UNIT 3

Complex Numbers

From Bombelli's square root of a negative number to the fundamental theorem of algebra

Image: Own rendering

By the end of this unit, I should be able to answer these questions:

- How can a formula that passes through $\sqrt{-121}$ arrive at the real answer 4?
- What are complex numbers, and why do the ordinary laws of arithmetic still hold for them?
- Why does multiplying by i turn the plane through a right angle, and what does multiplying by $1+i$ do?
- Where do the five fifth roots of unity lie, and why do they add up to zero?
- What does the fundamental theorem of algebra say, and why does it follow that every real polynomial is a product of linear and quadratic factors?
- What happens to the point 0 when I iterate $z \mapsto z^2 + c$ for $c = -1$, for $c = i$, and for $c = 1$?

In 1572 an engineer from Bologna published a calculation that should not have worked. He applied Cardano's formula to a cubic equation whose answer he already knew, found that the formula asked him for the square root of a negative number, carried on as if such a number existed, and arrived at the right answer. It took two more centuries before anyone could say what he had been computing with. This unit builds the numbers he needed, shows that they obey every law of arithmetic, draws them, and uses them to finish what Unit 2 started, the proof that every polynomial equation has a complete set of solutions. It ends with a picture of where those solutions lie.

1 Cardano's formula meets a negative square root

Unit 2 derived Cardano's formula for the depressed cubic. For an equation of the form $x^3 + px + q = 0$ with real coefficients p and q, one solution is

$$x = u + v, \qquad D = \left(\tfrac{q}{2}\right)^2 + \left(\tfrac{p}{3}\right)^3,$$
$$u = \sqrt[3]{-\tfrac{q}{2} + \sqrt{D}},$$
$$v = \sqrt[3]{-\tfrac{q}{2} - \sqrt{D}}, \tag{1}$$

where the two cube roots u and v are chosen so that $3uv = -p$. The formula handled $x^3 + 6x = 20$ without trouble, because with $p = 6$ and $q = -20$ the expression under the square root, $(-10)^2 + 2^3 = 100 + 8 = 108$, is positive. Rafael Bombelli (1526-1572) tried it on an equation of the same shape whose answer can be guessed.

WORKED EXAMPLE Bombelli's equation $x^3 = 15x + 4$

The equation $x^3 = 15x + 4$ has the solution $x = 4$, because $4^3 = 64$ and $15 \cdot 4 + 4 = 64$. Dividing $x^3 - 15x - 4$ by $x - 4$ leaves $x^2 + 4x + 1$, whose roots are $-2 \pm \sqrt{3}$ by completing the square (Unit 1), so all three roots are real. Written as $x^3 + px + q = 0$, the equation has $p = -15$ and $q = -4$, and the quantity D under the square root in Equation (1) is

$$D = (-2)^2 + (-5)^3 = 4 - 125 = -121.$$

The formula therefore asks for

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}},$$

and no real number has -121 as its square. Bombelli's idea was to write $\sqrt{-121} = 11\sqrt{-1}$, to treat $\sqrt{-1}$ as a symbol whose square is -1, and to compute with it by the ordinary rules. He then guessed that a cube root of $2 + 11\sqrt{-1}$ might have the same shape with small whole numbers, tried $2 + \sqrt{-1}$, and checked the guess by cubing. Since $(2 + \sqrt{-1})^2 = 4 + 4\sqrt{-1} + (\sqrt{-1})^2 = 3 + 4\sqrt{-1}$, the cube is

$$\begin{aligned}(2 + \sqrt{-1})^3 &= (3 + 4\sqrt{-1})(2 + \sqrt{-1}) \\ &= 6 + 3\sqrt{-1} + 8\sqrt{-1} + 4(\sqrt{-1})^2 = 2 + 11\sqrt{-1}.\end{aligned}$$

In the same way $(2 - \sqrt{-1})^3 = 2 - 11\sqrt{-1}$, and the product of the two cube roots is $(2 + \sqrt{-1})(2 - \sqrt{-1}) = 4 - (-1) = 5 = -p/3$, as the condition $3uv = -p$ requires. Adding them gives

$$x = (2 + \sqrt{-1}) + (2 - \sqrt{-1}) = 4,$$

so the formula does deliver the root $\underline{\underline{x = 4}}$, and the symbol $\sqrt{-1}$ cancels out on the way.

Bombelli called the new quantities "più di meno" (plus of minus) and "meno di meno" (minus of minus), wrote down the rules for multiplying them, and used them in his Algebra of 1572 (Figure 1) because they worked, while admitting that the whole thing seemed to rest on sophistry, that is, on reasoning that looks clever but is not sound. The calculation raises two questions: what kind of object $\sqrt{-1}$ is, and why the ordinary rules of algebra should apply to it. The next section answers both.


L'ALGEBRA
PARTE MAGGIORE
DELL'ARIMETICA
DIVISA IN TRE LIBRI
DI RAFAEL BOMBELLI
DA BOLOGNA.
Nouamente posta in luce.
IN BOLOGNA
Nella stamperia di Giouanni Rossi
M D L X X I I.
Con Licentia delli RR. VV. del Vesc. & Inquisit.

Figure 1. The title page of Rafael Bombelli's L'Algebra, Bologna, 1572, the book with the first rules for computing with square roots of negative numbers. Credit: Wikimedia Commons, public domain.

2 The complex numbers

2.1 Pairs of real numbers

The honest way to introduce a new number is to say exactly what it is and how it is added and multiplied, and then to check the laws. A **complex number** is a pair (a, b) of real numbers, and the pairs are added and multiplied as follows.

KEY RESULT The complex numbers

$$(a, b) + (c, d) = (a + c,\ b + d),$$
$$(a, b) \cdot (c, d) = (ac - bd,\ ad + bc).$$

The set of all such pairs with these two operations is denoted $\mathbb{C}$. The pair $(a, 0)$ behaves exactly like the real number a, because $(a, 0) + (c, 0) = (a + c, 0)$ and $(a, 0) \cdot (c, 0) = (ac, 0)$, and the pair $i = (0, 1)$ satisfies $i^2 = (0 \cdot 0 - 1 \cdot 1,\ 0 \cdot 1 + 1 \cdot 0)$, which is $(-1, 0) = -1$.

