial is a product of two others. Something much stronger is true: C_n cannot be factorised at all into polynomials with rational coefficients of lower degree. Gauss proved this for prime n (it follows from *Eisenstein's criterion*, which you will meet in a first algebra course), and Kronecker and Dedekind for general n . This *irreducibility* is what makes the cyclotomic polynomials the “primes” among polynomials related to roots of unity, and it is the starting point for the Galois theory of roots of unity.

The thinking skill

Classify, then build

Sort the objects in a problem into classes with a single, well-defined invariant, and build structure class by class.

Here the invariant was the order of a root of unity. Because each root has exactly one order, the classes do not overlap, and $x^n - 1$ splits as a product over them. The same pattern (a unique invariant giving a disjoint decomposition) underlies prime factorisation and the classification of finite abelian groups.

Try this next

- 1 **Check the factorisation.** Verify that $(x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1) = x^6 - 1$, and that $x^{12} - 1$ factorises as in Figure 6.1.1.
- 2 **Twice a prime.** Show that if p is an odd prime then $C_{2p}(x) = C_p(-x)$. Check this for $p = 3$.
- 3 **Totients.** Show that $\varphi(p^k) = p^k - p^{k-1}$ for a prime p , and that $C_{p^k}(x) = C_p(x^{p^{k-1}})$. Use this to find C_9 .
- 4 **A longer one.** HARDER Using $x^{15} - 1 = C_1 C_3 C_5 C_{15}$, show that $C_{15}(x) = x^8 - x^7 + x^5 - x^4 + x^3 - x + 1$.

Further reading

- **Ian Stewart, *Galois Theory*** (CRC Press). The standard undergraduate introduction; cyclotomic polynomials and the regular polygon appear early.
- **Harold Davenport, *The Higher Arithmetic*** (Cambridge University Press). Includes roots of unity and their arithmetic.
- **Marcus du Sautoy, *Finding Moonshine*** (Fourth Estate). A popular account of symmetry and group theory, including the story of Galois.

6.2 Gauss's 17-gon

The question

Let $y = \cos \phi + \cos 2\phi$, where $\phi = \frac{2\pi}{5}$. Verify by direct substitution that y satisfies the quadratic equation $2y^2 = 3y + 2$ and deduce that the value of y is $-\frac{1}{2}$.

Let $\theta = \frac{2\pi}{17}$. Show that

$$\sum_{k=0}^{16} \cos k\theta = 0.$$

If $z = \cos \theta + \cos 2\theta + \cos 4\theta + \cos 8\theta$, show that the value of z is $-(1 - \sqrt{17})/4$.

The hook

On 30 March 1796, an eighteen-year-old student at Göttingen made the first entry in a mathematical diary he would keep for the rest of his life. He had discovered how to construct a regular 17-sided polygon with ruler and compasses. The Greeks had known how to construct regular polygons with 3, 4, 5 and 15 sides, and doubles of these, and for two thousand years nobody had added a single new one. Carl Friedrich Gauss later said that this discovery decided him to become a mathematician rather than a philologist, and he is said to have asked for a 17-gon to be carved on his tombstone.

The heart of the construction is the second half of this question: the sum of four particular cosines equals $\frac{\sqrt{17}-1}{4}$, a number involving only a square root. Lengths built from square roots can be constructed with ruler and compasses, and with three more steps of the same kind Gauss could construct $\cos \frac{2\pi}{17}$ itself.

Before you start



Ask yourself Why does the first part, about $\cos \frac{2\pi}{5}$, appear in a question about the 17-gon?

It is a warm-up, using the same idea on a smaller polygon. For the pentagon, the sum $\cos \phi + \cos 2\phi$ satisfies a quadratic with rational coefficients, so it can be found by solving a quadratic. For the 17-gon, the sum z of four cosines will turn out to satisfy a quadratic too. The trick in both cases is to multiply cosines together and use

$\cos A \cos B = \frac{1}{2}(\cos(A+B) + \cos(A-B))$ to turn products back into sums of cosines of the same angles.

Working it through

STAGE 1 The pentagon

? **Ask yourself** Expand $y^2 = (\cos \phi + \cos 2\phi)^2$, turn each product into a sum, and use $5\phi = 2\pi$ to reduce angles like 3ϕ and 4ϕ .

$$y^2 = \cos^2 \phi + 2 \cos \phi \cos 2\phi + \cos^2 2\phi = \frac{1}{2}(1 + \cos 2\phi) + (\cos 3\phi + \cos \phi) + \frac{1}{2}(1 + \cos 4\phi).$$

Since $5\phi = 2\pi$, $\cos 3\phi = \cos(2\pi - 2\phi) = \cos 2\phi$ and $\cos 4\phi = \cos \phi$. So

$$y^2 = 1 + \frac{1}{2}(\cos \phi + \cos 2\phi) + (\cos \phi + \cos 2\phi) = 1 + \frac{3}{2}y,$$

that is, $2y^2 = 3y + 2$. ■ The roots of $2y^2 - 3y - 2 = (2y+1)(y-2) = 0$ are $y = 2$ and $y = -\frac{1}{2}$. But $y = 2$ would need both cosines to equal 1, which they do not, so $y = -\frac{1}{2}$. ■

STAGE 2 The seventeenth roots of unity sum to zero

$\sum_{k=0}^{16} \cos k\theta$ is the real part of $\sum_{k=0}^{16} \omega^k$, where $\omega = e^{i\theta}$ is a primitive 17th root of unity. That geometric series is $\frac{\omega^{17}-1}{\omega-1} = 0$. ■

So $\sum_{k=1}^{16} \cos k\theta = -1$. Since $\cos(17-k)\theta = \cos k\theta$, the terms pair up, and

$$\sum_{k=1}^8 \cos k\theta = -\frac{1}{2}.$$

STAGE 3 Splitting the cosines into two groups

? **Ask yourself** Let $z = \cos \theta + \cos 2\theta + \cos 4\theta + \cos 8\theta$, and let z' be the sum of the other four cosines, $\cos 3\theta + \cos 5\theta + \cos 6\theta + \cos 7\theta$. What is $z + z'$?

By Stage 2, $z + z' = -\frac{1}{2}$. If we can also find zz' , then z and z' are the roots of a quadratic.

Pause and try Compute zz' . Expand it as 16 products $\cos a\theta \cos b\theta$, turn each into a sum of two cosines, and count how often each of $\cos \theta, \cos 2\theta, \dots, \cos 8\theta$ appears. (Reduce angles using $\cos k\theta = \cos(17-k)\theta$ and $\cos(-k\theta) = \cos k\theta$.)

Each product gives $\cos a\theta \cos b\theta = \frac{1}{2}(\cos(a+b)\theta + \cos(a-b)\theta)$, so zz' is $\frac{1}{2}$ times a sum of 32 cosines. Working through the 16 pairs with $a \in \{1, 2, 4, 8\}$ and $b \in \{3, 5, 6, 7\}$: for

example $a = 1, b = 3$ gives $\cos 4\theta$ and $\cos 2\theta$; $a = 8, b = 7$ gives $\cos 15\theta = \cos 2\theta$ and $\cos \theta$. Carrying this out for all sixteen pairs, each of $\cos \theta, \dots, \cos 8\theta$ appears exactly four times. Therefore

$$zz' = \frac{1}{2} \times 4 \times \sum_{k=1}^8 \cos k\theta = 2 \times \left(-\frac{1}{2}\right) = -1.$$

STAGE 4 The quadratic

So z and z' are the roots of $t^2 + \frac{1}{2}t - 1 = 0$, that is, $4t^2 + 2t - 4 = 0$, with roots $t = \frac{-1 \pm \sqrt{17}}{4}$. Which is z ? Numerically $\cos \theta \approx 0.932$, $\cos 2\theta \approx 0.739$, $\cos 4\theta \approx 0.092$ and $\cos 8\theta \approx -0.983$, so $z \approx 0.78 > 0$. Hence

$$z = \frac{-1 + \sqrt{17}}{4} = -\frac{1 - \sqrt{17}}{4}. \quad \blacksquare$$

Looking back

- **Products to sums.** The product formula for cosines turned y^2 and zz' into sums of the same cosines, so the algebra closed up.
- **Sum and product give a quadratic.** Knowing $z + z'$ and zz' is the same as knowing the quadratic with roots z and z' .
- **The right split matters.** Splitting the eight cosines into $\{1, 2, 4, 8\}$ and $\{3, 5, 6, 7\}$ made every index appear equally often in zz' . A random split would not work.

What's really going on

Gauss periods

Why $\{1, 2, 4, 8\}$? Modulo 17, the powers $3^0, 3^1, 3^2, \dots$ of 3 run through all the non-zero residues:

$$1, 3, 9, 10, 13, 5, 15, 11, 16, 14, 8, 7, 4, 12, 2, 6.$$

Taking every other one, $3^0, 3^2, 3^4, \dots$, gives $1, 9, 13, 15, 16, 8, 4, 2$, which are $\pm 1, \pm 2, \pm 4, \pm 8$ modulo 17: exactly the angles in z . The other half gives $\pm 3, \pm 5, \pm 6, \pm 7$, the angles in z' . Gauss called the sums of ω^k over such classes *periods*. Because $17 - 1 = 16 = 2^4$, the residues can be halved four times, and at each stage the periods satisfy quadratic equations whose coefficients involve the periods of the previous stage. Four quadratics, so four square roots, lead from the rational numbers down to $\cos \frac{2\pi}{17}$:

$$\begin{aligned} \cos \frac{2\pi}{17} = \frac{1}{16} & \left(-1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} \right. \\ & \left. + 2\sqrt{17 + 3\sqrt{17} - \sqrt{34 - 2\sqrt{17}}} - 2\sqrt{34 + 2\sqrt{17}} \right). \end{aligned}$$

Every step involves only square roots, and square roots can be constructed with ruler and compasses (using similar triangles or a semicircle), so the 17-gon is constructible.

