A Brief History of Number Theory

Gudbrand Tandberg
University of Oslo

May 12, 2015

Contents

Foreword

This paper is my final project for the course *MAT 2000: Project in mathematics*, for the spring semester of 2015 at the Univerity of Oslo. I would like to extend my gratitude to professor Arne B. Sletsjøe for help and guidance along the way. I would also like to thank Walter D. Morris and Nora Tandberg for comments that greatly improved the manuscript.

The main inspiration and source for choice of topics has been the the excellent book *"An Adventurer's Guide to Number Theory"* by Richard Friedberg [12]. A lot of background material has been collected from the articles and lectures [20], [2], [17], [1], [16], [15] and I would like to thank the authors of these for their great work. Finally I would like to thank the curators and translators of the Euler Archive for making accessible the collected works of this great man.

Gudbrand Tandberg, May 2015

1 Introduction

The history of mathematics is a rich and intricate tapestry of interwoven threads reaching back to the beginning of mankind. The purpose of this text is to trace some of these threads in order to attain a deeper understanding of the development of ideas involved. The threads we will follow are those pertaining to the art and science of *Number Theory*, the branch of mathematics devoted to the study of the natural numbers: 1, 2, 3, ...
We will necessarily have to limit our scope to only a handful of mathematicians and results, but we hope that the selection presented is adequate to represent the full development of the subject. Furthermore we have not selected just anyone - the mathematicians we will meet are none other than Euclid, Diophantus, Fermat, Euler and Gauss - these are truly groundbreaking mathematicians.

The format of this text will be that of an informal narrative interlaced with sections of mathematical source material. We will present the mathematics itself as close as possible to the raw sources, but some freedom will be taken to ease the transitions or simplify notation when needed. Hopefully, by this method, the reader will gain a deeper understanding of not only the mathematical truths themselves, but of the gradual process by which they were acquired. On a parallel track we see how advances in notation play an integral part in the development of new and necessary concepts.

2 Euclid

The Great Geometer

We begin our story 2300 years ago, in Alexandria. At the time, Alexandria was the intellectual and cultural center of the Greek civilisation, in a period called the Hellenistic period - the final bloom of ancient Greek civilisation. Euclid was the greatest geometer in Alexandria, and he worked at the great Library of Alexandria, the world's largest collection of written knowledge at the time. Very little is known about Euclid's life, but luckily, his writing lives on. His most famous work *"The Elements"* [19] has served as the standard textbook in elementary mathematics from his death all the way into the 19th century (although with some gaps). The Elements is a mighty thirteen volume work, written in a style not at all foreign to modern students of mathematics. A book begins with definitions, axioms and common notions, and follows through with propositions, demonstrations and scholion (short explanatory discussions). Starting from very little Euclid was able to deduce an incredible amount of truths about geometry and numbers in the course of the volumes.

As all who have seen further, Euclid too stood on the shoulders of giants. Euclid should not only be seen as the great geometer, but also the great compiler. The collected works of the Greek classical era were accessible to him at the Library of Alexandria. Thales, Pythagoras, Democritus, Plato and many many others had contributed to the adolescent art of mathematics before him. Nonetheless, Euclid's Elements is a very fitting point to enter the story we are about to embark on.

The Alchemy of Numbers

Let's have a look at some of Euclid's mathematics. We begin with an excerpt from book VII in the Elements, where Euclid shifts focus from plane geometry to number theory. The definitions laid forth by Euclid in this section will in fact be all the definitions we need for the remainder of this text, which is a remarkable manifest to his deep understanding of numbers.

> BOOK VII.
>
> DEFINITIONS.
>
> 1. An **unit** is that by virtue of which each of the things that exist is called one.
>
> 2. A **number** is a multitude composed of units.
>
> 3. A number is **part** of a number, the less of the greater, when it measures the greater;
>
> 4. but **parts** when it does not measure it.
>
> 5. The greater number is a **multiple** of the less when it is measured by the less.
>
> 11. A **prime number** is that which is measured by an unit alone
>
> 12. Numbers **prime to one another** are those which are measured by an unit alone as common measure.
>
> 13. A **composite** number is that which is measured by some number.
>
> 22. A **perfect number** is that which is equal to its own parts.

Only some minor syntax corrections are needed for this to be in modern form; numbers *prime to one another* are what we now call relatively prime, and a *part* of a number we would call a divisor. It is evident that Euclid imagines numbers quite differently to most modern day mathematicians. It seems that when he thinks about numbers, he thinks about length. This is fitting to a geometer. Using words like *measure* and *parts* support this view; numbers to Euclid could be represented by line segments (entities having *length* but not *breadth*). But the bridge from geometry to number theory, i.e. from arbitrarily long line segments into to segments of only certain lengths is apparent only when you impose that a *number* has to be composed of *units*, thereby restricting the permissible lengths to integers. So the idea of a number line must not have been too foreign to Euclid. Whereas in geometry, lengths could take on any values, i.e. fractional or irrational, number theory is the study of whole numbers and the structure they possess. And, as Euclid knew, to understand this structure, one must study the 'alchemy of numbers' - how some numbers can be divided into smaller numbers, and some can not, and the different operations we have for combining numbers.

The Euclidean Algorithm

We now move on to some of the propositions in Book VII. We select a few that prepare for some of the themes of this text.

Proposition 1.

Two unequal numbers being set out, and the less being continually subtracted in turn from the greater, if the number which is left never measures the one before it until an unit is left, the original numbers will be prime to one another.

For, the less of two unequal numbers AB, CD being continually subtracted from the greater, let the number which is left never measure the one before it until an unit is left;

I say that AB, CD are prime to one another, that is, that an unit alone measures AB, CD.

For, if AB, CD are not prime to one another, some number will measure them.

Let a number measure them, and let it be E; let CD, measuring BF leave FA less than itself, let AF measuring DG leave GC less than itself, and let GC, measuring FH, leave an unit HA.

Since, then, E measures CD, and CD measures BF, therefore E also measures BF.

But it also measures the whole BA; therefore it will also measure the remainder AF.

But AF measures DG; therefore E also measures DG.

But it also measures the whole DC, therefore it will also measure the remainder CG.

But CG measures FH; therefore E also measures FH.

But it also measures the whole FA; therefore it will also measure the remainder, the unit AH, though it is a number: which is impossible.

Therefore no number will measure the numbers AB, CD; therefore AB, CD are prime relative to one another.

Q.E.D.

It takes some getting used to Euclid's writing style, but the effort is worthwhile. The notion of 'primeness relative to one another', which this proposition deals with, is in many ways a part of the essence of many of the arguments we will see later. Note that in his proof, Euclid terminates the process after two iterations, ending with the unit HA. This diminishes the generality of his proof slightly, but with the notation Euclid had to work with it is also completely understandable.

As an illustration of Euclid's test for relatively primeness we provide some examples. Take for instance the numbers 85 and 54. Subtracting 54 from 85 we get 31, which does not measure 54. Therefore, we continue subtracting; 54 minus 31 is 23, which again does not measure the number before it; 31. Subtracting 23 from 31 we get 8 which does not measure 23, so we subtract 8 from 23 twice to get 7, which does not divide 8, so the final remainder is 1, and we may conclude that 85 and 54 are prime to one another. If, on the other hand, we had taken 69 and 15, we would after one step have a remainder of 9, which does not measure 15, so we continue. Subtracting 9 from 15, we get a remainder of 6, which does not measure 9. Subtracting 6 from 9 gives us 3, which does in fact measure 6. So 69 and 15 are not relatively prime. Note also that at the end of applying this algorithm we arrived at the number 3, which is the *greatest common measure* of 69 and 15. Euclid follows up on this idea in the next proposition.

Proposition 2.

Given two numbers not prime to one another, to find their greatest common measure.

Let AB and CD be the two given numbers not relatively prime. It is required to find the greatest common measure of AB and CD.

If now CD measures AB, since it also measures itself, then CD is a common measure of CD and AB. And it is clear that it is also the greatest, for no greater number than CD measures CD.

But, if CD does not measure AB, then, when the less of the numbers AB and CD being continually subtracted from the greater, some number is left which measures the one before it.

For a unit is not left, otherwise AB and CD would be relatively prime, which is contrary to the hypothesis.

Therefore some number is left which measures the one before it.

Now let CD, measuring BE, leave EA less than itself, let EA, measuring DF, leave FC less than itself, and let CF measure AE.

Since then, CF measures AE, and AE measures DF, therefore CF also measures DF.

But it measures itself, therefore it also measures the whole CD.

But CD measures BE, therefore CF also measures BE. And it also measures EA, therefore it measures the whole BA.

But it also measures CD, therefore CF measures AB and CD. Therefore CF is a common measure of AB and CD.

I say next that it is also the greatest.

If CF is not the greatest common measure of AB and CD, then some number G, which is greater than CF, measures the numbers AB and CD.

Now, since G measures CD, and CD measures BE, therefore G also measures BE.

But it also measures the whole BA, therefore it measures the remainder AE.

But AE measures DF, therefore G also measures DF. And it measures the whole DC, therefore it also measures the remainder CF, that is, the greater measures the less, which is impossible.

Therefore no number which is greater than CF measures the numbers AB and CD.

Therefore CF is the greatest common measure of AB and CD.

The procedure used to find the greatest common measure of two numbers is today called the Euclidean Algorithm, and is the first and one of the most important non-trivial numerical algorithms still in use. Had Euclid considered 1 a number (1 is not a multitude of units!), propositions 1 and 2 could really be combined into one, because the greatest common measure of numbers that are relatively prime is 1. The basic step of the algorithm is that of continually subtracting a smaller number from a greater, until a remainder which is still smaller than the smallest of the first two numbers is left. In symbols, if we start off with numbers A and B, B being the smaller, then Euclid asserts that there exist numbers Q and R, such that $A - BQ = R$, where R is less than B. Alternatively, $A = BQ + R$. This representation of any number A, with respect to some given number B, will be paramount for understanding the alchemy of numbers.

Factoring Multitudes

In modern terms, the next two propositions concern how numbers can be factored, and how prime numbers fit into the factorisations. A noteworthy trait of Euclid's writing is the *mathematical rigour* with which he demonstrates his propositions. Each proposition builds only on earlier propositions and known axioms. Definitions are used accurately to move the ideas forwards. Nowhere before Euclid do we find this rigour in mathematical writing, and not until centuries later would we see it again!

Perhaps one of Euclid's most important assertions is that of the fundamental importance of prime numbers. In Book IX, prop. 20, Euclid proves that there is an infinity of primes. In the following propositions he proves some simpler facts about primes.

Proposition 30.

If two numbers by multiplying one another make some number, and any prime number measure the product, it will also measure one of the original numbers.

Let the two numbers A and B multiplied by one another make C, and let any prime number D measure C.

I say that D measures one of the numbers A or B.

Let it not measure A.

Now D is prime, therefore A and D are relatively prime.

Let as many units be in E as the times that D measures C.

Since then D measures C according to the units in E, therefore D multiplied by E makes C.

Further, A multiplied by B also makes C, therefore the product of D and E equals the product of A and B.

Therefore D is to A as B is to E.

But D and A are relatively prime, relatively prime numbers are also least, and the least measure the numbers which have the same ratio the same number of times, the greater the greater and the less the less, that is, the antecedent the antecedent and the consequent the consequent, therefore D measures B.

Similarly we can also show that, if D does not measure B, then it measures A. Therefore D measures one of the numbers A or B.

Q.E.D.

Proposition 32.

Any number is prime or is measured by some prime number.

Let A be a number.

I say that A either is prime or is measured by some prime number.

If now A is prime, then that which was proposed is done.

But if it is composite, then some prime number measures it.

Therefore any number either is prime or is measured by some prime number.

Q.E.D.

It is quite remarkable that even 2300 years ago, these seemingly self-evident facts about numbers are found worthy of proving. This shows a deep respect for the care one must take in order to derive *true*, *consistent* mathematics.

Combining these propositions with some simple arguments yields what we now call the unique factorisation property. This is one of the central pillars of number theory, and one of the most often used facts in proofs of elementary theory of prime numbers.

