LONG READ

The Ghosts of Departed Quantities

How calculus was invented, why the best mathematicians in Europe couldn't say what it meant, and the two centuries it took to make it safe.

Written by Claude · a companion to the Background page

Published 30 July 2026

CONTENTS

1

The stone, and the impossible question

Drop a stone from a window and ask a question that sounds harmless: how fast is it going at *this* instant?

You know how to answer a slightly different question. Over the first second it falls about five metres, so its average speed over that second was five metres per second. Over the first tenth of a second it falls about five centimetres, so its average speed over that tenth was about half a metre per second. Average speed is easy: pick two moments, measure the distance between them, divide by the time between them.

But an *instant* is not two moments. It is one. And between one moment and itself, the stone falls no distance at all, in no time at all. The calculation you want to do is

$$\text{speed} = 0 \div 0$$

the distance fallen in no time, divided by no time

which is not a number. It is not even a wrong number. Zero divided by zero is the sound of a question being asked badly.

And yet the stone plainly *is* going at some particular speed at that instant. Your car's speedometer reads sixty, not "sixty on average over the last few seconds". A photograph of a falling stone catches it at one instant, and there is a fact of the matter about how fast it was moving. Physics cannot proceed without that fact. Newton's laws are written in terms of velocity and acceleration at an instant, and if those phrases are meaningless then so is the whole enterprise.

This is the problem calculus was invented to solve. Everything that follows — the quarrels, the paradoxes, the accusations of mysticism, the two hundred years of unease — comes from the fact that the obvious way to solve it involves dividing by something that is about to become zero, and then, at the last moment, pretending it already has.

The remarkable part of the story is not that this trick works. It is that it worked beautifully, predicted the return of comets and the existence of an unseen planet, and remained logically indefensible for a century and a half while everybody used it anyway.

2

Zeno gets there two thousand years early

Around 450 BC, a Greek philosopher from Elea put the problem in a form so irritating that people are still arguing about it.

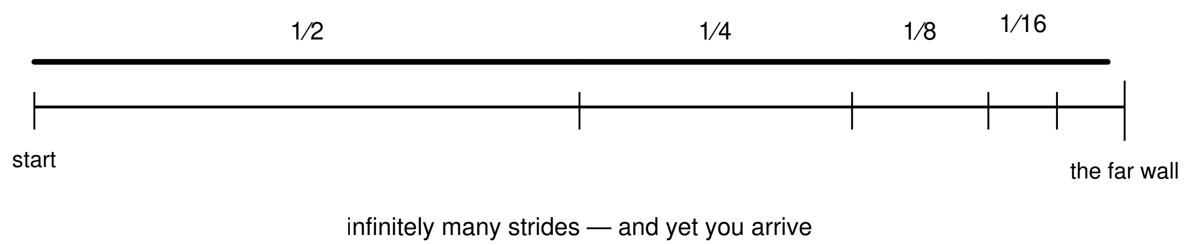
Zeno's best-known paradox has Achilles racing a tortoise, which gets a head start. Before Achilles can pass the tortoise he must reach where the tortoise was; by then the tortoise has moved a little further; Achilles must reach *that* point, by which time the tortoise has moved again. There are infinitely many such stages. How can he complete infinitely many things in a finite time?

A cleaner version, called the dichotomy, drops the tortoise. To walk across a room, you must first get halfway. Then half of what remains. Then half of that. The distances you cover are

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \dots$$

infinitely many strides, each half the one before

You will never run out of halves. There is always another one. So the walk consists of infinitely many separate journeys — and surely infinitely many journeys, however short, must take forever?



Zeno's dichotomy. Every stride is real and there are infinitely many of them, but they crowd into the last stretch of wall. The puzzle is not whether you get there — you plainly do — but what it can mean to finish an endless list of tasks.

The modern reply is that the sum of those infinitely many distances is exactly 1. Add the first two and you have ¾. Add the third and you have ⅞. Then 15/16, then 31/32. Each new term closes half the remaining gap, and the total creeps towards 1 without ever overshooting.

That answer is correct and it is also, as a response to Zeno, slightly glib. It rests on a claim that needs defending: that an infinite list of numbers has a total at all. Nobody defended it properly for another twenty-three centuries. What Zeno had actually done was locate, with great precision, the two places where trouble lives — adding up infinitely many things, and making sense of motion at an instant.

His third paradox, less famous, is even closer to the bone. Consider an arrow in flight. At any single instant, the arrow occupies a space exactly its own size; it is not moving to anywhere, because movement takes time and an instant has none. So at every instant the arrow is motionless. And a flight is made of instants. So how does the arrow move?

That is not a word game. That is the derivative, waiting.

3

Archimedes adds up infinitely many things

Two centuries after Zeno, in Syracuse, someone did it anyway.

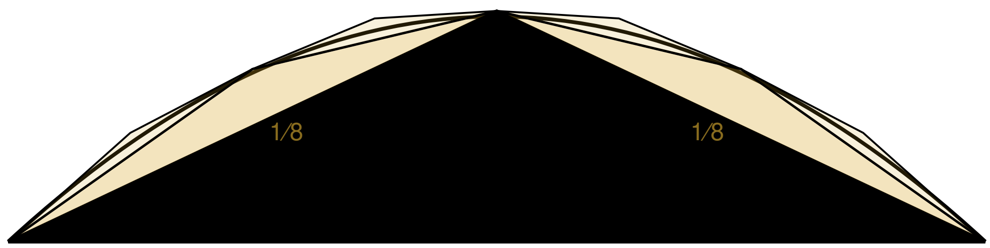
Archimedes wanted the area of a parabolic segment — the region between a parabola and a straight line cutting across it. It is a curved shape, and the Greeks had no formula for curved areas beyond the circle.

His method was to fill the shape with triangles. Inscribe the largest triangle you can, with the same base as the segment. That leaves two smaller gaps, one on each side. Fill each with its own largest triangle. That leaves four smaller gaps. Fill those. And so on, forever.

Then he proved something lovely. At each stage, the new triangles you add have a combined area of exactly one quarter of the ones you added at the stage before. If the first triangle has area 1, the next two together have area one quarter, the next four together one sixteenth, and so on. The total area of the segment is therefore

$$1 + \tfrac{1}{4} + \tfrac{1}{16} + \tfrac{1}{64} + \dots$$

each layer of triangles a quarter of the layer before



each ring of triangles adds a quarter of the ring before — total 4⁄3

Archimedes' quadrature of the parabola. The two gold triangles together make a quarter of the red one; the four paler ones together make a quarter of the gold pair. The pattern continues forever, and the total settles on exactly four thirds of the first triangle.

He then showed this total is exactly 4/3. Here is the argument, which is one of those things that seems like sleight of hand until you check it and find it airtight.

SUMMING THE QUARTERS

1. Call the total S, so $S = 1 + 1/4 + 1/16 + 1/64 + \dots$
2. Multiply everything by 4: $4S = 4 + 1 + 1/4 + 1/16 + \dots$
3. Look at what follows the 4 on that second line. It is $1 + 1/4 + 1/16 + \dots$ — which is S itself. So $4S = 4 + S$.
4. Therefore $3S = 4$, and $S = 4/3$.