The rule for the product looks arbitrary until the pair (a, b) is written as $a + bi$, which is allowed because $(b, 0)(0, 1) = (0, b)$ by the product rule, so that $(a, b) = (a, 0) + (b, 0)(0, 1)$. Then the rule is what expanding $(a + bi)(c + di)$ by the distributive law gives when i^2 is replaced by -1:

$$\begin{aligned}(a + bi)(c + di) &= ac + (ad + bc)\, i + bd\, i^2 \\ &= (ac - bd) + (ad + bc)\, i.\end{aligned}$$

From now on every complex number is written as $z = a + bi$. The number a is the **real part** $\Re z$ and the number b the **imaginary part** $\Im z$; both are real numbers. Two complex numbers are equal exactly when both parts agree. The word "imaginary" is a historical accident. A pair of real numbers is no more imaginary than a point on a map. The symbol i is a particular pair, and $i^2 = -1$ is a consequence of the multiplication rule, which is why Bombelli's computations were right.

The laws of arithmetic that Unit 1 listed for the real numbers (the commutative and associative laws of addition and multiplication, the distributive law, the roles of 0 and 1, and the existence of additive and multiplicative inverses) all hold in $\mathbb{C}$. Most of them follow at once from the corresponding laws for real numbers, because each component of a sum or product is a real expression in the components; for example, the commutative law of multiplication holds because $ac - bd = ca - db$ and $ad + bc = cb + da$. The associative law of multiplication is a longer but equally routine check, which the reader is invited to carry out once for $(1 + i)(2 - i)$ and i in both bracketings. The zero is $0 + 0i$, the one is $1 + 0i$, and the additive inverse of $a + bi$ is $-a - bi$. Only the multiplicative inverse needs an idea, and it comes from the conjugate below.

WORKED EXAMPLE A product and a quotient

To illustrate the rules, take $z = 3 + 4i$ and $w = 1 - 2i$. The product is

$$\begin{aligned}zw = (3 + 4i)(1 - 2i) &= 3 - 6i + 4i - 8i^2 \\ &= 3 - 2i + 8 = 11 - 2i.\end{aligned}$$

For the quotient z/w, multiply numerator and denominator by $1 + 2i$, the denominator with the sign of its i-part reversed, which turns the denominator into a real number, since $(1 - 2i)(1 + 2i) = 1 - 4i^2 = 5$:

$$\begin{aligned}\frac{3 + 4i}{1 - 2i} &= \frac{(3 + 4i)(1 + 2i)}{(1 - 2i)(1 + 2i)} \\ &= \frac{3 + 6i + 4i + 8i^2}{5} = \frac{-5 + 10i}{5} = -1 + 2i.\end{aligned}$$

As a check, $(-1+2i)(1-2i) = -1+2i+2i-4i^2 = 3+4i$. The results are $\underline{\underline{zw = 11-2i}}$ and $\underline{\underline{z/w = -1+2i}}$.

The powers of i repeat with period 4, since $i^1 = i$, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, and then the cycle $1, i, -1, -i$ starts again. Any power i^n is therefore found from the remainder of n on division by 4. For example, $i^{2026} = i^2 = -1$, because $2026 = 4 \cdot 506 + 2$.

2.2 The complex plane, the conjugate, and the modulus

A pair of real numbers is a point in the plane, and this is the picture that made complex numbers respectable. The number $z = a + bi$ is drawn as the point with coordinates (a, b), or as the arrow from the origin to that point (Figure 2). The horizontal axis carries the real numbers and is called the real axis, and the vertical axis carries the multiples of i and is called the imaginary axis. Addition of complex numbers is addition of arrows, head to tail, as for vectors.

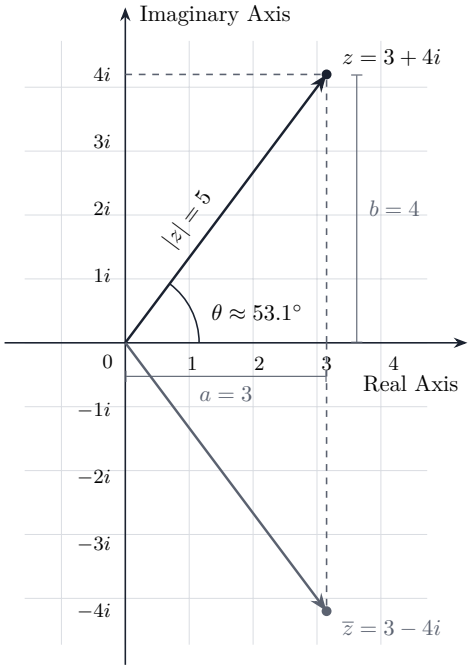


Figure 2. The complex plane with the number $z = 3 + 4i$ drawn as an arrow. Its conjugate $\overline{z} = 3 - 4i$ is the mirror image in the real axis, its modulus $|z| = 5$ is the length of the arrow by Pythagoras, and its argument θ is the angle from the positive real axis, about $53.1°$ here, since $\tan\theta = 4/3$.

Three quantities belong to every complex number $z = a + bi$. The **conjugate** $\overline{z} = a - bi$ is the mirror image of z in the real axis; the name goes back to Augustin-Louis Cauchy (1789-1857), who introduced it in 1821. The **modulus** $|z| = \sqrt{a^2 + b^2}$ is the length of the arrow, a non-negative real number, and for a real number it is the ordinary absolute value. The **argument** $\arg z$ is the angle between the positive real axis and the arrow, measured anticlockwise; Section 3 uses it. The conjugate and the modulus are tied together by one identity that does most of the work in this subject:

$$\begin{aligned} z\overline{z} = (a+bi)(a-bi) &= a^2 - abi + abi - b^2 i^2 \\ &= a^2 + b^2 = |z|^2. \end{aligned}$$

This is why multiplying by $1 + 2i$ cleared the denominator in the last example.

KEY RESULT Conjugate, modulus, and the inverse

$$z\overline{z} = |z|^2, \qquad \frac{1}{z} = \frac{\overline{z}}{|z|^2} \quad \text{for } z \neq 0,$$
$$\overline{z+w} = \overline{z} + \overline{w}, \qquad \overline{zw} = \overline{z}\,\overline{w}.$$

For $z = a + bi$ the conjugate is $\overline{z} = a - bi$ and the modulus is $|z| = \sqrt{a^2 + b^2}$. Every non-zero complex number has an inverse, so division is always possible in $\mathbb{C}$, as it is in $\mathbb{Q}$ and $\mathbb{R}$.