Perfect Numbers

We end our exposition of Euclid with a proposition regarding perfect numbers. Just from Euclid's wording of the definition of perfect numbers - *the sum of its own parts* - we get a feeling of why they should be called just that.

Proposition 32.

If as many numbers as we please beginning from an unit be set out in double proportion, until the sum of all becomes prime, and if the sum multiplied into the last make some number, the product will be perfect.

Isn't that a neat way of thinking about summing and multiplying numbers? The proposition looks simply wonderful in modern notation as well:

$$2^{n-1}(2^n - 1) \textit{ is perfect whenever } 2^n - 1 \textit{ is prime.}$$

We shall meet this expression again in a later chapter, as we shall see, it will kindle ideas in important minds many years to come. Before that we turn our attention to another great Greek mathematician, one of the last great contributors to mathematical development in the western world before the Dark and Middle Ages.

3 Diophantus

The Great Arithmetician

Diophantus lived about 400 years later than Euclid, also in Alexandria. He is best known as the author of the thirteen volume work *"Arithmetica"*, of which only 6 volumes have survived to the present day. The Arithmetica is a book of problems and solutions in algebra and arithmetic. Nearly all of his roughly 130 surviving short exercises can be simply phrased in modern algebraic notation, as systems of determinate or indeterminate equations in one unknown. Two or more unknowns was not possible for Diophantus, but he often achieved remarkable results with only one unknown by choosing the unknown very cleverly.

The notation Diophantus used did of course not have our convenient modern form, nonetheless the notation had come a long way since Euclid. The historian G.H. Nesselman observes that we can distinguish three historical stages of development of mathematical notation; rhetoric, syncopated and symbolic algebra. Euclid belongs to the first group, and Diophantus is a typical example of the second, which is essentially rhetorical, but we now find for often-recurring operations and quantities certain abbreviational symbols. Diophantus utilised symbols similar to ours to denote fractions, multiplication and addition. He also was amongst the first to employ a symbol for the word equals. Therefore Diophantus took a fundamental step from the verbal algebra of Euclid towards symbolic algebra.

A Practical Mind

Numbers to Diophantus have a perceivably different nature than to that of Euclid. Although he gives a very similar definition; *"all numbers are made up of some multitude of units, so that it is manifest that their formation is subject to no limit"*, he is much more likely to use specific numbers. While Euclid would *never* use specific numbers for his demonstrations, Diophantus *always* used specific numbers. From this it is not unreasonable to infer that Diophantus had more practical goals in mind when it came to using numbers, though we need not dwell too long on what these uses could be. Diophantus has no objections to fractional answers to problems. In fact he often gives fractional answers to problems for which he could have found answers without fractions. He also does not mind large numbers appearing in fractions. For example, in one problem he arrives at the answer $\frac{4993}{784}, \frac{6729}{784}, \frac{22660}{784}$.

However, the freedom to use fractions actually narrowed Diophantus' interests instead of widening them. He paid no attention to problems of divisibility. Nowhere in the Arithmetica is there mention of prime numbers, odd and even numbers or perfect numbers. Negative numbers were allowed in calculations, but they were not allowed in his solutions. If Diophantus arrived at the equation $x + 5 = 3$, he would say, *"This is impossible"*, and try to change the problem to get a positive answer.

One of Diophantus' chief concerns was with polygonal numbers, i.e. square, triangular, etc. These numbers had origins stretching back to Pythagoras and the Pythagoreans, a mystical cult following from around 500BC, centered around Pythagoras' metaphysical beliefs on mathematics, harmony and astronomy. The Pythagoreans were amongst the first ancient Greeks to discover interesting connections between polygons and numbers. Of course, the classic Pythagorean Theorem is such an example, as is the fact that summing consecutive numbers always results in triangular numbers. Pythagoras also knew that summing consecutive odd numbers always resulted in a square number. Hypiscles, who wrote about 170 B.C., is twice mentioned by Diophantus as the author of a definition of a polygonal number, also stating that the n'th a-gon was given by the formula

$$\frac{1}{2}n(2 + (n-1)(a-2)). \tag{1}$$

Mathematical Species

Diophantus divided numbers into different species, these being the unit, an unknown quantity, and consecutive powers of this unknown quantity, from the second to the sixth inclusive. He uses the terms *square, cube, square-square, square-cube* and *cube-cube* to refer to these different species and introduces special symbols to denote them; Δ^Y, K^Y, $\Delta^Y\Delta$,

ΔK^Y, and $K^Y K$ respectively . Diophantus lastly defines the reciprocals of these species, with names and symbols derived thereof.

"It is from the addition, subtraction or multiplication of these numbers or from the ratios which they bear to one another or to their own sides respectively that most arithmetical problems are formed [...] and each of these numbers [...] is recognised as an element in arithmetical inquiry. But the number which has none of these characteristics, but merely has in it an indeterminate multitude of units is called 'number' and its sign is ς. And there is also another sign denoting that which is in-variable in determinate numbers, namely the unit, the sign being M with $\circ$ superposed, thus $\overset{\circ}{M}$."

Diophantus goes on to explain one of our well-known elementary school results in arithmetic, along with the associated notation:

"A minus multiplied by a minus makes a plus; a minus multiplied by a plus makes a minus; and the sign of a minus is a truncated ψ turned upside down, thus Λ."

It sounds even more euphonic translated literally, thus: "*a wanting multiplied by a wanting makes a forthcoming*". As an interesting aside, we note that the situation where *wanting* numbers are allowed in calculations but not in answers is completely analogous to the situation of Cardano a thousand years later using his *impossible* (imaginary) quantities in calculations to arrive at permissible answers to particular polynomial equations.

Diophantus proceeds:

"It is well that one who is beginning this study should have acquired practice in the addition, subtraction and multiplication of the various species. He should know how to add positive and negative species with different coefficients to other species, themselves either positive or negative or likewise partly positive and partly negative, ..."

As a showcase of this notation consider decrypting the following expression (numbers are represented by corresponding Greek letters, coefficients come *after* their corresponding species and all the negative terms in an expression are placed together after all the positive terms.)[1]:

$$K^Y \alpha \varsigma \eta \Lambda \Delta^Y \epsilon \overset{\circ}{M} \alpha.$$

[1] Answer: $x^3 + 8x - 5x^2 - 1$

Problems and Solutions

The following quote illustrates Diophantus’ main goal in arithmetic[2]:

”This should be the object aimed at in framing the hypotheses of propositions, that is to say, to reduce the equations, if possible, until one term is left equal to one term.”

THE ARITHMETICA
BOOK I
PROBLEMS

1. To divide a given number into two having a given difference.

Given number 100, given difference 40.

Lesser number required x. Therefore

$$2x + 40 = 100,$$

$$x = 30.$$

The required numbers are 70, 30.

Today we would formulate and solve the equations $x + y = 100$, $x - y = 40$ to arrive at $y = 30, x = 70$, but still with only one variable this is no problem for Diophantus. This short and simple format is used in all his problems:

2. To divide a given number into two having a given ratio.

Given number 60, given ratio 3:1.

Two numbers $x, 3x$. Therefore $x = 15$.

The numbers are 45, 15.

As the mathematician Herman Henkel said, *”not the slightest trace of a general, comprehensive method is discernible; each problem calls for some special method which refuses*

[2] Some remarks regarding the translation are in order. We follow the translation due to T.L. Heath [14], in which both the notation and wording has been adapted somewhat. We follow this route instead of looking closer at the source, as the following example illustrates the difficulties this would invoke: where we would write $(12+6n)/(n^2-3)$, Diophantus says *”a sixfold number increased by twelve, which is divided by the difference by which the square of the number exceeds three”*

to work even for the most closely related problems. For this reason it is difficult for the modern scholar to solve the 101st problem even after having studied 100 of Diophantus' solutions".

Diophantus was very interested in perfect squares:

8. To divide a given square number into two squares

Given square number 16.

x^2 one of the required squares.

Therefore $16 - x^2$ must be equal to a square.

Take a square of the form $(mx - 4)^2$, m being any integer and 4 the number which is the square root of 16, e.g. take $(2x - 4)^2$, and equate it to $16 - x^2$.

Therefore $4x^2 - 16x + 16 = 16 - x^2$, or $5x^2 = 16x$, and $x = \frac{16}{5}$.

The required squares are therefore $\frac{256}{25}, \frac{144}{25}$.

27. To find two numbers such that their sum and product are given numbers.

Necessary condition. The square of half the sum must exceed the product by a square number.

Given sum 20, given product 96.

$2x$ the difference of the required numbers.

Therefore the numbers are $10 + x, 10 - x$.

Hence $10 - x^2 = 96$.

Therefore $x = 2$, and the required numbers are 12, 8.

Had the product in problem 27 been 95 instead of 96, the unknown would have been $x = \sqrt{5}$, which of course is impossible. It is understandable therefore why Diophantus was so interested in perfect squares, and in fact almost all his problems pertain to them. A typical problem, easy for Diophantus, would be to find three squares equally spaced.

In Book III, Diophantus solves problems of finding values which make two linear expressions simultaneously into squares or cubes. In Book IV, he finds rational powers between given numbers. He also noticed that numbers of the form $4n + 3$ cannot be the sum of two squares. Diophantus also appears to know that every number can be written as the sum of four squares, but it is unlikely that he proved it. Let us just say that there are plenty of impressive solutions in the Arithmetica, and let this small taste of Diophantus' mathematics suffice as a showcase of the state of mathematics at the end of classical civilisation.

The End of Classical Civilisation

Not long after the death of Diophantus, the Roman Empire - at that time the ruler of Alexandria - crumbled. The Library of Alexandria was torn apart and burned by Christian mobs trying to consolidate the power of Christianity by attempting to eradicate pagan influence and power. This had the disastrous effect of sweeping out of existence hundreds of thousands of rolls of papyrus containing the ideas and thoughts of thousands of scholars from all corners of the ancient world. The end to the Library of Alexandria marks the end of our survey of the mathematics of classical civilisation. It would take more than a thousand years for the subjects and ideas studied at the Library of Alexandria to take root and be properly built upon on the European continent.

4 Fermat

Europe and the Renaissance

It is inaccurate to say that mathematical research stagnated completely in the period after Diophantus. Even though the Roman empire crumbled, leading with it a period of cultural and economic deterioration in the parts of the world it had conquered, most parts of Europe had not been a part of the of classical Greek and Roman civilisations. The Middle Ages were the beginning of civilisation in Europe. The people of Europe - Norsemen, Goths, Franks - had been wanderers a short time before when the Romans taught them how to live an orderly life and stay in one place. The early Middle Ages were for them a period of activity and development, a rise from barbarism rather than a fall from glory. In medieval times, most people had no time for booklearning. The Church kept it going, but in Latin only. Hardly anyone in Europe knew Greek, and the greatest books of the Hellenistic period only gradually became known in the twelfth and thirteenth centuries. In the fourteenth century Europe was ruled by the plague, which slowed everyone down. But after that Europeans were ready to take on the Greek heritage, no longer as a monument to be admired, but as a a foundation to be continued. This is called the Renaissance or rebirth of Europe, although it was more like a coming of age.

The great Eastern civilisations had kept it going through the Middle Ages, developing such notions as *al*gebra, *al*gorithms and *al*chemy, and some of these discoveries did find their way into Europe. Some first noteworthy mathematical strides in Europe were taken by the Pisan Leonardo di Fibonacci working on arithmetical problems, and the French Jewish scholar Levi ben Gershon establishing connections to combinatorics and counting. From the fourteenth century on, a group of competitive Italians led by Cardano, Tartaglia and Ferrari made important progress in the solving of algebraic equations. Their mathematics was a key ingredient in forming the new civilisation; bankers, architects, merchants, all had many reasons to value mathematical progress. The source of inspiration and knowledge to the great thinkers of the Renaissance were the classical Greek texts. Philosophy, entertainment, politics, architecture and of course mathematics was ”taught” directly from the ancient Greeks to the European scholars. Consequently, Euclid and Diophantus certainly were beginning to be widely read in Europe at this time.