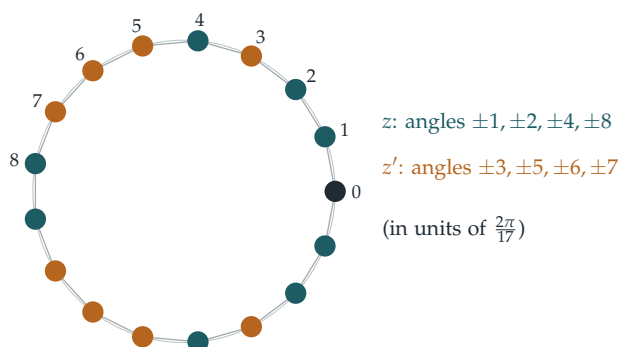


Figure 6.2.1 The vertices of the regular 17-gon, split into Gauss's two periods. The teal vertices are at angles $\pm 1, \pm 2, \pm 4, \pm 8$ (in units of $\frac{2\pi}{17}$), and the sum of their x -coordinates is $2z = \frac{-1+\sqrt{17}}{2}$; the amber vertices give $2z' = \frac{-1-\sqrt{17}}{2}$.

Which polygons can be constructed?

Gauss claimed, and Pierre Wantzel proved in 1837, that a regular n -gon can be constructed with ruler and compasses exactly when n is a power of 2 times a product of distinct *Fermat primes*, primes of the form $2^{2^m} + 1$. The known Fermat primes are 3, 5, 17, 257 and 65537, the Fermat numbers of Section 1.1. (Fermat guessed that all such numbers are prime, but Euler showed that $2^{32} + 1$ is divisible by 641, and no further Fermat primes have ever been found.) So the 257-gon and the 65537-gon are constructible; in the nineteenth century Johann Hermes spent about ten years working out the construction of the 65537-gon, and his manuscript is kept at Göttingen. The 7-gon and 11-gon are not constructible, because 7 and 11 are not Fermat primes ($7 - 1 = 6$ and $11 - 1 = 10$ are not powers of 2), and neither is the 9-gon, because $9 = 3^2$ repeats the Fermat prime 3. Wantzel's proof uses the theory of field degrees from Section 6.3.

The thinking skill

Find a quadratic by computing sum and product

To find a quantity, pair it with a natural partner, compute their sum and their product, and solve the resulting quadratic.

The art is in choosing the partner so that both sum and product can be computed. Here the partner came from the symmetry of the residues modulo 17, and repeating the idea four times gave Gauss his construction.

Try this next

- 1 **The heptagon.** (From STEP II, 1995.) Let $\alpha = e^{2\pi i/7}$. Show that $\eta = \alpha + \alpha^2 + \alpha^4$ and $\eta' = \alpha^3 + \alpha^5 + \alpha^6$ satisfy $\eta + \eta' = -1$ and $\eta\eta' = 2$. Deduce that $\cos \frac{2\pi}{7} + \cos \frac{4\pi}{7} + \cos \frac{8\pi}{7} = -\frac{1}{2}$, and that $\sin \frac{2\pi}{7} + \sin \frac{4\pi}{7} + \sin \frac{8\pi}{7} = \frac{\sqrt{7}}{2}$, explaining how you decide the sign.
- 2 **The pentagon.** Show that $\cos \frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}$, and use it to describe a ruler-and-compasses construction of a regular pentagon.
- 3 **The other period.** Check numerically that $z' = \cos 3\theta + \cos 5\theta + \cos 6\theta + \cos 7\theta = \frac{-1-\sqrt{17}}{4}$.
- 4 **Generators.** Show that the powers of 3 modulo 17 give every non-zero residue, but the powers of 2 do not. Which residues do the powers of 2 give?

Further reading

- **Ian Stewart, *Galois Theory*** (CRC Press). Chapter on ruler-and-compass constructions and the regular 17-gon.
- **Richard Courant & Herbert Robbins, *What Is Mathematics?*** (Oxford University Press). Chapter 3 is a beautiful elementary account of constructions, including the 17-gon.
- **Daniel Kehlmann, *Measuring the World*** (Quercus). A comic novel about Gauss and Alexander von Humboldt, for light relief.

6.3 A Polynomial for $\sqrt{2} + \sqrt{3} + \sqrt{5}$

The question

(i) Show that, if $(x - \sqrt{2})^2 = 3$, then $x^4 - 10x^2 + 1 = 0$.

Deduce that, if $f(x) = x^4 - 10x^2 + 1$, then $f(\sqrt{2} + \sqrt{3}) = 0$.

(ii) Find a polynomial g of degree 8 with integer coefficients such that $g(\sqrt{2} + \sqrt{3} + \sqrt{5}) = 0$. Write your answer in a form without brackets.

(iii) Let a , b and c be the three roots of $t^3 - 3t + 1 = 0$.

Find a polynomial h of degree 6 with integer coefficients such that $h(a + \sqrt{2}) = 0$, $h(b + \sqrt{2}) = 0$ and $h(c + \sqrt{2}) = 0$. Write your answer in a form without brackets.

(iv) Find a polynomial k with integer coefficients such that $k(\sqrt[3]{2} + \sqrt[3]{3}) = 0$. Write your answer in a form without brackets.

The hook

$\sqrt{2}$ is a root of $x^2 - 2$, a polynomial with integer coefficients. So is $\sqrt{3}$, of $x^2 - 3$. Is $\sqrt{2} + \sqrt{3}$ also the root of some polynomial with integer coefficients? What about $\sqrt{2} + \sqrt{3} + \sqrt{5}$, or $\sqrt[3]{2} + \sqrt[3]{3}$?

Numbers that are roots of polynomials with integer coefficients are called *algebraic*, and this question shows, by hands-on elimination, that sums of algebraic numbers are algebraic. The degrees that come out, 4, 8, 6, 9, are not accidents. They measure how “complicated” each number is, and that measurement, the *degree of a field extension*, is what finally settled the ancient Greek problems of doubling the cube and trisecting the angle.

Before you start



Ask yourself To find a polynomial with $\sqrt{2}$ as a root, you square $x = \sqrt{2}$ to get rid of the square root. How might you get rid of two different square roots, one at a time?

Isolate one square root and square, which removes it; then isolate the other and square

again. Each squaring may double the degree. That is the whole strategy of the question: eliminate the radicals one at a time.

Working it through

STAGE 1 Two square roots

If $(x - \sqrt{2})^2 = 3$, then $x^2 - 2\sqrt{2}x + 2 = 3$, so $x^2 - 1 = 2\sqrt{2}x$. Squaring removes the remaining square root: $(x^2 - 1)^2 = 8x^2$, that is,

$$x^4 - 10x^2 + 1 = 0. \quad \blacksquare$$

And $x = \sqrt{2} + \sqrt{3}$ satisfies $(x - \sqrt{2})^2 = (\sqrt{3})^2 = 3$, so $f(\sqrt{2} + \sqrt{3}) = 0$. $\blacksquare$

STAGE 2 Three square roots

? **Ask yourself** Let $x = \sqrt{2} + \sqrt{3} + \sqrt{5}$, so $x - \sqrt{5} = \sqrt{2} + \sqrt{3}$ is a root of f . Write $f(x - \sqrt{5}) = 0$ in the form $P(x) + \sqrt{5}Q(x) = 0$. How do you then eliminate $\sqrt{5}$?

Expand using $(x - \sqrt{5})^2 = x^2 + 5 - 2\sqrt{5}x$ and $(x - \sqrt{5})^4 = (x^2 + 5 - 2\sqrt{5}x)^2 = x^4 + 30x^2 + 25 - 4\sqrt{5}x(x^2 + 5)$:

$$f(x - \sqrt{5}) = (x^4 + 30x^2 + 25) - 4\sqrt{5}x(x^2 + 5) - 10(x^2 + 5) + 20\sqrt{5}x + 1.$$

Collecting, $P(x) = x^4 + 20x^2 - 24$ and $Q(x) = -4x^3 - 20x + 20x = -4x^3$. So $x^4 + 20x^2 - 24 = 4\sqrt{5}x^3$. Squaring,

$$(x^4 + 20x^2 - 24)^2 = 80x^6,$$

and expanding gives

$$g(x) = x^8 - 40x^6 + 352x^4 - 960x^2 + 576. \quad \blacksquare$$

(Check: $(x^4 + 20x^2 - 24)^2 = x^8 + 40x^6 + 352x^4 - 960x^2 + 576$, using $400 - 48 = 352$, and subtracting $80x^6$ gives g .)

STAGE 3 A cubic and a square root

? **Ask yourself** If $x = a + \sqrt{2}$, then $x - \sqrt{2}$ is a root of $t^3 - 3t + 1$. Substitute, separate the terms with and without $\sqrt{2}$, and square.

$(x - \sqrt{2})^3 = x^3 - 3\sqrt{2}x^2 + 6x - 2\sqrt{2}$, so

$$(x - \sqrt{2})^3 - 3(x - \sqrt{2}) + 1 = (x^3 + 3x + 1) - \sqrt{2}(3x^2 - 1) = 0.$$

Hence $(x^3 + 3x + 1)^2 = 2(3x^2 - 1)^2$. Expanding, $x^6 + 6x^4 + 2x^3 + 9x^2 + 6x + 1 = 18x^4 - 12x^2 + 2$, so

$$h(x) = x^6 - 12x^4 + 2x^3 + 21x^2 + 6x - 1. \quad \blacksquare$$

The same argument works for $b + \sqrt{2}$ and $c + \sqrt{2}$, since only the fact that a is a root of $t^3 - 3t + 1$ was used. (The other three roots of h are $a - \sqrt{2}$, $b - \sqrt{2}$ and $c - \sqrt{2}$, which came in when we squared.)

