Is mathematics cumulative?

Hugh Parsonage

17th November 2008

Contents

Preliminary

This is a report for a stand-alone ASC in the philosophy of mathematics. I would like to thank and acknowledge Dr Jason Grossman for his thoughtful contributions, helpful guidance and warm encouragement

Introduction

Today, any educated person is familiar with some form of mathematics. And mathematics students are certainly familiar with its philosophy: mathematics is the "language" of science. Synonymous with logic in some senses and a rich extension of it in others, students learn its strong, unassailable logical foundation. All students know that all theorems, once proved, whether from Euclid thousands of years ago to propositions proved in the past year, will be regarded as true forever. Indeed, logic, truth, and mathematics are often regarded as indistinguishable: Ayer felt compelled to regard statements in mathematics as not merely demonstrably true, but tautological.(1, p. 33)

Every student of it knows that mathematics is not like science as described by Popper(8) or Kuhn.(5) It does not seek to disprove the contemporaneous paradigms. There are no "mathematical revolutions", where the mathematics of the past is radically and completely changed because a discovery shows it to be false. There is no thought that the veracity of mathematical statements is connected to their utility or their empirical exactitude. (One does not believe that $12 \times 12 = 144$ because most of the time we have a twelve-by-twelve square array of apples it contains one hundred and forty-four apples.)

What I hope this essay will prove is that these notions of mathematics are naïve or wrong. The historical record has numerous examples of false mathematical statements that have been accepted as theorems and the development of mathematics is not uniform and calm.

This essay is concerned with the history of mathematics as a study in mathematical thought. It will turn to such questions as: What is the significance of pseudotheorems in judging the firmness of mathematical statements? To what extent do the axiological components of mathematical thought and writings destroy or defend the mathematics itself? Is the gradual acceptance of natural numbers, the number zero, the integers, the rational numbers, the real numbers, the complex numbers, and the hypercomplex numbers coherent with a cumulative mathematics or does the initial (and persistent) denial of the existence of these sets deny mathematics any such quality? Are non-Euclidean geometries merely additional geometries to Euclidean geometry, or does the former contradict the latter? Are bad proofs inherent to mathematics or are they not to be counted? Was the published theorem that a function has a derivative at every point at which the function is continuous just a result of uncharacteristically poorly-wording or is the flaw inherent to mathematics? When it is written in history that a proof is, say, "not rigorous by modern standards", what does this mean for the quality of mathematics? Is saying all theorems around today are rigorous (by modern standards) vacuous or profound?

As a study in mathematical thought, we should not be swayed by ideology or idealizations of mathematics that do not stand the scrutiny of history. To do this, it will be useful to distinguish between normative and descriptive accounts of mathematics. Put simply, a **normative** account of something states what it *ought* to be; a **descriptive** account of something states what it *is* and *how* it has been. A purely normative historical account of mathematics in this essay would be useless. Indeed, if we define mathematics as the body of knowledge dealing with quantity, space, and set theory that is the result of perfect deductive logic *alone*, then mathematics *is* cumulative and this essay is a boring ipso facto. What is (I hope) more interesting is how mathematics has actually operated. For it has operated far from perfectly.

It will also be important to ensure that our history of mathematics does not become Whiggish. **Whig history** is a mode of studying the people and societies of the past as if they were inevitably converging to the present or that actors of the past were striving for today's ideals. In the context of mathematical history, I am urging caution in reading, say, the 14th century theory of infinite series expansion by Madhava of Sangamagrama as a foundation for Fourier analysis in the 19th century or the mathematical analysis of today – Madhava developed infinite series with scant regard for the Defence Signals Directorate or iPods. I am also urgin

I will begin this essay with a closer read of the question and then discussing four paths in the history of mathematics, looking at the branches of analysis, algebraic number theory, and geometry, as well as the growth of the various number systems.

Chapter 1

What is the question asking?

1.1 Definition of mathematics

Wikipedia, in its English article “Mathematics”, defined **mathematics** as

> the body of knowledge and academic discipline that studies such concepts as quantity, structure, space, and change.

Wikipedia’s definition was a bit grammatically insecure (a “body of knowledge” can “study something?), so I edited it:

> **Mathematics** is the body of knowledge and academic discipline arising from the study of such concepts as quantity, structure, space, and change.

Today, 16 November 2008, Wikipedia’s definition had changed to:

> **Mathematics** is the academic discipline, and its supporting body of knowledge, that involves the study of such concepts as quantity, structure, space, and change.

This is the definition I will use for this essay.

1.2 Notions of "cumulative" in bodies of knowledge

1.2.1 Absolutely cumulative

For the sake of extremism, let us characterize a body of knowledge as **absolutely cumulative** if every statement that has ever been declared to be a theorem is true.

1.2.2 Strictly cumulative

Consider mathematics throughout history. Certainly what is called mathematics is larger than what it was when it began. Let B_t be the set of all theorems T accepted into the corpus at time t. Then a body of knowledge is **strictly cumulative** if $B_t \subset B_{t^*}$ for all $t \leq t^*$. The subtle difference between this definition and the one for absolutely cumulative is in the characterization of T as "a theorem at time t". This is meant to exclude theorems from naive or malicious amateur mathematicians who in writing "**Theorem 1.1** $2 + 2 = 73$" deny mathematics ever being absolutely cumulative. But if a bad proof makes a conjecture a theorem and this statement makes it into a 'well-respected' mathematics journal or finds acceptance amongst mathematicians of a 'high-regard', then mathematics is not strictly cumulative. As my scare-quotes suggest, the characterization of mathematics as *strictly cumulative* will be fuzzy. As we will discover, however, the fuzziness will not be a problem.

1.2.3 Weakly cumulative

Our next definition is motivated by the wish to credit bodies of knowledge that may have flaws at the boundaries but whose core is cumulative. Let B_t be defined as before. Then a body of knowledge is **weakly cumulative** if there is a **core to the body of knowledge**, C_t, with $C_t \subset B_t$ for all t such that C_t is strictly cumulative.

Of course, this definition is very weak. One could choose $C_T = \varnothing$ and prove theology or, indeed, a random statement generator weakly cumulative. To this, the obvious rejoinder is that the larger the size of C_T the firmer or stronger the accumulation of the body of knowledge.

1.2.4 Properly cumulative

In our discussion on the number systems, we accepted that the apparent stubbornness of mathematicians in the face of higher number systems was more a noble scepticism. Nevertheless, mathematicians were wrong in some sense by discriminating against negative numbers and by terming pejoratively $\sqrt{-1}$ 'imaginary'. When I say that the mathematicians of the past were "wrong", I am not taking aim at their tastes, but at their lack of ability to see the truth when it was presented to them. Just as a *weakly cumulative* body of knowledge can remain so by offering the empty set as their eternal success, so too could an improperly cumulative body of knowledge reject almost everything that was offered in order to preserve its integrity. This is what motivates our next definition.

We say that a body of knowledge is **properly cumulative** if every human submission to it is correctly decided as true or false by the community that owns the body of knowledge. It is worth noting that this judgement can only be made retrospectively.

1.2.5 Figuratively cumulative

Of course, there is the commonplace notion of cumulative meaning to grow into a heap or mass. For instance, a cathode in a galvanic cell may accumulate positive ions and a mass may form. An analogy can be drawn between mathematics and the mass accumulating around such a cathode: the latest accumulation is possible because of the previous accumulation and it proceeds in a locally chaotic (at the microscopic scale) but globally steady (at the scale of the cell) way. We could say that a body of knowledge is **figuratively cumulative** if it resembles this image.

In this sense of cumulative, mathematics could well be cumulative. It would allow for mathematics to lose a few theorems (as a galvanic cell might lose some mass) as long as the overall progress is one of accumulation. However, allowing mathematics to be called cumulative in this sense would be misleading. The common wisdom is that all mathematics is true and that theorems are ever-present. In the light of the history, this thesis cannot be maintained. It is disingenuous to call mathematics "cumulative" when it is only figuratively cumulative.

1.3 How to measure the size of mathematics

In order to give a meaningful answer to the question "Is mathematics cumulative?" we need to know what cumulative means, but we also need to know by what measure such an accumulation would be said to take place.

1.3.1 Different ways of characterizing theorems

True/False

If we mean a *theorem* in the descriptive sense of "the statement in a recognized mathematics text headed **Theorem**", a theorem can be either true or false. Of course, the term 'false theorem' to most is a contradiction: a statement can only be a theorem if it is true.

Nevertheless, the distinction is important to the superquestion. If we only define theorems to be true statements, the answer to the question "Is mathematics cumulative?" is trivial. Furthermore, if the following appeared in a mathematics text,

> **Theorem 1.3.1.** *Let x, y, z be the lengths of the sides of a right-angled triangle, with z the length of the hypotenuse. Then*
>
> $$x^2 + y^2 = z2.$$

most would probably not object to calling this a theorem, with the typo implicitly corrected. Accepting a theorem as true is not really affirming what is written is true.