The Prince of Amateurs

One of those who cherished his copy of Diophantus’ Arithmetica highly, and who is the next great mathematician we shall dwell on is Pierre Fermat, the pivotal figure in the Renaissance of mathematics in the 17th century. Pierre was in fact not a mathematician at all, but first and foremost a lawyer working in the Parlament of Toulouse. He led a sedate, pleasant life, was productive as a legal fellow, and privately had a great interest in and love for mathematics. This is the reason why he is often referred to as the ’Prince of Amateurs’. Pierre owned a copy of the Arithmetica, translated into Latin by Bachet in 1621. He studied the work frivolously, jotting down his many comments in the narrow margins of his copy[3]. Fermat did not publish formally, so it is only through these marginal comments and his letters to acquaintances that his results are known. And Fermat communicated many new and interesting propositions in his letters, in diverse fields such as optics, probability and analytic geometry.

Unturned Stones

Fermat’s main interest however was in number theory, which was still not a subject studied in its own right. A slight problem with Fermat’s mathematics was that he almost never shared his *proofs*. Perhaps it was to retain a sense of ownership or secrecy, or to tease his colleagues, or perhaps sometimes it was because he had not really proven them at all. In any case, he left a great deal of stones unturned for his successors to study. The most well known example of such a stone is what he proposes in his comment in the margin next to problem 2, Book III of the Arithmetica.

”It is impossible to partition a cube into two cubes or a fourth power into two fourth powers, or generally any power higher than the second into two powers of the same degree. I have discovered a truly wonderful proof of this, which, however, this margin is too narrow to contain.”

[3] *Fun fact:* Fermat was not the only one to scribble in Diophantus’ margin. A particularly amusing example came from the Byzantine scholar John Chortasmenos (1370-1437) who had written ”Thy soul, Diophantus, be with Satan because of the difficulty of your theorems” in one of the margins.

This is called Fermat's last theorem, and withstood attempts of proof by Euler, Dirichlet, Gauss and probably every other great mathematician since. Teo hundred years after Fermat's death, a proof had been published of every one of the propositions Fermat had claimed to have proven, except this one. It was finally proven in 1994 by Andrew Wiles using very advanced methods indeed.

Dissecting Species

Like Diophantus, Fermat also had a deep interest in squares, cubes, powers in general, and sums thereof. The key step Fermat took was to combine Euclid's interest in and approach to whole numbers, including primes and divisibility, and Diophantus' interests in squares, sums of squares, and equations in general asserting truths about certain 'species' of numbers. This combination was the crucial step in developing the general framework in which elementary number theory is best set. In keeping with this spirit, let us see what sort of results Fermat discovered in this subject.

A special type of numbers, which have been studied ever since Pythagoras, are whole numbers x, y, z such that $x^2 + y^2 = z^2$. Such so-called *Pythagorean triples* are the sides of some right triangle. Euclid knew a method of generating these triples, and Fermat was well aware of this. But Fermat, like Diophantus, was more interested in taking apart triangles than building them. An important result was to find out which numbers are in fact the hypotenuse of some primitive[4] right triangle, and in general; what sorts of factors do different sums of powers admit? As a marginal note on problem 19 in Book III of Arithmetica, Fermat writes:

A prime number of the form $4n + 1$ is the hypotenuse of a right-angled triangle in one way only, its square is so in two ways, its cube in three, its biquadrate in four ways, and so on ad infinitum.

The sister result to the previous one is this: (sent by Mersenne to Descartes, on March 22 1638, as having been proven by Fermat.)

A number of the form $4n - 1$ is neither a square nor the sum of two squares, either in integers or fractions.

Diophantus had already proven this for fractions[5], but Fermat took it even further in a letter of August 1640 to Roberval [18, p. 202-203]:

If a given number is divided by the greatest square which measures it, and the quotient is measured by a prime number of the form $4n - 1$, the given number is neither a square, nor the sum of two squares whether integral or fractional.

[4] Primitive means the lengths of the two shorter legs are relatively prime

[5] Numbers of the form $4n - 1$ are of course also in the form $4k + 3$

to which Fermat appended the comment:

I confess to you frankly that I have found nothing in the theory of numbers which has pleased me so much as the proof of this proposition, and I shall be glad if you will try to discover it, if only for the purpose of showing me whether I think more of my discovery than it deserves.

As an example, let the given number be 84. The greatest square which measures it is 4, and the quotient is 21 which is measured by 3 or by 7, both 3 and 7 being of the form $4n - 1$. I say that 84 is neither a square nor the sum of two squares either integral or fractional.

Now let the number be 77. The greatest square which measures it is 1, and the quotient is 77 which is here the same as the given number and is measured by 11 or 7, each of these numbers being of the form $4n - 1$. Therefore 77 is neither a square nor the sum of two squares, either in integers or fractions.

Investigations into Pythagorean triples lead Fermat to ponder expressions of the form $x^2 + 2y^2$, $x^2 + 3y^2$, and others like them. It was generally desired to set these forms equal to squares or cubes or such things. After discovering that something was *possible*, Fermat would go on to see in *how many ways* it was possible. Fermat was easily able to find a number which is the hypotenuse of 367 different right-angled triangles and no more, a remarkable yet seemingly useless achievement.

Another impressive unproven result of Fermat, stated in a comment to problem 29, Book IV of the Arithmetica [18, II p. 433], is the famous result that

Every number is a square or the sum of two, three or four squares.

Its generalisation, which is that: (letter to Pascal of 25 September, 1654[18, p. 313])

every number is made up of one, two or three triangles, of one two, three or four squares, of one two, three, four or five pentagons, and so on infinitum,

was first proven by Cauchy two centuries later. Fermat implies that he in fact has a proof, and adds to this that

to arrive at this it is necessary -

(1) To prove that every prime number of the form $4n + 1$ is the sum of two squares, e.g. 5, 13, 17, 29, 37, etc.;

(2) Given a prime number of the form $4n + 1$, as 53, to find by a general rule, the two squares of which it is the sum.

(3) Every prime number of the form $3n + 1$ is of the form $x^2 + 3y^2$, e.g. 7, 13, 19, 31, 37, etc.

(4) Every prime number of the form $8n+1$ or $8n+3$ is of the form x^2+2y^2, e.g. 11, 17, 19, 41, 43, etc.

(5) There is no rational right-angled triangle in whole numbers the area of which is a square.

This will lead to the discovery of many propositions which Bachet admits to have been unknown to him and which are wanting in Diophantus. I am persuaded that, when you have become acquainted with my method of proof in this kind of proposition, you will think it beautiful, and it will enable you to make many new discoveries, for it is necessary, as you know, that "multi pertransent ut augeatur scientia[6]"

It is as if he is deliberately teasing his colleagues by sharing only some of his knowledge.

Related to the problem of taking apart Pythagorean triples is what is now called *Pell's equation*; $x^2 - Dy^2 = 1$, where D is not a square. The problem of finding whole number solutions to this equation had been studied by the Pythagoreans, Archimedes, Diophantus, Brahmagupta and others, and finally by Fermat. He rediscovered the ancient problem, and was the first to show that the equation *always* has an unlimited number of solutions.

In a letter of February 1657 [18, p. 333-334], Fermat asks Frénicle for a general rule for finding, *"when any number not a square is given, squares which, when they are serpectively multiplied by the given number and unity is added to the product, give squares"*. If, says Fermat, Frénicle cannot give a general rule, will he *"give the smallest value of y which will satisfy the equations $61y^2+1=x^2$ and $109y^2+1=x^2$?"*

The Analyst

We have already seen that Fermat was equally at home with proving affirmative, or existence results and negative, or non-existence results. The Greeks however were *synthesists*, meaning they liked to make right triangles, and take square roots. They were devastated when they arrived at solutions which were impossible, like the square root of two. A Greek mathematician proving that something was impossible was like an artist claiming that a painting cannot be painted. Not so much Fermat however: he was an *analyst.* He often claimed that things were impossible, and used a very special technique to do it. The technique, which was invented by the Greeks (although they did not use it very much), and later revived by Fermat, is called *infinite descent.* It essentially consists of generating an infinite descending sequence of natural numbers having a certain quality. Since an infinite descending sequence of natural numbers is impossible, such a quality would be impossible

[6] Francis Bacon: "Through many science will expand"

to have for any number. Fermat used this method a lot for negative results, and he was even able to use it for affirmative results. He writes [18, p. 432]:

It was a long time before I was able to apply my method to affirmative questions because the way and manner of getting at them is much more difficult than that which I employ with negative theorems. So much so that, when I had to prove that every prime number of the form $4n + 1$ *is made up of two squares, I found myself in a pretty fix. But at last a certain reflection many times repeated gave me the necessary light, and affirmative questions yielded to my method, with the aid of some new principles by which sheer necessity compelled me to supplement it. This development of my argument in the case of these affirmative questions takes the following line: if a prime number of the form* $4n + 1$ *selected at random is not made up of two squares, there will exist another prime number of the same sort but less than the given number, and again a third still smaller and so on, descending ad infinitum, until you arrive at the number 5 which is the smallest of all numbers of the kind in question and which the argument would require not to be made up of two squares, although, in fact, it is so made up. From which we are obliged to infer, by reductio ad absurdum, that all numbers of the kind in question are in consequence made up of two squares.*

Anyone who has thought long and hard about a difficult problem will recognise what Fermat means by *"a certain reflection many times repeated"*. Here we once again see how Fermat shares some, but not all of his knowledge.

A Challenge to Mathematicians

It often seems that Fermat is well aware of his place in number theory - that he knows that he can see further than anybody else in his time. This is illustrated by the following passage, taken from a letter exhibiting a challenge to English mathematicians [18, II p. 334-335]:

There is hardly any one who propounds purely arithmetical questions, hardly anyone who understands them. Is this due to the fact that up to now arithmetic has been treated geometrically rather than arithmetically? This has indeed generally been the case both in ancient and modern works; even Diophantus is an instance. For, although he has freed himself from geometry a little more than others have in that he confines his analysis to the consideration of rational numbers, yet even there geometry is not entirely absent, as is sufficiently proved by the Zetetica of Vieta, where the method of Diophantus is extended to continuous magnitude and therefore to geometry. Now arithmetic has, so to speak, a special domain of its own, the theory of integral numbers. This was only lightly touched upon by Euclid in his Elements, and was not sufficiently studied by those who followed him (unless, perchance, it is contained in those Books of Diophantus of which the ravages of time have robbed us); arithmeticians have therefore now to develop it or restore it. To arithmeticians *therefore, by way of lighting up the road to be followed, I propose the following theorem to be proved or problem to be solved*[7]*. If they succeed in discovering the proof or solution, they will admit that questions of this kind are not inferior to the more celebrated questions in geometry in respect of beauty, difficulty or method of proof.*

Wishful Mathematical Thinking

Most of the time, Fermat was either right, or proven right at a later time. But not always. As a demonstration that even Fermat was mistaken at times, and as a reminder of the caveats of wishful mathematical thinking, some remarks regarding the so-called Fermat primes are in order. Fermat had the idea that every number of the form $2^{2^n}+1$ is prime[8], having computed and shown the primality of the numbers resulting from $n = 0, 1, 2, 3, 4$. As he writes in a October 1640 letter to Frénicle [18];

But I clearly admit to you (because as above I will make you aware that I am not capable of attributing to myself more than I know, I will speak with the same frankness about that which I do not know) that I still cannot prove the exclusion of all divisors in this lovely proposition which I have sent to you and which you have confirmed for me, touching the numbers 3, 5, 17, 257, 65537, etc. For, although I have reduced the exclusion to most of the numbers and also have a probable reason for the rest, I still cannot prove with necessity the truth of this proposition, although I still do not doubt to this hour that I have done it already. If you have assured the proof, you will oblige me in communicating it to me; because, after this, nothing would hinder me in these matters.

This was later proven by Euler to fail in the case $n = 5$. Imagine spending the day trying to figure out by direct calculation whether $2^{2^5}+1$ is prime or not! (It turns out that $2^{32}+1$ is divisible by 641).