The two rules for the conjugate of a sum and of a product say that conjugation can be done before or after the arithmetic, and they are checked by expanding both sides. For the product, with $z = a+bi$ and $w = c+di$, the left side is $(ac-bd)-(ad+bc)i$ and the right side is $(a-bi)(c-di) = (ac-bd)-(ad+bc)i$ as well. Applied repeatedly, the product rule gives $\overline{z^n} = \overline{z}^{\,n}$ for every positive integer n, a fact that Section 5 turns into a statement about the roots of real polynomials.

CHECK YOUR UNDERSTANDING Question 1

What is $1/(1+2i)$?

(A) $1-2i$
(B) $(1+2i)/5$
(C) $(-1+2i)/5$
(D) $(1-2i)/5$

The solution is at the end of the unit.

EXCURSION *Two centuries from sophistry to a plane*

Bombelli's Algebra, printed in Bologna in 1572, the year of his death, states the rules for "più di meno" in words, among them the rule that plus of minus times plus of minus makes minus, which is $i \cdot i = -1$. René Descartes (1596-1650) called them imaginary in 1637, and the name has remained in use ever since. Leonhard Euler (1707-1783) computed with them freely and proposed the letter i for $\sqrt{-1}$ in a paper of 1777.

The picture came last. In 1797 the Norwegian surveyor Caspar Wessel (1745-1818) presented a paper to the Royal Danish Academy in which a complex number is a directed line segment in the plane and multiplication adds the angles; printed in Danish in 1799, it stayed unread for a century. The Swiss bookkeeper Jean-Robert Argand (1768-1822) rediscovered the plane in 1806. Carl Friedrich Gauss (1777-1855) used the plane in his proofs and in 1831 gave the numbers the name they carry today, complex numbers. By then the question what $\sqrt{-1}$ is had an answer nobody could object to: it is the point one step up the vertical axis.

3 Polar form and de Moivre's formula

3.1 Multiplication as rotation and stretching

An arrow in the plane is fixed by its length and its direction as well as by its two coordinates. If $z = a+bi$ has modulus $r = |z|$ and argument θ, then $a = r\cos\theta$ and $b = r\sin\theta$, because $\cos\theta$ and $\sin\theta$ are by definition the coordinates of the point at angle θ on the circle of radius 1, and the arrow of z is that

point stretched by r (Figure 2). Hence

$$z = r(\cos\theta + i\sin\theta). \tag{2}$$

This is the **polar form** of z. The argument is determined only up to whole turns, which means that two angles differing by a whole number of turns describe the same arrow, since adding 2π (or $360°$) to θ does not move it; when one value is needed, the one in the interval $(-\pi, \pi]$ is taken. The number 0 has modulus 0 and no argument.

In polar form, multiplication becomes geometry. To multiply two complex numbers, multiply their moduli and add their arguments.

KEY RESULT Multiplication in polar form

$$z = r(\cos\alpha + i\sin\alpha),$$
$$w = s(\cos\beta + i\sin\beta),$$
$$zw = rs\big(\cos(\alpha+\beta) + i\sin(\alpha+\beta)\big).$$

In words, $|zw| = |z|\,|w|$ and $\arg(zw) = \arg z + \arg w$ (up to whole turns). Multiplying by w therefore turns the plane through the angle $\arg w$ and stretches it by the factor $|w|$, and division by $w \neq 0$ does the opposite: $|z/w| = |z|/|w|$ and $\arg(z/w) = \arg z - \arg w$.

PROOF Proof of the multiplication rule

Multiplying out the two polar forms with the rule of Section 2 gives

$$zw = rs\big[(\cos\alpha\cos\beta - \sin\alpha\sin\beta)$$
$$+i(\sin\alpha\cos\beta + \cos\alpha\sin\beta)\big].$$

The two brackets are exactly the addition formulas of trigonometry, $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$ and $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$. Hence $zw = rs(\cos(\alpha+\beta) + i\sin(\alpha+\beta))$, which is a complex number of modulus rs and argument $\alpha+\beta$, and this completes the proof. The rule for division follows by applying the product rule to w and z/w.

Two special cases show what the rule means. The number i has modulus 1 and argument $90°$, so multiplying by i turns every arrow through a right angle anticlockwise and changes no length, and multiplying by i twice turns the plane through half a turn, which is multiplication by -1. This is the geometric meaning of $i^2 = -1$. The number $1+i$ has modulus $\sqrt{2}$ and argument $45°$, so multiplication by $1+i$ turns the plane through $45°$ and stretches it by $\sqrt{2}$. Figure 3 shows what this does to the square with corners 0, 1, $1+i$, and i. Unit 4 describes the same turn and stretch by a 2×2 matrix, which gives a second way of building $\mathbb{C}$.

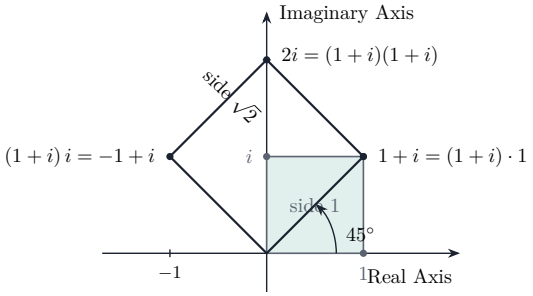
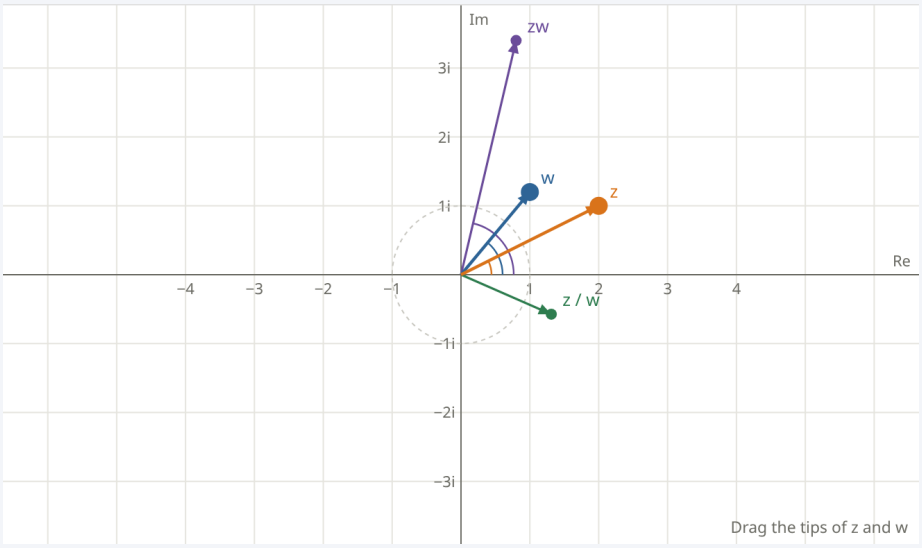


Figure 3. Multiplication by $1+i$ acting on the square of side 1 with corners 0, 1, $1+i$, and i (light). The image (dark) is the square turned anticlockwise through $45°$ and enlarged by the factor $\sqrt{2} = |1+i|$, with corners 0, $1+i$, $2i$, and $-1+i$.