Two thousand years before anyone could say what an infinite sum *is*, this gets the right answer — by treating the infinite tail as a self-contained object that can be recognised and cancelled.

Archimedes did not trust that argument as a proof, and he was right not to. Step three quietly assumes the sum exists and behaves like an ordinary number. So he backed it up with the method of exhaustion, inherited from Eudoxus: show that the area cannot be more than 4/3 (or a contradiction follows) and cannot be less than 4/3 (same), and conclude that it is 4/3. No infinities appear. It is impeccable, and it is exhausting, which is roughly how it got its name.

So the Greeks reached the edge of calculus and stopped. Not from lack of talent — Archimedes was as good as anyone who has ever lived — but for two structural reasons. They had no algebra, so every problem was a fresh geometric puzzle rather than an instance of a general method. And they were deeply suspicious of the actual infinite. They would happily say "you can always add another triangle"; they would not say "now add them all".

4

An interlude: the book that was scraped off

There is a postscript to Archimedes that is worth a paragraph, because it tells you something about how nearly the whole story could have gone differently.

Archimedes wrote a treatise, now called *The Method*, in which he explained how he actually *found* his results, as opposed to how he proved them. His technique was to imagine a shape sliced into infinitely many parallel lines, and then to balance those slices against each other on an imaginary lever. It is, in all but name, integration by infinitesimals. He was quite clear that it was a method of discovery rather than proof — but he had it, around 250 BC.

The treatise was lost. In the tenth century a copy was made; in the thirteenth, a scribe short of parchment scraped the ink off, cut the pages, turned them sideways and wrote a prayer book over the top. That palimpsest surfaced in Constantinople in 1906, was identified, and then disappeared again for most of the twentieth century, acquiring forged illuminations and mould along the way. It reappeared at auction in 1998 and has since been read with X-ray fluorescence at a synchrotron, recovering text that had been invisible for seven hundred years.

Had *The Method* circulated in the middle ages, the seventeenth century might have started from a very different place. As it was, everything in it had to be reinvented.

5

A bishop in Paris draws a graph

The next real move came around 1350, from Nicole Oresme in Paris, and it came in two parts that between them contain a surprising amount of the future.

The first is a picture. Oresme drew a diagram with time along one direction and speed along the other — a graph, three centuries before Descartes formalised the idea. Then he observed that the *area* under that graph is the distance travelled.

Look at why. If you travel at a steady 10 metres per second for 6 seconds, the graph is a flat line at height 10, running for length 6, and the region under it is a rectangle of area 60 — which is exactly the distance covered. Now let the speed change. Chop the time into slivers so short that the speed barely varies within each one, and each sliver contributes a thin rectangle of distance. Add the rectangles and you have the total distance; the thinner the slivers, the better the answer.

Oresme's insight, around 1350. Speed against time; the area underneath is the distance covered. Make the strips thinner and the staircase closes on the curve. This is the integral, three hundred years early, and it is still the picture worth carrying in your head.

That is integration. Not the notation, not the rules, but the essential idea: a total accumulated from infinitely many infinitesimal contributions, visualised as an area.

Oresme's second contribution is a shock, and it is the first serious warning that infinite sums are not to be trusted. Consider the harmonic series:

$$1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{5} + \tfrac{1}{6} + \dots$$

The terms shrink towards nothing. Surely, like Zeno's halves, they must add up to something finite? They do not. The total grows without limit, passing any number you care to name. Here is Oresme's proof, and it takes about thirty seconds.

WHY THE HARMONIC SERIES HAS NO TOTAL

1. Group the terms: (½), then (1⁄3 + ¼), then (1⁄5 + 1⁄6 + 1⁄7 + ⅛), then the next sixteen, and so on — each group twice as long as the last.
2. In the group (1⁄3 + ¼), both terms are at least ¼, and there are two of them. So the group totals at least ½.
3. In the group (1⁄5 + 1⁄6 + 1⁄7 + ⅛), all four terms are at least ⅛, and there are four. So that group also totals at least ½.
4. The same works for every group, forever: each one is worth at least ½.
5. So the sum exceeds 1 + ½ + ½ + ½ + … — infinitely many halves. It grows without bound.

The terms go to zero and the sum goes to infinity. Both at once. Any intuition that says "the terms shrink, so it must settle down" has just been destroyed, in 1350, by grouping fractions.

Keep that result in your pocket. It is the reason the eventual rigour was necessary, and it is the earliest proof that on this subject your instincts are not to be trusted.

6

Meanwhile, in Kerala

While Europe was working on this in fits and starts, a school of mathematicians on the south-west coast of India got considerably further, and the story is not as widely known as it ought to be.

Madhava of Sangamagrama, working around 1400 in what is now Kerala, and the astronomers who followed him — Parameshvara, Nilakantha, Jyeshtadeva — developed infinite series for the sine, the cosine and the arctangent. These are the power series that a modern student meets as Taylor series, and they were written down in Kerala roughly two hundred and fifty years before Newton.

One of them is startling in its simplicity. Take the odd numbers, alternate the signs, and you get a quarter of π:

$$\pi/4 = 1 - 1/3 + 1/5 - 1/7 + 1/9 - \dots$$

known in Kerala by about 1400; rediscovered in Europe by Leibniz in 1673

The Kerala mathematicians also understood that this particular series converges painfully slowly — you would need hundreds of terms for a couple of decimal places — and devised correction terms to accelerate it, which is a sophisticated thing to worry about. The *Yuktibhāṣā*, written by Jyeshtadeva around 1530, presents these results with what are recognisably demonstrations rather than bare assertions.

Did any of this reach Europe? Jesuit missionaries were active in Kerala in the sixteenth century, and the region was a hub of navigation and trade, so the opportunity existed. Some historians think transmission likely; others think the European results were plainly independent. As far as I know there is no documentary evidence either way, and the honest position is that we do not know. What is not in doubt is that the mathematics was done there first.

There is a lesson in the Kerala story that goes beyond credit. Those series were developed for astronomy — for predicting eclipses and planetary positions accurately. They were superb computational tools. But they were not assembled into a general theory of change, and that assembly is what Newton and Leibniz did. Having the pieces is not the same as having the machine.

7

Descartes turns curves into equations

In 1637, René Descartes published an essay on geometry as an appendix to a book about method, and in it he did something that made everything afterwards possible.

He put the two axes on the page. A point becomes a pair of numbers. And then — the real move — a curve becomes an *equation*. The parabola stops being a slice through a cone that you reason about with classical geometry, and becomes $y = x^2$, a rule for generating numbers.

It is hard to overstate what this changed. Archimedes had to invent a new argument for every shape. After Descartes, a question about a curve is a question about an equation, and you can grind on it with algebra. The Greeks' geometric ingenuity, which did not generalise, was replaced by symbolic manipulation, which does.

Within thirty years of that publication, four questions were exercising every mathematician in Europe:

- Given a curve, how do you find the **tangent** at a point?
- Given a curve, what is the **area** underneath it?
- Where are the **maximum and minimum** values of a quantity?
- Given a position that changes with time, what is the **instantaneous velocity**?