I claim that what is involved in accepting a theorem is not accepting exactly what has been written down, but accepting the *thoughts* of the author as expressed in the actual text. This is not a widely-accepted view and a defence of it is beyond the scope of this essay.

Interesting/Uninteresting/Trivial

Suppose we do not accept that false statements are theorems, there is still the issue of whether trivial or even uninteresting theorems should be counted with equal weight (or even whether they ought to be counted at all). Mathematicians already delineate between **Theorems** (which are important nontrivial statements), **Corollaries** (which are important statements, but easily follow from nearby Theorems), **Propositions** (which are uninteresting, unimportant statements that may or may not easily follow from other Theorems and Propositions), and **Lemmata** (which are unimportant statements from which important Theorems follow).

The answer to which to count is not easy. For example, how many theorems are in the following selection?

Proposition 1. *The projection $\pi : Y \times Z \to Y$ given by $\pi(y, z) = y$ for all $(y, z) \in Y \times Z$ is continuous.*

Corollary 1.3.2. *If $f : X \to Y \times Z$ is continuous at $x \in X$ then $f_1 : X \to Y$ and $f_2 : X \to Z$ are continuous, where $f(x) = (f_1(x), f_2(x))$.*

Proof. $f_1 = \pi \circ f$. ■

Another problem arises when we consider that many propositions are used to build to a single theorem which encompasses them all. One example of this is in the definition of the Lebesgue integral. The definition usually follows four stages: defining the integral for simple functions, then for bounded functions supported on a set of finite measure, then for non-negative functions, then for all functions. At each stage, authors (such as Stein (9)) prove that the integral thitherto defined is linear, additive, monotone, and obeys the triangle inequality. The final proposition encompasses the previous propositions, yet relies on them being true.

Conclusion

My conclusion is that **mathematics**, as a *body of knowledge*, does not count as separate theorems which do not add to the body of knowledge. Thus, corollaries and theorems which are incorporated in higher theorems ought not to be counted.

Throughout this essay, I have taken to a descriptive account of mathematics and in doing this I have taken a loose definition of what constitutes a theorem: a theorem is something which is called a theorem by the contemporaneous mathematicians. This avoids the difficult question of what (even descriptively) is the essence of a theorem. A more rigorous follow-up of the essay would be able to answer this question.

Chapter 2

Blemishes in the history of mathematics

2.1 Analysis

One of the most important concepts in mathematical analysis* is the continuous function. Leonard Euler (1707 – 1783) introduced the notion of the mathematical function and the notion of continuity was implicit already. To Euler's contemporaries, "continuous" was understood to mean "analytic almost everywhere" in today's terminology.†

This led to several problems and inadequacies. One of these was the calculation of derivatives of continuous functions. Since continuity was equivalent to analyticity, every continuous function had a power series expansion.

$$f(x) = \sum_{n=1}^{\infty} a_n x^n.$$

*We will restrict ourselves to the discussion of *real* analysis.

†**Continuous** nowadays means small enough changes in the input will give changes in the output as small as desired. **Analytic** functions are necessarily infinitely-differentiable. So "analytic" is a stronger notion that "continuous". (For example the function $y = |x|$ is continuous but not analytic. It is, however, analytic almost everywhere.)

Joseph-Louis Lagrange (1736 – 1813), in his *Theories des Fonctions Analytiques* of 1797 defined the derivative of a continuous function as the second coefficient of the power series. However, no formal definition of continuous was attempted. This lack of definition resulted in some theorems which, if stated literally, are false today. (For example, it was thought that a continuous function was differentiable everywhere or almost everywhere. Karl Weierstrass (1815 – 1897) disproved this by constructing counterexamples like

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cos(3^n x), \tag{2.1}$$

which is continuous everywhere but nowhere differentiable.(6, p. 246) See figure A on p. 46.)

Augustin-Louis Cauchy (1789 – 1857) in 1821 introduced a formal definition of continuity: Let $f : \mathbb{R} \to \mathbb{R}$ be a function of x that is bounded on a finite interval. Now, the value of f at x being $f(x)$, consider the difference in the value of f an ever decreasing increment α from x:

$$f(x+\alpha) - f(x)$$

This depends on α and x. Cauchy then defines the notion of continuity in the following way:

> Cela posé, la fonction $f(x)$ sera, entre les deux limites assignées à la variable x, fonction **continue** de cette variable, si, pour chaque valeur de x intermédiaire entre ces limites, al valeur numérique de la différence
>
> $$f(x+\alpha) - f(x)$$
>
> decroit indéfiniment avec celle de α.[‡]

[‡] The function $f(x)$ will be . . . **continuous** if, for each value of x between those bounds, the numerical value of the difference

$$f(x+\alpha) - f(x)$$

decreases indefinitely with α.(2, Translation from p. 68)

I have included Cauchy's original words because the definition he gives is vague. In particular, it does not distinguish between continuity and uniform continuity.

> A function $f : \mathbb{R} \to \mathbb{R}$ is **uniformly continuous** if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that
>
> $$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$
>
> for all $x, y \in \mathbb{R}$

> A function $f : \mathbb{R} \to \mathbb{R}$ is **continuous** if for every $\varepsilon > 0$ and for all $y \in \mathbb{R}$ there exists a $\delta > 0$
>
> $$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$
>
> for all $x \in \mathbb{R}$.

The subtle difference is that in the modern-day definition the value of δ is chosen with regard to the value of y. In modern terminology, the definition Cauchy gave was that of *uniform continuity*, which is a stronger notion of continuity: all uniformly continuous functions are continuous but not all continuous functions are uniformly continuous. (See figure 2.1.) The failure to grasp this subtle difference meant that

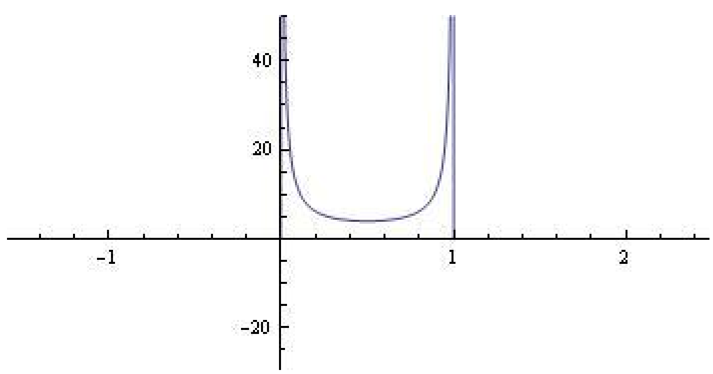


Figure 2.1: The graph of the function $y = \dfrac{1}{x(1-x)}$, on $(0, 1)$, which is a continuous but not uniformly continuous function.

many false theorems were spawned. One of them being in relation to multivariable functions:

> If the variables $x, y, z, \ldots$ have the fixed and determined quantities $X, Y, Z, \ldots$ as limits and if the function $f(x, y, z, \ldots)$ is continuous with respect to each variable $x, y, z, \ldots$ in the neighborhood of the system of values $X, Y, Z, \ldots, f(x, y, z, \ldots)$ will have the limit $f(X, Y, Z, \ldots)$.

The theorem is not true. The statement is only true if the function is *uniformly* continuous with respect to each variable.

2.1.1 Discussion

Cauchy's error is salient to this article because it is an example of an apparently sophisticated theorem that is wrong as stated. A possible way in which the cumulativeness of mathematics is rescued would be to deny that Cauchy's statement should have been admitted into the body of knowledge in the the first place. A defence of this position runs something like this: Cauchy's definition of continuity was used to prove a number of theorems (such as the multivariable function theorem), but the definition was too vague or "unmathematical", and thus not of the requisite standard needed to truly be a theorem.

Is Cauchy's statement unmathematical? The correct answer is, of course, another question: what is the essence of *mathematical* as opposed to *unmathematical* statements? The typical answer is that mathematical statements are *well-formed* and their terms *well-defined*. As our use of the prefix *well-* suggests, the word itself is not held to the same standard as its definiendum! To be fair, we will not quibble over the essence of a *well-formed* statement. However, arguing over the nature of well-defined terms as opposed to ill-defined statements is fair game. As a provisional definition, we could define a *well-defined* term as a term that is understood unambiguously by the readers or, equivalently, one that is strongly defined. For us, the contentious word is neither "unambiguously" nor "strongly" but "understood". A more descriptive–that is, historically valid—definition of a **well-defined** term is a term that is *thought* to be understood by its readers.