INTERACTIVE Multiplication in the complex plane



Drag the tips of the two arrows z and w and watch the product zw and the quotient z/w move with them: the length of zw is the product of the two lengths and its angle is the sum of the two angles. Switch to the roots to see the n-th roots of z, equally spaced on a circle whose radius is the n-th root of |z|.

Click the picture or this link to try it in the online version of this unit.

3.2 De Moivre's formula and Euler's formula

When a number of modulus 1 is multiplied by itself over and over again, the multiplication rule adds the same angle at every step and leaves the modulus at 1, so all the powers of such a number can be written down at once. The result is named after Abraham de Moivre (1667-1754), who used it from 1707 onwards.

THEOREM De Moivre's formula

For every real number θ and every positive integer n,

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

Combined with the modulus, $\big(r(\cos\theta + i\sin\theta)\big)^n = r^n(\cos n\theta + i\sin n\theta)$: the n-th power of z has modulus $|z|^n$ and argument $n\arg z$.

PROOF Proof of de Moivre's formula by induction

The proof is by induction on n. The base case $n = 1$ says that $\cos\theta + i\sin\theta$ equals itself. For the induction step, suppose the formula holds for some $n \geq 1$. Then

$$(\cos\theta + i\sin\theta)^{n+1} = (\cos n\theta + i\sin n\theta)(\cos\theta + i\sin\theta),$$

which equals $\cos(n\theta + \theta) + i\sin(n\theta + \theta)$ by the induction hypothesis and the multiplication rule, and this is $\cos(n+1)\theta + i\sin(n+1)\theta$, the formula for $n+1$. By the principle of induction the formula holds for every positive integer n, which completes the proof. It holds for negative integers as well, because $(\cos\theta + i\sin\theta)^{-1} = \cos(-\theta) + i\sin(-\theta)$, as multiplying the two numbers shows.

WORKED EXAMPLE The polar form of $1+i$ and the power $(1+i)^{10}$

The number $1+i$ has modulus $\sqrt{1^2+1^2}=\sqrt{2}$ and argument $45^\circ=\pi/4$, since its arrow is the diagonal of the square of side 1. Its polar form is therefore $1+i=\sqrt{2}\,(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4})$. By de Moivre's formula,

$$\begin{aligned}(1+i)^{10} &= (\sqrt{2})^{10}\left(\cos\frac{10\pi}{4}+i\sin\frac{10\pi}{4}\right)\\ &= 32\left(\cos\frac{5\pi}{2}+i\sin\frac{5\pi}{2}\right).\end{aligned}$$

The angle $5\pi/2$ is one full turn plus a quarter turn, so the cosine is 0 and the sine is 1, and $\underline{\underline{(1+i)^{10}=32i}}$. The shortcut $(1+i)^2=2i$, hence $(1+i)^{10}=(2i)^5=32\,i^5=32i$, confirms the result.

The expression $\cos\theta+i\sin\theta$ appears so often that it deserves a shorter name, and the name mathematics settled on is an exponential. Multiplication adds the arguments, exactly as multiplication of powers adds the exponents, and this is the reason for the notation.

KEY RESULT Euler's formula

$$e^{i\theta}=\cos\theta+i\sin\theta,$$

so that $z=re^{i\theta}$ for $r=|z|$ and $\theta=\arg z$.

Taken in this unit as the definition of $e^{i\theta}$, the formula has the properties an exponential should have: $e^{i\alpha}e^{i\beta}=e^{i(\alpha+\beta)}$ by the multiplication rule, $(e^{i\theta})^n=e^{in\theta}$ by de Moivre's formula, $e^{i0}=1$, $|e^{i\theta}|=1$, and $\overline{e^{i\theta}}=e^{-i\theta}$. The special value $\theta=\pi$ gives $e^{i\pi}=-1$.

An expert will object that $e^{i\theta}$ ought to be the exponential function e^x continued to complex arguments, and that the formula should then be a theorem. That is correct, and the proof compares the power series of the exponential, the sine, and the cosine, which needs calculus and lies beyond this unit. Everything this program uses is the definition together with the five properties just listed, and in those terms $re^{i\theta}$ is the polar form written compactly. The unit uses both notations from here on.

CHECK YOUR UNDERSTANDING Question 2

What does multiplying every complex number by $-i$ do to the plane?

(A) It turns the plane anticlockwise through a right angle

(B) It turns the plane clockwise through a right angle

(C) It reflects the plane in the real axis

(D) It turns the plane through half a turn

The solution is at the end of the unit.

4 Roots of unity and the cubic completed

4.1 The n-th roots of unity

The equation $z^2 = 1$ has the two solutions 1 and -1. How many solutions does $z^n = 1$ have in $\mathbb{C}$, and where are they? De Moivre's formula answers both questions. If $z = re^{i\theta}$, then $z^n = r^n e^{in\theta}$. This equals 1 exactly when two things hold at once: the modulus r^n is 1, which for a positive real number r means $r = 1$, and the angle $n\theta$ is a whole number of turns, which means $\theta = 2\pi k/n$ for some integer k. The values $k = 0, 1, \ldots, n-1$ give n different arrows, and $k = n$ gives the same arrow as $k = 0$ again.

KEY RESULT The n-th roots of unity

$$z^n = 1 \quad \Longleftrightarrow \quad z = \zeta_n^k = e^{2\pi i k/n}, \qquad k = 0, 1, \ldots, n-1,$$

where $\zeta_n = e^{2\pi i/n} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$. The n solutions are called the n-th **roots of unity**. They lie on the circle $|z| = 1$, for $n \geq 3$ at the corners of a regular polygon with n corners (a regular n-gon), one of them at 1, and for $n \geq 2$ their sum is 0.

