0 The Origins of Complex Analysis, and Its Challenge to Intuition

In a lecture in 1886, Leopold Kronecker asserted that the integers are made by God and all the rest is the work of Man (Gray [7]). If so, complex numbers are certainly one of humanity's most intriguing mathematical artefacts. For centuries they have been a wonder to mathematicians and philosophers alike. It took nearly 300 years from their first appearance in Girolamo Cardano's *Ars Magna* (The Great Art) to the publication of a formal definition that satisfies modern standards of rigour. Building on such foundations, the initiated reader might be forgiven for thinking that complex analysis must be an incredibly complicated theory. Yet here we come to a historical puzzle. Although it took nearly three centuries to obtain a satisfactory treatment of complex *numbers*, it then took less than a tenth of that time to complete a major part of complex *analysis*, which is far more sophisticated and extensive.

Obviously the numbers must come first, or there is nothing to do analysis with, but the timescale is surprising. A possible explanation is that setting up the foundations adequately involved deep problems of a philosophical nature: it took a long time to come to grips with them, but once the 'breakthrough' had occurred, the further development was easy by comparison.

History suggests otherwise.

0.1 The Origins of Complex Numbers

Cardano's celebrated *Ars Magna* of 1545 is one of the most important early algebra texts. Diophantus's *Arithmetica* of about 250 discussed the solution of equations and introduced a rudimentary form of algebraic notation. Muhammad al-Khwarizmi's *Al-kitab al-mukhtasar fi hisab al-gabr wa'l-muqabala* (The Compendious Book on Calculation by Completion and Balancing) appeared around 820. Its translation into Latin as *Liber Algebrae et Almucabola* gave us the word 'algebra'. Al-Khwarizmi's discussion was verbal, with no symbols but occasional diagrams.

Cardano introduced a systematic algebraic notation, very different from what we use today. He used this to present the newly discovered solutions of cubic and quartic equations. His book contained the solution of cubics discovered by Scipione del Ferro around 1500, and independently by Niccolo Fontana (nicknamed 'Tartaglia', the stammerer) around 1535. The high point of the text is the solution of quartic equations found by Cardano's student Lodovico Ferrari. The tangled tale of alleged duplicity and public

controversy that accompanied these discoveries can be found in Stewart [19, 20] and other historical sources.

Ars Magna also discussed the simultaneous equations

$$x + y = 10$$
$$xy = 40$$

and obtained a solution (in modern notation) of the form

$$x = 5 + \sqrt{-15} \qquad y = 5 - \sqrt{-15}$$

Cardano gave no interpretation for the square root of a negative number, but he did observe that, on the assumption that the quantities obey the usual algebraic rules, we can check that they satisfy the equations. His attitude to the discovery was dismissive: 'So progresses arithmetic subtlety, the end of which ... is as refined as it is useless.'

In the same book he observed that applying Tartaglia's formula to the cubic equation

$$x^3 = 15x + 4 \tag{0.1}$$

leads to the solution

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

in contrast to the obvious answer $x = 4$.

In both instances there was a conflict between the intuition about numbers that mathematicians had built up over the years, and the formal behaviour of the symbolic manipulations that Cardano was carrying out. It took centuries for mathematicians to extend the number concept and develop a refined intuition in which Cardano's observations make sense. The first step happened not long after, however. Raphael Bombelli (1526–73) suggested a way to reconcile the two solutions of (0.1) by manipulating the 'impossible' roots *as if they are ordinary numbers*. Since

$$(2 \pm \sqrt{-1})^3 = 2 \pm \sqrt{-121}$$

Cardano's expression becomes

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

and the 'impossible' root is just the familiar root in a complex disguise. Bombelli's work was the first hint that complex numbers can prove useful in solving real mathematical problems. But the message took a long time to sink in.

In *La Géometrie* (1637), René Descartes made the distinction between 'real' and 'imaginary' numbers, interpreting the occurrence of imaginaries as a sign that the problem concerned is insoluble, an opinion shared by Isaac Newton at a later date. However, this view sits uneasily with Bombelli's realisation that a formula involving complex numbers sometimes leads to a real solution, suggesting that the issue is not that simple.