I do not have much evidence as to whether Cauchy's definition was well-understood to its readers. Considering Cauchy's prominence as a mathematician and considering the numerous textbooks and mathematical papers that were published for him that included this definition, it is likely that the definition was well-understood. It is easy to see, under this weak assumption, that the defence of the cumulativeness of mathematics in this case fails, for the only remaining defence would rest upon the idea that mathematical statements are easy to spot now, or that we have some strong test to determine whether a mathematical statement is well-defined and well-formed. And that is not the case. So Cauchy's statement that f is continuous at x if $f(x+\alpha) - f(x)$ decreases indefinitely with α is, descriptively, as mathematical as the modern-day definition of continuity at a point.

The moral of the story of Cauchy is that we are no wiser than we have ever been to judge whether a statement is unmathematical or mathematical. Since seemingly mathematical statements have been wrong in the past, we simply cannot expect our beloved theorems to not be similarly exposed in the future. What is probably fair to say, though, is that mathematicians soon discover mistakes. (In this case, it was Weierstrass.)

2.2 Algebraic number theory

Number theory is the branch of mathematics concerned with the properties of numbers, in particular the integers. It is one of the oldest branches of mathematics. The Hellenistic Greeks were avid scholars of the subject and the most famous, Diophantus of Alexandria, from whose name we get *Diophantine equations*—indeterminate polynomial equations in $\mathbb{Z}$—which have been the subject of study up until the modern day. (Indeed, the Holy Grail of Diophantine equations, Matiyasevich's theorem, which states that Diophantine equations are in general unsolvable, was proved in 1970.)

Algebraic number theory is a major branch of number theory which studies algebraic structures on $\mathbb{Z}$ and more exotic numbers systems. One of the most famous unsolved problem in mathematics, Fermat's Last Theorem, had many wrong attempts that led to the development of algebraic number theory. We examine some of these attempts.

2.2.1 Fermat's Last Theorem

Fermat's Last Theorem is the statement that

Let n be an integer $n \geq 3$. Then $x^n + y^n = z^n$ has no solution in the nonzero integers.

In 1637, Pierre de Fermat (1601 – 1655) wrote in a margin of his copy of Diophantus' *Arithmetica* that he had found a "truly marvellous proof" of the proposition but that the margin was too small for it. Nevertheless, mathematicians for the next 350 years tried and failed to prove it. It is one of the most famous mathematical problems and probably attracted the largest number of misproofs. The misproofs themselves were generally quickly unmasked, as a proof was highly prized and jealously coveted. However, a few instances were not revealed and, while the proofs were actually easily repaired, the lemmas were not identified as flawed.

One particular example comes from Euler, "without a doubt the greatest mathematician of his time".(3, p. 39) In his textbook *Vollständige Anleitung zur Algebra*, the uniqueness of factorization was assumed in several proofs, even when the system he was dealing with did not admit it.

The "uniqueness of factorization" is the property of the fundamental theorem of arithmetic* in other domains. In $\mathbb{Z}[\sqrt{-5}]$, for example, it is not true. (The ring $\mathbb{Z}[\sqrt{-d}]$ is the set of all complex numbers $a + b\sqrt{-d}$ where $a, b \in \mathbb{Z}$. They have many interesting properties and have applications to factorization.)

The numbers 2 and 3 are prime and it can be shown that $1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ are also prime. Yet

$$6 = 2 \cdot 3 = (1 - \sqrt{-5})(1 + \sqrt{-5}).$$

That is, there is a number, six, that is not uniquely expressible as a product of

*In $\mathbb{Z}$, the **fundamental theorem of arithmetic** says that every number is expressible as the product of primes and (importantly) that this representation is unique (up to rearrangement of the primes). So $54 = 2^2 \times 3^2$ and there are no other primes $p_1, p_2, \ldots, p_n$ such that $54 = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_n^{\alpha_n}$.

primes. Euler was not aware of this and assumed that domains like $\mathbb{Z}[\sqrt{-5}]$ obeyed unique factorization. One of the more notable occasions in which this error was made was in his proof of Fermat's Last Theorem for $n = 3$.

In the proof, Euler relied on a lemma:

Lemma 2.2.1. *Let p, q be prime numbers. If $p^2 + 3q^2$ is a cube then there exist $a, b \in \mathbb{Z}$ such that $p = a^3 - 9ab^2$ and $q = 3a^2b - 3b^3$.*

Euler's proof of this lemma is based on the elegant but mistaken corollary of unique factorization that if $(x+y\sqrt{-3})(x-y\sqrt{-3})$ is a cube and if $(x+y\sqrt{-3})$ and $(x-y\sqrt{-3})$ are relatively prime—that is, have only 1 or -1 as common divisors—then $(x+y\sqrt{-3})$ and $(x-y\sqrt{-3})$ are also cubes. To illustrate this corollary, consider it in $\mathbb{Z}$, (or, if you're game, in any system where the fundamental theorem of arithmetic holds):

Proof. Suppose we have a cube c^3 for some $c \in \mathbb{Z}$. And suppose that $c^3 = ab$ for some $a, b \in \mathbb{Z}$. The claim is that if a and b are relatively prime then $a = r^3$ and $b = s^3$ for some $r, s \in \mathbb{Z}$. By the fundamental theorem of arithmetic,

$$c^3 = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_n^{\alpha_n}.$$

Since $c \in \mathbb{Z}$, by the FToA again $c = q_1^{\beta_1} q_2^{\beta_2} \dots q_m^{\beta_m}$, which implies that

$$c^3 = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_n^{\alpha_n} = q_1^{3\beta_1} q_2^{3\beta_2} \dots q_m^{3\beta_m}.$$

Since these expansions are unique it follows that $n = m$, $p_i = q_i$ (after rearrangement, which the FToA allows) and each α_i is a multiple of 3. Since $c^3 = ab$, by the FToA, ab must also be the product of primes with powers divisible by 3. Ex hypothesi, a and b have no common prime factor; thus, allowing rearrangement,

$$a = p_1^{3\beta_1} p_2^{3\beta_2} \dots p_j^{3\beta_j} \qquad b = p_{j+1}^{3\beta_{j+1}} \dots p_n^{3\beta_n}.$$

Therefore, a and b are each the product of primes with powers divisible by 3 and so are cubes, as required. ■

The proof is fairly simple to understand, even if it does require some subscripts on superscript text. The crucial aspect of the proof is the reliance on the uniqueness element of the FToA. Without it, the proof fails and a and b need not be cubes.

So Euler's proof was wrong. According to (3, p. 45), a "reasonable speculation about this erroneous proof is that his original method used a less imaginative [but logically sound] argument to prove ... [the lemma], and that it was only later that he had the elegant—but wrong—idea of proving this by using [unique factorization in $\mathbb{Z}[\sqrt{-3}]$".

2.2.2 Lamé

Another more spectacularly incorrect proof concerning Fermat's Last Theorem was by Gabriel Lamé (1795 – 1870). This was more spectacular partly because the proof was unsalvageable (he tried to cross an uncrossable part of the river, whereas Euler just picked the wrong stepping stone) but also because it was a proof for all $n \geq 3$ and not just for a particular n.

His proof, too, relied on unique factorization. The synopsis of the proof is as follows:(10, p. 5)

(1) Suppose $x^n + y^= z^n$ has an integer solution.

(2) We can factorize the left-hand side of $x^n + y^n = z^n$ as

$$(x + y)(x + \zeta y) \dots (x + \zeta^{n-1} y) = z^n.$$

where ζ is an n-th root of unity, $\zeta = \mathrm{e}^{2\pi \mathrm{i}/n}$.

(3) The factors $x + y$, $x + \zeta y$, ..., $x + \zeta^{n-1} y$ are relatively prime. Therefore, there

must be distinct elements $u_1, \ldots, u_n$, in what we now call $\mathbb{Z}[\zeta]$ such that

$$\begin{aligned} x + y &= u_1^n \\ x + \zeta y &= u_2^n \\ &\vdots \\ x + \zeta^{n-1} y &= u_n^n \end{aligned}$$

Since each of the factors are relatively prime, each of the u_i^n must be relatively prime, a contradiction, since this leads to an infinite descent.

Lamé presented a proof of Fermat's Last Theorem based on this synopsis at the Paris Academy in 1847. After he had made his presentation, Joseph Liouville (1809 – 1882) cast doubts on the problem. The gap, as he correctly pointed out it, was that the proof depended on the uniqueness of factorization of the factors (and thus the u_i^n) and that this, while true in $\mathbb{Z}$, was not obviously true in $\mathbb{Z}[\zeta]$. While most present accepted the logical validity of Liouville's argument, Lamé (and Cauchy, who was in attendance) believed that this was not a fatal flaw in the proof.

2.2.3 Discussion

Central to this story is the reliance the mathematician places in his intuition. Euler thought that $\mathbb{Z}[\sqrt{-3}]$ was enough like $\mathbb{Z}$ to justify the fundamental theorem of arithmetic for $\mathbb{Z}[\sqrt{-3}]$. For the question at hand, we have to ask: Is Euler's error typical of mathematics?