STAGE 4 Two cube roots

 **Pause and try** Let $x = \sqrt[3]{2} + \sqrt[3]{3}$. Cube both sides. Can you write the result in terms of x and a single cube root?

Using $(u + v)^3 = u^3 + v^3 + 3uv(u + v)$ with $u = \sqrt[3]{2}$, $v = \sqrt[3]{3}$:

$$x^3 = 2 + 3 + 3\sqrt[3]{6}x, \quad \text{so} \quad x^3 - 5 = 3\sqrt[3]{6}x.$$

Cubing again removes the last cube root: $(x^3 - 5)^3 = 27 \cdot 6x^3 = 162x^3$. Expanding $(x^3 - 5)^3 = x^9 - 15x^6 + 75x^3 - 125$,

$$k(x) = x^9 - 15x^6 - 87x^3 - 125. \quad \blacksquare$$

Looking back

- **Eliminate one radical at a time.** Isolate a radical, raise to the appropriate power, and repeat.
- **Separate rational and irrational parts.** Writing an expression as $P + \sqrt{5}Q$ and then using $P^2 = 5Q^2$ is the key move.
- **Squaring adds extra roots.** Each squaring brings in “conjugate” roots, such as $a - \sqrt{2}$ alongside $a + \sqrt{2}$. The final polynomial has all of them as roots.

What’s really going on

Fields and their degrees

The set $\mathbb{Q}(\sqrt{2})$ of all numbers $p + q\sqrt{2}$ with p, q rational is closed under addition, subtraction, multiplication and division (by non-zero numbers): it is a *field*. It is a two-dimensional “vector space” over $\mathbb{Q}$, with basis $1, \sqrt{2}$, so we say it has *degree 2* over $\mathbb{Q}$. Adjoining $\sqrt{3}$ gives $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, with basis $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ and degree 4; adjoining $\sqrt{5}$ as well gives degree 8 (Figure 6.3.1). The *tower law* says degrees multiply up a tower of fields. The degree of the field generated by a number equals the degree of its *minimal polynomial*, the lowest-degree polynomial with rational coefficients that it

satisfies. One can show that the polynomials f and g in this question are minimal, so $\sqrt{2} + \sqrt{3}$ has degree 4 and $\sqrt{2} + \sqrt{3} + \sqrt{5}$ has degree 8: in fact the single number $\sqrt{2} + \sqrt{3} + \sqrt{5}$ generates the whole degree-8 field.

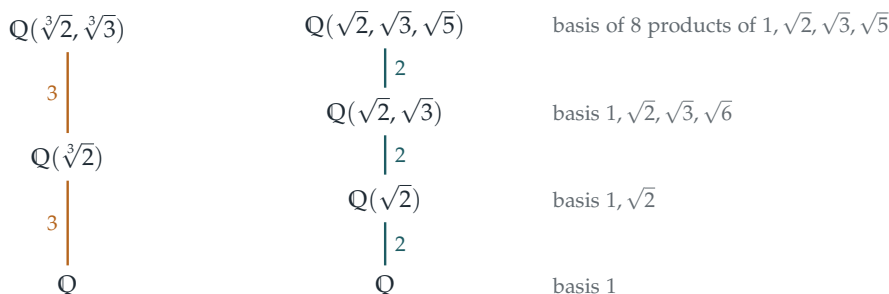


Figure 6.3.1 Towers of fields. Right: each square root doubles the degree, giving $2 \times 2 \times 2 = 8$, the degree of g . Left: each cube root triples it, giving $3 \times 3 = 9$, the degree of k . Degrees multiply up a tower; this *tower law* is the key to the impossibility proofs for ruler-and-compass constructions.

Doubling the cube

Ruler-and-compass constructions can only produce lengths obtained from rational numbers by repeatedly solving linear and quadratic equations. Each such step gives a field extension of degree 1 or 2, so by the tower law every constructible length lies in a field of degree a power of 2, and has degree a power of 2 itself. But $\sqrt[3]{2}$ has minimal polynomial $x^3 - 2$, of degree 3. Since 3 is not a power of 2, $\sqrt[3]{2}$ is not constructible, and the Delian problem of doubling the volume of a cube with ruler and compasses, posed around 430 BC, is impossible. Pierre Wantzel proved this in 1837, together with the impossibility of trisecting a general angle and the classification of constructible polygons (Section 6.2). Two thousand years of attempts were ended by an argument about degrees.

Algebraic numbers form a field

The method of this question shows, in examples, that sums of algebraic numbers are algebraic. In general this is proved with linear algebra: if α and β have degrees m and n , then all the numbers $\alpha^i \beta^j$ (with $i < m$, $j < n$) span a space of dimension at most mn , and the $mn + 1$ powers $1, \alpha + \beta, (\alpha + \beta)^2, \dots$ must be linearly dependent, which gives a polynomial of degree at most mn . Our degrees $4 = 2 \times 2$, $6 = 3 \times 2$ and $9 = 3 \times 3$ fit this pattern exactly. Numbers that are *not* algebraic are called *transcendental*; e and π are examples (Hermite, 1873, and Lindemann, 1882), and Lindemann's theorem that π is transcendental proved that squaring the circle is impossible too.

The thinking skill

Eliminate, then measure

To find a polynomial satisfied by an expression in radicals, remove the radicals one at a time by isolating and raising to a power.

Then ask what the degree tells you. The degree of the minimal polynomial is an invariant of the number, and comparing degrees is one of the most powerful ways to prove that something is impossible.

Try this next

- 1 **A primitive element.** Let $y = \sqrt{2} + \sqrt{3}$. Show that $y^3 = 11\sqrt{2} + 9\sqrt{3}$, and deduce that $\sqrt{2} = \frac{1}{2}(y^3 - 9y)$. Hence every number $p + q\sqrt{2} + r\sqrt{3} + s\sqrt{6}$ can be written as a polynomial in y .
- 2 **Mixed roots.** Find a polynomial of degree 6 with integer coefficients that has $\sqrt{2} + \sqrt[3]{3}$ as a root.
- 3 **Irreducibility.** Show that $x^4 - 10x^2 + 1$ has no rational roots and cannot be written as a product of two quadratics with integer coefficients. (*Hint:* try $(x^2 + ax + b)(x^2 - ax + c)$.) Deduce that $\sqrt{2} + \sqrt{3}$ has degree 4.
- 4 **Trisection.** Show that $\cos 3\theta = 4\cos^3 \theta - 3\cos \theta$, and deduce that $\cos 20^\circ$ is a root of $8x^3 - 6x - 1 = 0$. Assuming this cubic has no rational roots (check this), explain why a 60° angle cannot be trisected with ruler and compasses.

Further reading

- **Ian Stewart, *Galois Theory*** (CRC Press). Field extensions, the tower law and the classical construction problems, written for undergraduates.
- **Richard Courant & Herbert Robbins, *What Is Mathematics?*** (Oxford University Press). Chapter 3 proves the impossibility of the Greek problems with no prerequisites.
- **Robin Hartshorne, *Geometry: Euclid and Beyond*** (Springer). Ruler-and-compass constructions treated with modern algebra.

6.4 Solving the Quartic

The question

Let α , β , γ and δ be the roots of the quartic equation

$$x^4 + px^3 + qx^2 + rx + s = 0.$$

You are given that, for any such equation, $\alpha\beta + \gamma\delta$, $\alpha\gamma + \beta\delta$ and $\alpha\delta + \beta\gamma$ satisfy a cubic equation of the form

$$y^3 + Ay^2 + (pr - 4s)y + (4qs - p^2s - r^2) = 0.$$

Determine A .

Now consider the quartic equation given by $p = 0$, $q = 3$, $r = -6$ and $s = 10$.

- (i) Find the value of $\alpha\beta + \gamma\delta$, given that it is the largest root of the corresponding cubic equation.
- (ii) Hence, using the values of q and s , find the value of $(\alpha + \beta)(\gamma + \delta)$ and the value of $\alpha\beta$ given that $\alpha\beta > \gamma\delta$.
- (iii) Using these results, and the values of p and r , solve the quartic equation.

The hook

Every student learns the formula for solving quadratic equations. Formulas for cubics and quartics exist too; they were found in sixteenth-century Italy amid secret contests and bitter disputes over priority. Gerolamo Cardano published both in his *Ars Magna* of 1545, the cubic from Niccolò Tartaglia (who had sworn him to secrecy) and the quartic from his own student Lodovico Ferrari. Then, for nearly three hundred years, mathematicians searched in vain for a formula for the quintic.

This question shows the trick that solves the quartic, and it is a trick with a deep meaning. Out of the four roots, form three new quantities by pairing the roots up in the three possible ways. Those three quantities satisfy a cubic, which you can solve, and from them the original roots follow. The question *why* this works, why four roots can be reduced to three quantities but five roots cannot be similarly reduced, is what led Galois to invent group theory.

Before you start



Ask yourself How many ways are there to split four objects $\alpha, \beta, \gamma, \delta$ into two pairs?

Three: $\{\alpha\beta, \gamma\delta\}$, $\{\alpha\gamma, \beta\delta\}$ and $\{\alpha\delta, \beta\gamma\}$ (Figure 6.4.1). The three quantities in the question, $\alpha\beta + \gamma\delta$ and so on, correspond to these three pairings. Any rearrangement of the four roots just shuffles the three pairings among themselves, so any symmetric expression in the three quantities is a symmetric expression in the four roots, and can be written in terms of p, q, r, s . That is why they satisfy a cubic whose coefficients involve only p, q, r, s .

Working it through

STAGE 1 The missing coefficient



Ask yourself In a cubic $y^3 + Ay^2 + By + C = 0$, what is the sum of the roots in terms of A ? And what is $(\alpha\beta + \gamma\delta) + (\alpha\gamma + \beta\delta) + (\alpha\delta + \beta\gamma)$ in terms of the quartic's coefficients?