The number ζ_n is called a **primitive** n-th root of unity, because all the others are its powers. Multiplying two n-th roots of unity gives a third, because the arguments add and a sum of two multiples of $2\pi/n$ is again a multiple of $2\pi/n$, and Unit 5 returns to these n points as one of its first examples of a group. The claim about the sum has a two-line proof. Let $S = 1 + \zeta_n + \zeta_n^2 + \cdots + \zeta_n^{n-1}$. Multiplying S by ζ_n turns each term into the next one, $\zeta_n S = \zeta_n + \zeta_n^2 + \cdots + \zeta_n^{n-1} + \zeta_n^n$, and the new last term $\zeta_n^n = 1$ is the first term of S. The two sums therefore contain the same n terms, so $\zeta_n S = S$, which gives $(\zeta_n - 1)S = 0$. For $n \geq 2$ the factor $\zeta_n - 1$ is not zero, and therefore $S = 0$. Figure 4 shows the geometric version of the same statement: the five arrows of the fifth roots of unity, placed head to tail, close up into a regular pentagon and return to their starting point.

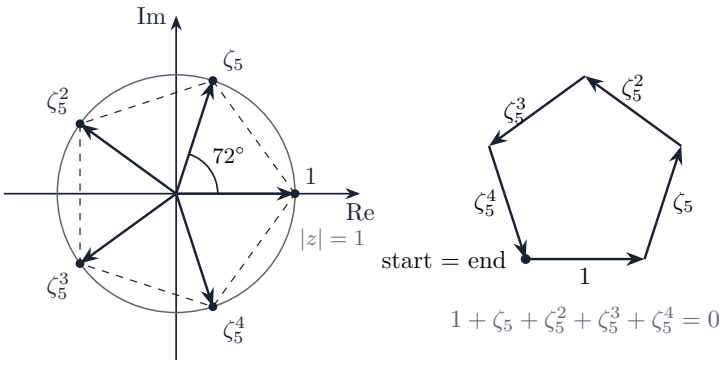


Figure 4. Left: the fifth roots of unity $1, \zeta_5, \zeta_5^2, \zeta_5^3, \zeta_5^4$ on the circle $|z| = 1$, at angles that are multiples of 72°. Right: the same five arrows placed head to tail close up to a regular pentagon, so their sum is 0.

WORKED EXAMPLE The fifth roots of unity in decimals

The fifth roots of unity are $\zeta_5^k = \cos(72°k) + i\sin(72°k)$ for $k = 0, 1, 2, 3, 4$. With $\cos 72° = 0.3090$, $\sin 72° = 0.9511$, $\cos 144° = -0.8090$, and $\sin 144° = 0.5878$, they are

$$\begin{aligned}
\zeta_5^1 &= 0.3090 + 0.9511\,i,\\
\zeta_5^4 &= 0.3090 - 0.9511\,i,\\
\zeta_5^2 &= -0.8090 + 0.5878\,i,\\
\zeta_5^3 &= -0.8090 - 0.5878\,i,
\end{aligned}$$

together with $\zeta_5^0 = 1$. The roots with $k = 1$ and $k = 4$ are conjugates, and so are those with $k = 2$ and $k = 3$, which is why the pairs are written under each other. The imaginary parts cancel in pairs, and the real parts add up to $1 + 2(0.3090) + 2(-0.8090)$, which is $1 + 0.6180 - 1.6180 = 0$. The sum of the five roots is therefore $\underline{\underline{0}}$, as the key result promised.

CHECK YOUR UNDERSTANDING Question 3

How many of the twelfth roots of unity are also cube roots of unity?

(A) 12
(B) 4
(C) 3
(D) 1

The solution is at the end of the unit.

The cube roots of unity appear in the cubic formula and are used again in Unit 8, so they get their own symbol. With $\omega = \zeta_3 = e^{2\pi i/3}$, Table 1 lists the three roots of $z^3 = 1$ in both forms. The identity $1 + \omega + \omega^2 = 0$ is the case $n = 3$ of the sum rule, and $\overline{\omega} = \omega^2$ follows from the table.

Root	Rectangular Form	Polar Form
1	1	e^{0}
ω	$-\frac{1}{2} + \frac{\sqrt{3}}{2}\,i$	$e^{2\pi i/3}$
ω^2	$-\frac{1}{2} - \frac{\sqrt{3}}{2}\,i$	$e^{4\pi i/3}$

Table 1. The three cube roots of unity, with $\omega = e^{2\pi i/3}$. They satisfy $\omega^3 = 1$, $\omega^2 = \overline{\omega}$, and $1+\omega+\omega^2 = 0$.

4.2 The n-th roots of any complex number

The same method finds the n-th roots of any non-zero complex number $w = \rho e^{i\alpha}$, where the Greek letter ρ (rho) stands for the modulus of w. The equation $z^n = w$ with $z = re^{i\theta}$ requires $r^n = \rho$ and $n\theta = \alpha + 2\pi k$, so the solutions are

$$z_k = \rho^{1/n}\, e^{i(\alpha + 2\pi k)/n}, \qquad k = 0, 1, \ldots, n-1. \tag{3}$$

Here $\rho^{1/n}$ is the ordinary positive n-th root of the positive number ρ. Every non-zero complex number therefore has exactly n different n-th roots, spaced evenly on the circle of radius $\rho^{1/n}$; each of them is the first one, z_0, multiplied by a power of ζ_n, because multiplying by ζ_n adds $2\pi/n$ to the argument and leaves the modulus unchanged.

WORKED EXAMPLE The cube roots of $8i$

The number $8i$ has modulus 8 and argument $\pi/2$, so $8i = 8e^{i\pi/2}$. By Equation (3) its cube roots have modulus $8^{1/3} = 2$ and arguments $(\pi/2 + 2\pi k)/3$ for $k = 0, 1, 2$, which are $\pi/6$, $5\pi/6$, and $3\pi/2$, or $30°$, $150°$, and $270°$. With $\cos 30° = \sqrt{3}/2$ and $\sin 30° = 1/2$ the three roots are

$$z_0 = 2e^{i\pi/6} = \sqrt{3} + i, \qquad z_1 = 2e^{5i\pi/6} = -\sqrt{3} + i,$$
$$z_2 = 2e^{3i\pi/2} = -2i.$$

The last one is the easiest to check, since $(-2i)^3 = -8i^3 = -8(-i) = 8i$. For the first, $(\sqrt{3}+i)^2 = 3 + 2\sqrt{3}\,i - 1 = 2 + 2\sqrt{3}\,i$, and multiplying by $\sqrt{3}+i$ gives $2\sqrt{3} + 2i + 6i + 2\sqrt{3}\,i^2 = 8i$. The cube roots of $8i$ are $\underline{\underline{\sqrt{3}+i,\ -\sqrt{3}+i, \text{ and } -2i}}$.