There are historical facts surrounding the faux pas that should be illuminated. Firstly, Euler's assertion that $x^3 + y^3 = z^3$ has only the trivial solutions is correct. Secondly, and more importantly, Euler had correctly proved that $\mathbb{Z}[\sqrt{-3}]$ was a unique factorization domain in earlier works. So Euler may have elected not to focus on the more subtle parts of a proof of a proposition he knew to be correct. However, his repetition of the incorrect proof in later works (including textbooks) suggests that he was unaware of the fallacy of the proof.

There is some evidence that mathematicians in the early 19th century were aware of the flaws in Euler's proofs in his *Anleitung zur Algebra* (1770). Euler's prominence and almost heroic stature among mathematicians of the time meant, however, that these criticisms were not published.

Though Lamé's fall is perhaps more spectacular, it is less interesting for our purposes because Liouville picked up on it almost immediately. Mathematics' cumulativeness is not inconsistent with mathematicians making errors: it is only when their errors are accepted by the mathematics community that that thesis is vitiated. Lamé accepted the criticism at once and, although he did not believe that the proof was deeply flawed, his belief was not made with the force or conviction of a mathematical assertion.

2.3 Geometry

Geometry comes from the Ancient Greek word γεομετρια. geo-, 'Earth' + *metria*, 'measure'. Its etymology is significant in an account of its history: geometry's links with the measurement of the universe were both its blood and its poison.

Geometry was initially a body of practical rules and experience for practical engineering and construction work. The *Elements* by the Greek mathematician EUCLID (c. 300 BC), of which geometry is a major part, is a canonical work in geometry and pre-eminent in its success and influence as a textbook of mathematics. Euclid's approach made the rules and practical knowledge *theorems*, rigorously proven generalities using deductive logic from the axioms alone. These axioms were uncontroversial and included the axiom that "any two points can be joined by a straight line". Thus, theorems that followed from these axioms were regarded as absolutely true and mathematicians regarded *Elements* as "the model of rigor".(4, p. 1007) Isaac Barrow (1630 – 1677), an ardent defender of the geometry, gave eight reasons to support geometry's certitude: "the clearness of the concepts, the unambiguous definitions, the intuitive assurance and universal truth of its axioms, the clear possibility and easy imaginability of its postulates, the small number of its axioms, the clear conceivability of the mode by which magnitudes are generated, the easy order of the demonstrations, and avoidance of things not known."(4, p. 862)

The most compelling philosophical argument, however, came from Immanuel Kant (1724 – 1808). In his *Critique of Pure Reason* (1781) he rejected Hume's empiricism and argued there are synthetic a priori truths. He maintained that the human mind has intrinsic rules or modes for organizing the sensory inputs coming from the real world. He called these modes **intuitions**. This explains the popularity (or ubiquity) of mathematical objects: they are, in fact, prior to experience. Kant explicitly gives (Euclidean) geometry as a body of knowledge that is equivalent to these intuitions and their logical consequences—collectively called **a priori synthetic truths**.

Now is not the time to discuss whether Kant's account of a priori knowledge is persuasive or whether his attack on Hume's empiricism succeeds. What is important from an historical perspective is that mathematicians and physicists considered Kant's defence of Euclidean geometry persuasive at the time it was written. Euclidean geometry (together with Newtonian mechanics) was seen as equivalent to the intuitions of space and time.

But there were problems, though most were unaware of them. As Kline puts it in (4, p. 1007): "one should not rely on drawing a correct figure to determine the location of [the interior of a triangle]. But that is precisely what Euclid and mathematicians up to 1800 did." and "Euclidean geometry was supposed to have offered accurate proofs of theorems suggested intuitively by figures, but actually it offered intuitive proofs of accurately drawn figures." Yet the *Elements* "were generally taken to be the model of rigor." These criticisms of Euclid's rigor in his work are valid and serious, but most of his theorems were valid and his proofs reparable. One proposition, however, has a dubious history and warrants special mention.

2.3.1 Parallel postulate

A **postulate**, as opposed to a theorem, is something which is not a definition but is considered self-evident and, as such, is introduced without proof. It is often synonymous with *axiom*.[*] Euclid's *Elements* Book 1 begins with 23 definitions (of notions like *point* and *line*) and then sets the foundation with five postulates, the first four of which are:

1. A straight line segment (or interval) can be drawn by joining any two points.
2. An interval can be extended indefinitely in a straight line.
3. Given an interval, there exists a circle whose centre lies on one endpoint of the interval and whose radius is the interval.
4. All right angles are congruent.

The parallel postulate is the fifth postulate in Euclid's *Elements*. It roughly states that any straight lines will intersect unless they are parallel.

> *If a line segment intersects two straight lines forming two interior angles on the same side that sum to less than two right angles, then the two lines, if extended indefinitely, meet on that side on which the angles sum to less than two right angles.*

Being a postulate, Euclid attempted no proof. Nevertheless, the parallel postulate was not universally regarded as self-evident, though almost universally regarded as true. Because of this, it was widely believed that the postulate could be derived from the other four postulates.

The truth of the parallel postulate was widely believed and there are many other statements which were certainly regarded as self-evident that are, in fact, equivalent to the parallel postulate, including "squares exist" and "triangles can be big".[†] These statements are so intuitively obvious that many mistaken proofs were attempted in proving the parallel postulate.

---

[*] I would say that the difference is that an axiom is introduced either because it is self-evident or to enact the governing laws of the domain it is axiomatizing, whereas a postulate is something that is obviously true from experience. Einstein's postulate that the speed of light is constant is a famous example.

[†] More precisely, "there exists a quadrilateral whose interior angles are all right angles" and "there is no upper limit to the area of a triangle".

Ptolemy attempted a proof from the other postulates and theorems. However, his proof made too many assumptions, effectively begging the question. PROCLUS (c. 411 – 485) was one of the last Classical philosophers. He thought that the postulate should be deleted from the geometry as a postulate "for it is a theorem involving many difficulties, ... and it requires for the demonstration [proof] of it a number of definitions as well as theorems, and the converse is actually proved by Euclid himself as a theorem."(as quoted in (4) on p. 864) In addition, he made the observation that although two lines may converge upon each other, it does not follow that they intersect at a (finite) point in the plane. For instance, the hyperbolas formed by the graph of $xy = \pm 1$ converge but do not meet in the coordinate plane.

The most concerted effort to prove the postulate was made by Gerolamo Saccheri (1667–1733). Nadsīr-al-Dīn al-Tūsī (1201–1274) and John Wallis (1616 – 1703) had attempted to prove it before, but they assumed the intuitively obvious statements that are equivalent to the postulate, effectively begging the question. Saccheri's proof went something like this:

Consider a quadrilateral $ABCD$ with the angles A, B, C, D labelled cyclically. Let $\angle A = \angle B = 90°$ and let $AC = BD$. It easily follows that $\angle C = \angle D$. Now, if $\angle C = \angle D = 90°$ then the parallel postulate is true. Suppose not. That is, suppose that either $\angle C$ and $\angle D$ are acute or $\angle C$ and $\angle D$ are obtuse. Omitting the details, in the case of $\angle C$ and $\angle D$ acute, he obtained a contradiction. In the second case, he proved that if the parallel postulate was not true then two lines which do not intersect will have their common perpendicular at ∞.

Though there was no real logical barrier to accepting this, Saccheri found it so repugnant that he declared he had found a contradiction. He felt that Euclid's postulate had been proved, going so far as to write a book in his dying year entitled *Euclides ab Omni Naevo Vindicatus* (Euclid Vindicated from All Faults).

However, Saccheri was wrong: the parallel postulate was not a logical consequence of the other postulates.

2.3.2 Non-Euclidean geometry

Non-Euclidean geometry arose from the realization that Euclid's parallel postulate was independent of the other postulates and thus could be removed and replaced without threatening the internal consistency of the geometry. Georg Klügel (1739 – 1812) and Johann Heinrich Lambert (1727 – 1788) recognised that there can be alternative geometries to Euclidean. Klügel, writing in his dissertation of 1763, wrote that the "certain" quality afforded to Euclidean geometry is based on experience. Lambert published a similar paper to Saccheri's, but did not believe he had arrived at any contradiction, just a result at odds with experience. Saccheri, though he did not regard the consequences of his technique proper, nevertheless was instrumental in the technical development of non-Euclidean geometry. However, it was Gauss (1777 – 1855) who had the most significant realization: that non-Euclidean geometry was just as accurate in describing physical space as Euclid's geometry was. In 1799, having thought up until then that Euclidean geometry was the only true geometry, he wrote to Bolyai and appeared to be sceptical that the parallel postulate was dependent on the others and that other geometries (that is, statements that use different postulates) were logically sustainable. He makes himself clear in his letter to Olbers in 1817 (4, quoted in p. 872) that Euclidean geometry is no longer necessarily the geometry of physics and that "we must place geometry not in the same class with arithmetic, which is purely a priori, but with mechanics".