The sum of the roots of the cubic is $-A$. The sum of the three quantities is the sum of all six products of pairs of $\alpha, \beta, \gamma, \delta$, which for the quartic $x^4 + px^3 + qx^2 + rx + s$ is q . So $A = -q$.

STAGE 2 The specific cubic

With $p = 0, q = 3, r = -6, s = 10$, the cubic is

$$y^3 - 3y^2 + (0 - 40)y + (120 - 0 - 36) = y^3 - 3y^2 - 40y + 84 = 0.$$



Ask yourself Try small divisors of 84.

$y = 2$ gives $8 - 12 - 80 + 84 = 0$. Dividing, $y^3 - 3y^2 - 40y + 84 = (y - 2)(y^2 - y - 42) = (y - 2)(y - 7)(y + 6)$. The largest root is 7, so $\alpha\beta + \gamma\delta = 7$. ■

STAGE 3 Pairs of roots

$(\alpha + \beta)(\gamma + \delta) = \alpha\gamma + \alpha\delta + \beta\gamma + \beta\delta$, which is the sum of all six pair-products minus $\alpha\beta + \gamma\delta$, that is, $q - 7 = 3 - 7 = -4$.

Also $\alpha\beta \cdot \gamma\delta = \alpha\beta\gamma\delta = s = 10$. So $\alpha\beta$ and $\gamma\delta$ have sum 7 and product 10: they are the roots of $t^2 - 7t + 10 = (t - 2)(t - 5)$. With $\alpha\beta > \gamma\delta$, $\alpha\beta = 5$ and $\gamma\delta = 2$. ■

STAGE 4 Solving the quartic



Pause and try Find $\alpha + \beta$ and $\gamma + \delta$. You know their sum (from p) and their product (from Stage 3); to decide which is which, use r .

$\alpha + \beta + \gamma + \delta = -p = 0$ and $(\alpha + \beta)(\gamma + \delta) = -4$, so $\{\alpha + \beta, \gamma + \delta\} = \{2, -2\}$. The coefficient r gives the sum of the four triple products: $\alpha\beta\gamma + \alpha\beta\delta + \alpha\gamma\delta + \beta\gamma\delta = -r = 6$. This can be written as $\alpha\beta(\gamma + \delta) + \gamma\delta(\alpha + \beta) = 5(\gamma + \delta) + 2(\alpha + \beta)$. If $\alpha + \beta = 2$ this is $-10 + 4 = -6$; if $\alpha + \beta = -2$ it is $10 - 4 = 6$. So $\alpha + \beta = -2$ and $\gamma + \delta = 2$.

Now α, β are the roots of $t^2 + 2t + 5 = 0$, giving $-1 \pm 2i$; and γ, δ are the roots of $t^2 - 2t + 2 = 0$, giving $1 \pm i$. The four roots are

$$x = -1 \pm 2i, \quad x = 1 \pm i. \quad \blacksquare$$

Check: $(x^2 + 2x + 5)(x^2 - 2x + 2) = x^4 + 3x^2 - 6x + 10$.

Looking back

- **Symmetric functions.** Every step used the relations between roots and coefficients: sums and products of roots.
- **Reduce the degree.** A quartic problem became a cubic (the *resolvent cubic*), and then two quadratics.
- **Use every piece of information.** p and q gave the sums and products; s split $\alpha\beta$ from $\gamma\delta$; r decided which pair went with which sum.

What's really going on

Why the trick works: permutations of roots

In 1770 Joseph-Louis Lagrange asked why the methods for cubics and quartics work, and found the answer in the *permutations* of the roots. There are 24 ways to rearrange four roots. Under all of them, the three quantities $y_1 = \alpha\beta + \gamma\delta$, $y_2 = \alpha\gamma + \beta\delta$, $y_3 = \alpha\delta + \beta\gamma$ are merely shuffled among themselves: each rearrangement of the four roots induces a rearrangement of the three pairings (Figure 6.4.1). So they satisfy a polynomial of degree 3 with coefficients symmetric in the roots. In modern language, there is a group homomorphism from the symmetric group S_4 (all 24 permutations) onto S_3 (the 6 permutations of the pairings), and that is why a quartic can be reduced to a cubic.

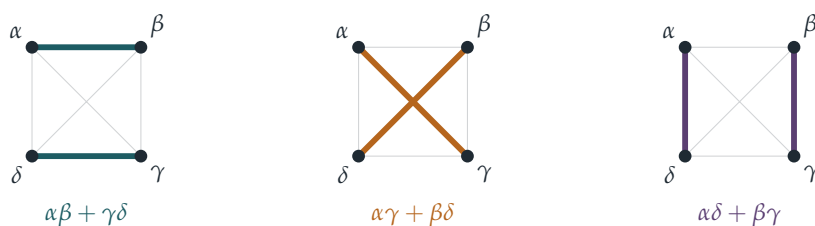


Figure 6.4.1 The three ways to split four roots into two pairs, and the corresponding roots of the resolvent cubic. Any rearrangement of $\alpha, \beta, \gamma, \delta$ permutes these three pictures among themselves, which is why the three quantities satisfy a cubic with coefficients expressible in terms of the quartic's.

The quintic

For five roots, Lagrange looked for a similar quantity taking only a few values under permutations of the roots, and could not find one that helped. In 1799 Paolo Ruffini, and in 1824 Niels Henrik Abel (completing the proof), showed that there is no general formula for the roots of a quintic using only arithmetic and radicals. A few years later Évariste Galois explained exactly which polynomials can be solved by radicals: those whose *group of symmetries of the roots* can be broken down in a particular way into simpler pieces. The symmetric group S_4 can be (via S_3 and smaller groups), but S_5 cannot, because it contains a “simple” piece of size 60 that cannot be broken down further. Galois wrote up his theory in a hurry the night before he was killed in a duel in 1832, at the age of twenty. It is now one of the jewels of a mathematics degree, usually met in the third year.

Solving cubics

The cubic itself can be solved by a similar trick: write a depressed cubic $x^3 - 3px + q$ as a combination of two perfect cubes, which reduces it to a quadratic. That is the method of the 2011 question in the exercises, and it is equivalent to Cardano's formula. When the cubic has three real roots, the formula still requires complex numbers along the way, the phenomenon that forced mathematicians to take complex numbers seriously (see the introduction to Chapter 5).

The thinking skill

Symmetry reduces complexity

If some quantities built from the roots are only permuted among themselves when the roots are permuted, they satisfy a polynomial of lower degree whose coefficients you can compute.

This is the seed of Galois theory: understand an equation by understanding the symmetries of its roots.

Try this next

- 1 **Check the cubic.** Verify that $y_1y_2 + y_2y_3 + y_3y_1 = pr - 4s$, where $y_1 = \alpha\beta + \gamma\delta$ etc. (Expand and use $\sum \alpha = -p$, $\sum \alpha\beta\gamma = -r$, $\alpha\beta\gamma\delta = s$.)
- 2 **The cubic as a sum of cubes.** (From STEP III, 2011.) Show that if $q^2 \neq 4p^3$ then $x^3 - 3px + q = a(x - \alpha)^3 + b(x - \beta)^3$, where α, β are the roots of $pt^2 - qt + p^2 = 0$. Hence show that one root of $x^3 - 24x + 48 = 0$ is $x = \frac{2(2 - 2^{1/3})}{1 - 2^{1/3}}$, and check this numerically.
- 3 **Another quartic.** Use the method of this section to solve $x^4 - 4x^3 + 7x^2 - 6x + 2 = 0$. (One root of the resolvent cubic is a small integer.)
- 4 **Counting permutations.** Of the 24 permutations of $\alpha, \beta, \gamma, \delta$, find all those that send the pairing $\{\alpha\beta, \gamma\delta\}$ to itself. Show that there are 8 of them. (They form a subgroup of S_4 .)

Further reading

- **Mario Livio, *The Equation That Couldn't Be Solved*** (Simon & Schuster). The story of the quintic and the birth of group theory, for a general reader.
- **Ian Stewart, *Why Beauty Is Truth: A History of Symmetry*** (Basic Books). From Babylonian equations to Galois and beyond.
- **Ian Stewart, *Galois Theory*** (CRC Press). The undergraduate text, with a historical introduction covering exactly this material.

6.5 Differentiation from Axioms

The question

$\diamond$ is an operation which takes polynomials in x to polynomials in x ; that is, given a polynomial $h(x)$ there is another polynomial called $\diamond h(x)$. It is given that, if $f(x)$ and $g(x)$ are any two polynomials in x , the following are always true:

- (i) $\diamond(f(x)g(x)) = g(x)\diamond f(x) + f(x)\diamond g(x)$,
- (ii) $\diamond(f(x) + g(x)) = \diamond f(x) + \diamond g(x)$,
- (iii) $\diamond x = 1$,
- (iv) if λ is a constant then $\diamond(\lambda f(x)) = \lambda \diamond f(x)$.

Show that, if $f(x)$ is a constant (i.e., a polynomial of degree zero), then $\diamond f(x) = 0$.

Calculate $\diamond x^2$ and $\diamond x^3$. Prove that $\diamond h(x) = \frac{d}{dx}(h(x))$ for any polynomial $h(x)$.

The hook

What is differentiation? At school it is defined by a limit, $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$, and then you learn rules: the derivative of a sum is the sum of the derivatives, the product rule, and so on. This question turns that upside down. Forget limits entirely. Suppose all you know about an operation is that it obeys the sum rule, the product rule and the rule for constants, and that it sends x to 1. Then, the question shows, it *must* be differentiation, at least on polynomials.

This is the algebraist's view of calculus: an operation is characterised by the rules it obeys, not by a formula. Operations obeying the product rule are called *derivations*, and they turn up throughout algebra, geometry and physics. Surprisingly, the same idea is at work inside the software that trains modern machine-learning systems.