4.3 Cardano's formula completed

Section 1 left two questions open. A complex number has three cube roots, so Cardano's formula offers more than one value, and the equation $x^3 = 15x + 4$ has two further roots that the formula seemed not to produce. By Equation (3), the cube roots of $2 + 11i$ are the first one multiplied by the powers of ω, namely $2 + i$, $\omega(2 + i)$, and $\omega^2(2 + i)$, and in the same way the cube roots of $2 - 11i$ are $2 - i$, $\omega(2 - i)$, and $\omega^2(2 - i)$. The rule $3uv = -p = 15$, that is $uv = 5$, decides which pairs belong together. Since $\omega \cdot \omega^2 = 1$, the product $\omega(2 + i) \cdot \omega^2(2 - i)$ is $(2 + i)(2 - i) = 5$, whereas $\omega(2 + i) \cdot \omega(2 - i)$ would be $5\omega^2$, so the pairs are $(2 + i, 2 - i)$, $(\omega(2 + i), \omega^2(2 - i))$, and $(\omega^2(2 + i), \omega(2 - i))$. The three sums $u + v$ are the three roots of the cubic. The first sum is 4. For the second, $\omega(2 + i) = \frac{1}{2}(-1 + \sqrt{3}\,i)(2 + i)$, which multiplies out to $\frac{1}{2}\big((-2 - \sqrt{3}) + (2\sqrt{3} - 1)i\big)$, and its partner $\omega^2(2 - i)$ is its conjugate, because $\omega^2 = \overline{\omega}$ and conjugation respects products, so the sum is twice the real part, $-2 - \sqrt{3}$. The third pair gives $-2 + \sqrt{3}$ in the same way. All three roots of $x^3 = 15x + 4$ come out of the formula, and all three are real, although every one of them was computed through $\sqrt{-121}$.

The example is typical. Whenever a cubic with real coefficients has three different real roots, the quantity under the square root in Cardano's formula is negative, and the formula cannot avoid complex numbers; this situation is called the **casus irreducibilis**, Latin for the irreducible case. A theorem proved by Pierre Wantzel (1814-1848) in 1843, whose proof lies beyond this unit, says that when such a cubic has rational coefficients and no rational root, none of its roots can be written with real radicals alone, that is, with square roots, cube roots, and higher roots of real numbers, so no cleverer formula avoids the detour through $\mathbb{C}$. Bombelli's equation escapes the theorem only through its rational root 4. The real numbers are not enough to solve real cubics, which is the strongest argument for $\mathbb{C}$ that the sixteenth century could have had.

5 The fundamental theorem of algebra

5.1 Every polynomial has a root

Unit 2 proved that a polynomial of degree n has at most n roots and promised that over $\mathbb{C}$ the count is exact.

THEOREM The fundamental theorem of algebra

Every polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$ of degree $n \geq 1$ with complex coefficients has at least one root in $\mathbb{C}$. Consequently f factors completely into linear factors,

$$f(x) = a_n(x - r_1)(x - r_2)\cdots(x - r_n),$$

and has exactly n roots in $\mathbb{C}$ when each root is counted as often as its factor appears, that is, with multiplicity.

The second sentence follows from the first by the factor theorem of Unit 2, whose proof used only division with remainder and therefore works for complex coefficients as well. If r_1 is a root of f, then $f(x) = (x - r_1)g(x)$ with $\deg g = n - 1$; if $n - 1 \geq 1$, the theorem gives a root r_2 of g, and so on, until a constant is left, which must be the leading coefficient a_n.

Jean le Rond d'Alembert (1717-1783) published the first serious attempt at a proof in 1746, which contained gaps: among them, he assumed that the modulus $|f(z)|$ attains a smallest value somewhere, a fact that is true but was not proved until the following century. Gauss gave a proof in his doctoral dissertation of 1799 and criticized all earlier attempts; by modern standards his first proof has a gap as well. The shortest of the classical proofs, published by Argand in 1814, repairs d'Alembert's idea; it too takes the smallest value for granted, and that step was secured only by the analysis of the 1870s. Every proof uses more than algebra, because the theorem depends on a property of the real numbers

that algebra alone does not see, namely that they have no gaps.

A full proof is beyond this unit, but the idea of one proof can be seen in a picture. Consider the polynomial $f(z) = z^n + (\text{terms of lower degree})$, and watch where f sends a circle $|z| = R$ as the point z runs around it once. For a very large R the leading term dominates, and z^n runs n times around the origin, because its argument is n times the argument of z; the lower terms disturb this curve a little but cannot pull it across the origin, because for large R their combined size is far smaller than R^n, so the image curve still winds n times around 0. For a very small R the image stays close to the constant $f(0)$ and, if $f(0) \neq 0$, does not wind around 0 at all. As R shrinks from very large to very small, the image curve changes continuously, and the number of times it winds around 0 can only change when the curve passes through 0. Somewhere on the way, therefore, some z has $f(z) = 0$. Making the words "continuously" and "winds around" precise belongs to analysis and lies beyond this program.

5.2 Consequences for real polynomials

Most polynomials that arise in practice have real coefficients, and for them the theorem says something about the shape of the complex roots.

KEY RESULT Complex roots of real polynomials come in conjugate pairs

If f has real coefficients and $f(z) = 0$, then $f(\overline{z}) = 0$.

A root that is not real therefore brings its conjugate along, with the same multiplicity, and the factor $(x-z)(x-\overline{z}) = x^2 - 2\Re(z)\,x + |z|^2$ has real coefficients, because $z+\overline{z} = 2\Re(z)$ and $z\overline{z} = |z|^2$. Every polynomial of degree at least 1 with real coefficients is a product of real linear factors and real quadratic factors without real roots, and every real polynomial of odd degree has at least one real root.