John Wallis [25] represented a complex number geometrically in his *Algebra* of 1685. On a fixed line the real part of the number was measured off (in the direction given by its sign); then the imaginary part was measured off at right angles, Figure 1. But this idea was largely forgotten.
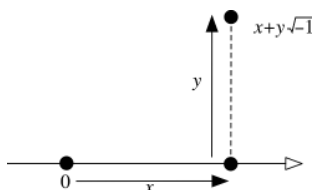
Figure 1 Wallis's geometric representation of a complex number.

In 1702 John Bernoulli was evaluating integrals of the form

$$\int \frac{\mathrm{d}x}{ax^2 + bx + c}$$

by partial fractions. Using the philosophy that complex numbers can be manipulated like real ones, he wrote the integrand as

$$\frac{1}{ax^2 + bx + c} = \frac{A}{x - \alpha} + \frac{B}{x - \beta}$$

(using modern notation) where α, β are the roots of the quadratic denominator, and obtained the integral in the form

$$A \log(x - \alpha) + B \log(x - \beta)$$

His bold decision to use the same method when the quadratic had no real solutions led to logarithms of complex numbers. But what were they? Both Bernoulli and Leibniz used the method, and by 1712 they were engaged in controversy. Leibniz asserted that the logarithm of a negative number is complex, while Bernoulli insisted it is real. Bernoulli argued that, since

$$\frac{\mathrm{d}(-x)}{-x} = \frac{\mathrm{d}x}{x}$$

it follows by integration that $\log(-x) = \log(x)$. Leibniz, on the other hand, insisted that the integration was correct only for positive x. Once again, formal calculations that seemed sensible were in conflict with intuition.

Leonhard Euler resolved the controversy in favour of Leibniz in 1749, pointing out that integration requires an arbitrary constant

$$\log(-x) = \log(x) + c$$

a point that Bernoulli had ignored. By formally manipulating expressions involving complex numbers, Euler derived a host of theoretical relations, including the famous formula of 1748:

$$\mathrm{e}^{\mathrm{i}\theta} = \cos \theta + \mathrm{i} \sin \theta \tag{0.2}$$

Putting $\theta = \pi$ we find

$$\mathrm{e}^{\mathrm{i}\pi} = -1 \tag{0.3}$$

a fantastic relation that blends the three mathematical symbols e, i, and π in one surprising equation. The formula (0.3) is widely referred to as Euler's formula, although he never published it explicitly. He did publish (0.2), of which it is a simple corollary, and this is also known as Euler's formula. However, a formula equivalent to (0.2) had been found earlier by Roger Cotes in 1714.

Extending the theory of logarithms to the complex case by defining

$$\log z = w \text{ if and only if } e^w = z$$

we obtain other intriguing results. Formal manipulation gives

$$e^{\log z + m\pi i} = e^{\log z}(e^{\pi i})^m = z \cdot (-1)^m$$

For an even integer $m = 2n$ this gives

$$e^{\log z + 2n\pi i} = z$$

So $\log z + 2n\pi i$ is also a logarithm of z: the complex logarithm is *many-valued*. For an odd integer $m = 2n + 1$ we have

$$e^{\log z + (2n+1)\pi i} = -z$$

whence

$$\log(-z) = \log z + (2n + 1)\pi i$$

This resolves the Leibniz–Bernoulli controversy: if x is real and positive, then $\log(-x)$ must be complex.

As mathematicians refined their intuition to encompass complex numbers, everything started to fit together and make sense. The theory of complex numbers grew ever more fascinating. What was lacking was an interpretation that explained precisely what these entities are – a formal counterpart to the newly extended intuitions.

In 1797 Caspar Wessel published a paper in Danish describing the representation of a complex number as a point in the plane. It went almost totally unnoticed until a French translation was published a hundred years later. Meanwhile the idea was attributed to Jean-Robert Argand, who wrote it up independently in 1806. Since that time the geometric interpretation of complex numbers has commonly become known as the Argand diagram.