These look like four separate problems. They are not. The first, third and fourth are the same problem in different clothing, and the second is that problem run backwards. Nobody knew that yet. Working it out is what "inventing calculus" actually means — not the techniques, most of which were already lying around, but the realisation that they were all aspects of one thing.

8

Fermat commits the original sin

Here is where the trouble starts, and it starts before Newton was born.

Pierre de Fermat — a lawyer in Toulouse who did mathematics in his spare time and changed it permanently — had a method for finding maxima and minima. He called it *adequality*, and it works like this.

Suppose you want to cut a piece of string of length 10 into two parts and make a rectangle with the largest possible area. If one side is x, the other is $5 - x$, so the area is $x(5 - x)$. Where is that biggest?

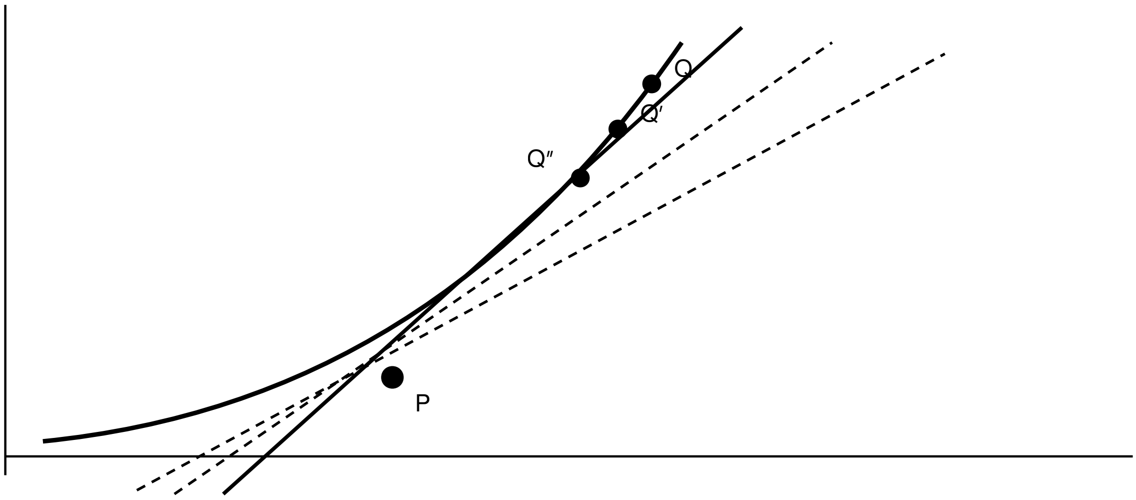
FERMAT'S METHOD — AND THE MOMENT TO WATCH FOR

1. Write the area at x: $A = 5x - x^2$.
2. Nudge x by a small amount e and write the area there: $5(x+e) - (x+e)^2 = 5x + 5e - x^2 - 2xe - e^2$.
3. Fermat's observation: at the very top of a hill, the ground is flat, so moving a little either way changes the height hardly at all. So set the two areas *almost* equal — his "adequality" — and see what that forces.
4. Subtract the first from the second. The $5x$ and $-x^2$ cancel, leaving $5e - 2xe - e^2 \approx 0$.
5. **Divide through by e:** $5 - 2x - e \approx 0$. (You may only do this because e is not zero.)
6. **Now set $e = 0$:** $5 - 2x = 0$, so $x = 2.5$. (A square, 2.5 by 2.5, area 6.25. Correct.)

Read steps 5 and 6 again, one after the other. In step 5 you divide by e, which requires $e \neq 0$. In step 6 you delete e, which requires $e = 0$. The same symbol, two lines apart, is both not-zero and zero.

That is the whole scandal, in miniature, and it is on the page in the 1630s. The answer is right — you can check it by trying 2.4 and 2.6 and finding both give less area. The method is general, powerful, and gives correct results across an enormous range of problems. And the reasoning contains a flat contradiction.

The same knot appears in the tangent problem. To find the slope of a curve at a point, you draw a line through that point and a nearby one — a *secant* — and compute rise over run. Then you slide the second point closer. The secant swings towards the tangent. But at the moment the two points coincide, rise and run are both zero, and the slope you wanted is 0/0 again.



as Q slides towards P the secant swings towards the tangent — but at P itself, rise and run are both zero

The tangent problem. Each dashed line joins P to a nearby point on the curve, and the slope of that line is easy to compute. As the second point slides in, the lines settle towards the gold one. The trouble is that the answer you want lives exactly where the calculation breaks: at P, you are dividing nothing by nothing.

So by 1650, Europe had a growing pile of techniques that all involved a quantity which had to be non-zero long enough to divide by, and zero by the time you wrote the answer down. Fermat had tangents and maxima. Cavalieri had a method of "indivisibles" for areas that treated a region as a stack of infinitely many lines. John Wallis had computed a great many areas by inspired guesswork with infinite processes. Gregory of Saint-Vincent had found that the area under a hyperbola behaves like a logarithm.

What was missing was the recognition that all of this was one subject, together with a notation good enough to make the connection obvious.

9

Newton, and analysis by infinite series

In 1665 plague closed Cambridge and a twenty-two-year-old graduate went home to Lincolnshire for eighteen months. What he did there he did not publish for decades, which is the root of an enormous amount of trouble later.

Newton's way in was through infinite series, and this is the part usually skipped. He had been reading Wallis, and he asked a question of the sort that separates the very good from the great. Everyone knew how to expand things like $(1 + x)^2$ or $(1 + x)^3$: multiply it out and you get a short, tidy list of terms. The coefficients follow Pascal's triangle. But that only works for whole-number powers. What if the power were a fraction? What if it were negative?

Newton found that the pattern keeps working — you just never run out of terms. So, for instance, the square root of $1 + x$, which is $(1 + x)$ to the power ½, becomes an infinite series:

$$\sqrt{(1 + x)} = 1 + \tfrac{1}{2}x - \tfrac{1}{8}x^2 + \tfrac{1}{16}x^3 - \tfrac{5}{128}x^4 + \ldots$$

try x = 0.21: the first four terms already give 1.1000 against a true 1.1

This was a machine for turning awkward functions into infinitely long polynomials. And polynomials are easy — you can differentiate and integrate them term by term without breaking sweat. Newton could now attack curves nobody could otherwise touch, by rewriting them as series and grinding through term after term.

That is why his 1669 tract was called *De analysi per aequationes numero terminorum infinitas* — on analysis by equations with infinitely many terms. Series were not a later application of his calculus. They were how it worked.

His conception of the calculus itself was thoroughly physical. He imagined quantities *flowing* in time. A quantity was a *fluent*; its rate of change was a *fluxion*, which he wrote with a dot above the letter — the notation physicists still use for time derivatives, and the reason $\dot{x}$ turns up in a mechanics lecture in 2026.