Before you start



Ask yourself Rule (i) is the product rule. How could the product rule tell you anything about a constant? (Think of a constant as a product.)

The number 1 equals 1×1 . Applying the product rule to 1×1 relates $\diamond 1$ to itself, and that will force it to be zero. In general, the trick with axioms like these is to apply them

to cleverly chosen simple inputs.

Working it through

STAGE 1 Constants



Ask yourself Apply rule (i) with $f = g = 1$. Then use rule (iv) for any other constant.

$\diamond 1 = \diamond(1 \cdot 1) = 1 \cdot \diamond 1 + 1 \cdot \diamond 1 = 2\diamond 1$, so $\diamond 1 = 0$. For any constant λ , rule (iv) gives $\diamond \lambda = \diamond(\lambda \cdot 1) = \lambda \diamond 1 = 0$. ■

STAGE 2 Powers

$\diamond x^2 = \diamond(x \cdot x) = x \diamond x + x \diamond x = 2x$, using rule (iii). Then

$$\diamond x^3 = \diamond(x^2 \cdot x) = x \diamond x^2 + x^2 \diamond x = x \cdot 2x + x^2 = 3x^2.$$



Pause and try Prove by induction that $\diamond x^n = nx^{n-1}$ for every positive integer n .

The case $n = 1$ is rule (iii). If $\diamond x^n = nx^{n-1}$, then

$$\diamond x^{n+1} = \diamond(x^n \cdot x) = x \diamond x^n + x^n \diamond x = x \cdot nx^{n-1} + x^n = (n+1)x^n. \quad \blacksquare$$

STAGE 3 All polynomials

A general polynomial is $h(x) = a_0 + a_1x + \cdots + a_mx^m$. By rule (ii) (applied repeatedly) and rule (iv),

$$\begin{aligned} \diamond h(x) &= \diamond a_0 + a_1 \diamond x + a_2 \diamond x^2 + \cdots + a_m \diamond x^m \\ &= a_1 + 2a_2x + \cdots + ma_mx^{m-1} = h'(x). \quad \blacksquare \end{aligned}$$

Looking back

- **Axioms determine the operation.** Four rules, with no mention of limits, pinned down $\diamond$ completely on polynomials.
- **Generators.** Every polynomial is built from constants and x using addition and multiplication. Knowing what $\diamond$ does to the building blocks, and how it interacts with the building operations, determines it everywhere.
- **Induction along the construction.** The proof followed the way polynomials are built: powers by repeated multiplication, then sums.

What's really going on

Derivations

An operation D on an algebraic system that is additive and satisfies the product rule $D(fg) = fD(g) + gD(f)$ is called a *derivation*. The question shows that on polynomials, a derivation that kills constants is completely determined by $D(x)$: in fact $D(h) = h'(x)D(x)$, so the only freedom is the choice of the polynomial $D(x)$. Rule (iii) fixes $D(x) = 1$. Derivations appear across mathematics: in geometry, a *vector field* on a surface can be defined as a derivation on the functions on that surface (it differentiates functions in the direction of the field); in algebra, derivations of a structure measure its “infinitesimal symmetries”; in physics, the commutator $[A, B] = AB - BA$ of quantum mechanics is a derivation in each argument.

Calculus without limits

The algebraic derivative works in settings where limits make no sense at all. Over a finite field, such as arithmetic modulo a prime p , there are no limits, but the formal derivative of a polynomial is defined by the same rules and still detects repeated roots: a polynomial f has a repeated root exactly when f and f' have a common factor. That is how one shows, for example, that $x^n - 1$ has no repeated roots when $p \nmid n$, a fact needed to define the cyclotomic polynomials of Section 6.1 over finite fields. The same idea is used in coding theory and cryptography.

Automatic differentiation

Here is a modern application. Invent a number ε with $\varepsilon^2 = 0$ but $\varepsilon \neq 0$, and do arithmetic with *dual numbers* $a + b\varepsilon$. Then $(a + b\varepsilon)(c + d\varepsilon) = ac + (ad + bc)\varepsilon$: the ε -part follows exactly the product rule. So for any polynomial,

$$f(x + \varepsilon) = f(x) + f'(x)\varepsilon,$$

and evaluating f at the dual number $x + \varepsilon$ computes the value and the derivative at the same time, exactly, with no limits and no rounding error from small differences (Figure 6.5.1). This is called *forward-mode automatic differentiation*. Its cousin, reverse mode, is the algorithm (*backpropagation*) used to train neural networks with billions of parameters: every time such a system learns, it is applying the product and chain rules, as derivations, step by step.

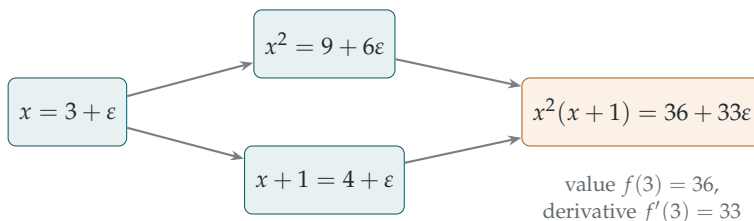


Figure 6.5.1 Automatic differentiation with dual numbers ($\epsilon^2 = 0$) for $f(x) = x^2(x + 1)$ at $x = 3$. Each step follows the rules of the question: $(3 + \epsilon)^2 = 9 + 6\epsilon$, and $(9 + 6\epsilon)(4 + \epsilon) = 36 + (9 + 24)\epsilon$. The coefficient of ϵ is $f'(3) = 3 \cdot 9 + 2 \cdot 3 = 33$, computed exactly.

The thinking skill

Characterise by properties

Describe an operation by the rules it obeys, then show that the rules determine it uniquely.

This is the axiomatic method, central to university algebra. Its power is that anything you prove from the axioms applies to every operation that satisfies them, whether it is differentiation of polynomials, of functions on a surface, or something not yet invented.

Try this next

- 1 A different derivation.** Suppose $\diamond$ satisfies rules (i), (ii) and (iv), but $\diamond x = x$ instead of 1. Show that $\diamond h(x) = x h'(x)$. What does $\diamond$ do to x^n ?
- 2 The quotient rule.** Extend $\diamond$ to rational functions f/g , assuming rule (i) still holds. By applying rule (i) to $g \cdot \frac{f}{g} = f$, derive the quotient rule.
- 3 Repeated roots.** Show that if $(x - a)^2$ divides a polynomial f , then $x - a$ divides f' . Use this to show that $x^n - 1$ has no repeated roots.
- 4 Dual numbers.** Use dual numbers to compute $f(2)$ and $f'(2)$ for $f(x) = (x^2 + 1)^3$, and check with the chain rule.

Further reading

- **Peter Cameron, *Introduction to Algebra*** (Oxford University Press). Rings, polynomials and fields at first-year level, with the formal derivative.
- **Andreas Griewank & Andrea Walther, *Evaluating Derivatives*** (SIAM). The standard

reference on automatic differentiation. Advanced.

- **Michael Nielsen, *Neural Networks and Deep Learning*** (free online). Chapter 2 explains backpropagation clearly.

6.6 Unions of Subgroups

The question

Give a careful argument to show that, if G_1 and G_2 are subgroups of a finite group G such that every element of G is either in G_1 or in G_2 , then either $G_1 = G$ or $G_2 = G$.

Give an example of a group H which has three subgroups H_1, H_2 and H_3 such that every element of H is either in H_1, H_2 or H_3 and $H_1 \neq H, H_2 \neq H, H_3 \neq H$.

The hook

This is one of the shortest questions in the book, and one of the most characteristic of university mathematics. It is about *groups*, the mathematical language of symmetry, and it asks something that sounds like a puzzle: can a group be covered by two smaller groups sitting inside it? The answer is no, and the proof is three lines long. But three smaller groups can do it, which is a surprise.

Group theory was on the STEP syllabus until 2019 and is now a staple of the first year at university. This question needs only the definitions, which are given below, and it is a perfect first taste of how group theorists think: a short, clever argument that uses nothing but the axioms.

Before you start

A *group* is a set G with an operation (written as multiplication, ab) such that:

- ▶ the operation is *associative*: $(ab)c = a(bc)$;
- ▶ there is an *identity* element e with $ea = ae = a$ for every a ;
- ▶ every a has an *inverse* a^{-1} with $aa^{-1} = a^{-1}a = e$.

Examples: the whole numbers under addition; the non-zero real numbers under multiplication; the rotations of a square under composition. A *subgroup* is a subset K of G that is a group in its own right under the same operation: it contains e , and if $a, b \in K$ then $ab \in K$ and $a^{-1} \in K$.



Ask yourself If K is a subgroup and $a \in K$, but $b \notin K$, can ab be in K ?

No. If $ab \in K$, then $b = a^{-1}(ab)$ would be a product of two elements of K , and so in K .

This tiny observation is the whole proof.

Working it through

STAGE 1 Two subgroups cannot cover

Suppose, for a contradiction, that $G = G_1 \cup G_2$ but neither G_1 nor G_2 is all of G .



Ask yourself Since $G_1 \neq G$, there is an element not in G_1 . Where must it be? Similarly for G_2 . Now multiply the two elements together.

Choose $a \in G$ with $a \notin G_1$; since every element is in G_1 or G_2 , $a \in G_2$. Choose $b \in G$ with $b \notin G_2$; then $b \in G_1$. Consider the product ab . It lies in G_1 or in G_2 .

- If $ab \in G_1$: since $b \in G_1$, $a = (ab)b^{-1} \in G_1$. Contradiction.
- If $ab \in G_2$: since $a \in G_2$, $b = a^{-1}(ab) \in G_2$. Contradiction.

Either way we reach a contradiction, so $G_1 = G$ or $G_2 = G$. ■

Notice that we never used the fact that G is finite. The result holds for every group.