Another pioneer of the theory of complex numbers was Carl Friedrich Gauss. In his doctoral dissertation of 1799 he addressed a problem that had concerned mathematicians since the early eighteenth century. Initially it had been widely believed that, just as the solutions of real quadratic equations could lead to new 'complex' numbers, so would solutions of equations with complex coefficients lead to even more kinds of new numbers. But Jean d'Alembert (1717–83) conjectured that complex numbers alone suffice. Gauss confirmed this in the 'fundamental theorem of algebra' – every polynomial equation has a complex root. At first he proved it in the purely real form that any real polynomial factorises into linear and quadratic factors, avoiding explicit use of imaginaries; later he treated the general case. By 1811 he viewed the complex numbers as points in the plane, saying so in a letter to Friedrich Bessel. In 1831 he published full

details of his representation of complex numbers, which had begun to acquire an air of respectability.

In 1837, nearly three centuries after Cardano's use of 'imaginary numbers', William Rowan Hamilton published the definition of complex numbers as ordered pairs of real numbers subject to certain explicit rules of manipulation. (In the same year Gauss wrote to Wolfgang Bolyai that he had developed the same idea in 1831.) At last this placed the complex numbers on a firm *algebraic* basis.

0.2 The Origins of Complex Analysis

Unlike the gradual emergence of the complex *number* concept, the development of complex *analysis* seems to have been the direct result of the mathematician's urge to generalise. It was sought deliberately, by analogy with real analysis. However, the mathematicians of the period tended to assume that everything in real analysis must automatically be meaningful in the complex case, so the main question must be how 'the' complex version behaves. That there might not *be* a complex version, or several alternatives, was seldom appreciated, as the controversy over $\log(-x)$ illustrates.

As noted above, there are early traces of analytic operations on complex functions in the work of Bernoulli, Leibniz, Euler, and their contemporaries.

In his 1811 letter to Bessel, Gauss shows that he knew the basic theorem on complex integration around which complex analysis was subsequently built. In real analysis, when we integrate a function f between limits a and b, to get

$$\int_a^b f(x)\mathrm{d}x$$

the limits fully specify the integral. But in the complex case, where a and b represent points in the plane, it is also necessary to specify a definite path from a to b, and to 'integrate along the path'. The question is: to what extent does the value of the integral depend on the chosen path?

Gauss says:

I affirm now that the integral $\int f(x)\mathrm{d}x$ has only one value even if taken over different paths, provided $f(x)\dots$ does not become infinite in the space enclosed by the two paths. This is a very beautiful theorem whose proof $\dots$ I shall give on a convenient occasion.

It seems the occasion never arose. The crucial step of publishing a proof of this result was taken in 1825 by the man who was to occupy centre stage during the first flowering of complex analysis: Augustin-Louis Cauchy. After him, this result is called 'Cauchy's Theorem'. In Cauchy's hands the basic ideas of complex analysis rapidly emerged. For a complex function to be differentiable, it must have a very specialised nature: its real and imaginary parts must satisfy certain properties called the Cauchy–Riemann Equations. He showed that contour integrals of differentiable functions have the property noted privately by Gauss. Further, if an integral is computed along a path that winds round points where the function becomes infinite, Cauchy showed how to compute this integral using the 'theory of residues'. The latter requires no more than the calculation of a

constant, called the 'residue' of the function, at each exceptional point, and knowing how many times the paths winds around that point. The precise route of the path does not matter at all – only how it winds round these exceptional points.

Power series turned out to be important in the theory, and other workers extended these ideas. Pierre-Alphonse Laurent introduced 'Laurent series' involving negative powers in 1843. In this formulation, near an exceptional point z_0, a differentiable function is expressed as a sum of two series

$$f(x) = [a_0 + a_1(z - z_0) + \cdots + a_n(z - z_0)^n + \cdots]$$
$$+ [b_1(z - z_0)^{-1} + \cdots + a_n(z - z_0)^{-n} + \cdots]$$

The residue of $f(z)$ at $z = z_0$ is then just the coefficient b_1. Using the theory of residues, the computation of complex integrals often proved to be far simpler than could ever have been dreamed.