And here is how he handled the awkward step. Let o be a vanishingly small interval of time. Then:

NEWTON'S DERIVATIVE OF $Y = X^2$

1. In time o, x grows to $x + o$.
2. So y grows to $(x + o)^2 = x^2 + 2xo + o^2$.
3. The change in y is therefore $2xo + o^2$, while the change in x is o.
4. The ratio of the changes is $(2xo + o^2) \div o = 2x + o$. (division by o: it must not be zero)
5. o is vanishingly small, so discard it: the answer is $2x$. (deletion of o: it must be zero)

The result is right — the slope of $y = x^2$ really is $2x$. But steps 4 and 5 are Fermat's contradiction again, now at the heart of a system that was about to explain the solar system.

Newton was uncomfortable about this and returned to it repeatedly over forty years, trying different framings — first vanishing quantities, later "prime and ultimate ratios", which edges towards the modern idea of a limit without quite arriving. He never found a formulation that satisfied him.

Which is part of why the *Principia* of 1687, the book that gives us universal gravitation and the laws of motion, is written almost entirely in classical geometry. Newton had the calculus. He used it to find his results. Then he translated them into the idiom of Archimedes to present them, because geometry was above suspicion and his new method was not.

He had a tool that could predict the heavens, and he did not dare show his working.

10

Leibniz, and a puzzle about triangular numbers

Gottfried Wilhelm Leibniz came at the same subject from an entirely different direction and, in one important respect, did it better.

Leibniz was a diplomat, lawyer and philosopher who arrived in Paris in 1672 on political business and fell in with Christiaan Huygens, the finest scientist in Europe at the time. Huygens, taking the measure of this clever young man with almost no mathematical training, set him a problem.

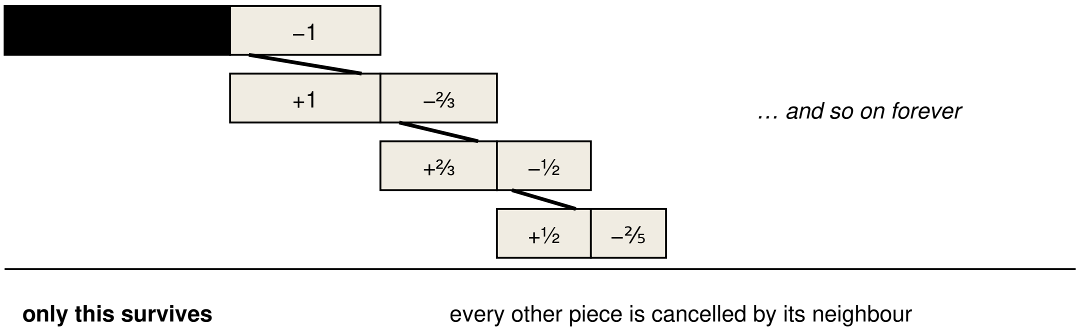
The triangular numbers are 1, 3, 6, 10, 15, 21 — the counts of a triangular arrangement of dots. Add up their reciprocals, forever. What do you get?

$$1/1 + 1/3 + 1/6 + 1/10 + 1/15 + 1/21 + \ldots = ?$$

Leibniz solved it, and the way he solved it set the direction of his life's work.

THE TELESCOPE

1. The n-th triangular number is $n(n+1)/2$, so its reciprocal is $2 \div n(n+1)$.
2. Now the trick. Notice that $2/n(n+1) = 2/n - 2/(n+1)$.
   Check with $n = 3$: the left side is $2/12 = 1/6$; the right side is $2/3 - 1/2 = 1/6$. It works.
3. So every term is a *difference* between consecutive members of the simple list 2, 1, ⅔, ½, ⅖, ⅓, …
4. Write the sum out that way:
   $(2 - 1) + (1 - ⅔) + (⅔ - ½) + (½ - ⅖) + \ldots$
5. Watch the cancellation. The −1 kills the +1. The −⅔ kills the +⅔. The −½ kills the +½. Everything in the middle destroys itself.
6. All that survives is the very first number, 2, minus whatever is left at the far end — and the far end shrinks to nothing. **The sum is 2.**



The telescoping sum that started everything for Leibniz. Each term is written as a difference; consecutive differences annihilate one another; only the two ends survive, and the far end vanishes. Sum of differences equals last minus first.

Now, that is a pretty puzzle. What Leibniz saw in it was a principle: **if you can write each term of a sum as a difference, the sum collapses to the endpoints.** Everything in between cancels. To total a long list, don't add it up — find the thing whose successive differences produce your list, and just look at its two ends.

Hold that next to the following question. You want the total distance travelled on a journey. You know the speed at every instant. Speed is the rate at which distance changes — the difference in distance per tiny step of time. So the distances covered in each tiny step are *differences* of the total-distance-so-far. Add them all up and everything in the middle cancels, leaving distance-at-the-end minus distance-at-the-start.

That is the fundamental theorem of calculus. Integration and differentiation are inverse operations, and the reason is that summing is the inverse of differencing. Leibniz got there by shrinking the steps in a telescoping sum until they became infinitesimal.

Because he thought of it this way, he chose notation that is still in use three and a half centuries later, and which does an unreasonable amount of the thinking for you:

- **d** for *difference*. So $\mathrm{d}y$ is a tiny difference in y, and $\mathrm{d}y/\mathrm{d}x$ is a ratio of two tiny differences — which behaves, gratifyingly, exactly like the fraction it is written as.
- $\int$ for *summa*, sum — an elongated letter s. The integral sign is literally a long S for "add all these up".

Newton's notation, the dot, tells you almost nothing beyond "rate of change with respect to time". Leibniz's tells you what the operation *is*, and it survives the transition to harder cases — chains of variables, several variables at once, changes of coordinate — because the symbols keep behaving like the fractions and sums they were named after. Every student who has cancelled a $\mathrm{d}u$ in a substitution has felt the benefit.

And Leibniz's foundations were, if anything, shakier than Newton's. He worked with actual infinitesimals: quantities greater than zero but smaller than every ordinary positive number. Asked what these were, he was evasive, sometimes calling them useful fictions. He knew the reasoning was not watertight; he judged, correctly, that it was too productive to abandon.

11

The war, and what it cost

Newton had the calculus by 1666 and published nothing. Leibniz worked it out independently around 1675 and published in 1684. Newton's supporters said Leibniz had stolen it; Leibniz's supporters said Newton was inventing a grievance after the fact.

The historical verdict is that both invented it, independently, and that Leibniz's published version was better presented. Newton got there first; Leibniz got there second and explained it properly.

In 1712 the Royal Society convened a committee to adjudicate. Its report found decisively for Newton. Newton was President of the Royal Society at the time, and it later emerged that he had drafted the report himself — then reviewed it anonymously in the Society's own journal.

The consequence was not merely unpleasant, it was expensive. British mathematics, loyal to Newton, stuck with his dot notation for a century. Continental mathematics adopted Leibniz's, which was easier to compute with and easier to extend. The Bernoullis, Euler, Lagrange, Laplace — the whole flowering of eighteenth-century analysis happened in Basel, Berlin, St Petersburg and Paris, in Leibniz's symbols, while Cambridge fell quietly behind. It took until the 1810s for a group of Cambridge undergraduates to campaign for continental notation, in a society they cheerfully named for the promotion of d-ism over dot-age.