Gauss realized that the parallel axiom was independent of the other axioms of Euclidean geometry. Thus, it is possible to introduce a contradictory axiom without inconsistency. He believed that this meant that Euclidean geometry was not the geometry of physics and so was not in the same class as arithmetic. This was a strange conclusion to draw, considering that arithmetic (and other equally rigorous branches of mathematics like algebra and analysis) had no logical foundation and were just as intuitive in their structure as geometry. That is not to say that Gauss was wrong to segregate geometry from arithmetic, only that it is curious he did not consider whether these other branches were equally artificial. Indeed, "[t]he history of non-Euclidean geometry reveals in a striking manner how much mathematicians are influenced not by the reasoning they perform but by the spirit of their times."(4, p. 880) Saccheri had reached the same conclusions as Gauss, Lobatchevsky, and Bolyai, but Saccheri rejected the conclusions as absurd whereas GLB accepted them. The consistency of non-Euclidean geometry was never proved. Bolyai and Lobatchevsky believed it "because their trigonometry was the same as for a sphere of imaginary radius and the sphere is part of Euclidean geometry."(4, p. 880) This history also demonstrates that mathematicians were not motivated by their own intellectual curiosity; they were not just crazily omitting axioms and seeing what the machine spat out. The non-Euclidean geometries arose from the centuries of work on the parallel axiom. Since Euclidean geometry was the canvas on which physics developed its laws, mathematicians wanted to make sure their beloved physicists were relying on truths.(4)

2.3.3 Discussion

If mathematics is cumulative then geometry is cumulative. Does the history of geometry support the hypothesis that geometry is a cumulative branch? To begin with, we consider a simpler question: is non-Euclidean geometry a contradiction of Euclidean geometry or a mere addition to it?

To answer this sensibly, we need to understand that what is meant by 'Euclidean geometry' now is different to what was meant before the 19th century. Quite understandably, in light of the history, there was no thought that Euclidean geometry was a subset of some family of geometries. 'Euclidean geometry' was simply 'geometry', and Euclid happened to have been the earliest extant axiomatization of it. Geometry and Euclidean geometry were equivalent in mathematics. Thus, it is premature to assert that non-Euclidean geometry is not a contradiction of Euclidean geometry because they exist in independent logical frameworks, even if we believe (as I do) that they do. It is premature because Euclidean geometry was believed to be the only consistent geometry. So while the theorems of Euclidean geometry remained while non-Euclidean geometry was founded, its axiology and the axiological matters of fact in geometry were torn asunder. In particular, the idea that the theorems in (Euclidean) geometry were essential a priori truths about the physical world was shown to be unfounded in the 19th century (and absolutely demolished in the 20th by Einstein).

But were these "axiological matters of fact" part of mathematics? Kant's philosophy may have been popular among pre-Gaussian geometers, but did they think, when they believed that Euclidean geometry was the geometry of space and time, that they were doing mathematics, or merely remarking on aesthetics? In other words, was the pure mathematics agnostic about the physical status of Euclidean geometry?

No. Saccheri's demonstration is telling. His misproof of the parallel axiom is based on his aversion to the idea of a common perpendicular of two lines happening at ∞. Yet this idea was not rejected by Gauss. The later acceptance of a theorem is only evidence for cumulativeness if the mathematics' community was agnostic or unaware of the proposition beforehand. In this case, it was aware and it was rejected; equivalently, a false theorem was accepted.

Perhaps Saccheri was logically repulsed by the theorem accepted by Gauss because he was working in the Euclidean domain? If he was, then he was confused, because he was trying to establish whether the Euclidean domain was consistent, and it would be begging the question to conclude that it was on the basis of itself. If he were not working in the Euclidean domain, then his assertion is plain wrong and the geometric corpus of knowledge is no longer cumulative.

The history of geometry shows that, up until the 20th century, the mathematical, axiological, and physical ideas from geometry were inheterogeneously mixed. Geometers threw all the components in with each other and labelled physical facts as mathematical theorems. They were confused. But confusion is not a defence here, as it may have been with Cauchy's fable, because they were not aware of the confusion – indeed, they thought (thanks to Kant) that their bouillabaisse géometrique was logically necessary.

2.4 Number

God invented the integers. Everything else is the work of man.

Krönecker (1823 – 1891)

The history of numbers is the history of stubborn mathematicians. Today, it is difficult to imagine why mathematicians were so reluctant to accept complex numbers, let alone real numbers or negative numbers and the number zero! Krönecker's oft-quoted salute to the constructed character of numbers could easily have been a tongue-in-cheek reference to the sloth of Man's acceptance of the higher number systems.

2.4.1 Beyond $\mathbb{N}$

Greece

Two of the prominent schools of mathematics were the Pythagorean school and the School of Eudoxus.

The Pythagoreans were concerned with natural numbers only. They were familiar with fractions of natural numbers (the rational numbers) as these were common in commerce and engineering, but these were considered "labourers' tools", and so beneath the Pythagoreans. Finding order in the universe was part of the Pythagoreans' ideology and it coloured how they did mathematics. Their antipathy towards non-natural numbers, combined with the profound mysticism of the school, led to some strange consequences. The legend of Hippasus discovery that $\sqrt{2}$ was αρρητος ("irrational" in today's terminology) being treated as heretical and being thrown overboard is probably apocryphal, but is a useful illustration of the Pythagorean ideals. To be accurate, Pythagoreans accepted that $\sqrt{2}$ was irrational, but would not accept such magnitudes as "numbers".

The School of Eudoxus introduced the notion of magnitudes to resolve the problem of incommensurability. Magnitudes were not numbers but stood for entities and could vary continuously. Euclid mentions magnitudes in fifth book of *Elements*. But neither Euclid nor Eudoxus gave any formal axiomatization of the operations on magnitudes and it is not clear whether they allowed operations on magnitudes.

Hindu-Arabic mathematics

The simplest irrational numbers are the algebraic integers that are solutions to quadratic equations: such as $\sqrt{2}$ or $\dfrac{1+\sqrt{5}}{2}$. Hindu mathematicians started to be comfortable with operating with surds using correct procedures, and in the 12th century Bhāskara developed the following rule.

$$\sqrt{a}+\sqrt{b}=\sqrt{(a+b)+2\sqrt{ab}} \qquad ((4, \text{p.}185))$$

The distinction between the Greeks and the Hindus was that the Greeks were more dogmatic and exclusive, whereas the Hindus were less sophisticated in accepting rules by analogy without any logical justification for why these rules held. Nevertheless, their work and persistence allowed mathematics to progress (though things may have turned out differently if they had been wrong).

The Islamic mathematicians Omar al-Khayyami (c. 1048–1122) and Nadsīr-al-Dīn al-Tūsī (1201–1274) declared separately that every commensurable or incommensurable ratio may be called a number. "It is fairly certain that the Hindus did not appreciate the significance of their own contributions. The few good ideas they had, such as separate symbols for the numbers from 1 to 9, the conversion to base 10, and negative numbers, were introduced casually with no realization that they were valuable innovations."(4, p. 190) The key shifts from the Greek tradition to the Hindu-Arabic system of arithmetic was the acceptance of irrational numbers and the emphasis on arithmetic, rather than geometry, in algebra. Further, while the Hindus did use negative numbers, they were initially suspicious of them. For instance, if a problem had two distinct solutions—a positive number and a negative number—the positive number was regarded as more correct.

16th-17th Centuries

Around the end of the 15th century, zero was accepted as a number.(4, p. 251) And irrational numbers were used without hesitation. François Viéte (1540–1603) discovered and proved the infinite product

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots$$

showing too how accepted irrational numbers had become.(12)

Even so, the club of numbers was still quite exclusive. Wallis in his *Arithmetica Infinitorum* (1655) argued that since a/b grows larger and larger as b becomes smaller, and since $a/0$ is infinite, a/b must be larger than infinity when b is smaller than zero (i.e. negative). Even Euler made a similar faux pas. Antione Arnauld a 17th century theologian and mathematician questioned the existence of negative numbers. He said (correctly) that they led to the identity $-1 : 1 = 1 : -1$. Yet how could the ratio of a lesser to a greater be the same as the greater to the lesser? A contradiction.(4, p. 252)

The utility of irrationals was evident, but for them too there was still dispute over whether they were truly numbers. Michael Stifel (1486/87 – 1567), an Augustinian monk, argued in his *Arithmetica Integra* (1544) that because irrational numbers have an infinite decimal expansion with no repetitive structure and since "that cannot be called a true number which ... lacks precision", irrational numbers are not numbers.(4, p. 251) Today, irrational numbers can be made precise in $\mathbb{R}$ through equivalence classes of Cauchy sequences.*

18th Century

By 1700, the elements of $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ were known. Yet still, there was opposition to these numbers, even negative numbers were occasionally persecuted. Quoth English mathematician Baron Francis Masères:

*For an easily accessible article, see "Construction from Cauchy sequences" in http://en.wikipedia.org/wiki/Construction_of_the_real_numbers, accessed 12 November 2008.