Cauchy's definition of analytic ideas such as continuity, limits, derivatives, and so on, were not the same as those we use today. He based them on infinitesimal notions, which fell into disrepute in the late nineteenth century – though recent developments in 'non-standard analysis', and a new theory we present in Chapter 15, show that we may have been over-hasty in judging Cauchy's ideas. Moreover, Cauchy's concept of 'infinitesimal' was a variable quantity that approaches zero as closely as we please, not a fixed quantity. See Tall and Katz [24] for detailed discussion and educational implications.

A rigorous treatment was devised by Karl Weierstrass (1815–97) using definitions which are still regarded as fundamental, the 'epsilon-delta' formulation. Weierstrass founded his whole approach on power series. However, the geometric viewpoint was sorely lacking in his work (at least as published). This deficiency was remedied by far-reaching ideas introduced by Bernhard Riemann (1826–66). In particular, the concept of a 'Riemann surface', which dates from 1851, treats many-valued functions by splitting the complex plane into multiple layers, on each of which the function is single-valued. The crucial feature is how the layers join up topologically.

From the mid-nineteenth century onwards, the progress of complex analysis has been strong and steady, with many far-reaching developments. The fundamental ideas of Cauchy remain, now refined and clothed in more recent topological language. The abstruse invention of complex numbers, once described by our mathematical forebears as 'impossible' and 'useless', has become part of an aesthetically satisfying theory with eminently practical applications in aerodynamics, fluid mechanics, electronics, control theory, and many other areas.

Since the first edition of this book, formal theory has also evolved so that Cauchy's ideas of infinitesimals can be visualised as points on an extended number line, which we describe in our new Chapter 15.

0.3   The Puzzle

We return to our historical puzzle. *Why was the development of complex numbers so laboured and hesitant, whereas that of complex analysis was explosive?* We suggest

a possible answer (only personal opinion and thus open to dispute). It is somewhat different from the 'foundations + breakthrough' explanation offered earlier.

Looking at the early history of complex numbers, the overall impression is of countless generations of mathematicians beating out their brains against a brick wall in search of – what? A triviality. The definition of complex numbers as ordered pairs of points (x, y), or as points in the plane, was obtained over and over and over again. It is even implicit in Bombelli's work; it is there for all to see in Wallis's; it crops up again by way of Wessel, Argand, and Gauss. Morris Kline remarks on page 629 of [11]:

That many men – Cotes, de Moivre, Euler, and Vandermonde – really thought of complex numbers as points in the plane follows from the fact that all, in attempting to solve $x^n - 1 = 0$, thought of solutions . . . as the vertices of a regular polygon.

If the problem has such a simple solution, why was this not recognised sooner?

Perhaps the early mathematicians were not so much seeking a *construction* for complex numbers as a *meaning*, in the philosophical sense: 'what *are* complex numbers?' However, the development of complex *analysis* showed that the complex number concept was so useful that no mathematician in his right mind could possibly ignore it. The unspoken question became 'what can we *do* with complex numbers?', and once that had been given a satisfactory answer, the original philosophical question evaporated. There was no jubilation at Hamilton's incisive answer to the 300-year old foundational problem – it was 'old hat'. Once mathematicians had woven the notion of complex numbers into a powerful coherent theory, the fears that they had concerning the existence of complex numbers became unimportant, because mathematicians lost interest in that issue.

There are other cases of this nature in the history of mathematics, but perhaps none is more clear-cut. As time passes, the cultural world-view changes. What one generation sees as a problem or a solution is not interpreted in the same way by a later generation. It is worth bearing this in mind when thinking about the historical development of mathematics. To interpret history solely from the viewpoint of the current generation may easily lead to distortion and misinterpretation.

What this explanation omits is any discussion of *why* mathematicians lost interest in the meaning of complex numbers. And that leads to a question that sheds a different light on the historical development, which we now discuss.