Three Beautiful Propositions

As alluded to in a previous chapter, Fermat was also interested in Euclid's formula for perfect numbers. In a letter to Marsenne[9] [18, II, p.197-99], Fermat enunciated three propositions which were meant to facilitate the investigation of whether a given number of the form $2^n - 1$ is prime or not (this is, recall, the criterion for Euclid's' formula to give perfect numbers).

[7] He refers here to the problem of solving the general Pell's equation.

[8] A formula that only generates primes has yet to be found

[9] Prime numbers of the form $2^n - 1$ are usually called Marsenne primes, probably because Marsenne studied them. These numbers are very important in modern day encryption techniques based on large prime numbers

If we write on one line the exponents 1, 2, 3, 4, etc. of the successive powers of two underneath them respectively the numbers representing the corresponding powers of 2 diminished by 1 thus,

$$\begin{array}{cccccccccccc}
1, & 2 & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10, & 11, & \ldots n \\
1, & 3 & 7, & 15, & 31, & 63 & 127, & 255, & 511, & 1023, & 2047, & \ldots 2^n - 1
\end{array}$$

the following relations are found to subsist between the numbers in the first line and those directly below them in the second line.

(1) If the exponent is not a prime number, the corresponding number is not a prime number.

(2) If the exponent is a prime number, the corresponding number diminished by 1 is divisible by twice the exponent.

(3) If the exponent n is a prime number, the corresponding number is only divisible by numbers of the form $2nm + 1$. If therefore the corresponding number in the second line has no factors of this form, it has no integral factor.

As usual Fermat does not give his proofs, he merely adds: *”Here are three very beautiful propositions that I have found and proven without difficulty”*. The second proposition states exactly that ”If n is prime, then $2n$ divides $2^n - 2$”, but is equivalent to the even neater statement that ”If n is prime, then n divides $2^{n-1} - 1$”, which we will meet again shortly. As an example of the third proposition, consider the prime exponent 11. Now, $2^{11} - 1 = 2047$ is not prime, and according to the proposition, any factor must be of the form $22m + 1$. Sure enough, its factors are $23 = 1 \times 22 + 1$ and $89 = 4 \times 22 + 1$.

We have seen throughout this chapter that expressing numbers in the form $bq + r$, where b and r are fixed, $r < b$, are often used in propositions of Fermat. The idea of expressing numbers in this form goes back to Euclid and his algorithm, and is central to most of the work we shall see later. Then also we shall discover exactly *why* this technique is so fruitful.

Fermat's Little Theorem

The final result we will look at, and also the most relevant result for the continuation of this story is what is now called Fermat's Little Theorem. This result came to Fermat while studying Euclid's equation of perfect numbers, and is in fact the natural generalisation of proposition *(2)* above. In the letter to Frénicle of October 1640 he writes:

It seems to me after this that it is important for me to tell you the basis on which I apply the proofs of everything that concerns geometric progressions, such as:

Every prime evenly divides one of the powers minus one of any progression in which the exponent of the given power is a factor of the given prime number minus one; and after one has found the first power which satisfies this property, all those numbers having exponents that are multiples of the exponent of the first satisfy all of the same properties. [...] And this proposition is generally true for all progressions and for all prime numbers; the proof of which I would send you, if I were not afraid that it would be too long.

Nowadays, this proposition is usually phrased ”If p is prime, then p divides $a^{p-1} - 1$, whenever a and p are relatively prime”. If we read carefully however, we see that Fermat is actually saying *more*. He says that there is some factor of $p - 1$ to which one can raise a to this power and find the result diminished by one divisible by p. Now this factor may be $p - 1$ itself, in which case we have the modern formulation, but it may also be smaller, in which case we have found a smaller power to which a must be raised to have the result diminished by one divisible by p. Fermat does not mention the necessity of relatively primeness; perhaps he considered it too obvious. Fermat provides an example, with $a = 3$, he lists the exponents and powers in two rows:

$$\begin{array}{cccccc} 1, & 2 & 3, & 4, & 5, & 6 \\ 3, & 9 & 27, & 81, & 243, & 729 \end{array}$$

He takes the prime $p = 13$ and notes that p divides the third power of 3 minus one, i.e. 13 divides 26. Thus 3 is the least such exponent, but by the second part of his proposition, all exponents that are multiples of 3 will also have the same property, e.g. 13 divides $3^6 - 1 = 728$. We must also have that 13 divides $3^9 - 1$, and $3^{12} - 1$, the latter being what we now call Fermat’s Little Theorem.

This theorem proved to be very useful to Fermat and his successors in proving all sorts of facts about prime numbers and factoring hypotenuses. It would also serve as a starting point for an entirely new way to think about numbers. We have to wait until Gauss comes along before we can fully appreciate this new way of thinking, but in the meantime we will let Euler bulk up the theories of Fermat.

5 Euler

The Golden Age

And so finally we enter into the Golden Age of mathematics. It took the great genius of Fermat standing on the shoulders of Diophantus to open the door into modern number theory, and now we shall meet the great geniuses of the 18th and 19th centuries who first stepped through the door.

The 18th century in Europe was a period of intellectual, social, philosophical and political ferment. This time is often referred to as the Age of Enlightenment, for it was now that the ideas of the previous 100 years were first implemented on a broad scale. In academia, the new fields of calculus and mechanics began to influence thinking about the workings of the universe. Politically and philosophically, the ideas of John Locke, Immanuel Kant, Hobbes, and others would give rise to a notion of democracy that would ultimately overthrow the monarchial power structure on the European continent.

For the first time in Europe, science became widely available. Until the 18th century, scientific inquiry was pursued by a relatively small group of academics whose writings did not enjoy widespread circulation. Beginning in the late 17th century, two major novelties started a development in academia that would bring about a rapid democratisation of scientific knowledge. The first was the foundations of the Paris Academy and the Royal Society of London, two institutions whose primary purpose was to perform scientific research and report their conclusions to the public. Over the succeeding decades, several other institutions were founded on the model of these two, including the Berlin Academy and the St. Petersburg Academy. The second major development in academic life was the rise of scientific journals (e.g. *Philosophical Transactions*, *Mémoires of the Paris Academy*, and *Crelle's journal*). These new journals were distributed to a wide audience that included many outside the scientific community.

An Unrivalled Mathematical Genius

Undeniably, the most important scientist of this century was the Swiss Leonard Euler. In the broader field of science, Euler made groundbreaking contributions in fields as diverse as astronomy, ballistics, cartography, elasticity, fluid mechanics, hydraulics, magnetism, music theory, optics, mechanics and ship theory, as well as state projects on fire engines, machines, engineering and science education. Crowning all these interests was mathematics, the "Queen of Sciences", that stood closest to Euler's heart. For Euler, a humble, deeply religious man, there was a close connection between mathematics and God. Within mathematics, Euler studied arithmetic, number theory and analysis, or calculus, differential equations, integration, infinite series, complex analysis, classical geometry, differential geometry and graph theory. The 20th century science historian Clifford Truesdell has calculated that of all the mathematical and scientific work published during the whole of the 18th century, a full 25% was written by Euler.

A Distinguished Career

Euler was born in Basel in 1707. At first Euler was set to follow his father's footsteps as a pastor in the local Protestant Church, but chance had it that Euler's talent for mathematics was spotted early by the great Johann Bernoulli I[10]. At the age of fifteen, Euler received his master's degree in philosophy from the University of Basel. In his thesis he compared and contrasted the philosophical ideas of Descartes and Newton. In April 1727 with no employment prospects in Basel, he left to join the recently started Saint Petersburg Academy of Sciences. Johann Bernoulli's sons, Daniel I and Nikolaus II had been hired by the academy two years earlier as part of Peter the Great's westernisation of Russia, and they had recommended Euler for the first open position. He stayed in St. Petersburg from 1727 to 1741, working hard the whole time. Unfortunately, Peter the Great, who died two years before Euler arrived, had left no successor to his throne, and gradually the political climate in Russia took a turn for the worse, and in many places certain xenophobic Russians erupted in violence against foreigners. At the same time, Frederick the Great of Prussia offered Euler an increased salary as well as a safe haven at the academy of Berlin. During his Berlin years Euler was at the peak of his career, contributing around 380 articles and several landmark books on diverse scientific subjects. In a letter from this

[10] The Bernoulli family had dominated the mathematical scene in the second half of the seventeenth century, and several more talented Bernoullis would follow.

period to a friend he says that *"the king calls me his professor, and I think that I am the happiest man in the world"* [1]. In 1766 he left Berlin to take the post as director of the Saint Petersburg Academy, at that time under the rule of Catherine the Great, and a much safer place than when he last left. He stayed in St. Petersburg and remained active until his death in 1783.

Euler maintained a keen interest in number theory throughout his career. Many of his achievements in number theory came about as he was studying the writings of Fermat, written a century earlier. Very few discoveries had been made in the field of number theory since that time, so there were many open questions to tackle. For the rest of this chapter we shall present some of Euler's relevant work, and we will reflect upon the ideas, results and techniques developed along the way.

Truth Through Proof

During his first stay in St. Petersburg, Euler made many important discoveries in number theory which would aid him in his investigations for the rest of his life. One of his first major achievements in the field came about in a 1736 paper on prime numbers[4]. We quote generously from it here, so as to give the reader a feel for Euler's writing style.

> ## THE DEMONSTRATION OF CERTAIN THEOREMS REGARDING PRIME NUMBERS
>
> 1. Many arithmetical theorems, though without proofs, were once brought to light by Fermat which (if they were true) not only would contain exceptional properties of numbers, but also would greatly promote the science of numbers itself, which seems for the most part to exceed the limits of analysis. However, although the famous geometer claimed, concerning many theorems that he proposed, that he either could prove them or that he was at least certain of their truth: nevertheless he never produced proofs for them at any time, insofar as I am aware. But on the other hand, Fermat seems to have grasped a large part of his numerical theorems through induction, which indeed seems to be an almost unique method for bringing to light properties of this kind. However, I could also speak of how little induction on many examples can yield in this matter; which was nevertheless sufficient for Fermat himself for eliciting unique observations.
>
> I am speaking no less about that theorem, whose falsity I already pointed out several years ago, in which Fermat asserted that all numbers expressible in the form $2^{2^n} + 1$ are prime numbers. [...]
>
> 2. For this reason, all such properties of numbers, which rested upon induction alone,

I now judge to have uncertainty, until they are either supplied with clearly valid proofs or altogether refuted. [...] But now, as I have attained the firmest proofs of these theorems by my own method, there can be no more doubt concerning their truth. And on this account, in order to establish the truth of those theorems, which is a method in itself and which may even bear usefulness in other investigations of numbers, I have resolved to set forth my proofs in this paper.

3. Now, the proposition, which I am prepared to prove, is the following:

With a signified prime number p, the formula $a^{p-1} - 1$ can always be divided by p, unless a can be divided by p.

DEMONSTRATION

We begin by demonstrating that with any odd prime number p specified, whichever formula $2^{p-1} - 1$ always will be able to be divided by p.

For in place of 2, there may be put 1+1 and there will be

$$2^{p-1} = (1+1)^{p-1} = 1 + \frac{p-1}{1} + \frac{(p-1)(p-2)}{1 \cdot 2} + \frac{(p-1)(p-2)(p-3)}{1 \cdot 2 \cdot 3} + \text{etc.},$$

the number of terms of which series is $= p$ and hence is odd. With the first term removed, the series becomes

$$\begin{aligned} 2^{p-1} - 1 = \frac{p-1}{1} + \frac{(p-1)(p-2)}{1 \cdot 2} + \frac{(p-1)(p-2)(p-3)}{1 \cdot 2 \cdot 3} \\ + \frac{(p-1)(p-2)(p-3)(p-4)}{1 \cdot 2 \cdot 3 \cdot 4} + \text{etc.}, \end{aligned}$$

the number of which is $= p-1$ and therefore even. Therefore each two terms may be gathered into one sum, from which the number of terms is made twice as small; there will be

$$2^{p-1} - 1 = \frac{p(p-1)}{1 \cdot 2} + \frac{p(p-1)(p-2)(p-3)}{1 \cdot 2 \cdot 3} + \text{etc.},$$

the final term of which on account of the odd number p will be

$$\frac{p(p-1)(p-2)(p-3)...1}{1 \cdot 2 \cdot 3...(p-1)} = p.$$

But it appears the individual terms to be divisible by p; for since p shall be a prime number, and greater than any factor of the denominator, nowhere by division will it be able to be removed. On account of which if p were an odd prime, 2^{p-1} always will be able to be divided by that.