> ... they [negative roots of polynomial equations] serve only ... to puzzle the whole doctrine of equations. ...It were to be wished therefore that negative roots had never been admitted into algebra or were again discarded from it[.]

19th century

While irrational numbers were accepted and used without hesitation in all of the 19th century, there were still gaps in mathematicians' knowledge. Before Weierstrass, irrational numbers were regarded as the limit of an infinite sequence of rational numbers. However, this limit is not well-defined unless the limit is defined in a set containing irrationals and rationals. Cantor recognised this flaw and presented a solution. Cantor's solution (*Math. Ann.,* 5, 1883, 545-591) was to invent $\mathbb{R}$, the set of real numbers which contain all the rational and irrational numbers. He defined $\mathbb{R}$ to be (in modern terminology) a *complete, ordered field.* In this context, **complete** means that every Cauchy sequence converges. Equivalently, it means that every nonempty subset of $\mathbb{R}$ with an upper bound has a least upper bound. This is what distinguishes $\mathbb{R}$ from $\mathbb{Q}$. Consider the set of all rational numbers less than $\sqrt{2}$. This is certainly a nonempty subset of $\mathbb{R}$ with an upper bound (for instance, 1000 is bigger than all of the numbers). But there is no least upper bound.

2.4.2 Into $\mathbb{C}$

The complex numbers are numbers of the form $a + b\mathrm{i}$ where a and b are real numbers and $\mathrm{i} = \sqrt{-1}$. The square root of negative one is a difficult concept to understand even today: if you have a simple calculator and try to extract the square root of negative one, it will complain in the form of an error message! Yet they are extremely useful both within mathematics and in other scientific disciplines. The fact that complex numbers were originally considered impossible or literally "imaginary", but today have so much utility both in pure and applied mathematics can only be regarded as a miracle.

Today, the complex numbers are introduced by declaring the necessity of all

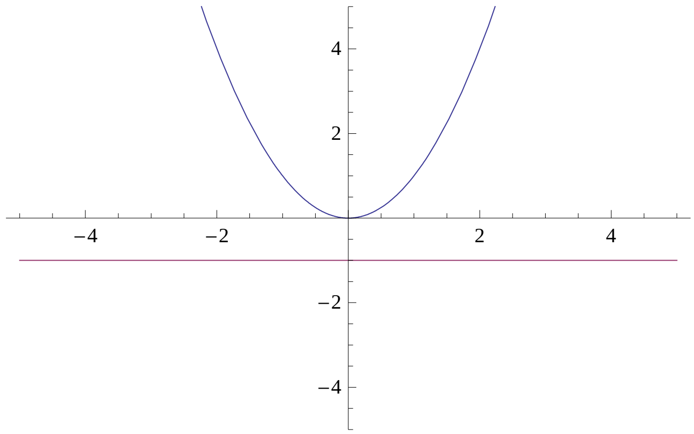


Figure 2.2: The parabola $y = x^2$ and the line $y = -1$. The imaginary solution $x = \mathrm{i}$ to the equation $x^2 = -1$ is the point at which the parabola and the line intersect. The pre-Classical mathematicians only ever encountered quadratic equations that had only complex solutions in geometrical problems like these. Since there is no solution to the geometrical problem, mathematicians saw no need to invent $\mathbb{C}$.

quadratic equations such as $x^2 + 1 = 0$ to have solutions. But this rhetoric, though it is undoubtedly a clear pedagogical device, was not persuasive in the history of algebra to convince mathematicians to employ them. The reason behind this apparent stubbornness was that geometry was the usual source of problems involving quadratic equations. A typical problem might be the intersection points of curves of conic sections and straight lines. Since there may be no intersection points (a problem corresponding to the equation $x^2 + 1 = 0$ is shown in figure 2.2), there was no need to look deeper for solutions. The mathematical problems which did motivate the beginning of acceptance of complex numbers were those involving cubic equations such as

$$y^3 = py + q.$$

The solution to this equation was developed by Niccolo Tartaglia and Gerolamo Cardano.

$$y = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}.$$

This solution involves complex numbers since p, q are arbitrary.(11, p. 257) This was known immediately and mathematicians began to experiment more with these strange numbers.

Descartes, in response to the growing use of complex numbers drew a clear distinction between real and imaginary roots of equations: real roots (even negative roots) are admissible because the equation can be transformed into one in which the roots are positive (and so more coherent when dealing with concepts like distance and duration) whereas complex roots do not have such a transformation, so they are not real (true) numbers but imaginary (fictitious) numbers.

Newton did not regard complex numbers as significant, because of their irrelevance to the physical problems he was concerned with. Though Leibniz worked with complex numbers (for instance in the reduction of imaginary expressions to real form), he professed no deep understanding of the operations governing them.(7) The muddiness and mystical curiosity involved in the European experience with complex numbers is illustrated by the passage by Liebniz:

> "The Divine Spirit found a sublime outlet in that wonder of analysis, that portent of the ideal world, that amphibian between being and non-being, which we call the imaginary root of negative unity." (4, p. 254)

Indeed, the Europeans "blundered" into complex numbers.(4, p. 253) One of the common mistakes was to apply surd rules to complex numbers (as the Hindus had to irrational numbers). For instance, a rule of surds says that $\sqrt{a}\sqrt{b} = \sqrt{ab}$. But to use this rule in $\mathbb{C}$ would imply that

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

and so $-1 = 1$. It is therefore with no surprise that some rejected complex numbers outright: Bombelli regarded the rules of operation governing complex addition, subtraction, multiplication, which are accepted today, as not merely useless but "sophistic".

2.4.3 Discussion

Numbers are what we as children were first introduced to in our mathematical education. So it may seem a bit contrary to question whether the concept of "number" is a mathematical concept. However, questioning whether the various idealizations of "number" were mathematical is not the same as questioning whether numbers are mathematical. What is important is how mathematicians have talked about numbers. Even today, the branch of mathematics called "number theory" deals almost exclusively with integers and prime numbers, yet we accept number systems as far removed from $\mathbb{Z}$ as the complex and hypercomplex numbers.

Again, as with our discussion on geometry, we have to distinguish between mathematicians' thoughts on mathematics and their submissions of statements to the body of knowledge of mathematics. When, say, Baron Francis Masères said he wished that negative roots of polynomial equations had never been admitted into algebra, we have to decide whether he is declaring *as a mathematician* that negative roots are not allowed, or merely expressing his taste. In contrast to geometry, I believe that mathematicians were indeed just making aesthetic judgements and did not believe that their statements were in the same class as arithmetic statements. Alternatively, even when their thoughts were expressed with the hope of changing mathematics, their attitude was one of caution in the face of new numbers, rather than denying their existential possibility. This attitude is typically mathematical and does not deny mathematics a cumulative character. That is, mathematicians' apparent discontent with negative numbers, with irrational numbers, and with complex numbers, was not because they believed that these new numbers were logically invalid, but rather because they were unsure how to use them safely in mathematical proofs.

The Hindus' confidence with surds and their assuredness in using irrational numbers in the same way as rational numbers was a fluke, and any similar bravado by a mathematician nowadays would be struck down as non-rigorous. Indeed, the identical application of those rules to $\mathbb{C}$ was unsound. The general respect for cautiousness that we have described, whether from fanatical Pythagoreans or Bombelli, was probably justified, even though the expression of this respect was misplaced.[†]

Nevertheless, while the body of knowledge concerned with numbers may have remained cumulative throughout the extension of the number systems, the motivating and persuasive forces involved were not the ones a normative account of mathematics would suggest, i.e. deductive validity. That is, although mathematics was cumulative in this chronology, in the sense that it only accepted a statement's truth when that statement's truth was not later vitiated, it was, perhaps, cumulative for the wrong reasons. This is the distinction between "properly cumulative" and other notions of cumulative that was made before our history took place.

Furthermore, it is to defend too much Descartes when he makes qualitative judgements discriminating against 'imaginary' solutions to equations. And it was certainly wrong to say that solutions outside a contemporaneously fashionable number system are invalid. It is, at best, very confused thinking to accept the logical validity of a solution but to exclude it from answers. True, often the problem was connected to some physical system where negative or imaginary solutions were not relevant; my criticism is that mathematicians denied these solutions when no physical system was in play.

Thus, in the case of number theory, while it may have been cumulative in the sense of "strictly cumulative" it was not "properly cumulative".

[†]The author is not suggesting Musean hypernumber theorists be executed.