0.4   Is Mathematics Discovered or Invented?

Students trying to understand new concepts are in a similar position to the pioneers who first investigated them. At any stage in our education, we build not just on our current knowledge, but on a variety of beliefs and intuitions that are often vague, and may not be consciously recognised. As a trivial example, children familiar with counting numbers may find it hard to adapt their thinking to negative numbers, or rational numbers. When faced with questions like 'what is 3 minus 7?' or 'what is 3 divided by 7', intuition based solely on whole numbers leads to the answer 'can't be done'. That makes it hard

to understand −4 or 3/7. In fact, these is not really trivial examples, because the world's top mathematicians, centuries ago, were just as confused by the question 'what is the square root of minus one?' Even their terminology – 'imaginary' – reveals how puzzled they were. Intuitively they considered numbers to be 'real' – not in the sense we now use to distinguish real from complex, but as direct representations of real measurements. The new objects behaved like numbers in many ways, but they seemed not to correspond directly to reality.

In such circumstances, it can be tempting to discard existing intuition completely. But it is more sensible to adapt the intuition to fit the new circumstances. It is much easier to do arithmetic with negative numbers or fractions if you remember how to do it with whole numbers; it is much easier to do algebra with complex numbers if you bear in mind how to do it with real numbers. So the trick is to sort out which aspects of existing intuition remain valid, and which need to be refined into a broader kind of understanding.

One way to approach this issue is to take seriously a question that is often asked but seldom answered satisfactorily: is mathematics discovered or invented? One answer is to dismiss the question, and agree that neither word is entirely appropriate; moreover, they are not mutually exclusive. Most discoveries have elements of invention, most inventions have elements of discovery. Galileo would not have discovered the moons of Jupiter without the invention of the telescope. The telescope could not have been invented without discovering that sand could be melted to make glass.

But leaving such quibbles aside, we can make a rough distinction between discovery, which is finding something that is *already there* but has not hitherto been noticed, and invention, which is a creative act that brings into being something that has not previously existed. There is a case to be made that in this sense, mathematicians invent new concepts but then discover their properties. For example, complex integration is all about 'paths' in the complex plane. Intuitively, a path is a line drawn by moving the hand so that the pencil remains in contact with the paper – no jumps. We might choose to formalise this notion as a continuous curve – the image of a continuous map from a real interval to the complex plane. We might be interested in how the pencil point moves along this curve, which requires the map itself, not just its image. Sometimes we might wish the path to be smooth – to have a well-defined tangent.

As it happens, we need all of these notions. Intuitively, they are all based on the same mental image. Formally, they are all very different. They have different definitions, different meanings, and different properties. A smooth path always has a meaningful length, for instance; a continuous path may not. The definitions we settle on in this book fit conveniently into the standard ideas of analysis, but they are not built into the fabric of the universe. We chose them, and by so doing we invent concepts such as 'path', 'curve', and 'smooth'.

On the other hand, once a concept has been invented, we cannot invent its *properties*. When we also invent the concept 'length', we discover that every smooth path has finite length. We cannot 'invent' a theorem that the length of a smooth path can be infinite. If we weaken 'smooth' to 'continuous', however, we can discover that infinite lengths are possible; indeed, 'length' need not have a sensible meaning at all. In short: invention

opens up new mathematical territory, but exploring it leads to discoveries. We may not know what things are present in the territory, but we do not get to choose them.

Sometimes – in fact, very often – we discover that our inventions have features that we neither expected nor intended them to have. We discover, perhaps to our dismay, that the image of a smooth path can have a right-angled corner, see Section 6.7. We did not expect that: a corner does not feel 'smooth'. But its possibility is a direct consequence of the definition we invented.

When this kind of thing happens, we have two choices. Accept the surprises as the price for having a nice, tidy definition; or rule them out by changing the definition – inventing a more comfortable alternative. In practice we often do both, by giving the alternative a different name. Here we could (and do) define a 'regular path' to be a smooth path $\gamma : [a, b] \to \mathbb{C}$ for which $\gamma'(t) \neq 0$ whenever $t \in [a, b]$. Now the image cannot have a sharp corner. On the other hand, every theorem about regular paths must now take account of the consequences of that extra condition. We also have to remember that some theorems may be valid for regular paths but not for smooth paths, and so on.