Next, we demonstrate that with p denoting a prime number, if $a^p - a$ can be divided by p, then the formula $(a+1)^p - a - 1$ will be able to be divided by the same p also.

For $(a+1)^p$ can be resolved into a series in the usual manner; there will be

$$(a+1)^p = 1 + \frac{p-1}{1}a + \frac{(p-1)(p-2)}{1 \cdot 2}a^2 + \frac{(p-1)(p-2)(p-3)}{1 \cdot 2 \cdot 3}a^3 + ... + \frac{p}{1}a^{p-1} + a^p,$$

of which the individual terms of the series are able to be divided by p except the first and the last, if indeed p were a prime number. On account of which $(a+1)^p - a^p - 1$ allows division by p; but this formula agrees with this $(a+1)^p - a - 1 - a^p + a$. But by hypothesis, $a^p - a$ can be divided by p, therefore also $(a+1)^p - a - 1$.

Therefore since, because on putting $a^p - a$ able to be divided by the prime number p, this formula too $(a+1)^p - a - 1$ allows division by p, it follows also $(a+2)^p - a - 2$, likewise $(a+3)^p - a - 3$ and generally $(a+b)^p - a - b$ to be able to be divided by p. But on putting $a = 2$, because $2^p - 2$, as we have now shown, can be divided by p, it is evident the formula $(b+2)^p - b - 2$ must be allowed to be divided by p, whatever whole number may be substituted in place of b. Therefore p passes through the formula $a^p - a$ and consequently also this $a^{p-1} - 1$, unless $a = p$ or by a multiple of p. And this has demonstrated the general theorem, that I undertook to discuss.

This is Euler's first proof of Fermat's Little Theorem. At once we become aware of some striking difference between the two men; in Euler's view, no proposition should be asserted true through what he calls *induction*, by which he means Fermat's habit of asserting the truth of conjectures based on 'inducting' from examples. In this insistence on absolute truth through proof, Euler is much like Euclid. Nonetheless, Euler recognises the usefulness of the method and utilises it often in his own work. He draws a much clearer line between proven truths and conjectures. In Euler's hands, mathematics, and in particular number theory, took a step towards the rigorous, precise art it is today. We see also that the writing style has come a long way since Fermat, and we have almost completely entered into the age of symbolic algebra. Still, no modern mathematician would honour this article of Euler as *elegant* or even very clear, but we must keep in mind that the method it uses is quite new. The Binomial Theorem, which this proof rests on, was incidentally discovered seven centuries earlier by the Chinese mathematician Chia Hsien, and then later rediscovered in Arabia and again by Newton in England. Finally we add that the method of proof Euler uses here is a primitive form of what we today call mathematical induction, albeit in an overly verbose and quite clumsy form.

Useful Truths

Regarding Euler's view on number theory, the following passage from a 1747 article [6] is enlightening:

Some time ago, the greatest geometers recognised that many distinguished properties lay hidden in the nature of numbers, the knowledge of which would increase the boundaries of mathematics in no moderate way. At first glance the theory of numbers seems only to deal with the elements of arithmetic, and hardly anything that requires any expertise or a method of analysis. But as it turns out, this theory has produced not only proofs of very difficult truths, but also statements whose certainty seems clear, even though proofs for them have yet to be found.

The famous geometer Fermat produced many theorems of the latter kind, whose truth, even though a proof was lacking, seemed no less certain. And this especially merits all attention, insofar as truths of this kind in pure mathematics exist, which seem true to us, even though we are unable to prove them. And this happens so much in arithmetic, which is considered to be the most basic and best known theory among all areas of mathematics, that it would be hard for me to deny that similar truths might be discovered in other areas.

Certainly in geometry, there is no proposition for which neither the truth nor falsity can be established by the firmest reasoning. This since the more abstruse a truth may be, the less evident an approach to its proof should be, we would not be able to deny that arithmetic especially, which deals itself with the nature of numbers, contains the most abstruse truths of all.

And even some of the greatest mathematicians have fallen short, because they judge truths of this sort to be entirely unfruitful and for that reason unworthy, and have given no effort towards their investigation. And more than that, as knowledge of every truth is a worthy matter in itself, even those which seem unrelated to popular use; we have seen that all truths, at least those which we are able to understand, are so greatly connected with one another, that we cannot consider any of them altogether useless without some rashness.

And so, even if a certain proposition seems to be this way[11]*, so that regardless of whether it turns out to be true or false, it would be of no benefit to us anyway, still the method itself, by which we would establish its truth or falsity, nevertheless may be useful in opening up the way for us to discover other, more useful truths. For that reason, I firmly believe that I have not uselessly expended work and effort in investigating the proofs of these certain*

[11] Euler almost certainly had Fermat's last theorem and others like it in mind here

propositions. Hence, this theory of divisors does not lack any use, but rather may at some time show a utility in analysis that cannot be scorned.

This is the writing of a true scientist; unbiased, adventurous and bold, caring only for Truth itself, and nothing less. As mentioned, Euler was a very religious man, and a hard worker. In his mind he was reading the writings of God, and no other past-time could be more worthwhile than this.

The Theory of Divisors

Even years before proving Fermat's Little Theorem, Euler put the result to use in many novel ways, all relating to what he calls 'the theory of divisors'. One of the first results of this kind was in showing that the fifth Fermat prime, $2^{32}+1$, is not a prime at all. He proved this in 1732 (*"Observations of a certain theorem of Fermat and on others regarding prime numbers"* [3]), although he does not at this point give any indication to how he discovered it, only stating that *"I have observed after working for many days that this number can be divided by 641, as it will be immediately clear to anyone who tries"*. Not until a paper of 1747 (*"Theorems on divisors of numbers"* [6]), does he elucidate the matter. We will have a look at some of the theorems leading up to this result shortly, in order to get a clearer understanding and to see what the proof refers to. For now, the key to understanding the main result is contained in Theorem 8:

Theorem 8

29. The sum of two numbers $a^{2^m} + b^{2^m}$ whose exponent is a power of two admits no divisors other than those of the form $2^{m+1}n + 1$.

Proof.

In the manner that we proved that all divisors of $a^2 + b^2$ are of the form $4n + 1$, and from there showed that divisors of $a^4 + a^4$ are of the form $8n + 1$, and those of the form $a^8 + b^8$ are of the form $16n + 1$; so it can be shown in the same way that $a^{16} + b^{16}$ are of the form $32n + 1$. From here on, we may further understand that $a^{32} + b^{32}$, $a^{64} + b^{64}$, and so on, can have no divisors other than those of the form $64n + 1$, $128n + 1$, etc. And so in general it is evident that there are no divisors of $a^{2^m} + b^{2^m}$ other than those of the form $2^{m+1}n + 1$.

Corollary 2.

31. Therefore, one investigating the divisors of numbers of the form $a^{2^m} + b^{2^m}$ would be

uselessly wasting his effort, if he wanted to try division by any prime numbers except those expressible in the form $2^{m+1}n+1$.

Scholion.

32. Fermat had held, even though he had confessed that he frankly was unable to prove it, that all numbers of the form $2^{2^m}+1$ are prime; and then elsewhere tried to resolve a very difficult problem, which involved finding a prime number greater than a given number. But from the last theorem, it is clear that unless the number $2^{2^m}+1$, it has no divisors other than those of the form $2^{m+1}n+1$.

And so since I wanted to examine the truth of this renowned claim of Fermat for the case of $2^{32}+1$, I managed a huge shortening of this, by not having to try division by any prime number except those in the form $64n+1$. And so with the problem thus reduced to this, I soon discovered that by setting $n=10$, the prime number 641 is a divisor of the number $2^{32}+1$, and so the well-known problem asking for a prime number larger than a given number, still remains unsolved.

To understand *why* only numbers of the form $2^{m+1}n+1$ need be checked as factors of 2^{2^m}, we need to have a look at some of the results leading up to this.

THEOREMS ON DIVISORS OF NUMBERS

Theorem 1.

1. If p is a prime number, then every number of the form $(a+b)^p-a^p-b^p$ is divisible by p.

Demonstration.

[*The proof is similar in argument to the one from '36*]

Theorem 2.

4. If either a^p-a or b^p-b are divisible by a prime p, then $(a+b)^p-a-b$ is also divisible by the same prime p.

Corollary 2.

6. If we assume $a^p - a$ is divisible by p, then $(a+1)^p - a - 1$ is also divisible by p; and in the same way, $(a+2)^p - a - 2$, and moreover $(a+3)^p - a - 3$, etc., and in general $c^p - c$, is divisible by p.

Corollary 3.

10. If p is a prime number, then every number of the form $a^{p-1} - 1$ is divisible by p, except those cases, in which the number a itself is divisible by p.

This last assertion is of course Fermat's Little. Though the method of proof is almost identical to his 1736 paper (based on the Binomial Theorem), the presentation is much more structured.

The neat thing is that the fact that Fermat's fifth 'prime' number is composite, follows from theorems regarding divisors of sums of powers, the exact subject studied by Diophantus and Fermat before. In using Fermat's Little Theorem, Euler is able to finally set the record straight. The theorems alluded to are concerned with divisors of sums of squares and sums of certain higher exponents.

Theorem 4.

11. If neither a nor b is divisible by a prime number p, then every number of the form $a^{p-1} - b^{p-1}$ is divisible by p.

Theorem 5.

16. The sum of two squares $a^2 + b^2$ can never be divided by any prime number of the form $4n - 1$, unless both a and b are divisible by $4n - 1$.

Demonstration.

If $4n - 1$ is a prime number, and neither a nor b is divisible by it, then $a^{4n-2} - b^{4n-2}$ is divisible by $4n - 1$ (11), and hence $a^{4n-2} + b^{4n-2}$ is not divisible by $4n - 1$, and for that reason neither is any factor of it. And since $4n - 2 = 2(2n - 1)$ is an even number times an odd number, $a^{4n-2} + b^{4n-2}$ has $a^2 + b^2$ as a factor; since it cannot be the case that $a^2 + b^2$ is any sum of two squares divisible by $4n - 1$.

Scholion.

20. It is easy to see that a number of the form $4n - 1$ can never be the sum of two

squares. Squares are either even or odd: the former have form $4a$, the latter $4b + 1$. For this reason, for the sum of two squares to be odd, from which we have $4 + 4b + 1$ or $4n + 1$, and thus no number of the form $4n - 1$ can be the sum of two squares.

Next, that the sum of two squares also does not admit any divisor of the form $4n - 1$, has always been affirmed by all the authors of the Diophantine method: but no one at all, as far as I know, has proven it, except Fermat, who however never published a proof of the truth that no number either of the form $4n - 1$ or divisible by a number of the same form can ever be the sum of two squares.

So we finally see how Euler could know that $2^{32} + 1$ was not prime. After this, similar results to those asserted in Theorem 5. are shown for exponents $4, 8$, and finally 2^m, as we have seen in Theorem 8. Note that in proving Fermat's conjectures, Euler uses no methods or tricks not known to Fermat. This does strengthen Fermat's claims to proof somewhat.

Many more results, similar to these in nature, were proven by Euler in the period from the 1750's to early 1770's. Amongst these were many we have seen earlier; in 1760 he published the paper *"Proof of a theorem of Fermat that every number of the form 4n+1 can be given as the sum of two squares"* [8]. Another, unpublished, was *"A solution to a problem of Fermat, on two numbers of which the sum is a square and the sum of their squares is a biquadrate, inspired by the Illustrious La Grange"* [11]. Euler also seriously considered the four-square theorem, first stated by Diophantus, for over forty years. Although it was Lagrange who first published a proof, Lagrange's proof was based mainly on lemmas provided earlier by Euler, and soon after Lagrange had published, Euler contributed his own, essentially simpler proof in the paper *"Proof that every Integer is a Sum of Four Squares"* [10].