Chapter 3

Conclusion

Mathematics is certainly not strictly cumulative

Mathematics is not strictly cumulative because there have been theorems in the history of mathematics that have been vitiated. It does not really matter how we determine whether a theorem has been accepted or not by the mathematical community. In Cauchy's case, it almost certainly had been, by any measure. As discussed, Cauchy's false proof cannot be attributed to a definition that is unusually or obviously ambiguous. The mathematics of today is not markedly different to the mathematics that Cauchy studied and contributed to. The other episodes in history that we have seen illustrate that mathematics is not, in its entirety, a cumulative body of knowledge, in any meaningful sense.

This may seem a harsh conclusion and deserves some discussion.

Chapter 4

Discussion

Many questions remain. Two of the most important that come to mind are "Is it possible to change mathematics so that it can be cumulative?" and "Although mathematics is not strictly cumulative, why is *so much* of it strictly cumulative?". For reasons of space, I will not attempt answers to these questions, but will finish by posing a question that ought to be answered before the others.

4.1 Why do we think mathematics is cumulative?

We have concluded that mathematics is not cumulative in any strong sense of the word. But why would we think mathematics to be cumulative in the first place? There are several reasons why mathematics is regarded as such:

Mathematics is based on deductively valid tools Ayer's remark that mathematics is tautological is shared by many who have an opinion on this subject. Mathematics is true because mathematics is true. The reason why mathematics is regarded as cumulative is because the parts of mathematics that have been shown to be flawed are ipso facto not parts of mathematics. The question-begging is not egregious, however, because similar fallacies are used to "justify" deductive logic.

The core mathematics is the first taught We are not taught, from the outset, controversial mathematical ideas. Our mathematical education begins with the tried-and-tested mathematical theorems. Not even in undergraduate education are students exposed to anything controversial. We are taught the strictly cumulative core of the weakly cumulative mathematics, to use my terminology.

Bad theorems have gone In physics and astronomy, it has been noted that if the fundamental constants of the universe had slightly different values the unverse would be radically different. For instance, if the strong nuclear force, the force that holds quarks and gluons together to form protons and neutrons, were a mere 2% stronger, the nature of nuclear fusion, which powers the stars, would alter stellar physics—and thus the nature of life—dramatically. The moral of this observation is that the universe appears to have been specially-adjusted to support life on Earth. A response to this amazement might be to attribute the universe's fine-tuning to God. Another response is the **anthropic principle**, which argues that humans should not be suprised at this apparent fine-tuning, for if the constants were so different that life could not exist, life could not have existed, and we would not be around to be astounded by it.

An analogy can be drawn between this view and a historical view of mathematics: we judge the veracity and rigor and robustness of theorems in mathematics by what remains; all the deductively invalid mathematics has not survived, and we are left astounded by how deductively valid mathematics has always been!

Glossary

C

$\mathbb{C}$ The set of all numbers $a + b\mathrm{i}$ where a, b are real numbers and $\mathrm{i} = \sqrt{-1}$.

F

FToA The **fundamental theorem of arithmetic**, which states that every natural number greater than 1 can be expressed uniquely as the product of primes, up to rearrangement.

N

$\mathbb{N}$ the (set of) natural numbers; the positive integers; the set $\{1, 2, 3, \dots\}$; the counting numbers.

Q

$\mathbb{Q}$ the rational numbers; the set of all numbers $r = p/q$ where p, q are integers.

R

$\mathbb{R}$ the real numbers; the unique complete Archimedean ordered field; the union of the rational numbers and the irrational numbers.

Z

$\mathbb{Z}$ — the set of integers; the union of zero and $\mathbb{N}$ with its additive inverses; the set $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

$\mathbb{Z}[\sqrt{-d}] = \mathbb{Z}[\delta]$ — the ring consisting of the subset of $\mathbb{C}$ of all numbers $a+b\sqrt{-d}$ where $a, b \in \mathbb{Z}$.

Bibliography

[1] A. J. Ayer. *Language, Truth, and Logic*. Dover Publications, 1952.

[2] Bruno Belhoste. *Augustin-Louis Cauchy: a biography*. Springer-Verlag, New York, 1991.

[3] Harold M. Edwards. *Fermat's Last Theorem: A genetic introduction to algebraic number theory*. Springer-Verlag, 1977.

[4] Morris Kline. *Mathematical Thought from Ancient to Modern Times*. Oxford University Press, 1990.

[5] Thomas S. Kuhn. *The structure of scientific revolutions*. University of Chicago, Ltd, Chicago, 3rd edition, 1996.

[6] François Le Lionnais. *Great Currents of Mathematical Thought: Mathematics: Concepts and Development*, volume I. Courier Dover Publications, 2004.

[7] R. B. McClenon. A contribution of Leibniz to the history of complex numbers. *The American Mathematical Monthly*, 30(7):369–374, 1923.

[8] K. Popper. *The Logic of Scientific Discovery*. Hutchinson, 1934, 1959.

[9] Elias M. Stein & Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press, 2005.

[10] Ian Stewart. *Algebraic number theory and Fermat's last theorem*. A K Peters, 3rd edition, 2002.

[11] John. Stillwell. *Mathematics and its history*. Springer, New YorkS, 2nd edition, 2002.

[12] Wikipedia. Viète's formula — wikipedia, the free encyclopedia, 2008. [Online; accessed 9-October-2008].

Index

Appendix A

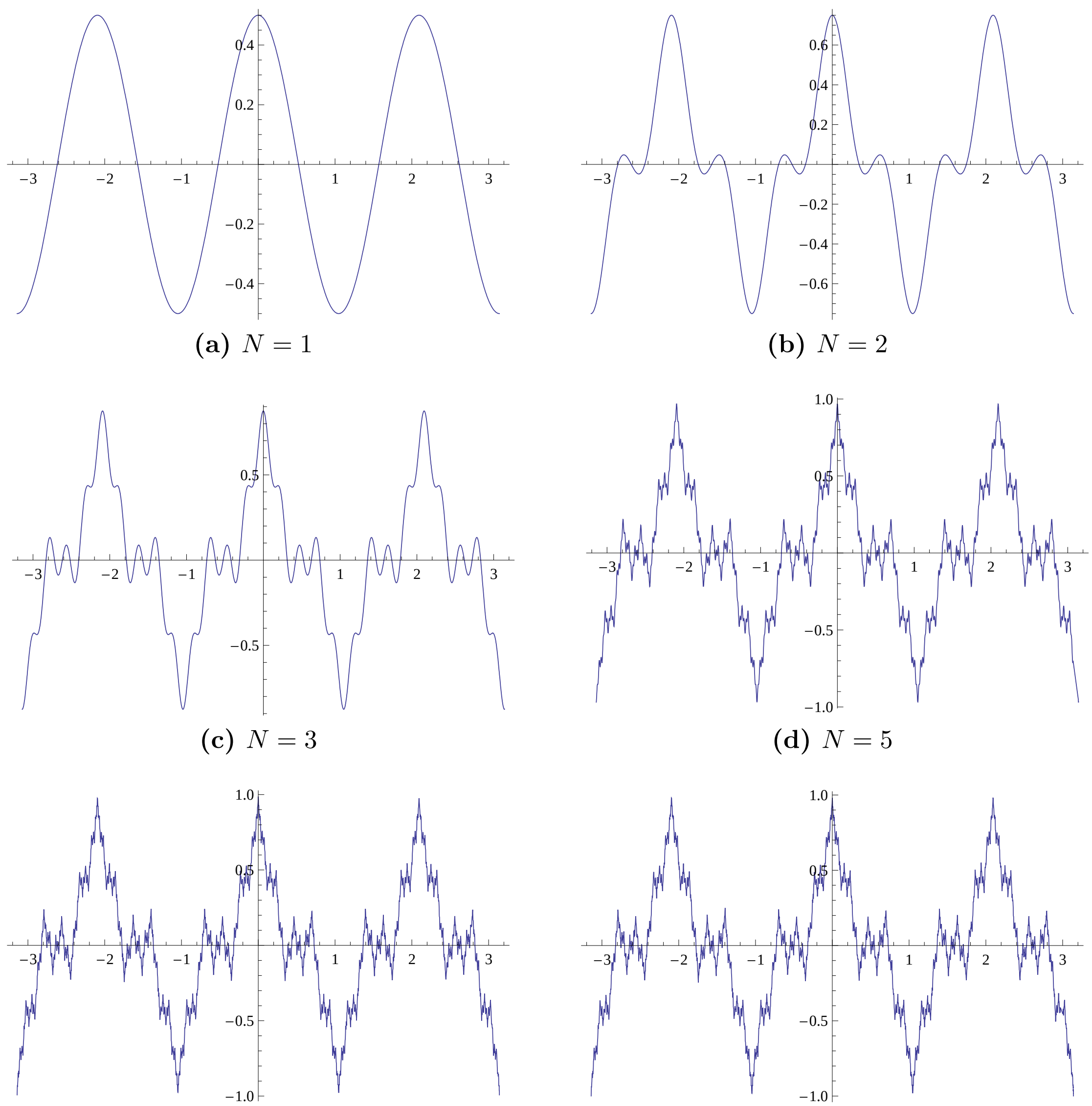
Everywhere continuous nowhere differentiable function

This appendix illustrates the example of the function at equation 2.1 on page 14 that is everywhere continuous but nowhere differentiable, discovered by Weierstrass in the 19th century. **Everywhere continuous** means (for a function $f : \mathbb{R} \to \mathbb{R}$), roughly speaking, that its graph can be drawn with a pencil from left to right without the pencil leaving the page. **Nowhere differentiable** means that the graph does not have a definite slope at any point. A moment's pause will illuminate why these characteristics of a graph seemed mutually exclusive to the mathematicians of the early 19th century.