PROOF Proof of the conjugate pair rule

Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ with real coefficients, and let $f(z) = 0$. Taking the conjugate of the equation $a_n z^n + \cdots + a_1 z + a_0 = 0$ and using that conjugation respects sums and products (Section 2) gives $\overline{a_n}\,\overline{z}^{\,n} + \cdots + \overline{a_0} = \overline{0} = 0$. A real number is its own conjugate, so $\overline{a_k} = a_k$ for every k, and the equation reads $f(\overline{z}) = 0$. Hence $\overline{z}$ is a root, which is what was to be shown. For the statement about multiplicities, divide f by the real quadratic $(x - z)(x - \overline{z})$; the quotient again has real coefficients, and the argument repeats on it. Finally, the roots of a real polynomial of odd degree are its real roots together with some number of conjugate pairs; the pairs account for an even number of roots, so at least one root is left over and is real, which completes the proof.

The quadratic formula shows the pairs explicitly. By completing the square (Unit 1), the roots of $x^2 + bx + c$ with real b and c are $-b/2 \pm \sqrt{b^2/4 - c}$, and when $b^2/4 - c$ is negative, the square root of the negative number $-(c - b^2/4)$ is $i\sqrt{c - b^2/4}$. The two roots are then

$$x = -\frac{b}{2} \pm i\sqrt{c - \frac{b^2}{4}},$$

a conjugate pair with real part $-b/2$. For example, $x^2 - 2x + 2$ has the roots $1 \pm i$, and indeed $(x - 1 - i)(x - 1 + i) = (x - 1)^2 + 1 = x^2 - 2x + 2$.

CHECK YOUR UNDERSTANDING Question 4

A polynomial of degree 7 with real coefficients has the roots $2+3i$ and $1-i$. What is the smallest number of real roots it can have, counted with multiplicity?

(A) 3
(B) 1
(C) 5
(D) 2

The solution is at the end of the unit.

WORKED EXAMPLE Factoring x^4+4 over the real numbers

The polynomial x^4+4 has no real root, since $x^4 \geq 0$ for every real x. Its complex roots are the fourth roots of $-4 = 4e^{i\pi}$, which by Equation (3) have modulus $4^{1/4} = \sqrt{2}$ and arguments $\pi/4$, $3\pi/4$, $5\pi/4$, and $7\pi/4$. Since $\sqrt{2}\cos(\pi/4) = \sqrt{2}\sin(\pi/4) = 1$, the four roots are $1+i$, $-1+i$, $-1-i$, and $1-i$, which are the corners of a square. They form the two conjugate pairs $1 \pm i$ and $-1 \pm i$, and each pair gives a real quadratic factor by the rule $(x-z)(x-\overline{z}) = x^2 - 2\Re(z)\,x + |z|^2$ with $|z|^2 = 2$, the first pair x^2-2x+2 and the second x^2+2x+2:

$$x^4+4 = (x^2-2x+2)(x^2+2x+2).$$

Multiplying out confirms it, since the product is $(x^2+2)^2 - (2x)^2 = x^4+4x^2+4-4x^2$. The factorization over $\mathbb{R}$ is $\underline{\underline{x^4+4 = (x^2-2x+2)(x^2+2x+2)}}$.

Where do the roots of polynomials lie? Figure 5 answers the question for one family, all monic polynomials $x^4+ax^3+bx^2+cx+d$ whose four coefficients are integers between -2 and 2; there are $5^4 = 625$ of them, one for each choice of the four coefficients from five values. Their $4 \cdot 625 = 2500$ roots, counted with multiplicity, fall on 1725 different points, and only 571 of those lie on the real axis. Three features of the picture are theorems. The cloud is symmetric about the real axis by the conjugate pair rule. No root lies farther than 3 from the origin. For $|z| > 3$ the factor $|z|-1$ is larger than 2, so the identity $|z|^4 - 1 = (|z|-1)(|z|^3+|z|^2+|z|+1)$ gives $|z|^4 > 2(|z|^3+|z|^2+|z|+1)$. The right side is at least $|az^3+bz^2+cz+d|$, because no coefficient is larger than 2 in size, so the term z^4 is larger in modulus than everything it would have to cancel, and $f(z)$ cannot be 0. And the real axis is crowded because 486 of the 625 polynomials have real roots (44 have four and 442 have two, counted with multiplicity), while 139 have none.

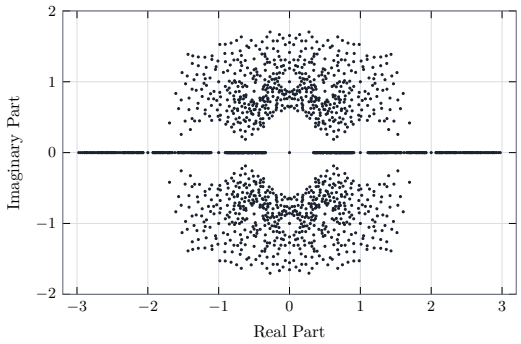


Figure 5. The roots of all 625 monic polynomials of degree 4 with coefficients in $\{-2,-1,0,1,2\}$, computed numerically and plotted in the complex plane. The cloud is symmetric about the real axis, and no root has modulus larger than 3. Own computation.

6 Iteration in the plane

6.1 The orbit of 0 under $z \mapsto z^2 + c$

A function from $\mathbb{C}$ to $\mathbb{C}$ can be applied to its own output, again and again. For a complex number c, consider the map $z \mapsto z^2 + c$ and apply it to the starting point 0, then to the result, and so on. The sequence $z_0 = 0$, $z_1 = c$, $z_2 = c^2 + c$, $z_3 = (c^2 + c)^2 + c$, and so on, is called the **orbit** of 0. For three values of c its first eight points are

$$\begin{aligned} c = -1: &\quad 0,\ -1,\ 0,\ -1,\ 0,\ -1,\ 0,\ -1,\ \ldots \\ c = 1: &\quad 0,\ 1,\ 2,\ 5,\ 26,\ 677,\ 458{,}330,\ \ldots \\ c = i: &\quad 0,\ i,\ -1+i,\ -i,\ -1+i,\ -i,\ -1+i,\ -i,\ \ldots \end{aligned}$$

where each term is the square of the one before plus c; for $c = i$, for example, $z_2 = i^2 + i = -1 + i$ and $z_3 = (-1 + i)^2 + i = -i$, and for $c = 1$ the eighth term is already 210,066,388,901, about 2×10^{11}.