We shall now have a look at two more of Euler's papers, one regarding sums of divisors, and the other regarding remainders of powers.

Sums of Divisors

As Euler was interested in everything Fermat had to say, perfect numbers were to him of great interest. Not surprisingly, Euler's discoveries in this subject tower over both Euclid's and Fermat's, and are worthy of attention. Ever since Euclid, mathematicians had known that Euclid's formula gave even perfect numbers whenever $2^n - 1$ is prime. Euler was the first to show the opposite; that every even perfect number arises from Euclid's formula. But Euler did not stop there. There are different kinds of numbers, related to perfect numbers called amicable numbers - two numbers that are each the sum of the other's parts, or proper divisors. The Greeks knew of, and were awestruck by, the pair 220 and

284. Two more pairs were discovered in the 9th century by the Arab mathematician Thābit ibn Qurra. One of these pairs was rediscovered by Fermat in 1638, and the other by his nemesis Descartes, two years later. No more pairs were discovered until Euler's 1736 paper [5], where he finds 30 more such pairs of amicable numbers. By the time of his death he had found 58 pairs!

The reason why Euler could find these pairs was because he was a master of recognising patterns, and looking at problems differently than before. For example, on a number of occasions, Eulers was led to look at 'the sum of the divisors' of a number as a mathematical function, and to study the properties of this function. This would prove a suitable starting point for studying perfect and amicable numbers, but even more importantly it would unravel many more mysteries at the heart of number theory. Let us have a look at a paper that was written in 1755 titled *"Discovery of a most extraordinary law of numbers, relating to the sum of their divisors* [7], in which Euler uses this function heavily.

DISCOVERY OF A MOST EXTRAORDINARY LAW OF NUMBERS, RELATING TO THE SUM OF THEIR DIVISORS

1. Mathematicians have searched so far in vain to discover some order in the progression of prime numbers, and we have reason to believe that it is a mystery which the human mind will never be able to penetrate. To convince ourselves so, we have only to cast our eyes on the tables of prime numbers, which some have taken the trouble to continue beyond 100,000, and we will notice at once that neither rule nor order reigns. This situation is all the more surprising since arithmetic gives us unfailing rules, by means of which we can continue the progression of these numbers as far as we wish, without however leaving us the slightest trace of any order. I believe myself also to be rather far from this goal, but I have just discovered a very strange law among the sums of the divisors of natural numbers, which at first glance would appear as irregular as the progression of the prime numbers, and which even seems to encompass it. This rule, which I am going to expand upon, is in my opinion all the more important because it is the sort of truth we can persuade ourselves of, without giving a perfect proof. Nevertheless, I will put forth such evidence that we might almost be able to imagine it as equivalent to a rigorous proof.

2. [...] Since the following thoughts will revolve around the sum of the divisors of each number, I will use a certain character to indicate this. The letter $\int$, which one employs in infinite analysis to indicate integrals, when put in front of a number, will mean the sum of its divisors. So $\int 12$ will signify the sum of all the divisors of 12, which is 1+2+3+4+6+12 = 28, so that $\int 12 = 28$. That fixed, we will see that $\int 60 = 168$ and $\int 100 = 217$. But since unity has no divisor other than itself, we will have

$\int 1 = 1$. Since the number 0 is divisible by every number, the value of $\int 0$ will be infinite. However, in what follows I will assign to it, for each instance put forward, a definite value appropriate to my design.

[*In the next two sections Euler shows the important lemma that $\int$ is multiplicative, i.e. $\int ab = \int a \int b$, and goes on to list the hundred first values of $\int n$*]

I do not doubt that when one looks at the progression of these numbers, one would nearly lose hope of discovering the least order in it, because the irregularity of the sequence of prime numbers is intermixed with it in such a way that it would at first seem impossible to indicate any law in the progression of these numbers without knowing that of the prime numbers. It even seems that there is more strangeness here than in the prime numbers.

5. Nevertheless, I have noticed that this progression follows a quite regular law, and that it is even the kind of progression that the geometers call recursive, so that we can always form each term from those preceding it, according to a constant rule. For if $\int n$ denotes an arbitrary term in this irregular progression, and $\int(n-1)$, $\int(n-2)$, $\int(n-3)$, $\int(n-4)$, $\int(n-5)$, etc., the preceding terms, I say that the value of $\int n$ is always formed from the preceding terms by following this formula:

$$\begin{aligned}\int n = \int(n-1) + \int(n-2) - \int(n-5) - \int(n-7)\\ + \int(n-12) + \int(n-15) - \int(n-22) - \int(n-26)\\ + \int(n-35) + \int(n-40) - \int(n-51) - \int(n-57)\\ + \int(n-70) + \int(n-77) - \int(n-92) - \int(n-100)\\ +\text{etc.}\end{aligned}$$

[*Euler goes on to explain how to use this formula, and gives several demonstrations of this using very large numbers*]

8. These examples that I have just developed will no doubt remove any scruple which one could still have about the truth of my formula. But one could be all the more surprised by this nice property, not seeing any connection between the composition of my formula and the nature of the divisors, the sum of which the proposition centres upon. The progression of the numbers 1, 2, 5, 7, 12, etc. appears not only to have no relation to the subject in question, but, seeing that the law of these numbers is interrupted and that they are a mixture of two different regular progressions, that is

> $$1, 5, 12, 22, 35, 51, \text{etc.} \qquad \textit{and} \qquad 2, 7, 15, 26, 40, 57, \text{etc.},$$
>
> it almost seems that such an irregularity would not find a place in analysis. Furthermore, the lack of a proof must in no small way increase the interest in this, seeing that it would be almost morally impossible to arrive at the discovery of such a property, without having been led there by a sure method, which might be able to take the place of a perfect proof. [...]. And although this investigation centres only on the nature of numbers, to which infinite analysis would not seem to have any applicability, it is nevertheless by means of differentiation and other detours that I was led to this conclusion. I would hope that one could find a shorter and more natural way to get there, and perhaps consideration of the route I followed will lead to it.

In the next three sections of the paper, Euler further explains how he discovered the incredible recurrence relation for the values of $\int n$. The method is quite ingenious, and is based on manipulations of the infinite product

$$(1-x)(1-x^2)(1-x^3)\cdots .$$

Euler notes that when the product is expanded, the coefficients of the different powers form the sequence $1, 5, 12, 22, \ldots$. Upon inspection, it is seen that these numbers are in fact the pentagonal numbers, exactly the ones arising from setting $a = 5$, and letting n range over the natural numbers in Hypiscles a-gon formula (1). The numbers $2, 7, 15, 26, \ldots$ are called generalised pentagonal numbers and arise from setting n equal to successive negative integers in the same formula. Euler showed that these numbers too could be realised as the coefficients of an infinite product of polynomials. Euler then takes logarithms, differentiates and rewrites these expressions to connect them with the sums of divisors function. This quite remarkable connection is now called the Pentagonal Number Theorem, and his discovery would later instigate a new topic in number theory; the theory of *partitions*. In addition, Euler's completely new method of elegantly applying concepts of analysis to number theory, would ignite a whole new subfield of number theory - analytic number theory. This would later prove to be the most fruitful and important part of number theory in the 19th century, inspiring mathematicians such as Riemann, Dirichlet and Dedekind.

Residues of Powers

For the grand finale of our showcase of Euler's work in number theory, we present the paper containing his final proof of Fermat's Little Theorem. This proof is not based on the Binomial Theorem, but takes a completely new approach. The paper is from 1760 and bears the title *"Theorems on residues obtained by the division of powers"* [9]. The mathematics in this paper is both thorough and concise. Euler presents his findings almost as if they were a textbook, carefully proving a natural progression of results. We select a few along the way, and omit some straightforward proofs for brevity. Hopefully the reader will be able to fill in the gaps.

THEOREMS ON RESIDUES OBTAINED BY THE DIVISION OF POWERS

Theorem 1.

1. If p is a prime number and a is prime to p, then no term of the geometric progression $1, a, a^2, a^3, a^4, a^5, a^6$, etc. is divisible by the number p.

Demonstration

This follows from Book VII, Prop. 26 of Euclid, where it is demonstrated that if two numbers a and b are prime to p then too that their product ab is prime to p. Thus with a prime to p, by putting $b = a$, the square a^2 will be prime to p; and then in turn by putting $b = a^2$; likewise $b = a^3$, etc. Therefore no power of a will be divisible by the prime number p.

Corollary 1.

2. Therefore if each of the terms of the geometric progression

$$1, a, a^2, a^3, a^4, a^5, a^6, a^7, a^8, \text{etc.}$$

were divided by a prime number p, division never happens without a residue, but rather a residue arises from each term.

Scholion

3. I have resolved to carefully study the residues which emerge from the division of all the terms of the given geometric progression by the prime number p. First indeed, it is apparent from the nature of division that each of these residues will be less than the number p; also, no residue will be $= 0$, because no term is divisible by p. For if residues are produced which are greater than the number p, the way in which they may be reduced to less than it is clear from arithmetic. Thus the residue of $p + r$ equals the

residue r, and in general the residue $np + r$ returns the residue r; and if r is greater than p, this residue is reprised as $r - p$, or $r - 2p$, or $r - 3p$, etc., until a number less than p is reached. Therefore all the residues $r \pm np$ reflect the same residue as r. [...]

Corollary 2.

4. Because all the residues are integral numbers which are less than p, it follows that no more than $p - 1$ different residues can arise. Since the terms in the geometric series $1; a; a^2; a^3; a^4; a^5$; etc. consist of infinitely many different numbers, it is necessary that multiple terms exhibit the same residues.

The idea of considering only the $p - 1$ different residues instead of *all* the numbers in a given sequence, seems ancient, as it only depends on the Euclidean Algorithm, but put to use in this way the proof contains a subtle new idea. For example, in this way of thinking, *all* numbers that leave the remainder 3 when divided by 4 are melted down into one object, and if one can prove an assertion about this object, it will hold for all numbers of the form $4n + 3$. It is important therefore to understand the laws governing the algebra of these objects. Euler continues to exhibit simple properties of the residues of a given geometric sequence.

Theorem 2.

7. If the power a^μ divided by p has the residue r, and the power a^ν the residue s, the power $a^{\mu+\nu}$ will have the residue rs.

[*Many elementary facts flow from this result, and give rise to techniques for computing residues. Euler examplifies by finding the residue of* 7^{160} *when divided by* 641. *After this, things start to resemble Fermat's Little Theorem*]

Theorem 3.

12. If the number a is prime to p, and the geometric progression $1; a; a^2; a^3; a^4; a^5; a^6; a^7$; etc were formed, innumerably many terms occur in it which leave the residue 1 when divided by p, and the exponents of these terms constitute an arithmetic progression.

Theorem 7.

27. If a^λ is the minimum power of a which yields a residue = 1 when divided by the

number p, then all the residues which result from the terms of the geometric progression $1, a, a^2, a^3, ..., a^{\lambda-1}$, continued onto the power a^λ, will be mutually distinct.

Euler now carefully goes on to lead up to Theorem 13, which encapsulates exactly what Fermat wrote to Frénicle in his October, 1640 letter. He first argues that if a^λ is the minimum power which yields a residue 1 when divided by p, and λ is less than $p-1$, then it must also be less than or equal to $(p-1)/2$. The same argument is repeated, showing that if λ is less than $(p-1)/2$, then certainly it is less than or equal to $(p-1)/3$, etc. The argument consists of showing that we must be able to be split the terms of the geometric progression $1, a, a^2, \ldots, a^{p-1}$ into non-overlapping 'chains' of equal length[12].

Theorem 13.

48. If p were a prime number, and a^λ the minimum power of a which leaves unity when divided by p, the exponent λ will be a divisor of the number $p-1$.