On the following page, the graph of

$$f_N(x) = \sum_{n=1}^{N} \frac{1}{2^n} \cos(3^n x)$$

is shown for increasing N. The everywhere continuous but nowhere differentiable function is the (well-defined) limit of these functions as $N \to \infty$.



(a) $N = 1$ **(b)** $N = 2$

(c) $N = 3$ **(d)** $N = 5$

(e) $N = 7$ **(f)** $N = 10$

Figure A.1: The graphs of $f_N(x)$ for various increasing N on $[-\pi, \pi]$. For small N, the graphs are smooth. As $N \to \infty$, the graphs become gnarlier and approach the nowhere differentiable function while remaining continuous everywhere.

Appendix B

Euler's proof of Fermat's Last Theorem for $n = 3$

The proof went as follows:(3, p. 40)

> Assume w.l.o.g. that x, y, z are relatively prime. Suppose that x, y are odd and z is even. Then $x + y$ and $x - y$ are both even, say $2p$ and $2q$ respectively, with
>
> $$x = \frac{1}{2}(2p + 2q) = p + q \qquad y = \frac{1}{2}(2p - 2q) = p - q.$$
>
> When $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$ is expressed in terms of p and q it becomes:
>
> $$2p\left[(p + q)^2 + (p + q)(p - q) + (p - q)^2\right] = 2p(p^2 + 3q^2).$$
>
> Here p and q are of opposite parities—because $p + q$ and $p - q$ are odd—and they are relatively prime because any factor they had in common would divide both $x = p + q$ and $y = p - q$, and therefore could only be 1. Moreover, p and q can both be assumed to be positive. (If $x < y$ then x and y can be interchanged to give $q > 0$. The case $x = y$ is impossible because x, y, z are relatively prime and there is no cube root

of 2.) Therefore, if $x^3 + y^3 = z^3$ with x, y both odd then there exist relatively prime positive integers p, q of opposite parity such that

$$2p(p^2 + 3q^2) = \text{cube}.$$

The same conclusion can be reached if z is odd and x or y is even. In this case, the odd one, say y^3, can be moved to the left side

$$x^3 = z^3 - y^3 = (z - y)(z^2 + zy + y^2).$$

Then $z - y = 2p$, $z + y = 2q$, $z = q + p$, $y = q - p$, and

$$x^3 = 2p\left[(q + p)^2 + (q + p)(q - p) + (q - p)^2.\right]$$

which leads to the same conclusion

$$2p(p^2 + 3q^3) = \text{cube}.$$

where p, q are relatively prime positive integers of opposite parity.

The next step in the argument is roughly to say that $2p$ and $p^2 + 3q^2$ are relatively prime and to conclude that the only way that their product can be a cube is for each of them separately to be a cube. ... However, the statement that $2p$ and $p^2 + 3q^2$ are relatively prime is not quite justified. Since p and q have opposite parity, $p^2 + 3q^2$ is odd and any common factor of $2p$ and $p^2 + 3q^2$ would be a common factor of p and $p^2 + 3q^2$ and thus a common factor of p and $3q^2$. Since p and q are relatively prime, this implies that the only possible common factor is 3. But if $3 \mid p$ then clearly it also divides $p^2 + 3q^2$ and $2p, p^2 + 3q^2$ are not relatively prime. The proof therefore splits into two cases, the one in which 3 does not divide p and consequently $2p, p^2 + 3q^2$ are relatively prime, and the other in which 3 does not divide p. The first of these two cases will be considered first and the second will be treated as a simple modification of the first.

Assume therefore that 3 does not divide p and that $2p$ and $p^2 + 3q^2$ are consequently both cubes. Using the formula,

$$(a^2 + 3b^2)(c^2 + 3d^2) = (ac - 3bd)^2 + 3(ad + bc)^2$$

... one can find cubes of the form $p^2 + 3q^2$ by writing

$$\begin{aligned}(a^2 + 3b^2)^3 &= (a^2 + 3b^2)\left[(a^2 - 3b^2)^2 + 3(2ab)^2\right] \\ &= \left[a(a^2 - 3b^2) - 3b(2ab)\right]^2 + 3\left[a(2ab) + b(a^2 - 3b^2)\right]^2 \\ &= (a^3 - 9ab^2)^2 + 3(3a^2b - 3b^3)^2.\end{aligned}$$

That is, one way to find cubes of the form $p^2 + 3q^2$ is to choose a, b at random and to set

$$p = a^3 - 9ab^2 \qquad q = 3a^2b - 3b^3$$

so that $p^2 + 3q^2 = (a^2 + 3b^2)^3$. The major gap to be filled in Euler's proof is the proof that this is the *only* way that $p^2 + 3q^2$ can be a cube; that is, if $p^2 + 3q^2$ then there must be a, b such that p and q are given by the above equations. Euler bases this conclusion on the fallacious [and bold] argument that $\mathbb{Z}[\sqrt{-3}]$ is essentially the same as $\mathbb{Z}$. However, he could have used results in his other works to prove the results. Assuming this conclusion, nevertheless, the proof continues:

The expressions for p and q can be factored

$$p = a(a - 3b)(a + 3b) \qquad q = 3b(a - b)(a + b)$$

Of course, a and b are relatively prime because any factor they had in common would also divide both p and q contrary to assumption. Moreover,

$$2p = 2a(a - 3b)(a + 3b) = \text{cube}.$$

The parities of a and b must be opposite because otherwise p and q would both be even. Therefore $a - 3b$, $a + 3b$ are both odd and the only possible common factor of $2a$, $a \pm 3b$ are both odd and therefore of $a, \pm 3b$. Similarly, any common factor of $a + 3b$ and $a - 3b$ would be a factor of a and of $3b$. In short, the only possible common factor is 3. But 3 does not divide a because if it did it would be divide p, contrary to assumption. Therefore, $2a, a - 3b, a + 3b$ are relatively prime and all three of them must be cubes, say $2a = \alpha^3, a - 3b = \beta^3, a + 3b = \gamma^3$. Then $\beta^3 + \gamma^3 = 2a = \alpha^3$ and thus gives a solution of $x^3 + y^3 = z^3$ in smaller numbers than the original solution.

More specifically, $\alpha^3\beta^3\gamma^3 = 2a(a - 3b)(a + 3b) = 2p$, which is positive and a divisor of z^3 is z is even and a divisor of x^3 if x is even. In any case, then, $\alpha^3\beta^3\gamma^3$ is less than z^3. There is nothing to prevent α, β, or γ from being negative, but since $(-\alpha^3) = -\alpha^3$, negative cubes can be moved to the opposite side of he equation to become positive cubes and the resulting equation is of the form $X^3 + Y^3 = Z^3$ in which X, Y, Z are all positive and $Z^3 < z^3$. Therefore, the descent has been accomplished in the case where $3 \nmid p$.

Consider finally the case $3 \mid p$. Then $p = 3s$, say, and 3 does not divider q. Then $2p(p^2 + 3q^2) = 3^2 \cdot 2s(3s^2 + q^2)$. The numbers $3^2 \cdot 2s$ and $3s^2 + q^2$ are easily seen to be relatively prime, and therefore both of them are cubes. By the lemma to be proved [later in the text], $3s^2 + q^2$ can be a cube only if

$$q = a(a - 3b)(a + 3b) \qquad s = 3b(a - b)(a + b)$$

for some integers a, b. Since $3^2 \cdot 2s$ is a cube, $3^3 \cdot 2b(a - b)(a + b)$ is a cube and therefore $2b(a - b)(a + b)$ is a cube. The factors are easily seen to be relatively prime, $2b = \alpha^3$, $a - b = \beta^3$, $a + b = \gamma^3$, $\alpha^3 = 2b = \gamma^3 - \beta^3$ and an equation of the form $X^3 + Y^3 = Z^3$ with $Z^3 < z^3$ can be derived, all exactly as before.

In any case, then the existence of a cube which was the sum of two cubes would imply the existence of a smaller cube of the same type and is therefore impossible. All that remains to be done in order to complete this proof is to show that if p and q are relatively prime integers such that $p^2 + 3q^2$ is a cube then there must be integers a and b such that $Rp = a^3 - 9ab^2$ and $q = 3a^2b - 3b^3$. [This is completed later in the text and is omitted here.]