STAGE 2 Three subgroups can



Pause and try Find a group that is the union of three subgroups, none of them the whole group. (*Hint*: try a small group in which every element is its own inverse.)

Take the *Klein four-group* $V = \{e, a, b, c\}$, with the rules $a^2 = b^2 = c^2 = e$ and $ab = ba = c$, $bc = cb = a$, $ca = ac = b$. (Concretely: the four symmetries of a non-square rectangle, namely doing nothing, rotating by 180° , and reflecting in either axis. Or the pairs (x, y) of bits 0, 1 under addition modulo 2.) Its three subgroups

$$H_1 = \{e, a\}, \quad H_2 = \{e, b\}, \quad H_3 = \{e, c\}$$

each have two elements, none is the whole group, and together they contain every element of V (Figure 6.6.1). ■

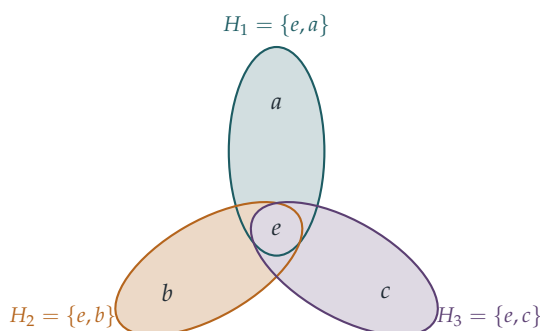


Figure 6.6.1 The Klein four-group $\{e, a, b, c\}$ is the union of its three subgroups of order 2, which overlap only in the identity e . By the first part of the question, no group can be the union of just two proper subgroups.

Looking back

- **Pick witnesses and combine them.** The proof chose an element outside each subgroup and multiplied them. This “take one from each and combine” move is extremely common in algebra.
- **Closure is powerful.** The only property of subgroups used was that they are closed under multiplication and inverses.
- **Check the hypotheses you need.** Finiteness was given but not used. Noticing that a hypothesis is unnecessary is part of understanding a proof.

What’s really going on

The language of symmetry

Groups arose in the nineteenth century from several directions at once: Galois’s study of permutations of the roots of polynomials (Section 6.4), Cauchy’s work on permutations, and the study of symmetries in geometry and crystallography. The abstract definition above, in which a group is anything satisfying the three axioms, was settled by the end of the century. Its power is generality: a theorem proved from the axioms, like the one in this question, holds simultaneously for groups of numbers, of matrices, of symmetries and of permutations. Today groups are everywhere: in particle physics (the standard model is built on symmetry groups), in chemistry (molecular symmetry determines spectra), in cryptography (elliptic curve groups, Section 1.9) and in the solution of Rubik’s cube.

How many subgroups does it take?

The question suggests a natural problem: what is the smallest number of proper subgroups whose union is the whole group? It is never 2, by the question. In 1926 Gaetano Scorza showed that a group is the union of three proper subgroups exactly when it has the Klein four-group as a “quotient”. In 1994 J. H. E. Cohn studied the general problem and conjectured several results, and in 1997 M. J. Tomkinson proved that there is no group for which this smallest number is seven, although every number from 3 to 6 occurs. Questions like this, simple to state and surprisingly subtle, are typical of group theory.

Vector spaces

A similar argument shows that a vector space over the real numbers (such as the plane, or three-dimensional space) can never be the union of finitely many proper subspaces: a plane is not the union of finitely many lines through the origin. But over a finite field it can be: the plane over the field with two elements, which is just the Klein four-group again, is the union of its three lines.

The thinking skill

Argue from the axioms

In abstract algebra, you often know nothing about the objects except the axioms they satisfy. The proof must use only those axioms.

This forces arguments to be short and clean, like the one here, and it makes them universally applicable. It takes some getting used to, and it is one of the most distinctive experiences of a first-year algebra course.

Try this next

- 1 Check the group.** Write out the multiplication table of the Klein four-group and check the group axioms. Show that it is not the same as the group $\{1, i, -1, -i\}$ under multiplication. (*Hint:* which elements are their own inverses?)
- 2 Cyclic groups.** A group is *cyclic* if every element is a power of a single element g . Show that a cyclic group is never the union of proper subgroups, however many. (*Hint:* which subgroup contains g ?)
- 3 The whole numbers.** Show directly that $(\mathbb{Z}, +)$ is not the union of the subgroups $2\mathbb{Z}$ (even numbers) and $3\mathbb{Z}$ (multiples of 3). Is it the union of any two proper subgroups?

- 4 **Lines in the plane.** **HARDER** Show that the plane $\mathbb{R}^2$ is not the union of finitely many lines through the origin. (*Hint*: each line contains at most one point of the parabola $y = x^2$ other than the origin.)

Further reading

- ▶ **M. A. Armstrong, *Groups and Symmetry*** (Springer). A beautifully illustrated first course, driven by geometric examples.
- ▶ **Nathan Carter, *Visual Group Theory*** (Mathematical Association of America). Groups through pictures, ideal before university.
- ▶ **Peter Neumann, Gabrielle Stoy & Edward Thompson, *Groups and Geometry*** (Oxford University Press). A classic Oxford first-year text.

6.7 The Quaternion Group

The question

The set S consists of ordered pairs of complex numbers (z_1, z_2) and a binary operation $\circ$ on S is defined by

$$(z_1, z_2) \circ (w_1, w_2) = (z_1 w_1 - z_2 w_2^*, z_1 w_2 + z_2 w_1^*).$$

Show that the operation $\circ$ is associative and determine whether it is commutative. Evaluate $(z, 0) \circ (w, 0)$, $(z, 0) \circ (0, w)$, $(0, z) \circ (w, 0)$ and $(0, z) \circ (0, w)$.

The set S_1 is the subset of S consisting of $A, B, \dots, H$, where $A = (1, 0)$, $B = (0, 1)$, $C = (i, 0)$, $D = (0, i)$, $E = (-1, 0)$, $F = (0, -1)$, $G = (-i, 0)$ and $H = (0, -i)$. Show that S_1 is closed under $\circ$ and that it has an identity element. Determine the inverse and order of each element of S_1 . Show that S_1 is a group under $\circ$. [You are not required to compute the multiplication table in full.]

Show that $\{A, B, E, F\}$ is a subgroup of S_1 and determine whether it is isomorphic to the group generated by the 2×2 matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ under matrix multiplication.

The hook

On 16 October 1843, walking along the Royal Canal in Dublin with his wife, William Rowan Hamilton had a flash of inspiration that he had been chasing for years. He had been trying to multiply triples of numbers, the way complex numbers multiply pairs, and failing. The answer, he suddenly saw, needed *four* numbers, and a multiplication in which the order matters: $ij \neq ji$. He carved the rules into the stone of Brougham Bridge:

$$i^2 = j^2 = k^2 = ijk = -1.$$

These are the *quaternions*. The question builds them from pairs of complex numbers, and finds inside them a small group of eight elements with a twisted multiplication. Quaternions are now used to describe rotations in three dimensions in every computer game, flight simulator and spacecraft guidance system.

Before you start

For the group definitions, see Section 6.6. Two more words: the *order* of an element g of a group is the smallest positive n with $g^n = e$; and two groups are *isomorphic* if there is a one-to-one correspondence between them that respects the operations, so that they

are “the same group with different labels”.

Ask yourself Think of the pair (z_1, z_2) as standing for a “number” $z_1 + z_2j$, where j is a new symbol with $j^2 = -1$ and $jw = w^*j$ for complex w . Multiply $(z_1 + z_2j)(w_1 + w_2j)$ using these rules. Do you get the formula for $\circ$?

$(z_1 + z_2j)(w_1 + w_2j) = z_1w_1 + z_1w_2j + z_2jw_1 + z_2jw_2j = z_1w_1 + z_1w_2j + z_2w_1^*j + z_2w_2^*j^2$, which is $(z_1w_1 - z_2w_2^*) + (z_1w_2 + z_2w_1^*)j$. Yes: the operation $\circ$ is quaternion multiplication in disguise.

Working it through

STAGE 1 Associativity via matrices

Checking $(a \circ b) \circ c = a \circ (b \circ c)$ directly is a long calculation. There is a slicker way.

Ask yourself Associate to (z_1, z_2) the complex matrix $M = \begin{pmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{pmatrix}$. What is the top row of the matrix product $M(z_1, z_2) M(w_1, w_2)$?

The top row of the product is $(z_1w_1 - z_2w_2^*, z_1w_2 + z_2w_1^*)$, exactly $(z_1, z_2) \circ (w_1, w_2)$, and the bottom row is $(-(z_1w_2 + z_2w_1^*)^*, (z_1w_1 - z_2w_2^*)^*)$, so the product has the same special form. So $M(a \circ b) = M(a)M(b)$. Matrix multiplication is associative, and M is one-to-one (you can read off z_1, z_2 from the top row), so $\circ$ is associative. ■

STAGE 2 Not commutative

$(0, 1) \circ (i, 0) = (0 \cdot i - 1 \cdot 0, 0 \cdot 0 + 1 \cdot i^*) = (0, -i)$, but $(i, 0) \circ (0, 1) = (0, i)$. So $\circ$ is not commutative.

The four special products follow directly from the definition:

$$\begin{aligned} (z, 0) \circ (w, 0) &= (zw, 0), & (z, 0) \circ (0, w) &= (0, zw), \\ (0, z) \circ (w, 0) &= (0, zw^*), & (0, z) \circ (0, w) &= (-zw^*, 0). \end{aligned}$$

STAGE 3 The group of eight

In quaternion notation, $A = 1$, $C = i$, $B = j$, $D = ij = k$, and E, G, F, H are their negatives. Every element of S_1 has the form $(u, 0)$ or $(0, u)$ with $u \in \{1, i, -1, -i\}$.

Ask yourself Use the four special products to show that S_1 is closed. What is the identity?