Theorem 14.

49. If p were a prime number, and a prime to p, then the power a^{p-1} will leave unity when divided by p.

Scholion.

Behold therefore a new demonstration of the extraordinary theorem, found before by Fermat, which differs greatly from that which I gave in the Comment. Acad. Petropol. Volume VIII. For there I called upon the expansion of the binomial $(a+b)^n$ into a series by means of the method of Newton, which reasoning seems quite remote from the proposition; here indeed I have demonstrated the same theorem from the properties of powers alone, by which this demonstration seems much more natural. In addition, other important properties about the residues of powers when divided by prime numbers may appear to us. For it is clear that if p were a prime number, then not only will the formula $a^{p-1} - 1$ be divisible by p, but it will also sometimes happen that the simpler formula $a^{\lambda-1} - 1$ will be divisible by p, and then for the exponent λ to be an aliquot part of the exponent $p-1$.

The above proof, simple as it seems, was actually more difficult at the time because the technique of counting residues was entirely unfamiliar. Not only unfamiliar, it was an early breath of a new wind that began to blow over all of mathematics only in the time of

[12] For the reader aquatinted with modern abstract algebra, this is a version of Lagrange's theorem for cyclic groups. The non-overlapping 'chains' of powers are cosets of $(\mathbb{Z}/p\mathbb{Z})^\times$!

Gauss.

At this point in our story, it seems that most of Fermat's results have finally been rediscovered and explained, and many of the questions studied in the emerging field of number theory are starting to fit into a bigger picture.

6 Gauss

Abstraction in Mathematics

In the 18th century, introducing new entities into mathematics was entirely unheard of. Mathematics was the science of quantity, and each number represented a quantity. One could not invent a new number or a new quantity, one could only find out new facts about the old ones. In the 19th century however, many profound advances were made through this very technique. From then on, mathematics became the science of form and relation. That is, any abstraction, whether number or residue or anything else, which is related to other abstractions according to definite patterns is an object for mathematical study. One of the most important mathematicians we have to thank for the early development of this process of abstraction in mathematics is Carl Friedrich Gauss.

A Regular 17-gon

Johann Carl Friedrich Gauss was born in 1777 to a working class family in Brunswick, Germany. An extraordinary example of a child prodigy (he was reportedly looking after his father's accounts on a regular basis by the age of 5), his talents were quickly recognised by the Duke of Brunswick, who sent him to the Collegium Carolinum at 15, and then to the prestigious University of Göttingen which he attended from 1795 to 1798. During these years Gauss' spent almost all of his time engulfed in mathematical inquiries, although he was not at that time planning to pursue a mathematical career. At age 17, a remarkable discovery changed his mind. He had discovered a way to construct a regular 17-gon using a compass and a straightedge. Construction problems of this kind had been of great interest ever since the ancient Greek geometers. Not only did he solve an interesting old problem, he provided deep mathematical insights that initiated a new area of study combining geometry, analysis and number theory (remarkably, the *reason* why a 17-gon is constructible is that $17 = 2^{2^2} + 1$ is a Fermat prime). From this discovery onwards, Gauss was to become considered the greatest mathematician of his time.

While studying at Göttingen, many more results were either discovered or rediscovered by Gauss. The same year he proved that every positive integer is the sum of at most three triangular numbers. As he wrote in his diary, $EUREKA!\ num = \Delta + \Delta' + \Delta''$. Gauss was also the first person to discover the regularity in the distribution of primes now known as the Prime Number Theorem.

Although he made contributions in almost all fields of mathematics, and is equally well-known to physicists and astronomers as he is to mathematicians, number theory was always Gauss' favourite area of study, and he is famous to have said that *"mathematics is the queen of the sciences, and the theory of numbers is the queen of mathematics"*.

Few, But Ripe

As we have seen in the past chapters, much was known in the field of number theory at Gauss' time. However, the results of Fermat, Euler, Lagrange and others were scattered throughout countless 'memoires' of countless academies, and rigorous proof had always come second to tantalising conjectures. This would all change with Gauss. In publishing his work, Gauss followed the motto *"Pauca sed matura"*[13] and he would not publish a result until it was complete and he was entirely satisfied with its presentation. In this way, he set the standard for the modern rigorous approach to mathematics, much like Euclid alluded to two millennia earlier. He wanted the proofs in his writing to be above reproach; at one point he wrote to a friend; *"I mean the word proof not in the sense of lawyers, who set two half proofs equal to a whole one, but in the sense of mathematicians, where* $\frac{1}{2}$*proof = 0, and it is demanded for proof that every doubt becomes impossible"*. As an adverse consequence of his perfectionism, much of his work was left unpublished, and a considerable amount of work was discovered only after his death. Mathematical historian Eric Temple Bell proposes that if Gauss had published all of his discoveries in a timely manner, he would have advanced mathematics by fifty years.

Higher Arithmetic

Gauss' first published work was the masterpiece *Disquisitiones Arithmeticae*[13]. The book, finished in 1796 when Gauss was only 17 years old, was published in 1801. It was the

[13] "Few, but ripe"

first comprehensive textbook in number theory ever written in a strict, rigorous Euclidean format. A remarkable thing about Disquisitiones is the way it combines the thoroughness of a textbook with the excitement of new discovery. This is quite unusual in the mathematical literature. However, the writing style is terse, polished, and devoid of motivation. Abel said, *"He is like the fox, who effaces his tracks in the sand with his tail"*. Gauss, in defence of his style, said, *"no self-respecting architect leaves the scaffolding in place after completing the building"*, and also, in the preface to the Disquisitiones, *"In several difficult discussions I have used synthetic proofs and have suppressed the analysis which led to the results. This was necessitated by brevity..."*. Perhaps not the best excuse for secrecy, Mr. Gauss; Euler would almost certainly not have approved of such an approach, although Fermat is probably snickering in his grave.

In the first few pages of Disquisitiones Gauss recounts all the known results in number theory at that time, while throughout the rest of the book, he presents an astounding amount of new and original discoveries in the field. We quote from the preface so as to give an idea of the subject matter:

" *The inquiries which this volume will investigate pertain to that part of mathematics which concerns itself with integers. I will rarely refer to fractions and never to surds. The Analysis which is called indeterminate or Diophantine and which discusses the manner of selecting from an infinite set of solutions for an indeterminate problem those that are integral or at least rational (and especially with the added conditions that they be positive) is not the discipline to which I refer but rather a special part of it just as the art of reducing and solving equations (Algebra) is a special part of universal Analysis. And as we include under the heading analysis all discussion that involves quantity, so integers (and fractions in so far as they are determined by integers) constitute the proper object of ARITHMETIC. However what is commonly called Arithmetic hardly extends beyond the art of enumerating for example by a decimal representation, and carrying out arithmetical operations). It often includes some subjects which certainly do not pertain to Arithmetic (like theory of logarithms) and others which are common to all quantities. As a result it seems proper to call this subject Elementary Arithmetic and to distinguish from it Higher Arithmetic which properly includes more general inquiries concerning integers. We consider only Higher Arithmetic in the present Volume.*

Included under the heading "Higher Arithmetic" are those topics which Euclid treated with elegance and rigour in Book VII ff., and they can be considered an introduction to this science. The celebrated work of Diophantus, dedicated to the problem of indeterminateness, contains many results which excite a more than ordinary regard for the ingenuity and proficiency of the author because of their difficulty and the subtle devices he uses, especially if we consider the few tools that he had at hand for his work. However these problems demand a certain dexterity and skillful handling rather than profound principles and, because the questions are too specialised and rarely lead to more general conclusions, Diophantus' *book seems to fit into that epoch in the history of Mathematics when scientists were more concerned with creating a characteristic art and a formal Algebraic structure than with attempts to enrich Higher Arithmetic with new discoveries. The really profound discoveries are due to more recent authors like those men of immortal glory P. de Fermat, L. Euler, L. Lagrange, A.M. Legendre (and a few others). They opened the door to what is penetrable in this divine science and enriched it with enormous wealth.*

Congruence

We shall now survey some of the contents of Disquisitiones Arithmetica. We limit ourself to the first few sections, where Gauss sums up most of the ideas contained in our analysis so far. In the first few pages we see how well Gauss organises and explains the work of past mathematicians in a unified, comprehensive manner.

Section I

Congruent numbers in general

1. If a number a divides the difference of the numbers b and c, b and c are said to be congruent relative to a; if not, b and c are *noncongruent.* The number a is called a modulus. If the numbers b and c are congruent, each of them is called a *residue* of the other. If they are noncongruent they are called *nonresidues.*

2. Given a, all the residues modulo m are contained in the formula $a + km$ where k is any integer. The easier of the propositions that we will demonstrate follow directly from this, as will be clear to the reader.

Henceforth we shall designate congruence by the symbol $\equiv$, joining to it in parentheses the modulus when necessary to do so; e.g. $-7 \equiv 15 \pmod{11}, -16 \equiv 9 \pmod{5}$.

[*here Gauss explains the concept of* least residue*, i.e. the residue with absolute value less than half the modulus, just as Euler described in* [8]]

5. Having established these concepts, let us establish the properties that follow from them.

Numbers that are congruent relative to a composite modulus are also relative to any divisor of the modulus

If many numbers are congruent to the same number relative to the same modulus, they are congruent to one another (relative to the same modulus).

This identity of moduli is to be understood also in what follows.

Congruent numbers have the same least residues; noncongruent numbers have different least residues.

6. *Given the numbers A, B, C, etc. and any other numbers a, b, c, etc. congruent to each other relative to any modulus whatsoever, i.e. $A \equiv a, B \equiv b, C \equiv c$, etc. then $A + B + C+$ etc. $\equiv a + b + c+$ etc.*

7. *If $A \equiv a$, then also $kA \equiv ka$*

8. *Given any numbers whatsoever A, B, C, etc. and other numbers a, b, c, etc. which are congruent to them, i.e. $A \equiv a, B \equiv b$, etc., the product of each will be congruent; i.e. ABC etc. $\equiv abc$ etc.*

These articles are of course all well-known elementary facts of arithmetic. Gauss' stroke of genius lies here in two simple novelties - the introduction of a new *symbol* and a new *word*. The importance of a new mathematical symbol is easily overlooked, but one shall not underestimate the influence of Gauss' notion of *congruence*, as described by the symbol $\equiv$. Not only does it allow the writer to succinctly convey a complex idea, it allows the writer a new means of *thinking* about the subject. For Euler, it was a nuisance to always have to worry about whether a number was between 0 and n, the modulus, and always to write sentences like "p divides the difference of the two numbers a and b". With the notion of congruence this trouble is eliminated, as it ceases to matter which number among all the congruent numbers relative to some modulus which is chosen as a representative. For the first time, Gauss is explicitly treating whole collections of objects as single entities, with precise rules and operations for their manipulations. Today, *set theory* is regarded as the very foundation of mathematical ideas.

As we move forwards in Disquisitiones, we will meet many familiar results presented in this new language.

Section II

CONGRUENCES OF THE FIRST DEGREE

13. THEOREM. *The product of two positive numbers each of which is smaller than a given prime number cannot be divided by this prime number.*

14. *If neither a nor b can be divided by a prime number p, the product ab cannot be divided by p.*

Euclid has already proved this theorem in his *Elements* (Book VII, No. 32). However we did not wish to omit it because many modern authors have employed vague computations in place of proof or have neglected the theorem completely, and because by this very simple case we can more easily understand the nature of the method which

> will be used later for solving much more difficult problems.
>
> 16. THEOREM. *A composite number can be resolved into prime factors in only one way.*
>
> *Demonstration.* It is clear from elementary considerations that any composite number can be resolved into prime factors, but it is tacitly supposed and generally without proof that this cannot be done in many various ways. Let us suppose that a composite number $A = a^\alpha b^\beta c^\gamma$, etc., with a, b, c, etc. unequal prime numbers, can be resolved in still another way into prime factors. First it is clear that in this second system of factors there cannot appear any other primes except a, b, c, etc., since no other prime can divide A which is composed of these primes. Similarly in this second system of factors none of the prime numbers a, b, c, etc. can be missing, otherwise it would not divide A (preceding article). And so these two resolutions into factors can differ only in that in one of them some prime number appears more often than in the other. Let such a prime be p, which appears in one resolution m times and in the other n times, and let $m > n$. Now remove from each system the factor p, n times. As a result p will remain in one system $m - n$ times and will be missing entirely from the other. That is, we have two resolutions into factors of the number A/p^n. One of them does not contain the factor p, the other contains it $m - n$ times, contradicting what we have just shown.