As we move from intuitive ideas to formal ones, we also refine our intuition so that it matches the formal theory better. Formal calculations start to make sense, not just as strings of symbols that follow from previous strings, but as meaningful statements that agree with our new intuitions. From this point of view, the history of complex analysis is the story of intuition co-evolving with an increasingly formal approach. This suggests that mathematicians lost interest in the meaning of complex numbers when they incorporated them into their intuitive assumptions and beliefs. With the apparent conflicts resolved by these refined intuitions, they were free to push the subject forward, no longer worried that it did not make logical sense.

When a mathematical area 'settles down' into a mature theory, there is a broad consensus that certain concepts provide the most convenient route through the material. These concepts then become standard – things like 'continuous', 'connected', and so on. They get taught in lecture courses and printed in books. We may start to feel that the standard definitions are the only reasonable ones. Even so, we are always free to work with different concepts if that seems sensible, or even to modify definitions while retaining the same name – though that can be dangerous. Today's concept of continuity is quite different from what it was in the time of Euler, but we use the same word; we just bear in mind that it now has a specific technical meaning. A historian reading Euler would need to be on their guard.

It is also worth remarking that many mathematical concepts seem more natural to us than others. Counting numbers are very natural (we even call them the 'natural numbers'). The number i was baffling for centuries (and was called 'imaginary' as a result). Our culture, our society, and even our senses, predispose us towards certain concepts. Euclid's points and lines correspond to early stages of the processing of images sent from the retina to the visual cortex. Newton's concept of acceleration being related to an applied force reflects the way our ears sense accelerations and make us 'feel' a push – a force.

It then becomes easy to imagine that mathematics somehow already exists in a realm outside the natural world. Even if humans invented numbers, in retrospect they seem

such a natural idea that surely they were just hanging around waiting to be invented. If so, that is more like discovery. This view is often called Platonism: the idea that mathematical concepts already exist in some ideal form in some kind of world outside the physical universe, and mathematicians merely discover how these ideal forms work. The contrary view is that mathematics is a shared human construct, *but* that construct is by no means arbitrary, because every new invention is made in the context of existing knowledge, and every new discovery must be logically valid.

A major theme of this book is that many apparently puzzling aspects of complex analysis can be made more intuitive by paying attention to the geometry of the complex plane (in a broad sense, including its topology). This brings one of the human brain's most powerful abilities, visual intuition, into play. For this reason, we draw a lot of pictures. However, a picture, and our visual intuition, can be misleading unless we examine the unstated assumptions that they involve. By doing so, we can refine out intuition and make it more reliable. For this reason, we do not just introduce important definitions and then deduce theorems that refer to them. We try to relate those definitions to intuition, to make the proofs easier to understand. Then we exhibit some of the positive results that arise, to convince you that the new concept is worth considering. And then . . . we show you that sometimes the formally defined concept does *not* behave the way intuition might suggest. Sometimes it turns out to be useful to strengthen the definition so that it matches intuition more closely. Sometimes we refine our intuition so that it matches the formal definition. Sometimes we can even do both, in which case we have to make some careful but useful distinctions.

The historical events sketched earlier in this chapter offer many examples of this process. The square root of minus one went from being a puzzling idea that seemed to have no meaning to one of the most important concepts in the whole of mathematics. Along the way, mathematicians' intuition for 'number' underwent a revolution. We can now to some extent short-circuit the historical debates – what were hang-ups then need not be hang-ups now – but when a new idea puzzles us, and doesn't seem to make sense until we finally sort it out, it is helpful to remember that the mathematical pioneers often experienced exactly the same feelings, for much the same reasons.

0.5   Overview of the Book

It is often useful to set the development of a mathematical theory in its historical context, but it is not always necessary to fight the historical battles again. In this text we give honour where we can to those pioneers who carved their way through uncharted mathematical territory. But more recent developments let us see the theory itself in a new light. To the modern ear the very *name* 'complex analysis' carries misleading overtones: it suggests complexity in the sense of complication. The older meaning, 'composite', was perhaps appropriate when the 'real part' of a complex number had a quite different status from that of the 'imaginary part'. But nowadays a complex number is a perfectly integrated whole. To think of complex analysis as if it were, so to speak, two copies of real analysis, is to place undue emphasis on the algebra at the expense of the geometry,