By the four special products, the product of two elements of the form $(u, 0)$ or $(0, u)$ is again of one of these forms, with first or second entry uv , uv^* or $-uv^*$. Since

$\{1, i, -1, -i\}$ is closed under multiplication, conjugation and negation, the product is in S_1 . The identity is $A = (1, 0)$: $(1, 0) \circ (w_1, w_2) = (w_1, w_2)$ and $(w_1, w_2) \circ (1, 0) = (w_1, w_2)$.

 **Pause and try** Find the order and inverse of each element.

A has order 1. $E \circ E = (-1, 0) \circ (-1, 0) = (1, 0)$, so E has order 2. For $C = (i, 0)$: $C \circ C = (-1, 0) = E$, so C has order 4 and $C^{-1} = C^3 = E \circ C = G$. For $B = (0, 1)$: $B \circ B = (-1 \cdot 1^*, 0) = (-1, 0) = E$, so B has order 4 and $B^{-1} = F$. Similarly $D \circ D = (-i \cdot i^*, 0) = (-i \cdot (-i), 0) = (-1, 0) = E$, so D has order 4 with inverse H . And G, F, H are the inverses of C, B, D , also of order 4. In summary:

element	A	B	C	D	E	F	G	H
inverse	A	F	G	H	E	B	C	D
order	1	4	4	4	2	4	4	4

With closure, associativity (Stage 1), an identity and inverses, S_1 is a group. ■ It is called the *quaternion group* Q_8 .

STAGE 4 A subgroup

$\{A, B, E, F\} = \{1, j, -1, -j\}$ consists of the powers of B : $B^1 = B$, $B^2 = E$, $B^3 = F$, $B^4 = A$. So it is closed and contains inverses: a subgroup, *cyclic* of order 4. The matrix $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ satisfies $J^2 = -I$, $J^3 = -J$, $J^4 = I$, so it generates a cyclic group of order 4 too. Matching $B^k \leftrightarrow J^k$ is an isomorphism. ■ (Indeed J is exactly $M(0, 1)$, the matrix of B .)

Looking back

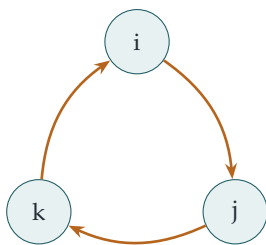
- **Represent by matrices.** Associativity, painful to check directly, was inherited from matrix multiplication.
- **Find the structure behind a formula.** The strange operation $\circ$ was quaternion multiplication, and seeing that made everything easier.
- **Orders and inverses describe a group.** The table of orders already shows Q_8 is unusual: one element of order 2 and six of order 4.

What's really going on

Beyond the complex numbers

The real numbers form a one-dimensional number system in which you can add, multiply and divide; the complex numbers form a two-dimensional one. Hamilton's

quaternions are four-dimensional, and you can still divide, but multiplication is no longer commutative (Figure 6.7.1). In 1878 Ferdinand Frobenius proved that these three, $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ (for Hamilton), are the only finite-dimensional “division algebras” over the reals with associative multiplication. Giving up associativity as well gives one more, the eight-dimensional *octonions*, discovered by John Graves and Arthur Cayley in the 1840s, and that is the end of the line.



With the arrows: $ij = k, jk = i, ki = j$

Against them: $ji = -k, kj = -i, ik = -j$

and $i^2 = j^2 = k^2 = -1$

Figure 6.7.1 Hamilton’s multiplication rules. Multiplying two of i, j, k in the direction of the arrows gives the third; multiplying against the arrows gives minus the third. The eight elements $\pm 1, \pm i, \pm j, \pm k$ form the quaternion group Q_8 of this question.

Rotations in three dimensions

The quaternions of modulus 1 describe rotations of three-dimensional space: the rotation through angle θ about a unit axis $\mathbf{u}$ corresponds to $q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2}(u_1i + u_2j + u_3k)$, acting on a vector $\mathbf{v}$ (written as $v_1i + v_2j + v_3k$) by $\mathbf{v} \mapsto q\mathbf{v}q^{-1}$. Composing rotations is multiplying quaternions. This is more efficient and numerically stable than using matrices or angles, and it avoids “gimbal lock”, a loss of one degree of freedom that the Apollo astronauts had to work around with their gyroscope-based navigation. The half-angle $\frac{\theta}{2}$ is also a clue to something deep: rotating by 360° gives $q = -1$, not 1, a fact that physicists see in the behaviour of electrons (spin).

Four squares

The modulus of a quaternion is multiplicative, $|pq| = |p||q|$, just like that of a complex number. Writing it out gives Euler’s *four-square identity*: a product of two sums of four squares is itself a sum of four squares. That is the key step in Lagrange’s theorem that every positive integer is a sum of four squares, which we prove in Section 7.8, the four-dimensional cousin of the two-squares result in Section 1.6.

The thinking skill

Represent abstract objects concretely

If an abstract operation is hard to reason about, find a concrete model, such as matrices, in which it becomes a familiar operation.

Properties like associativity then come for free. This idea, studying groups by representing them as matrices, grows into the large and beautiful subject of *representation theory*.

Try this next

- 1 Hamilton's rules.** Using the pairs, check that $C \circ B = D$ and $B \circ C = H$, that is, $ij = k$ and $ji = -k$.
- 2 Only one element of order 2.** Show that $E = -1$ is the only element of order 2 in Q_8 . Deduce that Q_8 is not isomorphic to the group of symmetries of a square (which has five elements of order 2).
- 3 Four squares.** Show that $\det M(z_1, z_2) = |z_1|^2 + |z_2|^2$. Using $\det(M_1 M_2) = \det M_1 \det M_2$, deduce that a product of two sums of four squares is a sum of four squares.
- 4 Subgroups.** Find all the subgroups of Q_8 . Show that every subgroup contains E unless it is $\{A\}$, and that Q_8 is the union of its three cyclic subgroups of order 4 (compare Section 6.6).

Further reading

- **John Conway & Derek Smith, *On Quaternions and Octonions*** (A. K. Peters). The geometry and arithmetic of quaternions and octonions. Advanced but fascinating.
- **John Baez, "The octonions", *Bulletin of the American Mathematical Society* 39 (2002)**; also freely available online. A wonderful survey; the introduction is readable now.
- **Andrew Hanson, *Visualizing Quaternions*** (Morgan Kaufmann). Quaternions for rotations, with applications to graphics.

6.8 A Group of Singular Matrices

The question

- (i) Let S be the set of matrices of the form

$$\begin{pmatrix} a & a \\ a & a \end{pmatrix},$$

where a is any real non-zero number. Show that S is closed under matrix multiplication and, further, that S is a group under matrix multiplication.

- (ii) Let G be a set of $n \times n$ matrices which is a group under matrix multiplication, with identity element $\mathbf{E}$. By considering equations of the form $\mathbf{BC} = \mathbf{D}$ for suitable elements $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ of G , show that if a given element $\mathbf{A}$ of G is a singular matrix (i.e. $\det \mathbf{A} = 0$), then all elements of G are singular. Give, with justification, an example of such a group of singular matrices in the case $n = 3$.

The hook

Every student learns that a matrix with zero determinant has no inverse. So a group of matrices, in which every element has an inverse, surely cannot contain any singular matrices?

It can. The catch is in the word “inverse”. In a group, the inverse of $\mathbf{A}$ is the element that combines with $\mathbf{A}$ to give the *group’s* identity, and nothing says the group’s identity has to be the identity matrix $\mathbf{I}$. This question builds a group in which every matrix is singular, and the identity element is a matrix with determinant zero. It is a lesson in reading definitions literally, and in the difference between an abstract structure and the familiar example you first met it in.

Before you start

For the definition of a group, see Section 6.6.



Ask yourself In part (i), the identity matrix $\mathbf{I}$ is not of the form $\begin{pmatrix} a & a \\ a & a \end{pmatrix}$. So what could the identity element of S be?

It must be some matrix $\mathbf{E}$ in S with $\mathbf{EX} = \mathbf{XE} = \mathbf{X}$ for every $\mathbf{X}$ in S . That is a condition

only on how $\mathbf{E}$ acts on matrices of this special form, not on all matrices. Find it by solving the equation.

Working it through

STAGE 1 Closure

Write $\mathbf{M}_a$ for the matrix with all four entries equal to a . Then

$$\mathbf{M}_a \mathbf{M}_b = \begin{pmatrix} a & a \\ a & a \end{pmatrix} \begin{pmatrix} b & b \\ b & b \end{pmatrix} = \begin{pmatrix} 2ab & 2ab \\ 2ab & 2ab \end{pmatrix} = \mathbf{M}_{2ab},$$

and $2ab \neq 0$ when $a, b \neq 0$. So S is closed. ■

STAGE 2 The group axioms



Ask yourself Matrix multiplication is associative. For the identity, solve $\mathbf{M}_e \mathbf{M}_a = \mathbf{M}_a$. For inverses, solve $\mathbf{M}_a \mathbf{M}_b = \mathbf{M}_e$.

Associativity is inherited from matrix multiplication. *Identity*: $\mathbf{M}_e \mathbf{M}_a = \mathbf{M}_{2ea} = \mathbf{M}_a$ for all a exactly when $e = \frac{1}{2}$; by the symmetry of the formula, $\mathbf{M}_a \mathbf{M}_{1/2} = \mathbf{M}_a$ too. So

$$\mathbf{E} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

is the identity of S . *Inverses*: $\mathbf{M}_a \mathbf{M}_b = \mathbf{M}_{2ab} = \mathbf{M}_{1/2}$ when $b = \frac{1}{4a}$, which is non-zero, so $\mathbf{M}_{1/(4a)}$ is the inverse of $\mathbf{M}_a$ in S . So S is a group. ■ Every element has determinant $a^2 - a^2 = 0$.