The three orbits show the two things that can happen. For $c = -1$ and $c = i$ the orbit stays inside a bounded region of the plane for ever, and for $c = 1$ it escapes to infinity. The **Mandelbrot set** is the set of all c for which the orbit of 0 stays bounded, so -1 and i belong to it and 1 does not. The picture on the title page of this unit shows it in black. A test that decides most cases quickly is that once some $|z_n|$ exceeds 2 and $|c| \le 2$, the orbit escapes, because the modulus of a sum is at least the difference of the two moduli, as two arrows placed head to tail show, so that $|z_{n+1}| \ge |z_n|^2 - |c| \ge |z_n|^2 - 2 > |z_n|^2 - |z_n| = |z_n|(|z_n| - 1) > |z_n|$, and the moduli keep growing by a factor larger than 1. The picture is made by running the orbit of 0 for every c on a grid and colouring c by the number of steps until $|z_n| > 2$, black if that never happens.

6.2 Newton's method in the plane

Unit 2 ran Newton's method on $x^3 - 2$ along the real line; the same iteration makes sense in the plane and divides it into regions in a second way. For $f(z) = z^3 - 1$ the formal derivative is $3z^2$, and the step of Newton's method is

$$z_{k+1} = z_k - \frac{z_k^3 - 1}{3z_k^2} = \frac{2z_k^3 + 1}{3z_k^2}.$$

The polynomial has three roots, the cube roots of unity 1, ω, and ω^2, and a starting point near any one of them is pulled towards it. For example, started at $z_0 = 1 + i$ the method produces $0.667 + 0.500i$, $0.579 - 0.127i$, $1.247 + 0.314i$, $1.009 + 0.114i$, $0.988 + 0.004i$, and $1.000 - 0.000i$ to three decimals after six steps, so it converges to the root 1. Figure 6 colours every starting point by the root it reaches. The three regions, called the **basins of attraction** of the three roots, are not separated by three simple curves. Wherever two basins seem to meet, the third one appears between them in a chain of ever smaller patches, and every point on the boundary of one basin is on the boundary of all three. The boundary is the Julia set of the Newton map, a name explained in the excursion below, and a starting point on it never settles down. The picture also shows the limits of the method: a starting point in one of the small patches converges to a root far away rather than to the nearest one, points near the boundary need many steps, and for degree 4 and higher the situation is worse still, as the spotlight below explains.

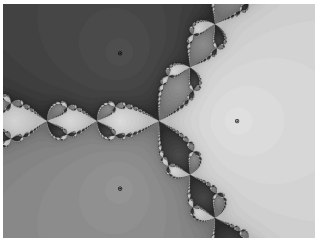


Figure 6. The basins of attraction of Newton's method for $z^3 - 1$ on $-2 \le \Re z \le 2$, $-1.5 \le \Im z \le 1.5$: each starting point is shaded by the root it converges to (lightest for the root 1), and the roots are marked by circles. Own computation.

EXCURSION *Fatou, Julia, and the first pictures*

The study of iterated maps of the plane began during the First World War, when the Académie des Sciences in Paris set iteration as the subject of its Grand Prix of 1918. Pierre Fatou (1878-1929) and Gaston Julia (1893-1978) independently developed the theory in long papers; Julia's won the prize, and Fatou's appeared in 1919 and 1920. They showed, without a single picture, that for a polynomial map the plane splits into a region of orderly behaviour, now called the Fatou set, and a boundary of chaotic behaviour, the Julia set, where the map is sensitive to the smallest change of the starting point.

The pictures came with computers. The first printed image of the set of Section 6 appeared in 1981, in a paper by two mathematicians who had drawn it in 1978 while studying a different problem, and Benoit Mandelbrot (1924-2010) made detailed images of it at IBM in 1980 and published the first of them that year. Its boundary is a thin set, one-dimensional in the topological sense (the sense in which a line is one-dimensional and a surface two-dimensional), yet Mitsuhiro Shishikura (Kyoto University) proved in a paper published in 1998 that it is as complicated as such a set can possibly be: its Hausdorff dimension, a measure of how much a set fills the plane, is 2, the value of a surface.

RESEARCH SPOTLIGHT

Curtis T. McMullen *Harvard University*

What is known, and what is still open, about iterating $z^2 + c$

Curtis McMullen, who received the Fields Medal in 1998 in part for his work on iteration in the plane, wrote a survey in 1994 for a general mathematical audience on what is known about the orbit of 0 and the Mandelbrot set. It defines the Julia set as the closure of the repelling periodic points of a map (points that return to themselves after some steps and push their neighbours away), defines the Mandelbrot set as Section 6 does, and is organized around one conjecture: for every c in the interior of the Mandelbrot set, the orbit of 0 is attracted to a periodic cycle, as it is for $c = -1$. This would follow from local connectivity, a property of the boundary conjectured by Adrien Douady (1935-2006) and John Hubbard (Cornell University), and the conjecture is still open in 2026. McMullen's separate theorem of 1987 says that no purely iterative algorithm finds a root of every polynomial of degree 4 or higher from almost every starting point. For cubics such an algorithm exists, but Newton's method is not it, because for some cubics whole regions of starting points cycle for ever. The survey's first two pages can be read after this unit.

Read the paper: McMullen (1994), Frontiers in complex dynamics, Bulletin of the American Mathematical Society 31, 155 to 172

KEY POINTS

- A complex number is a pair of real numbers $a + bi$ with $i^2 = -1$; all laws of arithmetic hold.
- The conjugate $\overline{z} = a - bi$ and the modulus $|z| = \sqrt{a^2 + b^2}$ satisfy $z\overline{z} = |z|^2$, so $1/z = \overline{z}/|z|^2$.
- In polar form, multiplication multiplies the moduli and adds the arguments, so $z^n = r^n e^{in\theta}$.
- The n-th roots of unity ζ_n^k solve $z^n = 1$, form a regular n-gon on $|z| = 1$, and add up to 0.
- The fundamental theorem of algebra says that every polynomial of degree $n \geq 1$ has n complex roots counted with multiplicity; for real coefficients the non-real roots come in conjugate pairs.
- The c for which the orbit of 0 under $z \mapsto z^2 + c$ stays bounded form the Mandelbrot set; the basins of attraction of Newton's method share one boundary.

Solutions to the questions in the text: 1-D, 2-B, 3-C, 4-B. The online version of this unit explains each answer.