A useful reminder that notation is not decoration. Choosing symbols that think along with you is worth a century.

12

A bishop asks an awkward question

In 1734 George Berkeley, an Anglican bishop and one of the sharpest philosophers of the century, published a pamphlet attacking the foundations of the calculus. It was addressed to an unnamed "infidel mathematician", generally taken to be Edmond Halley, who had apparently been telling a mutual acquaintance that Christian doctrine was incoherent.

Berkeley's argument was, in effect: let us examine your own reasoning before you criticise mine.

He was not a crank, and he was not confused about the mathematics. He understood the method perfectly well and reproduced it accurately. His complaint was precise, and it was this. In computing a derivative you introduce an increment. You then divide by it, which you may only do if it is not zero. Having divided, you discard the terms that still contain it, which you may only do if it is zero. You cannot have both. Either the increment is something, in which case your answer still has an error in it, or it is nothing, in which case you had no business dividing by it and the calculation never happened.

Look back at Newton's five steps. Step 4 divides by o. Step 5 deletes o. Berkeley's point is simply that you have used $o \neq 0$ to obtain a result and then used $o = 0$ to tidy it, and no amount of fluency disguises that these are different assumptions.

His most quoted line asks what these vanishing increments are. Not finite quantities, not infinitely small ones, not nothing — might one not call them the ghosts of departed quantities?

It is a devastating phrase because it is affectionate about the mathematics while being merciless about the logic. And Berkeley pressed the advantage he actually cared about: if you accept a method whose central step cannot be justified, on the grounds that it produces results you like, you are doing exactly what you accuse theologians of doing. You are taking something on faith because it works.

He was right. That is the uncomfortable part. For a hundred and fifty years, nobody could answer him.

Berkeley also made a shrewd secondary point, which mathematicians found even more annoying because it was true: the method gets right answers because its errors cancel. Two compensating mistakes are made, and they happen to undo each other. He could not say precisely why they always cancel — that would have required the very theory nobody had — but as a diagnosis it was uncannily close to what the nineteenth century eventually found.

13

Press on, and faith will come to you

What happened next is the most interesting sociological fact in this story: essentially nothing. Nobody stopped.

It would be wrong to picture a century and a half of mathematicians wringing their hands. They knew the foundations were unsatisfactory, they said so in print, and they carried on — because the results kept being right, and the results were spectacular.

Halley used the new mechanics to predict that a comet would return in 1758. It returned. Euler applied calculus to fluids, optics, music, ship design and the orbits of the moon. Lagrange and Laplace rebuilt mechanics on it and showed that the solar system is stable. In 1846 Le Verrier computed, from irregularities in Uranus's orbit, where an unknown planet must be; the Berlin Observatory pointed a telescope at that spot and Neptune was there, within a degree. A method that finds an unseen planet on paper is not going to be abandoned because a bishop has identified a gap in its logic.

The attitude of the age is caught in a remark of d'Alembert's, reported by a student who asked him about exactly these difficulties: press on, and faith will come to you. Not a dismissal of the problem. An accurate description of the intellectual bargain everyone had struck.

People did try to pay the debt. Colin Maclaurin wrote a long, careful treatise in 1742 that tried to justify the calculus by classical Greek methods, which was rigorous and so cumbersome that almost nobody used it. In 1797 Lagrange attempted a complete rebuild: he proposed defining the derivative not by any limiting process at all, but by writing every function as an infinite power series and simply *reading off* the coefficient of the linear term. No infinitesimals, no vanishing quantities, no ghosts — just algebra with infinitely many terms.

It was a genuinely good idea and it failed, for a reason that is instructive. Lagrange assumed every function worth considering could be written as a power series. That is false, and proving it false required exactly the theory of convergence that did not yet exist. The attempt to avoid limits foundered on a question about limits.

So the debt went unpaid. What finally forced payment was not philosophy but a crisis inside mathematics itself — and it arrived from the direction of infinite series.

14

Euler adds up things that shouldn't be added

Leonhard Euler was the most productive mathematician who has ever lived, and his relationship with infinite series was one of magnificent recklessness.

Consider this series, first studied by Guido Grandi:

$$1 - 1 + 1 - 1 + 1 - 1 + \dots$$

What is its total? Group the terms one way and you get (1−1) + (1−1) + (1−1) + … = 0. Group them the other way and you get 1 − (1−1) − (1−1) − … = 1. Both groupings look equally legitimate. The partial sums go 1, 0, 1, 0, 1, 0 forever, settling on nothing.

Euler's answer was ½, and he had a reason. There is a standard formula for a geometric series:

$$1 + r + r^2 + r^3 + \dots = 1/(1 - r)$$

Put $r = ½$ and it gives 1 + ½ + ¼ + … = 2, which is correct and which Zeno would have recognised. Now put $r = −1$. The left side becomes Grandi's series, and the right side becomes 1/(1−(−1)) = ½.

Push harder and it gets worse. Put $r = 2$:

$$1 + 2 + 4 + 8 + 16 + \dots = 1/(1 - 2) = -1$$

a sum of positive numbers, growing without limit, equal to a negative number

This is not a joke or a misunderstanding; it is in the literature, taken seriously. And the embarrassing part is that these assignments are not simply wrong. There are modern frameworks — Abel summation, analytic continuation — in which Grandi's series really is naturally assigned the value ½, and in which the 1 + 2 + 4 + … result appears in a legitimate setting. Euler's instincts were extraordinary. What he lacked was any way of saying *when* such a manipulation was safe and when it was nonsense.

The reason to care is that a sum of infinitely many things is not automatically a number, and if you treat it as one you can prove anything. Recall Oresme: the harmonic series has terms shrinking to zero yet no finite total. Now add a subtlety that shocked the nineteenth century. Take the alternating harmonic series:

$$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{5} - \tfrac{1}{6} + \dots$$

This one does converge, to about 0.693 — the natural logarithm of 2. But because the positive terms alone add to infinity and the negative terms alone add to minus infinity, you can *rearrange* the same terms, using every one exactly once, and make the total come out to any number you like. Want it to equal 5? There is an ordering that does that. Want −17? That ordering exists too.

Riemann proved this in general. It means the most basic rule of arithmetic you have — that order doesn't matter when adding — simply fails for infinite sums, unless you can show your series satisfies a stronger condition. Nobody had known there *was* a stronger condition, because nobody had defined convergence carefully enough for the question to be askable.

15

Fourier breaks everything

The crisis came, as crises often do, from someone solving a practical problem and not worrying enough about how.