STAGE 3 Singularity spreads



Pause and try Suppose $\mathbf{A} \in G$ is singular, and let $\mathbf{B}$ be any element of G . Write $\mathbf{B}$ as a product involving $\mathbf{A}$, and take determinants.

Let $\mathbf{A}^{-1}$ denote the inverse of $\mathbf{A}$ in the group G (not necessarily a matrix inverse). Then $\mathbf{C} = \mathbf{A}^{-1} \mathbf{B}$ is in G , and

$$\mathbf{AC} = \mathbf{AA}^{-1} \mathbf{B} = \mathbf{EB} = \mathbf{B}.$$

Taking determinants, $\det \mathbf{B} = \det \mathbf{A} \det \mathbf{C} = 0$. So every element of G is singular. ■ In particular, if one element of a matrix group is invertible as a matrix, all are; a matrix group is either entirely invertible or entirely singular.

STAGE 4 A 3×3 example



Ask yourself Can you generalise part (i) to 3×3 matrices? Or find something even simpler, using a matrix that “ignores” two of the three coordinates?

Take

$$G = \left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} : a \neq 0 \right\}.$$

The product of the matrices with corner entries a and b is the one with corner entry ab , so G is closed; multiplication is associative; the identity is the matrix with corner entry 1; the inverse of the one with corner a is the one with corner $\frac{1}{a}$. So G is a group, and every element has determinant 0. ■ (Another example: the 3×3 matrices with all nine entries equal to $a \neq 0$, whose identity has all entries $\frac{1}{3}$.)

Looking back

- **Read definitions literally.** A group’s identity is whatever element acts as an identity *within the group*; it need not be the familiar identity.
- **Solve for the structure.** The identity and inverses in S were found by solving equations, not by guessing.
- **Determinants are multiplicative.** That one property showed that singularity spreads from one element to all of them.

What’s really going on

Idempotents and hidden invertibility

The identity E of any matrix group satisfies $E^2 = E$: it is *idempotent*. Geometrically, an idempotent matrix is a *projection*: it maps every vector onto a subspace and leaves the vectors in that subspace alone. In part (i), E projects the plane onto the line $y = x$, and every matrix M_a first projects onto that line and then stretches along it by the factor $2a$ (Figure 6.8.1). On the line, the group acts by invertible stretches; it simply ignores the direction perpendicular to it. So S is really the group of non-zero real numbers under multiplication (via $M_a \mapsto 2a$), living on a line inside the plane. This is true in general: every group of matrices is, after a change of basis, a group of invertible matrices acting on the image of its identity element, padded with zeros.

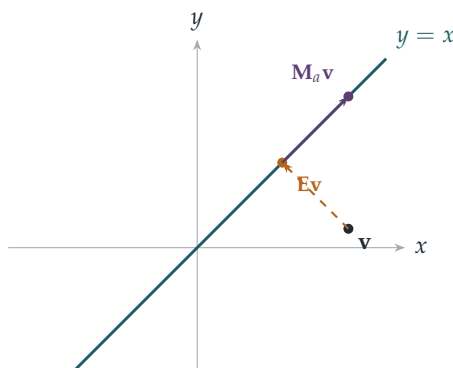


Figure 6.8.1 What the matrices in part (i) do. The identity element $\mathbf{E}$ projects every vector $\mathbf{v}$ onto the line $y = x$ (amber, dashed). Each $\mathbf{M}_a$ does the same projection and then stretches along the line by the factor $2a$ (plum; here $2a = \frac{16}{9}$). On the line itself, the group acts by ordinary invertible scaling.

Abstract and concrete

The first groups people studied were concrete: permutations, symmetries, invertible matrices. The abstract definition (a set with an operation satisfying three axioms) was distilled from them. This question shows the value of the abstract view: S and the group of non-zero reals under multiplication are *isomorphic*, the same group in different clothing, even though one consists of singular matrices and the other of numbers. Recognising when two different-looking structures are “the same” is one of the central skills of algebra, and the notion of isomorphism is how mathematicians make that precise.

The thinking skill

Question hidden assumptions about familiar objects

When a familiar object appears in an abstract definition, check which of its familiar properties the definition actually requires.

“The identity of a group of matrices is $\mathbf{I}$ ” sounds obvious but is not part of the definition, and dropping it led to a genuinely new example. Questioning assumptions like this is how many new mathematical structures are found.

Try this next

- 1 An isomorphism.** Show that the map $\mathbf{M}_a \mapsto 2a$ is a one-to-one correspondence from S to the non-zero reals that turns matrix multiplication into multiplication of numbers.

- 2 Not a group.** Show that the set of matrices $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ with $a \neq 0$ is closed under multiplication but is not a group. (*Hint:* look for an identity element. Is there a left identity? A right identity?)
- 3 Idempotents.** Find all 2×2 real matrices E with $E^2 = E$. Describe them geometrically.
- 4 All entries equal.** Show that the $n \times n$ matrices with all entries equal to $a \neq 0$ form a group under multiplication, and find its identity element.

Further reading

- **M. A. Armstrong, *Groups and Symmetry*** (Springer). Matrix groups and isomorphisms, with many geometric examples.
- **Gilbert Strang, *Introduction to Linear Algebra*** (Wellesley-Cambridge). Projections and idempotent matrices, clearly explained.
- **Joseph Gallian, *Contemporary Abstract Algebra*** (CRC Press). A friendly first course with many examples of unusual groups.

What We Have Learned

We began with the shift in algebra from solving equations to studying structures, and with three questions: what can be built from what, why the formulas for equations work, and what is gained by abstraction. Here is what the chapter's eight questions showed.

1. What can be built from what?

In Section 6.3 we eliminated radicals one at a time to find polynomials satisfied by $\sqrt{2} + \sqrt{3}$, $\sqrt{2} + \sqrt{3} + \sqrt{5}$ and $\sqrt[3]{2} + \sqrt[3]{3}$, of degrees 4, 8 and 9. Those degrees are invariants: the degree of the field a number generates. Because degrees multiply up a tower of fields, numbers reachable by square roots have degrees that are powers of 2. That single fact proves that the cube cannot be doubled ($\sqrt[3]{2}$ has degree 3) and that a 60° angle cannot be trisected.

In the other direction, Section 6.2 showed how something *can* be built. Splitting the 17th roots of unity into Gauss's periods, the sum $\cos \frac{2\pi}{17} + \cos \frac{4\pi}{17} + \cos \frac{8\pi}{17} + \cos \frac{16\pi}{17}$ satisfied a quadratic, with root $\frac{-1+\sqrt{17}}{4}$; three more quadratics lead to $\cos \frac{2\pi}{17}$, and so to the construction of the regular 17-gon. Underneath both sections were the cyclotomic polynomials of Section 6.1, the building blocks of $x^n - 1$, obtained by sorting the roots of unity by their order.

2. Why do the formulas for equations work?

Section 6.4 showed the secret of the quartic: from four roots, form the three quantities $\alpha\beta + \gamma\delta$, $\alpha\gamma + \beta\delta$, $\alpha\delta + \beta\gamma$, one for each way of pairing the roots. Any rearrangement of the roots merely shuffles these three, so they satisfy a cubic whose coefficients are known. Solving the cubic, then two quadratics, gave the roots. Lagrange saw that this works because of how permutations of four objects act on the three pairings; Galois saw that it fails for five roots because the group of permutations of five objects contains a piece that cannot be broken down. The formulas are shadows of group structure.

3. What is gained by abstraction?

The last four sections described objects by the rules they obey.

- In Section 6.5, four rules (additivity, the product rule, constants, $\diamond x = 1$) forced an operation to be differentiation, with no limits in sight. The same rules power automatic differentiation in modern software.
- In Section 6.6, a three-line argument from the group axioms showed that no group is the union of two proper subgroups, for every group at once.

- In Section 6.7, a strange operation on pairs of complex numbers turned out to be Hamilton’s quaternion multiplication, and representing it by matrices made associativity automatic.
- In Section 6.8, reading the definition of a group literally produced a group of singular matrices whose identity is not the identity matrix, and which is “really” the group of non-zero real numbers in disguise.

In each case, the abstract view explained something that calculation alone would have obscured.

4. A toolkit for algebra

Method	Where you used it
Classifying by a unique invariant	Roots of unity by order (6.1)
Sum and product give a quadratic	Gauss periods (6.2); the quartic (6.4)
Product-to-sum formulas	The 17-gon (6.2)
Eliminating radicals	Minimal polynomials (6.3)
Degrees and the tower law	Impossibility of constructions (6.3)
Symmetric functions of roots	The resolvent cubic (6.4)
Determining an operation by its axioms	Derivations (6.5)
Arguing from group axioms	Unions of subgroups (6.6)
Matrix representations	Quaternions (6.7)
Isomorphism	Singular matrix groups (6.8)

Where next

- Quaternions lead to Lagrange’s four-square theorem in Section 7.8.
- Matrices as algebraic objects, with their own rules, are the subject of Chapter 7.
- The arithmetic of number systems like the Gaussian integers, where factorisation can fail, was the theme of Section 1.7.

If you want to go further before university

- **Mario Livio, *The Equation That Couldn’t Be Solved***: the quintic, Galois and the birth of group theory.
- **Nathan Carter, *Visual Group Theory***: groups through pictures, the ideal first look.
- **M. A. Armstrong, *Groups and Symmetry***: a first university course, built on geometry.
- **Ian Stewart, *Galois Theory***: for when you are ready for the full story.

Galois died at twenty, after a duel, leaving a few dozen pages that took the mathematical

world decades to understand. What he saw was that an equation's deepest secrets are not in its coefficients but in the symmetries of its solutions. That insight, giving the same name to different things and studying the structure they share, now runs through all of mathematics. You have seen it at work in polynomials, polygons, quaternions and matrices; at university you will see it almost everywhere.