Here we have the fundamental theorem of arithmetic presented for the first time in the modern form. In the next few articles Gauss puts the theorem to use proving important and useful properties about divisors of numbers. He then goes on to explain one of his main goals in number theory:

> 25. Any expression containing two congruent quantities in the manner of an equation will be called a *congruence.* If it involves an unknown, it is said to be *solved* when a value (*root*) is found satisfying the congruence. Hence it is clear what is meant by *solvable and unsolvable congruences.* Obviously distinctions similar to those used when speaking of equations can be used here. Examples of *transcendental* congruences will occur below; with regard to *algebraic* congruences they will be divided according to the highest power of the unknown into congruences of the first, second and higher *degrees.* Similarly, many congruences involving many unknowns can be proposed, and we can treat of their *elimination.*

So the main goal for Gauss in number theory is to solve congruences, and this he does with extraordinary generality and ingenuity. He starts off with some recap of old results, beginning with systems of congruences of the first degree. Such systems were well-known, and the result known as the Chinese Remainder Theorem - finding a number with given residues with respect to given moduli - had been studied before him by Chinese and Indian mathematicians, as well as more recently by Fermat, Euler and Lagrange. When this is done, he goes on to solve polynomial congruences. This problem had been treated earlier by Lagrange, but the full consequences of the main result had to wait for Gauss to be discovered. The main result is stated as article 43:

> 43. *A congruence of the m'th degree*
>
> $$Ax^m + Bx^{m-1} + Cx^{m-2} + \text{etc.} + Mx + N \equiv 0$$
>
> *whose modulus is a prime number p which does not divide A, cannot be solved in more than m different ways, that is, it cannot have more than m noncongruent roots relative to p.*
>
> [...]
>
> This theorem was first proposed and demonstrated by Lagrange. It appears also in the dissertation of Legendre, Euler showed that the congruence $x^n - 1 \equiv 0$ can have no more than n different roots, and although the result is particular, the method which this illustrious mathematician used is easily adaptable to all congruences.

As Gauss alludes to, finding the roots of unity relative to a given modulus is a problem of utmost importance. This is, as we have seen, closely connected with Fermat's Little Theorem, and was part of the subject matter of Euler's paper *"On residues obtained by the division of powers"*. Once again Gauss follows the breadcrumbs left by Euler, the next section of Disquisitiones being titled *"Residues of Powers"*; basically a recount of Euler's paper, including Fermat's Little Theorem. These facts being established, Gauss begins to push the boundaries of the present day understanding.

Primitive Roots

Years earlier, Euler had introduced the function $\phi(n)$ to stand for the number of numbers less than n, relatively prime to n. Euler had used this function to prove a generalisation of Fermat's Little Theorem amongst others. In the next excerpt Gauss uses this function to study what Euler had called *primitive roots.*

> 52. Suppose we are given numbers which are to be made congruent to unity by raising to a power. We know that for the exponent involved to be of the lowest degree it must be a divisor of $p - 1$. The question arises whether all the divisors of $p - 1$ enjoy this

property. And if we take all numbers not divisible by p and classify them according to the exponent (in the lowest degree) which makes them congruent to unity, how many are there for each exponent? We observe first that it is sufficient to consider all positive numbers from 1 to $p-1$; for, manifestly, numbers congruent to each other have to be raised to the same exponent as its least positive residue. Thus we must find out how in this respect the numbers $1, 2, 3, ..., p-1$ should be distributed among the individual factors of the number $p-1$. In brief, if d is one of the divisors of the number $p-1$ (among these 1 and $p-1$ itself must be included), we will designate by ψd the number of positive numbers less than p whose d'th power is the lowest one congruent to unity.

53. To make this easier to understand, we give an example. For $p = 19$ the numbers $1, 2, 3, ...18$ will be distributed among the divisors of 18 in the following way:

1	1
2	18
3	7, 11
6	8, 12
9	4, 5, 6, 9, 16, 17
18	2, 3, 10, 13, 14, 15

Thus, in this case $\psi 1 = 1, \psi 2 = 1, \psi 3 = 2\psi 6 = 2, \psi 9 = 6, \psi 18 = 6$. A little attention shows that with each exponent there are associated as many numbers as there are numbers relatively prime to the exponent not greater than it. In other words in this case $\psi d = \phi d$. Now we will show that this observation is true in general.

[...]

55. There is a particular case of the preceding proposition which merits special attention. *There always exist numbers with the property that no power less than the $p-1$st is congruent to unity,* and there are as many of them between 1 and $p-1$ as there are numbers less than $p-1$ and relatively prime to $p-1$.

57. Along with Euler we will call numbers belonging to the exponent $p-1$ *primitive roots*. Therefore if a is a primitive root the least residues of the powers $a, a^2, a^3, \ldots a^{p-1}$ will all be different. It is then easy to deduce that among them we will find all the numbers $1, 2, 3, ...p-1$, since each has the same number of elements. This means that any number not divisible by p is congruent to some power of a. This remarkable property is of great usefulness, and it can considerably reduce the arithmetic operations relative to congruences in much the same way that the introduction of logarithms reduces the operations in ordinary arithmetic. We will arbitrarily choose some primitive root a as a *base* to which we will refer all numbers not divisible by p. And if $a^e \equiv b \pmod{p}$, we will call e the *index* of b.

Gauss' treatment of the subject of primitive roots and the following discussions show not only how successfully he has completed Euler's goal of understanding the multiplicative structure of the integers, but also shows how Gauss is able to effectively penetrate the heart of the matter. In one sense, a primitive element modulo some number is *all we need.* Every number can be generated using only this primitive root. Such *primitive elements* are hugely important in the branch of modern algebra pertaining to *groups* - collections of objects subject to operations analogous to multiplication or addition. From a modern perspective it is safe to say that the theory of groups wholly underpins Gauss' treatment and way of thinking in number theory.

Profound Results

It is clear that in Gauss' hands, the field of Higher Arithmetic, or elementary number theory as we may now call it, finally matured into an esteemed field of research, worthy of popular attention. In many ways, the manner in which Gauss treats the subject is closer to the abstract modern way of thinking than to the treatment given by Euler and Fermat. For example, comparison of their respective treatment of certain quadratic expressions show just how far the idea of generality had come in half a century. Euler, having solved the problem $x^2 + y^2$, takes a fresh start with $x^2 + 2y^2$, and then $x^2 + 3y^2$ and so on. He would often publish half-solved problems, and return to them years later with progress. Gauss on the other hand understood that the essence of the problem was to find the laws governing *all* forms at once, and would not publish until he had worked out the full theory. His discussion was so complete that it left virtually nothing to be done.

In Section IV of Disquisitiones; *Congruences of the second degree*, he treats expressions of the form $Ax^2 + 2ABxy + By^2$, or *quadratic forms* as he calls them. Lagrange[14] was the first to systematically study these objects, but of course, Gauss really worked out the details. Through his treatment, several of Fermat and Euler's results on primitive hypotenuses, Pythagorean primes, etc. flow as simple consequences of more profound results. But for Gauss, that is all they are: simple consequences. The main theorem of section IV is what is now called the law of quadratic reciprocity, or as Gauss refers to it privately, *"the Golden Theorem"*. It states the conditions for the solvability of any quadratic equation modulo a prime number. Throughout his life he gave 8 different proofs of the theorem.

Past, Present and Future

Alas, our survey of Gauss' work in number theory comes to an end here, but that is not to say that Gauss stops here. In fact, Gauss was at this time just getting started,

[14] The illustrious Lagrange properly deserves a chapter of his own in the present text due to his many exciting original results in number theory. , however brevity has necessitated its suppression.

both within the context of Disquisitiones Arithmetica, and also in his mathematical career (remember, Gauss is 17 years old at this point!). Gauss did not publish very much in number theory for the rest of his career, but still managed to remain a central figure through the encouragement of his students and fellow mathematicians. Amongst these were the important 19th century number theoreticians Dedekind, Riemann, Dirichlet, Möbius, Kummer and many more.

So here we are at the beginning of the 19th century, Gauss standing as a leader uniting past, present and future mathematics, exposing to the world the work of the past, and at the same time leading the way for future research. After working out a new machinery for doing mathematics, old problems could be revisited and easily solved from the vantage point of generality and abstraction. And so it progressed and spread far beyond number theory. By the mid-19th century, set theory was being used to solve some of the most difficult problems of algebra, to shed new light on calculus and infinite series and also to understand the working of mathematics itself. That story will, however, have to wait for another time, new mathematicians, new symbols and new ideas. This story - the story of the origins and development of number theory, ends here with Gauss.

References

[1] Ronald S. Calinger and John Glaus. Leonhard Euler 1707-1783 Switzerland's Foremost Scientific Expatriate. `http://eulerarchive.maa.org/`, april 2015.

[2] William Dunham. A tribute to Euler. `https://www.youtube.com/watch?v=fEWj93XjON0`, april 2015.

[3] Leonard Euler. Observations on a certain theorem of Fermat and on others concerning prime numbers. *Comm. Acad. Petrop.*, 6:103–107, 1738.

[4] Leonard Euler. A proof of certain theorems regarding prime numbers. *Comm. Acad. Petrop.*, 8:141–146, 1741.

[5] Leonard Euler. On amicable numbers. *Nova acta eruditorum*, pages 267–269, 1747.

[6] Leonard Euler. Theorems on divisors of numbers. *Comm. Acad. Petrop.*, 1:20–48, 1750.

[7] Leonard Euler. Discovery of an extraordinary law of numbers in relation to the sum of their divisors. *Biblioteque impartiale*, 3:10–31, 1751.

[8] Leonard Euler. Proof of a theorem of Fermat that every number of the form 4n+1 can be given as the sum of two squares. *Comm. Acad. Petrop.*, 5:3–13, 1760.

[9] Leonard Euler. Theorems on residues obtained by the division of powers. *Comm. Acad. Petrop.*, 7:49–82, 1761.

[10] Leonard Euler. Proof that every integer is a sum of four squares. *Nova acta eruditorum*, pages 48–69, 1777.

[11] Leonard Euler. A solution to a problem of Fermat, on two numbers of which the sum is a square and the sum of their squares is a biquadrate, inspired by the illustrious La Grange. *Comm. Acad. Petrop.*, pages 3–6, 1826.

[12] Richard Friedberg. *An Adventurer's Guide to Number Theory.* McGraw-Hill Book Company, 1968.

[13] Carl Friedrich Gauss. *Disquisitiones Arithmeticae.* Yale University Press, 1965.

[14] Thomas L. Heath. *Diphantus of Alexandria, a study in the history of greek algebra.* Cambridge University Press, 1910.

[15] Luke Mastin. 19th century mathematics - Gauss. `http://www.storyofmathematics.com/19th_gauss.html`, april 2015.

[16] Chris May. Carl Friedrich Gauss. `http://users.wfu.edu/kuz/Stamps/Gauss/Gauss.html`, april 2015.

[17] J. J. O'Connor and E. F. Robertson. Leonhard euler. `http://www-history.mcs.st-and.ac.uk/Biographies/Euler.html`, may 2015.

[18] Paul Tannery and Charles Henry. *Oevres de Fermat*. Gauthier-Villars et fils, 1891.

[19] Euclid translated by Heath Thomas L. *The Thirteen Books of Euclid's Elements, Books 1 and 2*. Dover Publications, Incorporated, 1956.

[20] Norman J. Wildberger. MathsHistory 13: The Number Theory Revival. `https://www.youtube.com/watch?v=dh_4dn8FvCY`, mar 2015.