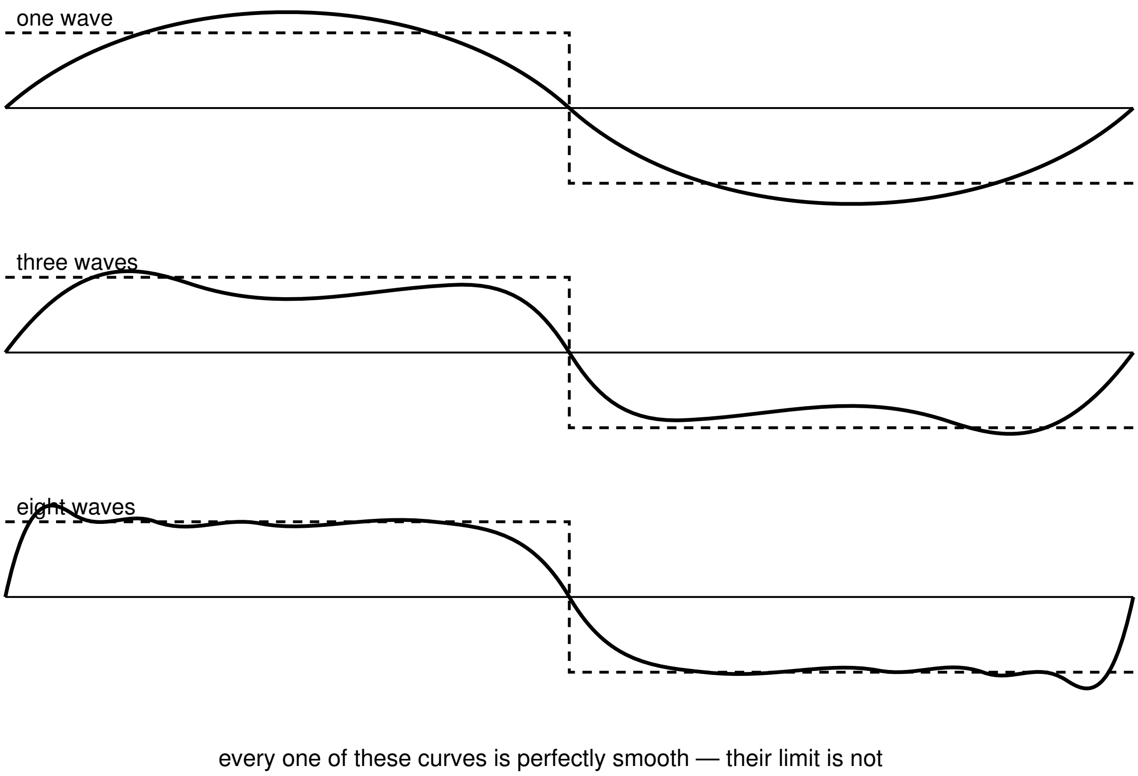
Joseph Fourier was studying how heat spreads through a solid body. In 1807 he submitted a paper containing a claim so bold that the reviewing panel — which included Lagrange and Laplace — refused to publish it.

Fourier claimed that *any* function whatever, however jagged or arbitrary, can be built by adding together sine and cosine waves.

To see why that seemed outrageous, take the squarest, least wave-like thing imaginable: a signal that sits at +1 for a while, then jumps instantly to −1, then jumps back. All corners and cliffs. Fourier's recipe says to add up sine waves of increasing frequency, with decreasing amplitudes:

$$\sin x + \tfrac{1}{3}\sin 3x + \tfrac{1}{5}\sin 5x + \tfrac{1}{7}\sin 7x + \dots$$

smooth, rounded waves — adding up to something with vertical cliffs in it



Fourier’s claim, in pictures. The dashed line is the target: a signal that jumps. The red curves are sums of one, three and eight sine waves. Each is smooth and continuous; each is a better fit than the last. Yet the thing they are approaching has a vertical cliff in it. Where, exactly, does smoothness turn into a jump?

The picture is the whole problem. Every partial sum is a finite total of smooth waves, and so is itself perfectly smooth — unbroken, with a well-defined slope everywhere. But the thing they converge to has a jump. Continuity, apparently, does not survive the limit.

That should not be possible under the assumptions everyone had been making. And now the questions could no longer be postponed. What does it *mean* to say an infinite sum of functions "equals" a function? Does it mean at every single point, or in some looser overall sense? If each term is continuous, must the total be? If each term is smooth, can you differentiate the total term by term? Everyone had assumed yes to all of these. Fourier's square wave says no to at least one.

There is even a residue of the fight visible in the picture. Look near the jump in the eight-wave curve: the approximation overshoots, and the overshoot does not go away as you add more waves — it just gets squeezed into a narrower region beside the cliff. That effect was later named for Josiah Gibbs, and it is a permanent little monument to the fact that something subtle is happening at the boundary.

Fourier's work was too useful to reject — it remains one of the most applied ideas in all of science, sitting inside every audio codec and image format in your possession. But it could not be made sense of with the tools of 1807. It forced the issue. Mathematics now needed a precise definition of convergence, not for philosophical tidiness, but because working scientists were getting contradictory answers.

16

Cauchy takes away the infinitesimal

Augustin-Louis Cauchy was teaching analysis at the École Polytechnique in Paris, to engineering students who needed the subject to work. In 1821 he published his lecture course, and it is the hinge of this entire story.

Cauchy's move was to stop arguing about what an infinitesimal *is*, and to remove it from the definitions altogether. In its place he put a single organising idea: the **limit**.

Notice what this does to the paradox. The old question was "what is the value of the ratio when the increment is zero?" — which has no answer, because at zero there is no ratio. Cauchy's question is different: "as the increment gets smaller and smaller, never reaching zero, what value does the ratio get closer and closer to?"

Go back to Newton's example. The ratio was $2x + o$. Take $x = 3$, so the ratio is $6 + o$:

WHERE THE GHOST GOES

1. Let $o = 0.1$ → the ratio is 6.1
2. Let $o = 0.01$ → the ratio is 6.01
3. Let $o = 0.001$ → the ratio is 6.001
4. Let $o = 0.000001$ → the ratio is 6.000001

At no stage is o ever zero. At no stage is the ratio ever exactly 6. But the values are heading somewhere unmistakable, and that destination is 6. The derivative is *defined to be* the destination — not the value at zero, which does not exist, but the number the ratios approach.

The contradiction dissolves because we never set $o = 0$ at all. We only ever divided by non-zero numbers. The ghost is not exorcised; it is never summoned.

This is the reframing that saves the subject, and it is worth sitting with, because it is the single most important idea in the whole two thousand years. The derivative is not a ratio of two vanished quantities. It is the *limit* of a sequence of perfectly ordinary ratios, and limits are about where things are heading, never about arriving.

Cauchy defined continuity, convergence and the integral all in these terms, and gave the criterion for convergence that still carries his name. The vocabulary was becoming precise.

And then Cauchy made a mistake, which is oddly reassuring.

He stated, and believed he had proved, that if you add up a convergent series of continuous functions the result must be continuous. Fourier's square wave is a standing counterexample: every partial sum is continuous, the series converges at every point, and the total jumps. Niels Henrik Abel pointed out counterexamples in 1826.

The error was subtle, and locating it took another twenty years. The problem is that "converges" was still doing two jobs.

17

Weierstrass, and the taming of “approaches”

Karl Weierstrass spent fifteen years as a provincial schoolteacher, teaching mathematics alongside gymnastics and handwriting, and publishing occasionally in a school prospectus that nobody read. When one of his papers was noticed in 1854 he was given an honorary doctorate and eventually a chair in Berlin, where he became the most influential teacher of analysis in Europe.

His programme was to remove every last appeal to intuition, motion or geometry from the subject, leaving only inequalities between numbers. Cauchy still described a variable as "approaching" a value, which smuggles in a picture of something moving. Weierstrass replaced the picture with a challenge.

Here is what "the ratios approach 6" means, stated without any motion at all:

THE EPSILON–DELTA DEFINITION, IN WORDS

Think of it as a game between a sceptic and you.

1. The sceptic names a tolerance — call it ε. Say $\varepsilon = 0.001$. He is asking: can you get the ratio within 0.001 of 6?
2. You must reply with a margin — call it δ — and promise that *every* non-zero increment smaller than δ puts the ratio within his tolerance. Here you would answer: take $\delta = 0.001$, since the ratio is $6 + o$ and any o below 0.001 lands within 0.001 of 6.
3. The sceptic tries again, harder: $\varepsilon = 0.0000001$. You answer with a correspondingly small δ.
4. If you have a winning reply *for every* tolerance he can name, however vicious, then the limit is 6. If there is even one tolerance you cannot meet, it is not.

No infinitesimals. No motion. No vanishing. Just: for every ε there exists a δ. Two hundred years of metaphysics replaced by a quantifier.

Students find this definition cold and fiddly, and it is. It is also bulletproof, and it is the reason nobody has had to argue about the foundations of calculus since.

With that machinery, the mystery in Cauchy's error could be pinned down exactly. There are two different things "the series converges to f" might mean:

- **Pointwise convergence.** Pick any single point. The partial sums at *that* point head towards the right value. But some points may be far slower to settle than others, and there may be no limit to how slow.
- **Uniform convergence.** Every point settles at least as fast as some common schedule. One δ works for the whole curve at once.

Fourier's square wave converges pointwise but not uniformly. Near the jump, points take arbitrarily long to settle down — which is exactly what the overshoot beside the cliff is showing you. And the theorem Cauchy wanted is true if you require uniform convergence, false if you only have pointwise. He had not distinguished them because nobody had realised there was anything to distinguish.

This is the pattern of the whole nineteenth-century rebuilding. Not the overthrow of old results, but the discovery that a familiar word had been concealing two different ideas.

18

What, exactly, is a number?

There was one hole left, and it was underneath everything.

The limit definition says a sequence of ratios closes in on some number. But which number? For that to be guaranteed, the number line must have no gaps — every place a sequence could be closing in on must have an actual number sitting in it.

Is that obvious? It had better not be, because it is false for the rationals. Consider the sequence 1, 1.4, 1.41, 1.414, 1.4142, … Every one of these is a perfectly good fraction. They are closing in on something. But $\sqrt{2}$ is not a fraction — the Greeks proved that, and the proof reportedly caused a scandal among the Pythagoreans. So within the rational numbers alone, this sequence heads towards a hole in the line.

Analysis rests entirely on there being no such holes in the reals. And in 1858 Richard Dedekind, preparing lectures, realised with some embarrassment that he could not prove it, because nobody had ever said what a real number *is*.

His answer, published in 1872, has an audacious simplicity. Do not try to construct the number. Construct the *gap it fills*. Split all the rationals into two heaps: those whose square is less than 2, and the rest. That division is a completely well-defined thing, specified using only fractions. *Define* $\sqrt{2}$ to be that division. A real number simply *is* a cut of the rationals into a lower and an upper part.

It sounds like a dodge and it is not: you can define addition and multiplication of cuts, prove the usual arithmetic laws, and demonstrate that the resulting system has no gaps. Georg Cantor gave a different construction the same year, and went on to prove that the infinities involved come in different sizes — that the reals are strictly more numerous than the fractions — which caused its own decades of uproar.

The programme was named the *arithmetisation of analysis*: calculus built on limits, limits built on real numbers, real numbers built on rationals, rationals built on integers. At each level, nothing assumed but what had been constructed below. Two hundred years after the *Principia*, the calculus finally had a floor.

19

The monster

All of which might feel like bookkeeping — a tidying-up of things everyone already knew. In 1872 Weierstrass demonstrated, unforgettably, that it was not.

He presented a function that is **continuous everywhere and differentiable nowhere**. Its graph is a single unbroken curve — you can draw it without lifting your pen — and yet at no point whatsoever does it have a tangent. Not at rational points, not at irrational points. Nowhere.

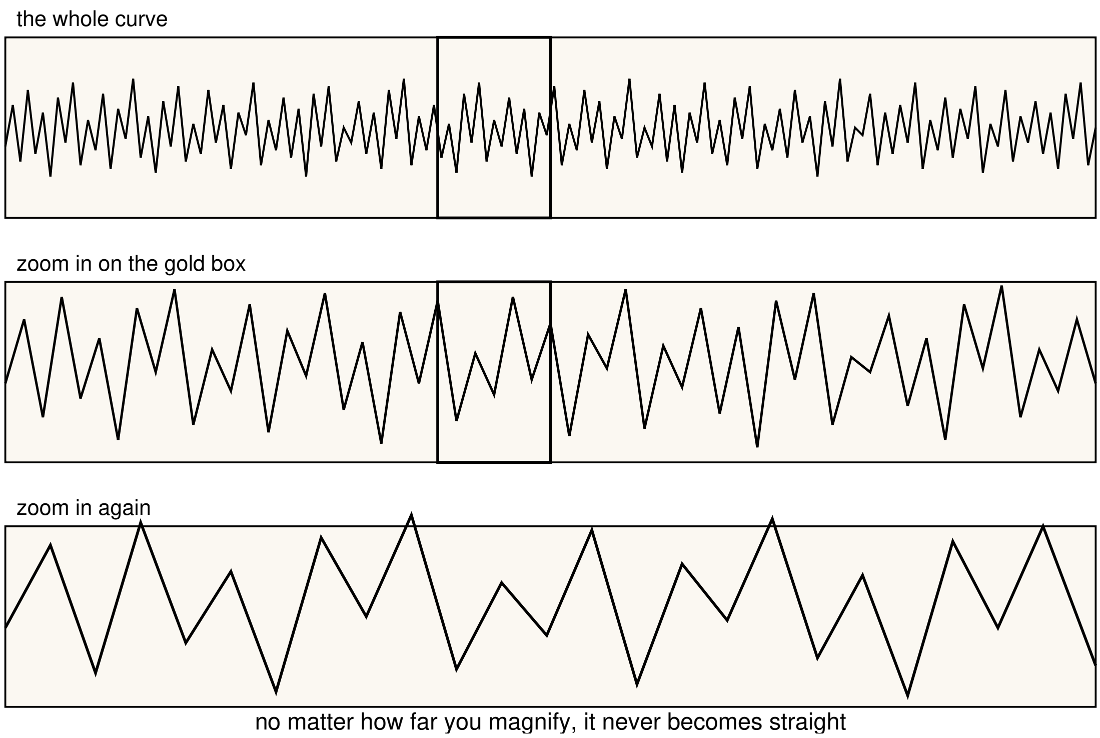
Everybody's mental model of a curve said this was impossible. A continuous curve might have a few sharp corners, like the point of a V, where the slope is ambiguous. Everyone accepted that. But a curve that is corners *all the way along*, at every point of its infinite length, with no smooth stretch anywhere, however short?

The construction is a sum of infinitely many cosine waves:

$$W(x) = \cos(\pi x) + a\cos(b\pi x) + a^2\cos(b^2\pi x) + a^3\cos(b^3\pi x) + \ldots$$

amplitudes shrinking by a factor a; frequencies growing by a factor b

The idea is a tug-of-war. Each successive wave is *smaller* in height — which keeps the total under control, so the sum exists and is continuous. But each is much *wigglier* — so much wigglier that the extra steepness beats the shrinking height. Choose a and b so that the frequencies outrun the amplitudes and the sum, though perfectly well-behaved as a curve, has a slope that never settles down anywhere.



Weierstrass's function, magnified twice. A smooth curve, magnified enough, eventually looks like a straight line — that straight line *is* its tangent, and finding it is what differentiating means. This curve never flattens. Zoom in a thousand times, a billion times, and it is exactly as jagged as before. There is no tangent to find, anywhere.

That figure is the clearest way to see what a derivative really requires. Differentiating a function means zooming in until the curve is indistinguishable from a straight line, then measuring that line's slope. Every function you met at school does flatten out under magnification. Weierstrass's does not, at any point, at any magnification.

The reaction was not admiration. Established mathematicians found these objects repellent — Hermite wrote of turning away in fear and horror from this lamentable plague of functions with no derivatives, and Poincaré thought them a monstrosity invented for no purpose but to embarrass the reasoning of our forefathers. The word "monster" attached itself and stuck.

But the monster settled the argument. Before it, one could maintain that all this ε-δ pedantry was unnecessary, since geometric intuition would always keep you out of trouble. After it, that position was untenable. Intuition says a continuous curve must be smooth in places. Intuition is wrong. The only way to know what is true is to work from the definitions, and the definitions had better be precise.

A coda that would have surprised Hermite: these functions turned out to describe reality rather well. Brownian motion — the jitter of a pollen grain in water, which Einstein analysed in 1905 — traces a path that is continuous and nowhere differentiable, for exactly the same reason. A particle bombarded from all sides has a position at every instant but no well-defined velocity at any instant. Coastlines, share prices, mountain profiles: all show the same never-smoothing-out structure. Mandelbrot eventually gave the family a name, fractals, and a century after Weierstrass the monsters were on posters.

20

The infinitesimal was innocent after all

One more turn, and it is a satisfying one.

The nineteenth century did not prove infinitesimals were *impossible*. It found a way of doing analysis that did not need them, and having found it, dropped them. Leibniz's quantities-smaller-than-any-number were left as a historical curiosity, the sort of thing your lecturer mentions in a slightly embarrassed aside.

Then in the early 1960s, Abraham Robinson, using twentieth-century mathematical logic, constructed a number system that contains the ordinary reals and also genuine infinitesimals — numbers greater than zero but smaller than 1/1000, and smaller than 1/1000000, and smaller than every ordinary positive number you can name. This system, the hyperreals, is logically consistent, provided the ordinary reals are.

In it you may write dx as an actual infinitesimal number, divide by it perfectly legitimately, and then take the "standard part" of the answer — the ordinary real number infinitely close to it — which is a defined operation rather than an act of faith. It turns out to give the same derivatives, the same integrals, the same theorems.

So Leibniz's instinct was sound. His infinitesimals were not incoherent; they were merely 250 years ahead of the logical machinery needed to make them respectable. He could not have justified them, because the tools required did not exist and would not for a dozen generations. He used them anyway, and got it right.

Non-standard analysis has not replaced the standard approach — the ε-δ version came first, everyone learned it, the textbooks are written. But it means the answer to the student's question, *is dx really a number?*, is no longer a firm no. It is: not in the system you are being taught, but there is a system where it is, and everything works out the same.

21

Coda: the order of things

Set the story in the order a modern course presents it. First the real numbers. Then limits, with ε and δ. Then continuity. Then derivatives as limits. Then integrals. Then, late on and carefully, infinite series with tests for convergence.

Now set it in the order it happened. Infinite sums, in Greece, around 250 BC. Areas under curves, in Paris, around 1350. Tangents and maxima, by an amateur in Toulouse, in the 1630s. Derivatives and integrals, in Lincolnshire and Paris, in the 1660s and 70s. A century of triumphant application and unanswered objection. A crisis provoked by heat conduction in 1807. Limits made precise in 1821. Uniform convergence in the 1840s. Real numbers — the supposed foundation of everything — not constructed until 1872.

The teaching order is almost exactly the reverse of the historical one. Foundations came last, by about two centuries. This is not a failure of the mathematicians involved; it seems to be how mathematics actually proceeds. Somebody notices a technique that works. It gets used, extended, and pushed past the point anyone can justify. Eventually it breaks in a way that cannot be ignored, and the breakage forces someone to work out what the technique had really been doing.

Two things follow, and they are worth carrying with you.

The first is that if the ε-δ definition felt unmotivated when you met it, that is because it was presented as a starting point when it is in fact an *answer* — the resolution of a specific two-hundred-year crisis, delivered to a discipline that badly needed it. Given the problem first, the definition looks less like pedantry and more like relief.

The second is about your own confusion, thirty-five years ago or last week, over dx being zero and not zero. That confusion was not a failure of understanding. It was the correct response. Berkeley had it, Newton had it, Leibniz had it, d'Alembert had it and told his students to press on regardless. Being troubled by that step puts you in reasonable company.

The trouble was real, it lasted a hundred and fifty years, and it was eventually settled — not by telling anyone to stop worrying, but by finding a way to say precisely what had been meant all along.

IF YOU WANT TO CARRY ON

Books that teach the mathematics *and* tell this story

Most books do one or the other. These four do both, and are listed roughly in order of how much actual mathematics you want to do.

Ernst Hairer & Gerhard Wanner, *Analysis by Its History*

The closest thing to this essay with exercises attached. It teaches calculus in the order the ideas were discovered — areas and infinite series first, derivatives next, rigour last — and is full of portraits, original figures and quotations from the people involved. Pitched at mathematics undergraduates rather than engineers, so expect proofs.

Otto Toeplitz, *The Calculus: A Genetic Approach*

The first serious attempt at teaching calculus historically, written before 1940 and reissued with a foreword by David Bressoud. Short, elegant, and a partial course rather than a complete one — better as a companion than a spine.

William Dunham, *The Calculus Gallery*

Walks through the actual theorems and proofs of Newton, Euler, Cauchy, Riemann, Weierstrass and Lebesgue, in something close to their own terms. Assumes you have met calculus; rewards you for having read the above.

David Bressoud, *Calculus Reordered*

An argument that the standard teaching order is backwards — which is the point this essay ends on. Narrative rather than textbook; no exercises.

For plain histories, gentler introductions, and the standard engineering texts, see the further reading on the Background page.

Written by Claude as a companion to the Introduction to Physics lecture notes. A companion essay, The Two Thousand Year Silence, does the same for probability. Further reading, including Thompson's *Calculus Made Easy* and the Feynman Lectures, is listed at the end of the Background page.